Optimal Policy for Behavioral Financial Crises*

Paul Fontanier⁺

November 21, 2022

Abstract

Should policymakers adapt their macroprudential and monetary policies when the financial sector is vulnerable to belief-driven boom-bust cycles? I develop a model in which financial intermediaries are subject to collateral constraints, and that features a general class of deviations from rational expectations. I show that distinguishing between the drivers of behavioral biases matters: when biases are a function of equilibrium asset prices, new externalities arise, even in models that do not have any room for policy in their rational benchmark. These effects are robust to the degree of sophistication of agents regarding their future biases. I show how time-varying leverage, investment and price regulations can achieve constrained efficiency. Finally, I show that greater uncertainty about the extent of behavioral biases in financial markets reinforces incentives for preventive action.

JEL Classification Numbers: D62, E44, E52, E70, G28

Keywords: Financial crises, Beliefs, Macroprudential policy, Monetary Policy

^{*}I am extremely grateful to my advisors Emmanuel Farhi, Sam Hanson, Andrei Shleifer, Jeremy Stein and Ludwig Straub for invaluable guidance and support. I also thank Nicholas Barberis, Francesca Bastianello, Jonathan Benchimol, Adrien Bilal, John Campbell, Gabriel Chodorow-Reich, Chris Clayton, Maria Fayos Herrera, Xavier Gabaix, Nicola Gennaioli, Robin Greenwood, Antoine Jacquet, Spencer Yongwook Kwon, Jean-Paul L'Huillier, Albert Marcet, Kenneth Rogoff, Alp Simsek, Larry Summers, Adi Sunderam, Jean Tirole, Luis Viceira, Jonathan Wallen, Michael Woodford, as well as participants at the 2019 Sloan-NOMIS Workshop on the Cognitive Foundations of Economic Behavior, the CU-RIDGE 2021 Financial Stability Workshop, the 2022 Theories & Methods in Macroeconomics Conference, the AQR Top Finance Graduate Award Conference at CBS, the Virtual Israel Macro Meeting, and numerous seminars for helpful comments. I am grateful to the Corps des Mines for supporting this research.

⁺Yale SOM. paul.fontanier@yale.edu & paulfontanier.github.io

1 Introduction

Should policymakers be concerned about asset price booms, and should they act preemptively before they burst? Historically the dominant paradigm among policymakers has relied on the idea that financial crises are "bolts from the sky," triggered by unpredictable and large negative shocks. Because private agents implicitly understand the riskiness of the activities they engage in, rapid growth in asset prices can only be supported by sound fundamentals and is not a cause for concern *per se*.¹ This contrasts sharply with the alternative, behavioral view of financial bubbles and crises that has been revived after the great financial crisis. Following in the footsteps of Minsky (1977) and Kindleberger (1978), researchers showed that factors such as credit growth and asset price booms successfully predict financial crises (Jordà, Schularick, and Taylor 2015). At the same time, survey data supports the idea that investors' beliefs are inconsistent with the Rational Expectations hypothesis, generally pointing towards the importance of extrapolation in financial markets (Gennaioli and Shleifer 2018).² In response, economists have developed a number of behavioral models of financial instability.³ Still, how policymakers should adapt their toolbox when financial instability is driven by systematic behavioral biases is largely an open question.

I tackle this question by constructing a model of financial crises in which agents display arbitrary deviations from rationality, and analyze optimal policy from the perspective of a social planner who recognizes that agents may have behavioral biases. I use this model to clarify three key normative questions surrounding the policy debate. First, which features of behavioral biases matter for welfare and should therefore be a concern for financial stability? Second, is there still room for intervention when the social planner and the market share the same beliefs, or when agents are sophisticated? And third, how should regulators incorporate incomplete information about behavioral biases when

¹This view has been articulated by, e.g. Gorton (2012) or Geithner (2014).

²Specifically, forecast errors made by market participants are reliably predictable ex ante, using for example forecast revisions as pioneered by Coibion and Gorodnichenko (2012).

³See Bordalo, Gennaioli, and Shleifer (2018), Greenwood, Hanson, and Jin (2019), Maxted (2020) and Krishnamurthy and Li (2020). These models are able to match moments that are inconsistent with rational frameworks, such as low credit spreads during the the run-up to financial crises.

contemplating early action?

I first show that welfare losses are driven by three key features of behavioral biases: (i) irrational optimism in booms if financial frictions might bind later on; (ii) future irrational pessimism during financial crises; and (iii) how equilibrium prices impact biases. I show how regulators can use leverage, investment, and price regulation (such as monetary policy) to improve welfare, and highlights which welfare effects are robust to the degree of sophistication of private agents and thus survive even if the planner shares the same beliefs as the market. Finally, I show that greater uncertainty about the precise extent of behavioral biases in financial markets increases the incentives for the planner to act early.

I present the model in Section 2. It features three periods and two types of agents: financial intermediaries and households. Financial intermediaries borrow by issuing deposits to households, and can invest in the creation of risky assets which can be thought of, e.g. as real estate or mortgage loans. At the heart of the model lies a financial friction: in the intermediate period, borrowing by intermediaries needs to be secured by posting these risky assets as collateral. The amount of borrowing available depends on the quantity of collateral available, and on the expectation of its future payoff. Such a friction, while keeping the economy away from the first-best, does not create any externality in a rational benchmark, and thus does not leave any room for policy.

The central element of the model is a general class of deviations from rationality in the formation of agents' expectations, which applies in all periods. I introduce a behavioral bias that shifts agents' perceived distribution of future dividends. The behavioral bias is allowed to depend on both fundamentals and asset prices. It is general enough to represent many psychological phenomena, while keeping the welfare analysis tractable. Crucially, being over-optimistic in booms regarding the prospects of the collateral asset is by implication being over-optimistic regarding the capacity of the financial sector to refinance itself. Behavioral biases in the asset market thus spread over the entire economy and distort all allocations. In order to study the robustness of my results, I allow agents to be potentially sophisticated: a parameter controls by how much agents realize that the market's future expectations will be biased.

I present the welfare analysis in Section 3, where a paternalistic social planner evaluates welfare using his own (rational) expectations. I start by fleshing out how behavioral factors and financial frictions interact to create uninternalized welfare effects. This analysis clarifies that irrational over-optimism in booms creates first-order welfare losses only when there is a chance that financial frictions bind in the future. Furthermore, it highlights how the predictable components of *future* behavioral biases formed inside a financial crisis also create losses and should be monitored when agents are not fully sophisticated. Indeed, if private agents tend to be over-pessimistic during financial crises, but neglect this future bias, they over-borrow in good times. If the social planner anticipates that future behavioral biases will be on the side of over-pessimism during an eventual financial crisis, there is a wedge between private expectations and those of the social planner. Here again, the interaction with financial frictions is crucial. Expected losses are greater when over-pessimism coincides with deeper financial crises: behavioral biases are tightening an already tight collateral constraint.

The welfare decomposition delivers a second key insight: precisely distinguishing between the drivers of these behavioral biases matters. When behavioral biases depend on current and past asset prices, new externalities arise. By borrowing and investing, agents influence the realization of current and future equilibrium prices, which can in turn alter the magnitude of behavioral biases. These effects, only present in the case of endogenous sentiment, are akin to pecuniary externalities but work through beliefs. For example, short-term borrowing lowers agents' net worth in a future crisis, which has a negative effect on future equilibrium prices. With endogenous sentiment such as price or return extrapolation, this fall in asset prices can trigger irrational pessimism, which tightens collateral constraints and deepens financial crises. Belief amplification thus creates an externality that calls for reducing leverage ex-ante: by increasing the net worth of intermediaries in a crisis, this policy supports asset prices, which in itself supports sentiment and thus relaxes the future collateral constraint.

I also uncover a novel effect, termed a *reversal externality*, that works through prices during the boom. When agents invest in risky assets in good times they bid up their prices. This can feed pessimism tomorrow by impacting the magnitude of behavioral biases in the future. For instance, if agents are simply extrapolating price changes, a high price in the past is a force that pushes agents towards irrational pessimism later. Hence an increase in prices today will cause a reversal tomorrow, which will tighten collateral constraints and prevent all intermediaries from rolling over their debt in a crisis.

Notably, these externalities can still be present even when private agents are fully sophisticated and the planner shares the same beliefs as the market. Even though financial intermediaries can be fully aware that the market will be irrationally panicking in a future financial crisis, their decisions are still privately optimal. Atomistic intermediaries cannot coordinate to collectively reduce their leverage or decrease asset prices in order to alleviate the effects of future pessimism. Only an intervention from the planner can solve these externalities, showing that naïvety or belief differences between policymakers and market participants are not key for these results.

How can one interpret these results of the model in terms of real-world policy? The tax on short-term borrowing can naturally be interpreted as capital structure regulation. If behavioral biases fluctuate along the business cycle, the optimal level of these restrictions is time-varying. My model thus calls for the use of counter-cyclical capital buffers.⁴ It shows that the time-variation should not only track the contemporaneous extent of over-optimism in financial markets, but should also consider how it will influence the *future* realizations of behavioral biases in eventual financial crises, as well as the expected impact of future prices on future biases. Similarly, to regulate the quantity of investment, regulators can rely on the implementation of Loan-to-Value (LTV) ratios. The optimal LTV limit should also be time-varying, and should closely track the same behavioral biases as do the counter-cyclical capital buffers. The presence of the reversal externality however calls for the use of a third instrment in order to control prices. I show that monetary policy can be used as a complementary tool: even when counter-cyclical capital buffers and LTV ratios can be flexibly adapted, an increase in the interest rate can be beneficial. By low-ering contemporaneous asset prices, monetary policy influences the future equilibrium

⁴Counter-cyclical capital buffers are at the center of the Basel III regulatory framework. My model shows how to optimally vary the levels of buffers when sentiment is fluctuating.

determination when sentiment is endogenous. The future price crash inside a financial crisis will be less severe, mitigating the reversal externality and relaxing collateral constraints. Such action does not require any information about biases in the boom phase, and is robust to the level of sophistication of agents. My model thus suggests that the concern for the central bank should not only be placed on whether prices are rational, but also on whether price booms will trigger further rounds of price extrapolation later on.⁵

It is however undeniable that identifying a bubble, or anticipating future pessimism, is intrinsically difficult since corresponding fundamentals are not observable. Indeed, the challenge for financial authorities of detecting contemporaneous irrationality in financial markets is a recurring argument from the advocates of the "wait-and-see" approach (Bernanke 2002). I acknowledge this issue but show that the intuition goes in the opposite direction. In Section 5, I allow the social planner to have an imprecise estimate of behavioral biases. The key result is that the strength of the desired ex-ante intervention on leverage is actually *increasing* in uncertainty. The more uncertainty there is about irrationality today, the more important it is to tighten leverage restrictions today. Intuitively, this is because sentiment interacts with financial frictions to create strong non-linearities: the costs of having intervened when it turns out that the price boom was entirely justified by sound fundamentals are dwarfed by the benefits of mitigating a possible sentimentdriven financial crisis. Similarly, I show that this intuition also goes through regarding price regulation in order to counter the reversal externality. Imagine that the regulator fears that high prices today could translate into over-pessimism in a future crisis, but is unsure of the strength of this mechanism. In that situation, the more uncertainty there is around this extrapolation force, the more the regulator wants to force lower prices in the boom.⁶

⁵An interesting example is the housing boom of the 2000s: while initial price increases in 2001-2003 may have been supported by fundamentals and low interest rates, it might have been the trigger for further irrational extrapolation down the road, resulting in adverse welfare consequences. If that is the case, my model suggests that an interest rate hike is warranted.

⁶Investment regulation, however, must be *relaxed* in the face of uncertainty, in order to encourage shifting resources to the crisis state.

