Optimal Policy for Behavioral Financial Crises*

Paul Fontanier†
Harvard University

January 12, 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. I build on this theory to examine policy implications. First, the policymaker should use counter-cyclical capital buffers and time-varying loan-to-value ratios. These restrictions must be strengthened in times of over-optimism, as well as when the regulator is concerned that over-pessimism will arise in a future crisis. Second, uncertainty about the precise extent of behavioral biases in financial markets increases the incentives for the planner to act early. Finally, when biases depend on asset prices, monetary policy optimally complements macroprudential policy by leaning against the wind even when these macroprudential tools are unconstrained. Conventional monetary policy however loses power in normal times when agents expect the central bank to lean against the wind in the future.

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

*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 Martin Aragoneses, Malcolm Baker, Nicholas Barberis, Robert Barro, Francesca Bastianello, Jonathan Benchimol, Adrien Bilal, John Campbell, Gabriel Chodorow-Reich, Veronica De Falco, Maria Fayos Herrera, Adriano Fernandes, Xavier Gabaix, Nicola Gennaioli, Robin Greenwood, Antoine Jacquet, Spencer Yongwook Kwon, Sean Lee, Jean-Paul L'Huillier, Omeed Maghzian, Albert Marcet, Mathilde Muñoz, Kenneth Rogoff, Giorgio Saponaro, Suproteem Sarkar, Javier Suarez, Larry Summers, Adi Sunderam, Jean Tirole, Luis Viceira, Jonathan Wallen, Michael Woodford and participants at the 2019 Sloan-NOMIS Workshop on the Cognitive Foundations of Economic Behavior, Harvard Macroeconomics Lunch, Harvard Finance Lunch, and the CU-RIDGE 2021 Financial Stability Workshop for helpful comments. I am grateful to the Corps des Mines for supporting this research. All errors are my own.

†Email: fontanier.p@gmail.com. Website: https://scholar.harvard.edu/fontanier

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), this research was motivated in part by the growing evidence that factors such as credit growth and asset price booms successfully predict financial crises (Jordà, Schularick and Taylor 2015).[2] The behavioral view has also been supported by the findings from surveys that investors' beliefs are inconsistent with the Rational Expectations hypothesis. Such evidence generally points to the importance of extrapolation in financial markets (Gennaioli and Shleifer 2018).[3] In response, economists have developed a number of behavioral models of financial instability.[4] 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 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, how much information does the regulator need about behavioral biases to warrant early action? And third, should monetary policy be part of the toolbox, in that central banks should intervene by raising interest rates when asset prices soar?

I show first 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 asset prices impact biases. I also show that uncertainty about the precise extent of behavioral biases in financial markets increases the incentives for the planner to act early. Finally, I show that monetary policy should lean against the wind when high asset prices in good times trigger irrational pessimism in future crises.

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

[2] This predictability has been further documented by Greenwood, Hanson, Shleifer and Sørensen (2020): while financial crises in normal times only happen with a probability of 7% within three years, this figure reaches as high as 40% once conditioning on rapid credit growth and asset price booms.

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

[4] See 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.

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.[5]

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, financial frictions make all equilibrium variables dependent on the asset's payoff during a financial crisis: 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.

Behavioral biases during crises also have a direct impact: excess pessimism about the future payoffs of the collateral asset directly tightens the borrowing constraint, amplifying the severity of the crisis. The stark difference between exogenous sentiment – when behavioral biases only depend on exogenous variables – and endogenous sentiment – when equilibrium asset prices enter the determination of behavioral biases – manifests itself during these sentiment-driven financial crises. When behavioral biases are positively linked to asset prices, a financial crisis provokes a fall in the price of collateral assets, leading to irrational pessimism, and further tightens the borrowing constraint. Furthermore, the fall in consumption feeds back to asset prices through the stochastic discount factor, creating a further rise in pessimism, which feeds back to consumption through the collateral constraint again, and so on. This spiral effect exists independently of whether the collateral constraint features financial amplification or not. I call this new phenomenon *belief amplification*.