Relation to the Literature: This paper is primarily motivated by the recent empirical evidence on credit cycles that revived the Minsky (1977) and Kindleberger (1978) narratives. This line of research started with Borio and Lowe (2002) showing that asset price growth and credit growth predict banking crises, stimulating research on the predictabililty of financial crises. Schularick and Taylor (2012) demonstrate that credit expansions forecast real activity slowdowns. Jordà et al. (2015) and Greenwood, Hanson, Shleifer, and Sørensen (2022) show that combining credit growth measures with asset price growth substantially increases the out-of-sample predictive power on a subsequent financial crisis. In a recent survey, Sufi and Taylor (2021) argue that "all told, the emerging historical evidence supports the existence of systematic behavioral biases in explaining credit cycles." Direct evidence of such biases comes from survey data: Bordalo et al. (2018) document the predictability of forecast errors regarding the Baa bond – Treasury credit spread.

My paper integrates these lessons into the literature on normative macro-finance. My framework follows from earlier work characterizing generic inefficiencies created by incomplete markets (Geanakoplos and Polemarchakis 1985; Greenwald and Stiglitz 1986). In my model, the amount of borrowing is limited by the expectation of the asset's future payoffs, a friction similar to Kiyotaki and Moore (1997). By contrast, most of the recent normative literature (as in Mendoza 2010, Bianchi 2011 and Jeanne and Korinek 2019) uses a collateral constraint that features instead the current price of the asset. This creates a pecuniary externality, since agents do not internalize how their ex-ante leverage decisions impact market prices tomorrow, and hence the aggregate borrowing capacity of the financial sector in the future. Dávila and Korinek (2018) offer a sharp analysis of this market failure.

There is also a vast literature showing that adaptive learning improve the fit of business cycles models to the data (see Gaspar, Smets, and Vestin (2010) for a survey), but such models usually deliver under-reaction. A notable exception is the work of Adam and Marcet (2011) and Adam, Marcet, and Beutel (2017), where subjective price dynamics lead to boom-bust dynamics in asset prices. Using this mechanism, Winkler (2020) builds a model where firms face financial constraints (leading to an amplification mechanism similar to the one in my paper) but does not study financial crises or optimal macroprudential and monetary policies. Caines and Winkler (2021) and Adam, Pfäuti, and Reinelt (2022) study monetary policy in the presence of subjective asset price dynamics also using learning models in standard New-Keynesian setups.

I end this section by focusing on the most closely related papers. First, Farhi and Werning (2020) analyze an environment with aggregate demand – rather than pecuniary – externalities, where agents extrapolate returns. Second, Dávila and Walther (2021) study an environment without financial frictions with general belief distortions during the boom, and characterize optimal leverage and monetary policies.⁷ Third, Caballero and Simsek (2020) study monetary policy when macroprudential policy is constrained, and agents have heterogenous beliefs.⁸ I build on the results of these three papers, and also complement them by showing how: (i) behavioral biases create powerful welfare effects even in a model without a market failure in its rational benchmark ; (ii) different types of biases lead to different forms of optimal intervention ; (iii) the eventual presence of pessimism in crisis is key ; (iv) the externalities created by biases are robust to the degree of sophistication of agents ; and importantly (v) uncertainty about the precise extent of biases in financial markets reinforce the motives for preventive intervention.

⁷I also contribute to this line of research by providing an alternative way of modeling general belief distortions. My proposal is simpler to use, especially for the welfare analysis, but at the cost of not being able to replicate the arbitrary distortions on the entire probability distribution used in Dávila and Walther (2021). For instance, Dávila and Walther (2021) can investigate how policy depends on whether agents are optimistic regarding left-tail or right-tail outcomes, a case my modeling choice cannot nest. However, it proves particularly convenient when I study the empirically relevant case where the social planner is uncertain about the precise extent of irrationality in financial markets.

⁸The proposal to use interest rate hikes to act early has been central to the policy debate on asset bubbles, even though it has often been resisted by policy makers (Greenspan 2002; Bernanke 2002). Bernanke and Gertler (2000) show in a conventional model that asset prices are relevant to monetary policy only to the extent that they may signal inflationary pressures. Woodford (2012) complements this analysis by demonstrating that if the probability of a financial crisis increases with the output gap, it may be necessary to conduct tighter policy. Gourio, Kashyap, and Sim (2018) study this problem quantitatively when the probability depends on the amount of inefficient credit. Barlevy (2022) shows that this may backfire if the boom is driven by a "speculation shock."

2 Model

This section presents the framework that will serve as the basis for the subsequent welfare analysis. The model is stylized in the tradition of the over-borrowing literature (Lorenzoni 2008), and financial intermediaries play a crucial role (He and Krishnamurthy 2013). To isolate the effects of behavioral biases, it features a borrowing constraint that does not create externalities in a rational equilibrium.⁹

2.1 Setup

Time is discrete, with three periods $t \in \{1, 2, 3\}$. There are two types of agents: financial intermediaries (or banks) and households. Both types are present in measure 1. There is a single good used both for consumption and for investment in the creation of a risky asset. The risky asset can only be held by financial intermediaries, and pays a stochastic dividend at times t = 2 and t = 3. The asset is also used as a collateral by financial intermediaries to issue deposits in period t = 2, and this constraint depends on the expectation of the future payoff of the asset. I define a "financial crisis" as a moment when the borrowing constraint of financial intermediaries binds at time t = 2.

Preferences: Bankers have log-utility in period t = 1 and t = 2, and linear utility in the last period:¹⁰

$$U^{b} = \mathbb{E}_{1}\left[\ln(c_{1}) + \beta \ln(c_{2}) + \beta^{2} c_{3}\right]$$
(1)

where c_t is the consumption of bankers at t, and β is the standard time discount factor. For simplicity, households (lenders) have linear utility throughout the three periods:

$$U^{h} = \mathbb{E}_{1} \left[c_{1}^{h} + \beta c_{2}^{h} + \beta^{2} c_{3}^{h} \right].$$

$$\tag{2}$$

⁹All my results go through with the same intuition when I perform the same analysis with a price-dependent collateral constraint that creates standard pecuniary externalities. See Online Appendix C and the discussion below.

¹⁰This functional form is adopted for simplicity. It allows for tractability in the equilibrium expressions during a financial crisis. Online Appendix H presents the analysis without linear utility in the ultimate period and with a general IES throughout. The results of the analysis and the intuitions are entirely similar, at the cost of unnecessary complexity.

Financial Assets: There are two financial assets in the economy: deposits and the risky asset. Financial intermediaries issue deposits $d_t \ge 0$ to households at time t, to finance their consumption and their investment in the risky asset. The price of the risky asset at time t is denoted by q_t . At time t = 1, financial intermediaries can create H units of the asset by paying a convex cost c(H).¹¹ The equilibrium price of the risky asset at t = 1, by no-arbitrage, is thus $q_1 = c'(H)$. This asset pays stochastic dividends z_2 and z_3 in future periods, drawn from independent cumulative probability distributions F_2 and F_3 , with support on \mathbb{R}^*_+ . Only intermediaries have the expertise required to hold risky assets.¹²

Financial Friction: At time t = 2, financial intermediaries face a collateral constraint: the amount they can borrow by issuing deposits must be secured by the risky asset, and is thus limited by the expectation of its future payoff. The collateral constraint takes the specific form:

$$d_2 \le \phi H \mathbb{E}_2[z_3] \tag{3}$$

where the parameter ϕ depends on the legal environment. The lower ϕ is, the less the bank is able to issue deposits to households in the intermediate period.

Here and in what follows, the expectation operator \mathbb{E}_t must not be interpreted as en objective (rational) expectation, but rather as agent's own subjective expectations. I specify belief distortions after laying out the rest of the model, in Section 2.2, but equation (3) makes clear that beliefs will alter the refinancing capacity of the financial sector.

I make one parametric assumption that guarantees that the equilibrium is not trivial.

Assumption 1. The financial friction parameter is small enough such that the collateral con-

¹¹Each intermediary has an indivisible real option to do business. Hence even faced with convex costs, an intermediary cannot multiply to operate at the infinitesimal scale.

¹²Assets thus never change hands in equilibrium. Although a rather stark assumption, it is consistent with He, Khang, and Krishnamurthy (2010), documenting that toxic MBS were always on the balance sheet of financial intermediaries during the 2008 financial crisis. Haddad and Muir (2021) provide further evidence suggesting that intermediaries are responsible for a large fraction of risk premium variation in various asset classes. This also allows me to sidestep "distributive" externalities (Dávila and Korinek 2018) that can lead to under- as well as over-borrowing in the rational benchmark.

straint is not always slack, and hence that financial crises are non-zero probability events:

$$\phi < \beta. \tag{4}$$

Constraints: Financial intermediaries' constraints are then as follows:

$$c_1 + c(H) + q_1 h_1 \le e_1 + d_1 + q_1 H \tag{5}$$

$$c_2 + d_1(1+r_1) + q_2h_2 \le d_2 + (z_2 + q_2)h_1 \tag{6}$$

$$c_3 + d_2(1+r_2) \le z_3 h_2 \tag{7}$$

$$d_2 \le \phi h_2 \mathbb{E}_2[z_3] \tag{8}$$

where *H* is the quantity of the asset created at t = 1, h_1 is the quantity intermediaries keep on their balance sheet at the end of period t = 1, and h_2 is the quantity of the risky asset held by financial intermediaries at time t = 2. In equilibrium, $h_1 = h_2 = H$ since households cannot hold the asset, and all intermediaries are identical. Financial intermediaries have an endowment e_1 in the initial period.

In order for financial intermediaries to always be able to repay their debt, I make the following parametric assumption, assumed to hold throughout the analysis.

Assumption 2. *The financial friction parameter is small enough such that:*

$$\phi \mathbb{E}_2[z_3] < \min z_3. \tag{9}$$

Throughout the paper, I make use of the marginal utility of consumption of financial intermediaries, $\lambda_t = 1/c_t$ in period 1 and 2, while $\lambda_3 = 1$ in the last period because of linear utility. A key object of interest, as in most models with financial frictions, is the *net worth* n_2 of financial intermediaries at t = 2, defined as:

$$n_2 = z_2 H - d_1 (1 + r_1). \tag{10}$$

Finally, the budget constraints of intermediaries are simply given by:

$$c_1^h - d_1 \le e_1^h \tag{11}$$

$$c_2^h - d_2 \le e_2^h + d_1(1 + r_1) \tag{12}$$

$$c_3^h \le e_3^h + d_2(1+r_2) \tag{13}$$

with the usual non-negativity constraints:

$$c_t^h \ge 0 \text{ and } c_t^h \ge 0, \text{ for } t \in \{1, 2, 3\}.$$
 (15)

We assume that households endowments e_t^h are large enough to guarantee positive consumption in every period.

Interpretation of the Environment: Financial intermediaries should be interpreted as levered financial institutions that are using short-term debt: commercial and investment banks, insurance companies, hedge funds, brokers, etc.

The risky asset is used as collateral for short-term debt by financial intermediaries. A favored interpretation is that *H* represents Mortgage-Backed Securities (MBS), complex products widely used in repo markets. In this case, the costs c(H) can be interpreted as securitization costs (e.g. search costs, legal fees, or the wages of structured traders).¹³ When concerns about the future value of these assets arise, collateral values fall, forcing the banking system to cut back on other activities in order to roll-over its short-term debt. The model thus seeks to capture a typical "run on repo" as the panic of 2007-2008 (Gorton and Metrick 2012). In the baseline model, banks against their collateral constraints can only reduce their consumption. In Online Appendix B I extend the model to allow for real production: intermediaries lend money to firms subject to pay a wage-in-advance constraint. Naturally, banks against their collateral constraint also have to reduce their

¹³Alternatively, one can picture intermediaries as firm/bank coalitions, where *H* represent C&I loans or projects funded by the intermediaries. *H* may also represent real estate held by the financial sector: the dividends are then simply rents coming from these operations.

lending activity to the real sector, which creates a real recession. The welfare results are identical in this case, only more involved, which is why I focus on the simpler case. But a "financial crisis" is meant to capture moments where intermediaries are under stress and cannot fulfill their role.

Remark 1 (Microfoundations of the Collateral Constraint). The specification of the collateral constraint in equation (3) can be obtained from the following microfoundations:

- i) Financial intermediaries lack commitment to repay in the final period ;
- ii) Financial intermediaries must take the decision of whether to default before observing the realization of z_3 ;
- iii) In the event of default, lenders can seize a fraction ϕ of the asset held by intermediaries.

These frictions lead lenders to only be willing to lend up to a fraction of the average future payoffs of the risky asset.¹⁴ While also realistic, this form of the collateral constraint allows me to fully isolate the effects of behavioral biases on welfare. Despite the presence of financial frictions, the equilibrium is constrained-efficient when expectations are rational.

A large part of the normative macro-finance literature, for this reason, uses an alternative formulation for financial frictions to obtain pecuniary externalities. Dávila and Korinek (2018) show that a *collateral externality* arises when the collateral constraint depends on the *current* price of the asset, as in:

$$d_2 \le \phi H q_2. \tag{16}$$

This type of collateral constraint is used for example by Farhi and Werning (2016), Bianchi and Mendoza (2018), and Jeanne and Korinek (2019). It can be microfounded by assuming that avoiding repayment requires diverting resources in the current period, and this is

¹⁴The model could perfectly be written with a collateral constraint of the form $d_2 < \phi H \min z_3$. This would relax the second assumption made for the micro-foundations: borrowers could default after observing the realization of z_3 . My conclusions would be unchanged because the behavioral bias in my model shifts min z_3 exactly as it shifts $\mathbb{E}_2[z_3]$. I keep the $\mathbb{E}_2[z_3]$ formulation as expressions are cleaner this way.

perfectly observed by lenders. Ottonello, Perez, and Varraso (2022) show that the quantitative predictions of both types of constraint are similar. Without taking a stance on which one is more realistic, I focus on the future payoff constraint in equation (8) since it cleanly isolates the effects of behavioral biases, and show in Online Appendix C the robustness of the results to this alternative formulation.¹⁵

2.2 Beliefs

Households and Intermediaries have homogenous beliefs. I allow for a class of deviations from rationality that strikes a balance between generality and tractability, as well as providing several different insights.

Recall that dividends are drawn from (objective) independent cumulative probability distributions: $z_2 \sim F_2(z)$ and $z_3 \sim F_3(z)$, with support on \mathbb{R}^*_+ . At time *t*, the subjective expectations of agents regarding period-*t* + 1 dividends are encoded in a scalar Ω_{t+1} .

Definition 1 (Behavioral Bias in Expectations). *A behavioral bias at time t is a function* Ω_{t+1} *that can depend on fundamental or prices:*

$$\Omega_{t+1}: \mathcal{I}_t = \{z_{t-i}, q_{t-i}\}_{i>0} \to \mathbb{R}$$

$$\tag{17}$$

At time t, agents believe that dividends paid in the next period are drawn from the cumulative distribution $F_{t+1}(z - \Omega_{t+1}(\mathcal{I}_t))$. They thus believe that the dividend paid in each state of the world will be $z_{t+1} + \Omega_{t+1}(\mathcal{I}_t)$ rather than z_{t+1} .

In other words, the behavioral bias Ω_{t+1} is a location shifter on expected dividends. In that respect, Ω_{t+1} exactly represents the *predictable component* at *t* of forecast errors realized at t + 1.¹⁶ A positive bias Ω_{t+1} means that agents are over-optimistic at time *t* regarding the prospects of dividends in the future. In this case, sentiment will be said to be high, or equivalently that markets are displaying "irrational exuberance" (Shiller

¹⁵Results are also similar if intermediaries are subject to a constraint in the initial period.

¹⁶See Cieslak (2018), Bordalo et al. (2018), or Ma, Ropele, Sraer, and Thesmar (2020) for empirical evidence on this predictability.

2015). A negative bias Ω_{t+1} means that agents are over-pessimistic at *t*. In this case, sentiment will be said to be low, or equivalently that markets are displaying "irrational distress" (Fisher 1932).

The generality of the approach comes from allowing the Ω biases to depend on several variables (z_{t-i} or q_{t-i}). This approach is particularly flexible for the subsequent welfare analysis, since it summarizes all possible distortions in a single quantity. The functional dependence on \mathcal{I}_t is, however, assumed to be independent of policy.¹⁷ Throughout the paper the biases Ω are kept general, highlighting the properties of sentiment that matter for welfare. It will be useful to flesh out specific examples to build intuition, however. I will focus on two functional forms that are common in the behavioral finance literature, and have been used to explain the credit cycle facts I reviewed in the introduction.¹⁸

Fundamental Extrapolation: This case captures models where investors extrapolate fundamentals, which here are $\{z_t\}$. Several influential papers use this class of models to explain a wide range of facts about asset prices, starting with Barberis, Shleifer, and Vishny (1998). This extrapolation can come from a variety of psychologically founded biases.¹⁹ I model fundamental extrapolation in reduced-form as:

$$\Omega_{t+1} = \alpha_z (z_t - z_{t-1}) \tag{18}$$

where α_z is a positive number. Because there is no fundamental realization at t = 1 or before, I assume that there are hypothetic values z_1 and z_0 driving initial sentiment. The bias at t = 1 about next period's payoff will thus be $\Omega_2 = \alpha_z(z_1 - z_0)$, while the bias in the intermediate period will be given by $\Omega_3 = \alpha_z(z_2 - z_1)$. A boom-bust cycle in the spirit

¹⁷I could also allow Ω to depend on other equilibrium variables, like debt or investment. This does not bring additional insight. The case where biases depend on policy could yield further insight and is left for future research.

¹⁸A particularly clear survey of this literature can be found in Barberis (2018). While the core of the paper focuses on these two cases, other behavioral models are nested by the Ω -formulation, such as sticky beliefs. I present and discuss several cases in Online Appendix F.

¹⁹Constraints on memory and cognition can make it difficult for agents to work with complicated models, as in Fuster, Hebert, and Laibson (2012), leading agents to excessively use recent data points. Bordalo et al. (2018) link extrapolative beliefs about fundamentals to the representativeness heuristic of Kahneman and Tversky (1972). In Rabin and Vayanos (2010), extrapolative beliefs stem from believing in the law of small numbers.

of Gennaioli and Shleifer (2018) is thus represented by fundamental realizations $z_1 > z_0$ (good news at t = 1) followed by $z_2 < z_1$.

Price/Return Extrapolation: While price extrapolation is aimed at explaining the same set of facts as fundamental extrapolation, it can have drastically different implications, and in particular in terms of policy as this paper will show. Early examples include models by DeLong, Shleifer, Summers, and Waldmann (1990), Hong and Stein (1999) and Barberis and Shleifer (2003).²⁰ Close to this paper, Farhi and Werning (2020) use return extrapolation in a model with aggregate demand externalities to study macroprudential and monetary policy. In the present paper, price extrapolation is modeled in reduced-form as:

$$\Omega_{t+1} = \alpha_q (q_t - q_{t-1}) \tag{19}$$

where, in period t = 1, we will postulate the existence of a hypothetic price q_0 that prevailed in the past and anchors agents' expectations. Crucially, agents' present and *future* beliefs now move with policies that influence asset prices (a potential channel for monetary policy, as studied in Section 4.3).²¹

Remark 2 (Agents' Beliefs and Collateral Constraint). I assumed above that all agents (borrowers and lenders) share the same behavioral biases Ω throughout the paper. In the micro-foundations of the collateral constraint proposed above, households are lending up to a limit where they think intermediaries would default. Hence, what matters in equation (3) are the beliefs of lenders, rather than borrowers. But since beliefs are homogenous,

²⁰Recent research leverages the use of survey data to motivate price or return extrapolation: McCarthy and Hillenbrand (2021) estimate that return extrapolation accounts for 23% of movements in the S&P500 index. Price and return extrapolation have been used by Barberis, Greenwood, Jin, and Shleifer (2018) to present a model of financial bubbles, while DeFusco, Nathanson, and Zwick (2017) apply it to the housing market.

²¹Notice that the bias in the baseline version of my model is always modeled as a shift in the subjective distribution of *dividends*, not *prices*. This is a reduced-form assumption, and such a bias appears for example when agents are learning from prices as in Chahrour and Gaballo (2021) or Bastianello and Fontanier (2022). Subjective prices are then also distorted since agents form expectations in a consistent manner, see the paragraph after equation (21). Online Appendix F.3 also discusses the case where the behavioral bias is distorting expected *prices*, while expectations of fundamentals stay rational (as in, e.g. Adam and Marcet 2011). In that case, some externalities survive only when the collateral constraint depends on prices, rather than expected payoffs.

the tightness of the collateral constraint will nevertheless depend on $\mathbb{E}_2[z_3 + \Omega_3]$.²²

2.3 Sophistication

Crucially for my results, agents are allowed to be biased in the initial period t = 1 as well as in the intermediate (crisis) period t = 2. It thus begs the question of whether agents realize that they, or the market, might be biased in the future. To flexibly show how my results change with sophistication or naïvety, I introduce a second parameter that controls the level of sophistication of agents, ζ :

Definition 2 (Sophistication Parametrization). *The parameter* $\zeta \in \mathbb{R}$ *captures the level of sophistication of agents. At time* t = 1, *agents have expectations regarding their future expectations such that:*

$$\mathbb{E}_1\left[\mathbb{E}_2[z_3]\right] = \mathbb{E}_1[z_3 + \zeta \Omega_3]. \tag{20}$$

When $\zeta = 0$, agents expect their future selves to have unbiased expectations regarding future dividends (*naïvety*). When $\zeta = 1$, agents understand that their future selves will have expectations biased by Ω_3 .²³ Agents are partially sophisticated when $\zeta \in]0, 1[$.

2.4 Equilibrium

Before formally defining the equilibrium concept used in the paper, it is informative to understand in greater details how agents form expectations about future variables given the specific form of biases I assumed in Definitions 1 and 2. One way to gain intuition is to inspect the usual pricing equation of the risky asset at *t*:

$$q_{t} = \beta \mathbb{E}_{t} \left[\frac{\lambda_{t+1}(z_{t+1} + \Omega_{t+1}, \zeta \Omega_{t+2})}{\lambda_{t}} \left(z_{t+1} + \Omega_{t+1} + q_{t+1} \left(z_{t+1} + \Omega_{t+1}, \zeta \Omega_{t+2} \right) \right) \right]$$
(21)

²²It would differ, however, in a model where the collateral constraint depends on the equilibrium price q_2 . Since intermediaries are the marginal pricers of the asset, the beliefs of borrowers would determine the tightness of the constraint in a crisis (see Simsek (2013) for an in-depth analysis). For simplicity beliefs are assumed to be homogenous in my model.

 $^{^{23}\}zeta$ thus measures the degree of inconsistency in the Law of Iterated Expectations. Implicitly, this formulation assumes that sophisticated agents understand how Ω_3 will be determined according to information available at t = 2.

where $q_{t+1}(z_{t+1} + \Omega_{t+1}, \zeta\Omega_{t+2})$ is the price that would prevail, at t + 1, in a environment where the state of the world realizes at $z_{t+1} + \Omega_{t+1}$, and future agents will price the asset believing that at t + 2 it will pay $z_{t+2} + \zeta\Omega_{t+2}$ rather than z_{t+2} . Similarly, agents can potentially realize that their marginal utility λ_{t+1} will be determined by biased agents: agents form marginal utility expectations taking into account that their future selves will receive dividends $z_{t+1} + \Omega_{t+1}$. But they also realize that, if they are constrained next period according to (3), their borrowing will be limited by $\mathbb{E}_{t+1}[z_{t+2} + \zeta\Omega_{t+2}]$. For instance, picture a fully sophisticated agent ($\zeta = 1$) that is over-optimistic today ($\Omega_{t+1} > 0$) but realize that everyone will be over-pessimistic tomorrow ($\Omega_{t+2} < 0$ some states of the world). This agent think that the asset is going to pay more than in reality. But the agent realize that his consumption will be more severely impacted in a crisis because of a tight collateral constraint caused by excessive pessimism.²⁴

Throughout the paper, I use a streamlined notation:

$$q_t = \beta \mathbb{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t} \left(z_{t+1} + \Omega_{t+1} + q_{t+1} \right) \right]$$
(22)

where the dependence of the stochastic discount factor λ_{t+1}/λ_t and of the price on the behavioral biases are kept implicit, for conciseness. We now proceed to formally define the equilibrium, and analyze it in more details in the rest of this section.

Definition 3 (Equilibrium Definition). Given the behavioral biases functions Ω_2 and Ω_3 , a sophistication parameter ζ , as well as initial values z_0 , z_1 , and q_0 , an equilibrium consists of a real allocation $\{c_1, c_2(z_2), c_2(z_2, z_3), c_1^h, c_2^h(z_2), c_2^h(z_2, z_3), H\}$, prices $\{q_1, q_2(z_2)\}$, and biases $\{\Omega_2(\mathcal{I}_1), \Omega_3(\mathcal{I}_2(z_2))\}$, such that: (i) markets clear; (ii) agents maximize (1) and (2) subject to (5)-(8) and (11)-(13); and (iii) biases are consistent with definitions 1 and 2.

2.5 Equilibrium Analysis

I solve for the equilibrium by backward induction, starting from the intermediate period.

²⁴Because what matters for agents is how tight the collateral constraint will be, and how the asset will be priced, their sophistication ultimately needs to be about whether the *market* will be biased, rather than just themselves.

Households: Households are passive throughout the three periods, and they pin down the rate of interest through their Euler equation:

$$\beta(1+r_t) = 1 \tag{23}$$

Financial Intermediaries at t = 2: Entering period t = 2 with a stock H of collateral assets, and debt d_1 to repay, financial intermediaries must decide on their borrowing and consumption levels. There are two separate cases.

No Crisis: When financial intermediaries are not constrained, their Euler equation simply sets consumption such that:

$$\lambda_2 = \frac{1}{c_2} = \mathbb{E}_2[\lambda_3] = 1 \tag{24}$$

because of the linearity of utility in the last period. The consumption level is thus independent of the price of the risky asset, and consequently of any behavioral bias. Finally the price of the collateral asset is simply given by:

$$q_2 = \beta \mathbb{E}_2[z_3 + \Omega_3] \tag{25}$$

Financial Crisis: When the collateral constraint is binding, the associated Lagrange multiplier on the constraint, κ_2 , is given by:

$$\kappa_2 = \lambda_2 - 1 > 0. \tag{26}$$

This directly quantifies the severity of the crisis: it encodes how far we are from the unconstrained equilibrium. The asset price comes from intermediaries' maximization which yields:

$$q_2 = \beta c_2 \mathbb{E}_2[z_3 + \Omega_3] + \phi(1 - c_2) \mathbb{E}_2[z_3 + \Omega_3]$$
(27)

where the last term is a *collateral premium* (holding marginally more of the asset is valuable since it relaxes financial constraints). Consumption, on the other hand, is directly coming from the budget constraint of financial intermediaries (6), since agents are against the

collateral constraint:

$$c_2 = z_2 H - d_1 (1 + r_1) + \phi H \mathbb{E}_2 [z_3 + \Omega_3].$$
(28)

This last expression makes clear that, unlike in the unconstrained case, behavioral biases have direct effects on real allocations in crises. Pessimism ($\Omega_3 < 0$) reduces the amount households are willing to lend to financial intermediaries, leading to a one-for-one fall in their consumption level c_2 . I distinguish the cases when Ω is exogenous or endogenous to clarify their differences, which will be crucial for welfare.

Exogenous Bias: When Ω_3 is exogenously set, the budget constraint equation is sufficient to obtain the consumption level in a crisis (as in a rational benchmark). Sentiment simply shifts consumption by a constant relative to the REE benchmark. It also has an effect on asset prices through the stochastic discount factor, c_2 , and the expectation of future dividends. But this drop in asset prices does not spill back to consumption. The left panel of Figure 1 illustrates this equilibrium determination when $\Omega_3 < 0$ with the solid line, while the rational case is depicted by the dotted line. Exogenous pessimism makes the pricing condition steeper, but consumption is pinned down independently.

Endogenous Bias: When the behavioral bias Ω_3 depends on equilibrium prices q_2 , however, the budget constraint is not enough anymore to determine the consumption level of financial intermediaries in a crisis. The equilibrium now requires solving for a fixed-point between the budget constraint and the pricing equation. This process is represented on the right panel of Figure 1. It illustrates the presence of a new feature that I call *belief amplification*.²⁵ Intuitively, a fall in net worth causes a fall in current consumption. This decreases the stochastic discount factor used by agents to price the risky asset, which in itself creates endogenous pessimism. This leads the price of the asset to fall further, which tightens the borrowing constraints of financial intermediaries by aggravating pessimism, and in turn creates a further fall in the price that leads to more pessimism. The

²⁵In a setup where the collateral constraint depends on current prices q_2 , this belief amplification channel compounds the traditional *financial amplification* mechanism. See Online Appendix C.

arrow on Figure 1 illustrates the further contraction in consumption c_2 due to this belief amplification.²⁶

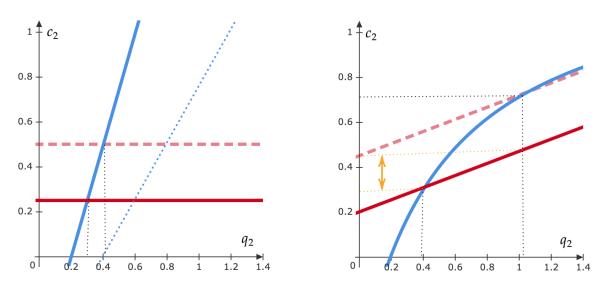


Figure 1: Graphical Illustration of Equilibrium Determination with exogenous sentiment at t = 2. The red line represents the budget constraint equation (28), and the blue line represents the pricing equation (27). The dotted line represents the rational pricing equation. The left panel illustrates how the equilibrium shifts after an exogenous shock to net worth n_2 (dashed line to solid line) when agents have an exogenous bias $\Omega_3 < 0$, and the blue dotted line depicts the case where $\Omega_3 = 0$. The right panel illustrates the same experiment but with an endogenous $\Omega_3(q_2) = \alpha(q_2 - q_1)$.

Financial Intermediaries at t = 1: The consumption Euler equation for financial intermediaries in the initial period is simply given by:

$$1 = \mathbb{E}_1 \left[\frac{\lambda_2}{\lambda_1} \right] \tag{29}$$

since financial intermediaries and households have the same time-preference parameter β . Collateral creation is driven by the pricing equation of intermediaries:

$$q_1 = c'(H) = \beta \mathbb{E}_1 \left[\frac{\lambda_2}{\lambda_1} (z_2 + \Omega_2 + q_2) \right].$$
(30)

²⁶As can be seen in Figure 1, the equilibrium is ensured to be unique in the exogenous sentiment case. This is not immediate anymore for endogenous sentiment. Online Appendix G shows how linear forms of price extrapolation also guarantee the uniqueness of a stable equilibrium. Complex non-linear forms of endogenous sentiment can however lead to multiple equilibria. Since this is not the focus of this analysis, for the rest of the paper I assume that belief distortions are not strong enough such that equilibrium uniqueness is guaranteed.

Because consumption inside a crisis, c_2 , depends directly on z_2 , agents with an optimistic bias $\Omega_2 > 0$ expect their future consumption to be higher than in reality. Accordingly, their Euler equation directly implies that an optimistic bias translates into overconsumption at t = 1 relative to the rational benchmark, financed by borrowing (so a higher d_1). Similarly, the gap between expected consumption and the realized one is driven, for the case of unsophisticated agents, by *future* sentiment, since an $\Omega_3 < 0$ at t = 2 leads to a tighter collateral constraint.

3 Welfare Decomposition

The paternalistic social planner evaluates welfare with two key distinctions relative to atomistic behavioral agents:

- 1. The social planner takes general equilibrium effects into account ;
- 2. The social planner uses its own expectations and knows that private agents are subject to a bias Ω_2 , and that they can be subject to a bias Ω_3 in the future.

I adopt the notation \mathbb{E}^{SP} to denote expectations formed according to this process, and use \mathbb{E} for the expectations of private agents.²⁷

3.1 Decomposition

One contribution of this paper is to precisely identify how behavioral biases impact welfare. I present a general decomposition in the spirit of Dávila and Korinek (2018), that fleshes out how a marginal increase in leverage or in investment leads to uninternalized welfare consequences, and classify the different channels. A key advantage of this approach is that the decomposition naturally determines which features of behavioral biases matter for financial stability, and need to be quantified by the regulator.

²⁷A non-paternalistic planner would have the same first-order conditions as private agents since there is no pecuniary externality.

I start by analyzing how changes in debt d_1 , investment H, and prices q_1 , each fixing all others variables at t = 1, affect the welfare of individual agents.

Proposition 1 (Uninternalized Effects). *The uninternalized first-order impact on welfare, when infinitesimally varying one aggregate variable while keeping the others constant, are given by:*

i) For short-term debt d_1 :

$$\mathcal{W}_{d} = \underbrace{\left(\mathbb{E}_{1}[\lambda_{2}] - \mathbb{E}_{1}^{SP}[\lambda_{2}]\right)}_{\mathcal{B}_{d}} - \underbrace{\mathbb{E}_{1}^{SP}\left[\kappa_{2}\phi H\frac{d\Omega_{3}}{dq_{2}}\frac{dq_{2}}{dn_{2}}\right]}_{\mathcal{C}_{d}}; \quad (31)$$

ii) For investment in collateral assets H:

$$\mathcal{W}_{H} = \underbrace{\left(\beta \mathbb{E}_{1}^{SP} \left[\lambda_{2}(z_{2}+q_{2})\right] - \lambda_{1}q_{1}\right)}_{\mathcal{B}_{H}} + \underbrace{\beta \mathbb{E}_{1}^{SP} \left[\kappa_{2}\phi H \frac{d\Omega_{3}}{dq_{2}} \left(\frac{dq_{2}}{dn_{2}}z_{2} + \frac{dq_{2}}{dH}\right)\right]}_{\mathcal{C}_{H}};$$
(32)

iii) And for prices q_1 :

$$\mathcal{W}_q = \beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{d\Omega_3}{dq_1} \right].$$
(33)

Proof. All proofs are provided in Online Appendix A.

Two classes of effects appear in this Proposition, which I now explore in turns.

3.2 Behavioral Wedges

The first type of effects (the terms \mathcal{B}_d and \mathcal{B}_H of equations 31 and 32) are *behavioral wedges*, as in Farhi and Gabaix (2020). They quantify *differences in beliefs* between the planner and the market. Take the behavioral wedge for leverage for instance. The strength of the behavioral wedge is driven by the difference in the expected severity of crisis. Indeed, because of the linearity of utility in the last period, the marginal utility of financial intermediaries is constant outside of a crisis, while even when optimistic agents expect a crisis they expect to withstand it with stronger capital buffers thanks to a payoff $z_2 + \Omega_2$ on

their holdings of risky assets, rather than just z_2 . Because of the strong non-linearity of the model, the behavioral wedge is a complex object. Nonetheless, an infinitesimal perturbation around the REE is enlightening (assuming Ω_2 and Ω_3 are small state-by-state):

Proposition 2 (Behavioral Wedge Approximation). If Ω_2 and Ω_3 are small state-by-state, the behavioral wedge for short-term debt, \mathcal{B}_d , can be approximated by:²⁸

$$\mathcal{B}_{d} \simeq \underbrace{-\Omega_{2} H \mathbb{E}_{1}^{SP} \left[\lambda_{2}^{2} \mathbb{1}_{\kappa_{2} > 0}\right]}_{(i)} + \underbrace{\phi H(\mathbf{1} - \boldsymbol{\zeta}) \mathbb{E}_{1}^{SP} \left[\Omega_{3} \lambda_{2}^{2} \mathbb{1}_{\kappa_{2} > 0}\right]}_{(ii)}.$$
(34)

The first term quantifies the welfare losses from contemporaneous irrationality at t = 1. It is negative when Ω_2 is positive, naturally implying that an additional unit of leverage is costly when agents are over-optimistic. Importantly the bias is multiplied by a measure of the expected severity of a future financial crisis, outlining that what affects welfare is not simply deviations from rationality, but their *interaction* with financial frictions.

The second term quantifies welfare changes emanating from the predictable behavior of future deviations from rationality, Ω_3 . Once again, predictable pessimism in the future is not enough to generate first-order welfare losses: this term is non-zero only when the *product* of sentiment with marginal utility in a crisis is non-zero. In other words, it is the *comovement* of irrationality with the health of financial intermediaries that is a cause of concern for the planner. An interesting case in point of equation (34) is that even if $\Omega_2 = 0$, welfare losses are possible because of the predictable behavior of *future* irrationality. Even if, on average, there is no deviation from rationality (i.e. $\mathbb{E}_1^{SP}[\Omega_3] = 0$), the possible covariance of Ω_3 with the health of financial intermediaries, λ_2 , creates a welfare loss from increasing leverage in period t = 1. This implies that is it not necessary for the social planner to know the current state of irrational exuberance to be justified to act preemptively: knowing that agents will be pessimistic in bad states of the world is enough. This insight, however, heavily depends on the degree of sophistication of agents. Indeed, equation (34) makes clear that the second part of this behavioral wedge disappears when agents are sophisticated ($\zeta = 1$): in this instance, agents realize that the crisis will

²⁸The same first-order analysis for investment can be found in Online Appendix A.4.

be worse because of over-pessimism in the future, and thus lower their leverage in the initial period accordingly.

3.3 Externalities

The second class of effects are externalities that works through *future beliefs* and *prices*. They are operative even though, as explained earlier, there is no externality in the rational benchmark. Even more surprisingly, they are effective even in the case of fully sophisticated agents as can be seen from the absence of ζ in these expressions. Let us examine in detail the terms composing this externality for leverage:

$$\mathcal{C}_d = -\mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{d\Omega_3}{dq_2} \frac{dq_2}{dn_2} \right].$$
(35)

The first term is the Lagrange multiplier κ_2 , again indicating that welfare losses are present only in the event of a binding financial friction at t = 2. The term ϕH then corresponds to the fact that this externality operates at the level of the friction that limits borrowing at t = 2. The derivative dq_2/dn_2 quantifies the change in asset prices implied by the change in short-term debt at t = 1: taking on more leverage mechanically lowers net worth in the future, which impacts equilibrium prices in the future (see Section 2). For now, all of these terms also exist in a rational world. The bold term, however, is specific to behavioral distortions and is thus zero in a rational counterfactual, making the expression zero in total. The fraction $d\Omega_3/dq_2$ measures how sentiment *inside* a financial crisis changes when equilibrium prices change.

This externality can be intuitively described as follows. Agents fail to internalize that, by increasing their leverage in good times, they lower asset prices tomorrow, which can make everyone in the economy more pessimistic. This pessimism, in turn, tightens the collateral constraint of financial intermediaries, preventing them to roll-over their debt as desired, and aggravating the financial crisis.

For this externality to exist it is necessary for the belief derivative to be non-zero. In other words, the collateral externality is operative if and only if behavioral biases at t = 2 are a direct function of equilibrium prices at t = 2. This means, for example, that any fundamental-based behavioral bias as in equation (18) will not feature such a market failure. In the natural benchmark of price extrapolation, as in equation (19), this derivative is simply $d\Omega_3/dq_2 = \alpha_q > 0$. This externality is then pushing the private solution towards excessive borrowing.²⁹

More surprisingly, the collateral externality for investment in Proposition 1 can in the opposite direction as the leverage one. Agents are not taking into account how a supplementary unit of collateral, by raising net worth next period, can support asset prices and thus consequently reduce pessimism. In turn, this ameliorates the borrowing capacity of the whole economy, thus improving welfare.³⁰ Irrational exuberance thus helps overcome the under-investment problem coming from financial frictions.³¹ In a model with a collateral constraint directly featuring q_2 , the rational benchmark features such a positive collateral externality. Irrational exuberance thus helps to alleviate this market failure. The welfare impact of an additional unit of investment is however unambiguously negative for large enough Ω_2 , since the behavioral wedge can be unboundedly negative.

The last part of Proposition 1 highlights another effect. In most models (rational models or models with exogenous sentiment) the two uninternalized effects for leverage and investment are enough to characterize efficiency. Indeed, once allocations are fixed the equilibrium level of prices has no effect on welfare. The problem is different, however, in the presence of endogenous sentiment since a new state-variable enters the optimal policy problem: the equilibrium level of asset prices today can enter the determination of

$$\frac{dq_2}{dn_2} = \frac{(\beta - \phi)\mathbb{E}_2[z_3 + \Omega_3]}{1 - (\phi + (\beta - \phi)(2c_2 - n_2))\frac{d\Omega_3}{dq_2}}.$$
(36)

²⁹Notice that when this externality exists because of endogenous sentiment, the price sensitivity dq_2/dn_2 that enters this expression is also magnified by *belief amplification*. A change in net worth in period t = 2 impacts equilibrium asset prices as:

A positive change in net worth leads to a change in price through the stochastic discount factor c_2 , which itself can alleviate pessimism, supporting asset prices in a feedback loop. This makes the price more sensitive to changes in net worth when $d\Omega_3/dq_2 > 0$.

³⁰In models where assets change hands in equilibrium, as in Lorenzoni (2008), the price of the asset is decreasing in the aggregate quantity, since outside investors are usually assumed to have a concave production function.

³¹This is reminiscent of Martin and Ventura (2016), where the presence of bubbles alleviates financing frictions.

future allocations, and thus the expected level of welfare:

$$\mathcal{W}_{q} = \beta \mathbb{E}_{1}^{SP} \left[\kappa_{2} \phi H \frac{d\Omega_{3}}{dq_{1}} \right].$$
(37)

This effect works through the interaction of financial frictions, *past* asset prices, and sentiment in a crisis. The intuition for this term is as follows. When private agents push up the price of the asset today, it might influence the formation of behavioral biases in the future. This is represented by the term $d\Omega_3/dq_1$. Typically, in our illustrative price extrapolation case where $\Omega_3 = \alpha(q_2 - q_1)$, this derivative is equal to $-\alpha$, a negative term. This change in sentiment at time t = 2 impacts the collateral limit for short-term debt d_2 , in proportion to ϕH , a positive quantity. It then impacts welfare if agents are against their borrowing constraint, i.e. if $\kappa_2 > 0$, since it directly alters the amount they can borrow. Succinctly, when agents bid up prices, this can feed pessimism tomorrow by increasing the anchor agents use to form expectations: an increase in prices today will cause a reversal tomorrow. I thus call this effect a *reversal externality*. This new force is independent of the current extent of behavioral biases in the initial period, and has important implication for the conduct of policy as I show next.³²

Finally, these externalities are robust to the degree of sophistication of agents, as can be seen from the absence of ζ in these expressions. As for regular externalities, agents do not internalize that their decisions at t = 1 can influence the determination of sentiment at t = 2 if sentiment depends on equilibrium variables like prices. Take the collateral externality for example. Market's participants can be sophisticated enough to realize that low prices in a future crises will create endogenous pessimism and tighten collateral constraints. But they cannot coordinate to reduce their aggregate leverage at t = 1, in order to sustain asset prices in a future bust and attenuate future pessimism. Only the regulator can internalize these effects, even though the regulator and the market share the same beliefs about how sentiment will manifest itself in the future.

³²Similarly, Schmitt-Grohé and Uribe (2016) show that with downward wage rigidity, past wages become a relevant state variable, motivating the planner to reduce real wages in booms. This effect is also present in Farhi and Werning (2020). In their model, a high price in the initial period can translate into over-pessimism when the ZLB hits. This creates a downward pressure on prices at the ZLB, which affects aggregate demand via a wealth effect.

Remark 3 (Sentiment during crises). A majority of the effects presented in Proposition 1 operates through the interaction of irrationality with the health of financial intermediaries during crises. Suggestive evidence supports the assumption that the two objects Ω_3 and $\lambda_2 \mathbb{1}_{\kappa_2>0}$ negatively comove. Figure 2 uses two proxies to construct time series for Ω_3 and for λ_2 . For the marginal utility of intermediaries λ_2 , I rely on He, Kelly, and Manela (2017) who compute an intermediary capital ratio (inversely proportional to λ_2 when agents have log-utility). For Ω_3 , I use the forecast errors made by stock market analysts on the long-run growth of stocks, a measure from Bordalo, Gennaioli, La Porta, and Shleifer (2020) which is directly constructed from survey data. Figure 2 shows how Ω_3 is consistently negative in crises.^{33,34}

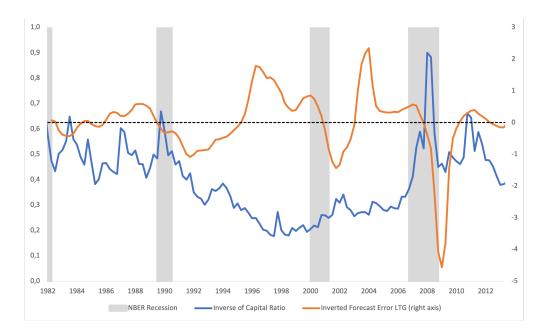


Figure 2: Time-series variation of proxies for λ_2 and Ω_3 . For the financial health of intermediaries λ_2 , I rely on He et al. (2017) which computes an intermediary capital ratio. The inverse of this capital ration is proportional to λ_2 when agents have log-utility, as in this model. For Ω_3 , I use the inverted forecast errors made by stock market analysts on the long-term growth of stocks, a measure of Bordalo et al. (2020)

³³In 2008, forecast errors Ω_3 crashed while the marginal utility of intermediaries λ_2 spiked, suggesting sizable welfare losses. But in events such as the dot-com bubble burst, pessimism was not accompanied by declines in the financial health of intermediaries. The theory I am developing suggests that these events are less of a concern for welfare.

³⁴See also Bordalo et al. (2018) and Maxted (2020) for other examples of over-pessimism. An earlier version (Fontanier 2022) also presents other alternatives to measure sentiment, leading to the same result.

4 **Optimal Policy**

4.1 Constrained Efficiency

I can now characterize the allocation the planner would like to implement in the presence of sentiment. A planner subject to the same constraints as agents, with prices determined by market-clearing as in the decentralized case, evaluates welfare using its own expectations and thus takes the previous uninternalized effects into account. These objects are crucial to characterize optimal policy in this setting.³⁵

I first start with a natural proposition: in order to achieve the second-best the planner makes agents internalize their uninternalized welfare effects. This is done by choosing taxes or subsidies that exactly cancel out the uninternalized effects described in Proposition 1.

Proposition 3 (Second-Best Policy). *The social planner achieves the second-best by imposing:*

- 1. A tax $\tau_d = -W_d / \lambda_1$ on short-term borrowing ;
- 2. A tax $\tau_H = -W_H/(\lambda_1 q_1^*)$ on the creation of collateral assets ;
- 3. A tax $\tau_q = \frac{q_1 q_1^*}{q_1^*}$ on the holding of collateral assets

where λ_1 is the marginal utility of financial intermediaries at time t = 1 evaluated at the desired allocation, q_1 is the price that would arise through market-clearing at the desired allocation without the holding tax, and q_1^* is the price such that $W_q = 0$ when evaluated at the desired allocation.

Proposition 3 is rather abstract, but makes two simple points. First, the calibration of macroprudential policy should be done by focusing on the key aspects of sentiment

³⁵The concept of constrained efficiency also restricts the analysis to a planner who takes financial frictions as given, following Hart (1975), Stiglitz (1982) and Geanakoplos and Polemarchakis (1985). It can be understood as answering the following question: can a planner subject to the same constraints as private agents improve on the market outcome? In particular, any direct intervention at t = 2 is proscribed. Online Appendix E allows for the simultaneous choice of ex-ante and ex-post policies. In particular, it shows that in my framework the possibility of intervention at t = 2 does not change the desirability of macroprudential interventions at t = 1 (Jeanne and Korinek 2020).

driving the uninternalized effects from the previous part, W_d and W_H : (i) the current extent of sentiment Ω_2 ; (ii) the future covariance of $(1 - \zeta)\Omega_3$ with λ_2 , conditional on Ω_2 ; (iii) the sensitivity of sentiment with respect to current and future prices. Second, when current asset prices impact future sentiment, three instruments are needed to achieve the second-best, and not only two.³⁶

How can one interpret the results of Proposition 3 in terms of real-world policy? The optimal taxes on debt and investment correspond to the usual instruments in the macro-prudential toolkit: capital requirements and Loan-to-Value (LTV) restrictions (Claessens 2014). This is not the case for the tax on holdings, designed to influence equilibrium price. I now explore the concrete policy lessons coming out of the analysis.

4.2 Implementation

Counter-cyclical Capital Buffers: The tax on short-term borrowing can naturally be interpreted as capital structure regulation. Proposition 3 thus provides the financial regulator with the features of behavioral biases that are necessary to quantify in order to optimally calibrate leverage restrictions. Because Ω_2 is a largely volatile object (see Figure 2), the optimal value of this macroprudential leverage tax is also time-varying. But importantly, the time-variation in τ_d should not only track Ω_2 , but also take into account the expected *future* realizations of Ω_3 (if agents are not fully sophisticated) as well as the expected impact of future prices on Ω_3 .

Most macroprudential regulations on capital structure are nonetheless set in terms of leverage limits rather than leverage taxes. Are a leverage tax and a leverage limit equivalent in this model? Weitzman (1974) showed that whether price or quantity regulation is more desirable depends on which one is more robust to changes in parameters. Here, when financial intermediaries are against the regulatory leverage limit, an increase in sentiment Ω_2 increases their incentives to take on more debt, but agents simply cannot

³⁶Achieving the second-best does not imply that the benchmark rational level of welfare is attainable. This is because, while the effects of over-optimism can be fully resorbed by appropriate prudential regulation, the presence of pessimism in crises makes intermediaries worse off even with optimal policy compared to the rational benchmark, since no intervention is possible at t = 2 to directly increase borrowing.

change their positions. Their leverage thus stays at the exact allocation desired by the social planner. This is not the case for debt taxes, as can be seen from equation (34): the tax needs to be calibrated at the exact level of Ω_2 to achieve the second-best. This intuition leads to the following proposition.

Proposition 4 (Leverage Limits Robustness). Leverage limits are more robust than leverage taxes to small movements in the behavioral bias around a positive Ω_2 .

The intuition can also be seen when sentiment moves downward, towards less exuberance. For small departures from an equilibrium with $\Omega_2 \ge 0$, movements in Ω_2 on the downside do not call for changing the allocation desired by the planner, because the pecuniary externality still needs to be corrected. A leverage limit thus stays binding for agents, while a leverage tax would force financial intermediaries to decrease their leverage below the socially desirable outcome.^{37,38}

This insight, however, does not imply that counter-cyclical restrictions are not desirable when a flat leverage limit is imposed. This is because, as explicit from Proposition 1, the behavior of future sentiment matters as much as the extent of contemporaneous irrational exuberance from the perspective of period t = 1. As long as the planner's estimate of Ω_3 given the information available is time-varying, and agents are not fully sophisticated, the leverage limit needs to be tightened or relaxed accordingly.

LTV Regulation: The second tax in Proposition **3** directly aims at regulating the *quantity* of risky investments. For this reason, this policy can be interpreted as loan-to-value (LTV) regulation. Importantly, the welfare analysis highlights again that the optimal LTV limit is time-varying, tracking the same behavioral biases as do leverage restrictions.

The crucial difference with counter-cyclical capital requirements lies in the time-variation required by variation in the expected impact of prices on sentiment. When the regulator

³⁷Note that quantity *caps* are enough only because Ω_2 is assumed to be positive. In the event that agents are too pessimistic initially, the regulator would need to subsidize borrowing and thus caps fail to achieve this objective.

³⁸The previous working paper Fontanier (2022) additionally shows that a leverage limit is robust to the introduction of belief heterogeneity in the model, while a leverage tax becomes less efficient. See Clayton and Schaab (2020) and Jeanne and Korinek (2020) for the disparities between price and quantity regulations in macroprudential policy.

is concerned that a future crash in prices will result in a greater sensitivity of sentiment with respect to prices in a crisis (all else equal), the optimal reaction is to tighten leverage restrictions more but to *relax* LTV ratios. Indeed, as explained in Section 3.3, the collateral externality for *H* calls for *higher* investment than in the decentralized equilibrium, in order to alleviate pessimism during crises by strengthening the net worth of the financial sector.

Price Regulation: The third tax in Proposition 3 does not have a simple relation to the current macroprudential toolbox, however. This is because my model is the first to high-light the need for an additional instrument that complements traditional macroprudential tools like counter-cyclical capital buffers and LTV ratios. From an abstract perspective, this instrument can be modeled as a tax on asset holdings. But the concrete goal is to directly manipulate asset prices through the demand for these assets. A direct tax on asset holdings, however, seems rather unrealistic to implement. A more natural candidate for this instrument is to use monetary policy.

4.3 Monetary Policy as Price Regulation

A large part of the "leaning vs. cleaning" policy debate revolves around the possible use of a monetary tightening to tame asset prices in the face of irrational exuberance. The conventional view holds that "monetary policy is not a useful tool for achieving this objective" (Bernanke 2002). Recent work challenged this perspective: Caballero and Simsek (2020) show that when traditional macroprudential policy is constrained, leaning against the wind with a monetary tightening is valuable. This occurs because the gains from a preventive tightening are first-order, while the losses from deviating from perfect inflation targeting are only second-order (assuming that the output gap can be perfectly closed). As anticipated, my model features a different channel through which monetary policy can affect welfare: the reversal externality.

I start by introducing rigidities in order for monetary policy deviations to have potential costs. Because aggregate demand is not the focus on this paper, this is done by following Farhi and Werning (2020): households supply labor and output is demanddetermined at t = 1 by assuming wages are fully rigid.³⁹ Households now have the following utility function:

$$U^{h} = \mathbb{E}_{1} \left[\left(\ln(c_{1}^{h}) - \nu \frac{l_{1}^{1+\eta}}{(1+\eta)} \right) + \beta c_{2}^{h} + \beta^{2} c_{3}^{h} \right]$$
(38)

which introduces curvature in consumption utility, and labor disutility in period t = 1. Firms produce using labor linearly, $Y_1 = l_1$. Wages are fully rigid and normalized to 1, causing workers to be potentially off their labor supply curve. This creates a role for monetary policy: the central bank can close the output gap by choosing the nominal rate of interest that brings workers back to their labor supply curve. The *labor wedge* $\mu_1 = 1 - vc_1^h l_1^\eta$ quantifies how far off are workers from their optimality condition. The labor wedge is positive when there is underemployment, and negative when there is overheating. Perfectly achieving natural employment means that $\mu_1 = 0$. Finally, Pareto weights are simply taken to be equal to the marginal utility of each type of agent at t = 1, in order to sidestep redistribution concerns.

A change in the nominal interest works through five different channels: (i) traditional aggregate demand ; (ii) credit ; (iii) investment ; (iv) current beliefs and (v) future beliefs. We can once again leverage the prior general welfare analysis.

Proposition 5 (Welfare Effects of Monetary Policy). *The total welfare effects, as evaluated through the central bank's expectations, of an infinitesimal interest rate can be expressed by:*

$$\frac{d\mathcal{W}_{1}}{dr_{1}} = \underbrace{\frac{dY_{1}}{dr_{1}}\mu_{1}}_{(i)} + \underbrace{\frac{dd_{1}}{dr_{1}}\mathcal{W}_{d}}_{(ii)} + \underbrace{\frac{dH}{dr_{1}}\mathcal{W}_{H}}_{(iii)} + \underbrace{\frac{d\Omega_{2}}{dq_{1}}\frac{dq_{1}}{dr_{1}}\left(\frac{dd_{1}}{d\Omega_{2}}\mathcal{W}_{d} + \frac{dH}{d\Omega_{2}}\mathcal{W}_{H}\right)}_{(iv)} + \underbrace{\frac{dq_{1}}{dr_{1}}\beta\mathbb{E}_{1}\left[\kappa_{2}\phi H\frac{d\Omega_{3}}{dq_{1}}\right]}_{(v)} (39)$$

³⁹In Farhi and Werning (2020), there is an aggregate demand externality because wages are also fully rigid at t = 2, when the economy hits the ZLB. In my model there is no ZLB at t = 2, thus no aggregate demand externality. The results in this section are thus complementary to those in Farhi and Werning (2020), and do not rely on the inability of the central bank to lower rates sufficiently in crises.

where $W_d = B_d + C_d$, the sum of the behavioral wedge and the collateral externality for leverage, and $W_H = B_H + C_H$, the sum of the behavioral wedge and the collateral externality for investment. The last term is proportional to W_q , the reversal externality (see Section 3 for details).

If the monetary authority is able to perfectly close the output gap and bring the economy to full employment, then it can achieve $\mu_1 = 0$ (and the perturbation is taken around the natural rate). As mentioned earlier, there is thus no first-order costs from deviating slightly from perfect inflation targeting. This expression thus embodies the idea in Stein (2021) that financial stability concerns loom large when unemployment is low (μ_1 close to zero), and should be negligible when unemployment is extremely high (μ_1 strongly positive).

This, however, does not necessarily imply that leaning against the wind is always desirable when the output gap can be closed, however. To see why, take the extreme case where the financial authority is able to adapt its leverage restrictions perfectly such that $W_d = 0$, and look at the simpler case where $d\Omega_3/dq_1 = d\Omega_2/dq_1 = 0$ such that channels (*iv*) and (*v*) disappear. The welfare effects are thus now given in this special case by:

$$\frac{d\mathcal{W}_1}{dr_1} = \frac{dH}{dr_1}\mathcal{W}_H \tag{40}$$

because investment is unambiguously decreasing in the interest rate r_1 , tightening is desirable only if $W_H < 0$, i.e. if the uninternalized welfare effects of marginally increasing the creation of collateral assets is negative. As fully explained in Section 3.3, this object is actually positive for small belief deviations and becomes negative only if irrational exuberance is large enough. In other words the central bank would only pursue leaning against the wind when facing large enough behavioral distortions.⁴⁰

Notice from equation (39) that the ability of the central bank to improve financial sta-

⁴⁰My framework also abstracts from other considerations that could argue against tightening in such a situation. For example, Martinez-Miera and Repullo (2019) show that tightening can increase risk-taking by financial institutions by shifting investment towards riskier firms. Allen, Barlevy, and Gale (2022) present a model where risk-shifting raises asset prices above fundamentals, but tighter monetary policy further decreases investment that is already underfunded. Caines and Winkler (2021) and Adam and Woodford (2021) are other examples where the central bank leans against the wind of high house prices to avoid excessive investment in housing. In my framework such welfare costs are primarily addressed using LTV/LTI tools that directly target investment inefficiencies.

bility largely depends on the reaction of beliefs to policy. Without the belief channels (iv) and (v), the potential efficacy of leaning against the wind rests on the ability to curb leverage directly by raising rates, dd_1/dr_1 . As emphasized by Werning (2015) and Farhi and Werning (2020), this is not a robust prediction of these models: it varies with the initial debt position as well as the shape of the utility function. To the contrary, the fact that increasing interest rates has a negative impact on asset prices is unambiguous in our models and is supported by robust empirical evidence (see e.g. Rigobon and Sack 2004 and Bernanke and Kuttner 2005). Thus if Ω_2 or Ω_3 depend directly on asset prices, leaning against the wind can have first-order benefits, by providing a supplementary instrument affecting equilibrium prices, and not only real allocations at t = 1.

These results also directly speak to the debate about time-varying macroprudential tools. A common argument for using monetary policy to rein in financial excess is that, practically, macroprudential policy cannot be quickly adapted to be synchronized in real-time.⁴¹ Inspecting Proposition 5, however, suggests that this is only part of the story. To focus on this question, assume: (i) fully unconstrained counter-cyclical capital regulation and (ii) fully unconstrained LTV regulation. Despite these assumptions, monetary policy still has an effect through prices and future behavioral biases.⁴²

Proposition 6 (Monetary Policy as Complement). When policymakers have access to unconstrained leverage and investment taxes, welfare changes evaluated around the equilibrium with optimal taxes are given by:

$$\frac{d\mathcal{W}_1}{dr_1} = \frac{dY_1}{dr_1}\mu_1 + \beta \mathbb{E}_1 \left[\kappa_2 \phi H \frac{d\Omega_3}{dq_1} \frac{dq_1}{dr_1}\right].$$
(41)

This particular case calls for leaning against the wind in order to lower *current* asset prices, which will then cure *future* pessimism in a possible crisis – a new channel for monetary policy.⁴³ Furthermore, such action does not necessarily require information

⁴¹See e.g. Dudley (2015) ; Caballero and Simsek (2020) and Stein (2021).

⁴²This is similar to Farhi and Werning (2020) when they consider the presence of biases during crises. The main difference is that my results do not depend on the inability of the central bank to change interest rates during the bust.

⁴³Barlevy (2022) notes too that monetary policy is effective only if it can cure optimism or pessimism.

about contemporaneous biases. A sharp increase in asset prices could be entirely due to fundamentals, but the planner can still have an incentive to make prices deviate from their rational value today to prevent irrational distress from happening. Finally, implementing such a policy allows for financial regulation to be adapted and *relaxed*. Indeed, by acting preventively the central bank makes the future realizations of pessimism less severe, thus directly reducing the size of behavioral wedge and of the collateral externality. Taking this into account leads the optimal macroprudential limit to be less strict, which raises welfare.⁴⁴

Finally, notice how the sophistication parameter ζ is naturally absent of Proposition 6. Sophisticated financial intermediaries realize very well that a high price today will translate into over-pessimism and tight collateral constraints in a future crisis, but cannot coordinate to reduce their buying pressure in order to cool off asset prices.

4.4 Small Deviations from Rationality

Suppose that we place ourselves at the REE constrained-efficient allocation. Agents are fully rational, so the planner has no reason to intervene. If we add an infinitesimal degree of irrationality, which forces cause first-order welfare losses? The answer comes by inspecting equations (31), (32), and (33). At the rational expectations constrained-efficient equilibrium, behavioral wedges are zero, so the only effects left are the collateral externalities and the reversal externality:

$$C_d = -\beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{d\Omega_3}{dq_2} \frac{dq_2}{dn_2} \right]$$
(42)

$$\mathcal{C}_{H} = \beta \mathbb{E}_{1}^{SP} \left[\kappa_{2} \phi H \frac{d\Omega_{3}}{dq_{2}} \left(\frac{dq_{2}}{dn_{2}} z_{2} + \frac{dq_{2}}{dH} \right) \right]$$
(43)

$$\mathcal{W}_{q} = \beta \mathbb{E}_{1}^{SP} \left[\kappa_{2} \phi H \frac{d\Omega_{3}}{dq_{1}} \right]$$
(44)

which, as explicated earlier, are only present when *future* sentiment is impacted by *current* and *past* asset prices, and there is a positive probability of a crisis in the future.

⁴⁴Evidently, this policy problem is also plagued with uncertainty. Section 5.2 shows that the incentives to tighten monetary policy are increasing in the uncertainty around the strength of the reversal force.

The fundamental intuition behind this result is that small changes in leverage due to fluctuating sentiment are not harmful to the first-order since agents are on the objective Euler equation. But anything that directly impacts the tightness of the collateral constraint in a crisis, where agents are *not* on their Euler equation, has a first-order impact on welfare by aggravating financial turmoil. This result draws attention to irrational distress during financial crises, while the literature has mostly focused on irrational exuberance during the build-up leading to the crash.⁴⁵

5 Sentiment Uncertainty

I so far assumed that the social planner had perfect information about behavioral biases. In a famous speech on asset price bubbles, Bernanke (2002) discussed the "identification problem" that naturally arises once the financial stability authority contemplates a proactive approach to bubbles. A natural question of practical importance is then whether my results are impaired in the presence of imperfect knowledge about behavioral biases.

The short answer is: no, to the contrary. Sentiment uncertainty reinforces motives for preventive action, in contrast with Brainard (1967)'s "attenuation principle." While recognizing that identifying a bubble is intrinsically difficult, this section shows that the widespread intuition that this uncertainty calls for more *laissez-faire* is actually erroneous.

5.1 Ω -Uncertainty

I start with the case where the regulator is uncertain about the level of irrational exuberance. To this end, I leverage the previous equilibrium and welfare analysis. Ω_3 is assumed to be certain and constant in the future.⁴⁶ Recall that private agents are shifting the entire distribution of future dividends by Ω_2 , believing that dividends will be $z_2 + \Omega_2$ instead

⁴⁵Of course irrational exuberance is also costly, as it triggers more frequent credit crunches. It is also possible that irrational distress is a direct function of past optimism, creating the same kind of reversal externality, but the first-order damages to welfare would not be directly attributable to irrational exuberance either. There is also the possibility that over-optimism has other effects on investment in the real sector, which can be costly as in Rognlie, Shleifer, and Simsek (2018).

⁴⁶The analysis for Ω_3 -uncertainty is presented in Online Appendix D. The results are identical.

of z_2 . The following assumptions make the analysis more convenient:

Assumption 3. All parameters of the model and of the probability density function $f_2(z_2)$ are common knowledge to private agents and the social planner, except possibly for its mean, \bar{z}_2 .

Assumption 4. Equilibrium prices at time t = 1 are strictly increasing in \bar{z}_2 .

These assumptions imply that, in the absence of sentiment, the social planner could simply infer the value of \bar{z}_2 by looking at equilibrium prices in period 1, q_1 . Let me now assume that the social planner's prior over sentiment is given by a uniform distribution:

$$w \sim \mathcal{U}\left[\bar{\Omega}_2 - \sigma_{\Omega}, \bar{\Omega}_2 + \sigma_{\Omega}\right] \tag{45}$$

where $\bar{\Omega}_2$ is the average level of sentiment according to the planner's prior, and σ_{Ω} controls the amount of uncertainty around it. By observing asset prices the planner can infer what agents believe the average future dividend is, so that the planner's posterior mean regarding future dividends becomes:

$$\bar{z}_2 = g_q^{-1}(q_1) - \bar{\Omega}_2 \tag{46}$$

Taking the uncertainty in its posterior into account, the planner first-order condition for short-term debt is now given by:

$$u'(c_1) = \frac{1}{2\sigma_{\Omega}} \int_0^\infty \left[\int_{-\sigma_{\Omega}}^{\sigma_{\Omega}} \frac{\partial \mathcal{W}_2}{\partial n_2} \left(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \omega_2 \right) d\omega_2 \right] f_2(z_2) dz_2$$
(47)

This expression contains all of the intuition for how sentiment uncertainty can reinforce or weaken the need for preventive leverage tightening.⁴⁷ Once deducing the average behavioral error $\overline{\Omega}_2$, the planner is uncertain about the exact distribution of the state of the world next period. It thus takes the distribution that agents use, but factors in the noise it attributes to their expectations. This leads the social planner to consider, for each realization z_2 , all values inside the segment $[z_2 - \sigma_{\Omega}, z_2 + \sigma_{\Omega}]$ as equally likely.

⁴⁷Note that $\overline{\Omega}_2$ is an argument of W_2 here, because the planner is using the same distribution agents are using. In the previous Section, I was using an equivalent notation where private agents use the same distribution as the planner but add their bias $\overline{\Omega}_2$. The two are of course equivalent, but in the present section this clarifies that the planner simply extracts the distribution used by private agents.

If expanding the set of possible behavioral biases, by increasing σ_{Ω} , increases the value of the expectations term, it means that sentiment uncertainty increases expected marginal utility. This, in turn, implies that the social planner wishes to reduce the leverage of agents today to get back to the optimality condition, by increasing initial marginal utility $u'(c_1)$, and by diminishing expected future marginal utility. Conversely, if enlarging the possible values of ω_2 decreases expected marginal utility, the social planner should relax leverage constraints compared to the absolute certainty case. Using the analysis of the equilibrium presented in Section 2.5, uncertainty about behavioral biases unambiguously calls for precautionary restrictions, as expressed in the following Proposition.

Proposition 7 (Ω_2 -Uncertainty and Leverage Restrictions). If the social planner believes that the behavioral bias at t = 1 can be expressed as $\overline{\Omega}_2 + \omega$, where ω is uniformly distributed on $[-\sigma_{\Omega}, \sigma_{\Omega}]$, and Ω_3 is constant state-by-state at t = 2, then the optimal leverage tax is increasing in σ_{Ω} , and the optimal investment tax is decreasing in σ_{Ω} .

The proof is rather involved (see Online Appendix A.8), but the intuition can be understood succinctly. The key is to notice that marginal welfare is a *convex* function, as shown in Figure 3. Intuitively, sentiment uncertainty adds terms to the expectation computed by the planner relative to the private solution, but the parts coming from intermediaries' optimism are more costly than the ones coming from pessimism.⁴⁸ The strong non-linearities associated with the interaction of sentiment with financial frictions make it attractive to tighten capital requirements in the face of uncertainty. Using the words of Yellen (2009), a "type 1" error is simply much less costly than a "type 2" error.

Proposition 7 also highlights that investment regulation behaves in a different way: an increase in σ_{ω} calls for more investment in *H* in the planner's problem relative to the private solution. This is because increasing uncertainty increases the incentive to shift consumption to the next period. Indeed, if there is a risk that agents are extremely overoptimistic and that a crisis will be extremely severe, it is even more valuable to hold an asset that is going to pay dividends, albeit low, in this state of the world. Concretely,

⁴⁸The same insights can be obtained if we were to consider uncertainty about the extent of sentiment *inside* a financial crisis. Furthermore, endogenous sentiment, for example in the form of price extrapolation, amplifies this effect by adding *more* curvature. These results are presented in Online Appendix D.

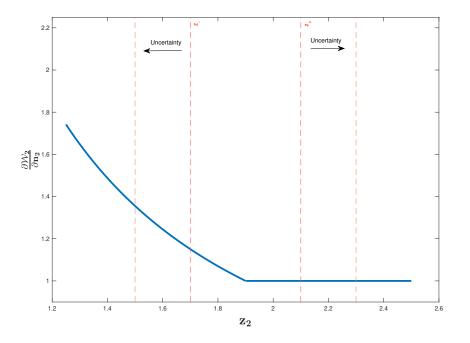


Figure 3: Non-linearities and Ω -uncertainty. This figure plots $\partial W_2 / \partial n_2$ against the fundamental realization z_2 . The thick dotted lines correspond to the range of values of z_2 where the expectations is taken. The thin dotted lines represent the widening of the range where the social planner computes expectations caused by the uncertainty on Ω_2 . The break arises at the crisis cutoff.

this means that in times of heightened uncertainty, the regulator should tighten countercyclical capital buffers but at the same time relax LTV ratios.

5.2 Reversal Uncertainty

The previous analysis shows how the regulator should adapt leverage and investment regulations in the face of sentiment uncertainty. The last natural question is how price regulation (and thus the eventual use of interest rates) should be adapted. The last proposition answers this interrogation unambiguously.

Proposition 8 (Reversal Uncertainty and Price Regulation). Assume that inside crises, the behavioral bias takes the form $\Omega_3 = \overline{\Omega}_3 - \alpha_q q_1$ with $\overline{\Omega}_3$ a constant, and the planner believes that α is uniformly distributed on $[\overline{\alpha}_q - \sigma_{\alpha}, \overline{\alpha}_q + \sigma_{\alpha}]$, where $\overline{\alpha}_q$ and σ_{α} are positive constants. Then the optimal interest rate at t = 1 is increasing in σ_{α} if the regulator has access to unconstrained leverage and investment regulations.

This proposition formalizes the following intuition: the regulator fears that high prices today could translate into over-pessimism inside a future crisis, but is unsure of the strength of the extrapolation. In that situation, the more uncertainty there is around this extrapolation mechanism, the more the regulator wants to lower prices in the boom. This is one again coming from a simple convexity insight: cases where the extrapolation parameter is strong are more costly because of the non-linearities typically found in financial crises.

6 Conclusion

Should financial regulators and monetary authorities try to mitigate the potential instabilities associated with irrational booms and busts? In this paper I provide a framework that allows for the rigorous analysis of this crucial policy question. I showed how leverage, investment and price regulations can achieve constrained efficiency in the presence of behavioral biases, even in an environment that does not feature any externality in its rational benchmark. Importantly, some of the effects uncovered depend directly on beliefs being a function of equilibrium prices, and are robust to the degree of sophistication of agents. Finally, I showed that adding uncertainty about the extent of behavioral biases in financial markets reinforced incentives for leverage regulation, as well as for the use of monetary policy to lean against the wind.

While the model can be extended along several dimensions, the results suggest a need for research on two specific dimensions. First, while empirical research has convincingly demonstrated that overreaction is a pervasive feature of financial markets, we have less certainty about its drivers. My paper shows that understanding what drives deviations from rationality will simultaneously advance our comprehension of what policy can and should do to deal with financial bubbles. Second, in my model the small number of periods obfuscates the timing subtleties faced by regulators. But we have little understanding over the dynamic build-up of sentiment, and over which horizon it is influenced by monetary policy and asset prices. Further empirical and theoretical research is needed to fully grasp the complex timing interactions between policy, crises, and behavioral biases.

References

- Adam, K. and A. Marcet (2011). Internal rationality, imperfect market knowledge and asset prices. *Journal of Economic Theory* 146(3), 1224–1252.
- Adam, K., A. Marcet, and J. Beutel (2017). Stock price booms and expected capital gains. *American Economic Review* 107(8), 2352–2408.
- Adam, K., O. Pfäuti, and T. Reinelt (2022). Subjective housing price expectations, falling natural rates and the optimal inflation target.
- Adam, K. and M. Woodford (2021). Robustly optimal monetary policy in a new keynesian model with housing. *Journal of Economic Theory* 198, 105352.
- Allen, F., G. Barlevy, and D. Gale (2022). Asset price booms and macroeconomic policy: A risk-shifting approach. *American Economic Journal: Macroeconomics*.
- Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 1, pp. 79–175. Elsevier.
- Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2018). Extrapolation and bubbles. *Journal of Financial Economics* 129(2), 203–227.
- Barberis, N. and A. Shleifer (2003). Style investing. *Journal of Financial Economics* 68(2), 161–199.
- Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics* 49(3), 307–343.
- Barlevy, G. (2022). On speculative frenzies and stabilization policy.
- Bastianello, F. and P. Fontanier (2022). Partial Equilibrium Thinking in General Equilibrium. *Working Paper*.
- Bernanke, B. S. (2002). Asset-price "bubbles" and monetary policy. *Speech before the New York Chapter of the National Association for Business Economics.*
- Bernanke, B. S. and M. Gertler (2000). Monetary policy and asset price volatility. *Working Paper*.
- Bernanke, B. S. and K. N. Kuttner (2005). What explains the stock market's reaction to federal reserve policy? *The Journal of finance* 60(3), 1221–1257.
- Bianchi, J. (2011). Overborrowing and systemic externalities in the business cycle. *American Economic Review* 101(7), 3400–3426.
- Bianchi, J. and E. G. Mendoza (2018). Optimal time-consistent macroprudential policy. *Journal of Political Economy* 126(2), 588–634.
- Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2020). Expectations of fundamentals and stock market puzzles. *NBER Working Paper*.
- Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *The Journal of Finance* 73(1), 199–227.

- Borio, C. E. and P. W. Lowe (2002). Asset prices, financial and monetary stability: exploring the nexus. *BIS working paper*.
- Brainard, W. C. (1967). Uncertainty and the effectiveness of policy. *The American Economic Review* 57(2), 411–425.
- Caballero, R. J. and A. Simsek (2020). Prudential monetary policy. *MIT Working Paper*.
- Caines, C. and F. Winkler (2021). Asset price learning and optimal monetary policy. *Journal of Monetary Economics* 123, 53–67.
- Chahrour, R. and G. Gaballo (2021). Learning from house prices: Amplification and business fluctuations. *The Review of Economic Studies 88*(4), 1720–1759.
- Cieslak, A. (2018). Short-rate expectations and unexpected returns in treasury bonds. *The Review of Financial Studies* 31(9), 3265–3306.
- Claessens, M. S. (2014). An overview of macroprudential policy tools. IMF Working Paper.
- Clayton, C. and A. Schaab (2020). Multinational banks and financial stability. *Working Paper*.
- Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy* 120(1), 116–159.
- Dávila, E. and A. Korinek (2018). Pecuniary externalities in economies with financial frictions. *The Review of Economic Studies* 85(1), 352–395.
- Dávila, E. and A. Walther (2021). Prudential policy with distorted beliefs. *Working Paper*.
- DeFusco, A. A., C. G. Nathanson, and E. Zwick (2017). Speculative dynamics of prices and volume. Technical report, National Bureau of Economic Research.
- DeLong, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Positive feedback investment strategies and destabilizing rational speculation. *the Journal of Finance* 45(2), 379–395.
- Dudley, W. (2015). Is the active use of macroprudential tools institutionally realistic? *Federal Reserve Bank of New York Working Paper*.
- Farhi, E. and X. Gabaix (2020). Optimal taxation with behavioral agents. *American Economic Review* 110(1), 298–336.
- Farhi, E. and I. Werning (2016). A theory of macroprudential policies in the presence of nominal rigidities. *Econometrica* 84(5), 1645–1704.
- Farhi, E. and I. Werning (2020). Taming a minsky cycle. Working Paper.
- Fisher, I. (1932). Booms and depressions: some first principles. Adelphi Company New York.
- Fontanier, P. (January 2022). Optimal Policy for Behavioral Financial Crises. *Harvard University Working Paper*.
- Fuster, A., B. Hebert, and D. Laibson (2012). Natural expectations, macroeconomic dynamics, and asset pricing. *NBER Macroeconomics Annual* 26(1), 1–48.

- Gaspar, V., F. Smets, and D. Vestin (2010). Inflation expectations, adaptive learning and optimal monetary policy. In *Handbook of monetary economics*, Volume 3, pp. 1055–1095. Elsevier.
- Geanakoplos, J. D. and H. M. Polemarchakis (1985). Existence, regularity, and constrained suboptimality of competitive allocations when the asset market is incomplete. *Cowles Foundation Discussion Papers* 764.
- Geithner, T. F. (2014). Stress test: Reflections on financial crises. Broadway Books.
- Gennaioli, N. and A. Shleifer (2018). A crisis of beliefs. Princeton University Press.
- Gorton, G. (2012). *Misunderstanding financial crises: Why we don't see them coming*. Oxford University Press.
- Gorton, G. and A. Metrick (2012). Securitized banking and the run on repo. *Journal of Financial economics* 104(3), 425–451.
- Gourio, F., A. K. Kashyap, and J. W. Sim (2018). The trade offs in leaning against the wind. *IMF Economic Review 66*(1), 70–115.
- Greenspan, A. (2002). Economic volatility. *Remarks by Chairman Alan Greenspan at a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.*
- Greenwald, B. C. and J. E. Stiglitz (1986). Externalities in economies with imperfect information and incomplete markets. *The quarterly journal of economics* 101(2), 229–264.
- Greenwood, R., S. G. Hanson, and L. J. Jin (2019). Reflexivity in credit markets. *NBER Working Paper*.
- Greenwood, R., S. G. Hanson, A. Shleifer, and J. A. Sørensen (2022). Predictable financial crises. *The Journal of Finance* 77(2), 863–921.
- Haddad, V. and T. Muir (2021). Do intermediaries matter for aggregate asset prices? *The Journal of Finance*.
- Hart, O. D. (1975). On the optimality of equilibrium when the market structure is incomplete. *Journal of economic theory* 11(3), 418–443.
- He, Z., B. Kelly, and A. Manela (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics* 126(1), 1–35.
- He, Z., I. G. Khang, and A. Krishnamurthy (2010). Balance sheet adjustments during the 2008 crisis. *IMF Economic Review* 58(1), 118–156.
- He, Z. and A. Krishnamurthy (2013). Intermediary asset pricing. *American Economic Review* 103(2), 732–70.
- Hong, H. and J. C. Stein (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *The Journal of finance* 54(6), 2143–2184.
- Jeanne, O. and A. Korinek (2019). Managing credit booms and busts: A pigouvian taxation approach. *Journal of Monetary Economics* 107, 2–17.

- Jeanne, O. and A. Korinek (2020). Macroprudential regulation versus mopping up after the crash. *The Review of Economic Studies* 87(3), 1470–1497.
- Jordà, Ò., M. Schularick, and A. M. Taylor (2015). Leveraged bubbles. *Journal of Monetary Economics* 76, S1–S20.
- Kahneman, D. and A. Tversky (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology* 3(3), 430–454.
- Kindleberger, C. P. (1978). *Manias, panics and crashes: a history of financial crises*. Palgrave Macmillan.
- Kiyotaki, N. and J. Moore (1997). Credit cycles. Journal of political economy 105(2), 211–248.
- Krishnamurthy, A. and W. Li (2020). Dissecting mechanisms of financial crises: Intermediation and sentiment. *NBER Working Paper*.
- Lorenzoni, G. (2008). Inefficient credit booms. *The Review of Economic Studies* 75(3), 809–833.
- Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020). A quantitative analysis of distortions in managerial forecasts. *NBER Working Paper*.
- Martin, A. and J. Ventura (2016). Managing credit bubbles. *Journal of the European Economic Association* 14(3), 753–789.
- Martinez-Miera, D. and R. Repullo (2019). Monetary policy, macroprudential policy, and financial stability. *Annual Review of Economics* 11, 809–832.
- Maxted, P. (2020). A Macro-Finance Model with Sentiment. Working Paper.
- McCarthy, O. and S. Hillenbrand (2021). Heterogeneous investors and stock market fluctuations. *Working Paper*.
- Mendoza, E. G. (2010). Sudden stops, financial crises, and leverage. *American Economic Review* 100(5), 1941–66.
- Minsky, H. P. (1977). The financial instability hypothesis: An interpretation of keynes and an alternative to "standard" theory. *Challenge* 20(1), 20–27.
- Ottonello, P., D. J. Perez, and P. Varraso (2022). Are collateral-constraint models ready for macroprudential policy design? *Journal of International Economics*, 103650.
- Rabin, M. and D. Vayanos (2010). The gambler's and hot-hand fallacies: Theory and applications. *The Review of Economic Studies* 77(2), 730–778.
- Rigobon, R. and B. Sack (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics* 51(8), 1553–1575.
- Rognlie, M., A. Shleifer, and A. Simsek (2018). Investment hangover and the great recession. *American Economic Journal: Macroeconomics* 10(2), 113–53.
- Schmitt-Grohé, S. and M. Uribe (2016). Downward nominal wage rigidity, currency pegs, and involuntary unemployment. *Journal of Political Economy* 124(5), 1466–1514.

- Schularick, M. and A. M. Taylor (2012). Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008. *American Economic Review* 102(2), 1029–61.
- Shiller, R. J. (2015). Irrational exuberance. Princeton university press.
- Simsek, A. (2013). Belief disagreements and collateral constraints. *Econometrica* 81(1), 1–53.
- Stein, J. C. (2021). Can policy tame the credit cycle? *IMF Economic Review* 69(1), 5–22.
- Stiglitz, J. E. (1982). The inefficiency of the stock market equilibrium. *The Review of Economic Studies* 49(2), 241–261.
- Sufi, A. and A. M. Taylor (2021). Financial crises: A survey. NBER Working Paper.

Weitzman, M. L. (1974). Prices vs. quantities. *The Review of Economic Studies* 41(4), 477–491.

- Werning, I. (2015). Incomplete markets and aggregate demand. NBER Working Paper.
- Winkler, F. (2020). The role of learning for asset prices and business cycles. *Journal of Monetary Economics* 114, 42–58.
- Woodford, M. (2012). Inflation targeting and financial stability. *Working Paper*.
- Yellen, J. L. (2009). A minsky meltdown: Lessons for central bankers. presentation to the 18th annual hyman p. In *Minsky Conference on the State of the US and World Economies "Meeting the Challenges of the Financial Crisis", Levy Economics Institute of Bard College, New York City, April, Volume 16.*