Behavioral biases during financial crises, especially biases that depend on asset prices, have critical implications for policy. I present the welfare analysis in Section 3, where a paternalistic social planner evaluates welfare using his own (rational) expectations, and recognizes that agents' expectations can be distorted in the future. I start by developing a general welfare decomposition, in the spirit of Dávila and Korinek (2018). The decomposition shows how behavioral factors and financial frictions interact to create first-order uninternalized welfare effects. This analysis clarifies that irrational over-optimism in booms creates welfare losses only when there is a chance that financial frictions bind in the future. Furthermore, it highlights how the predictable components of

[5] Externalities only arise when the price directly enters the collateral constraint. Thus, the equilibrium is constrained efficient when agents are rational and the price does not enter the constraint. See Ottonello, Perez and Varraso (2021), the discussion in Section 2.1, as well as Appendix C.

future behavioral biases formed inside a financial crisis also create losses and should be monitored. Indeed, if private agents tend to be over-pessimistic during financial crises, but neglect this future bias in good times, 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. I provide suggestive evidence that this is indeed the case: using empirical proxies for the tightness of the borrowing constraints of intermediaries, and forecast errors from survey data, I document that there is a strong negative comovement between these two measures during financial crises. Estimating these two objects does not necessarily require a quantification of contemporaneous irrationality, and is thus largely independent of the degree of over-optimism the social planner believes is present in good times.

The welfare decomposition delivers a second key insight. It shows that 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. This pecuniary effect is always operative, but in a rational case it does not affect welfare. Prices change, but since assets stay in the hands of intermediaries, allocations are unaffected.

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 second effect, called a *reversal externality*, that works through current prices. 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.

Notably, these externalities are still present even in the case where private agents are sophisticated and the planner shares the same beliefs. Atomistic intermediaries cannot coordinate in order to collectively reduce their leverage or decrease asset prices in order to alleviate the effects of future pessimism. 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. Only an intervention from the planner can solve these externalities, showing that belief differences between policymakers and market participants are not key for some of my results.

This decomposition has important implications for the conduct of optimal policy, which I develop in Section 4. The welfare decomposition implies that the second-best can be restored through intervention along three margins: (i) a tax on short-term borrowing, (ii) a tax on investment in collateral assets, and (iii) a policy that restrains asset price growth if the reversal externality is operative. Furthermore, my analysis provides the financial regulator with the properties of behavioral biases which need to be quantified in order to optimally calibrate these taxes: current irrational optimism, conditional expectation of future irrational pessimism inside a crisis, and the effect of asset prices on biases in the future.

It is however undeniable that identifying a bubble is intrinsically difficult since corresponding fundamentals are not observable. In his influential "Asset-Price Bubbles and Monetary Policy" speech, Bernanke (2002) forcefully exposed this issue and named it the "identification problem." 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. 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 sentiment-driven financial crisis.

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.[6] Furthermore, 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. Finally, capital buffers should be increased in times of heightened uncertainty about behavioral 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 endogenous sentiment nevertheless calls for the use of a third instrument in order to control asset prices and counter the reversal externality. Monetary policy is a natural candidate. I consider its optimal use in Section 6, using the insights obtained from the general welfare decomposition and adding nominal rigidities. I show that monetary policy can be used as a comple-

[6] Counter-cyclical capital buffers are at the center of the Basel III regulatory framework (Basel Committee on Banking Supervision 2011). My model shows how to optimally vary the levels of buffers when sentiment is fluctuating.

mentary tool.[7] Even when counter-cyclical capital buffers and LTV ratios can be flexibly adapted, an increase in the interest rate can be beneficial. By lowering contemporaneous asset prices, with endogenous sentiment monetary policy influences the future equilibrium determination. 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 contemporaneous biases. Fully rational prices today, which by definition are interest rate-sensitive, can still create behavioral biases in the future. 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.[8]

Systematically acting in this way can however have unintended consequences. Agents anticipating that the central bank will tighten monetary policy when asset prices soar weakens the central bank's traditional stimulus power. Indeed by cutting interest rates to achieve full employment, the central bank indirectly supports asset prices through the usual discount rate channel. This can cause agents to become over-optimistic regarding future prices through extrapolation. But agents now internalize that these high prices will be accompanied by an interest rate hike, inducing them to cut consumption and depress current aggregate demand through the substitution channel. When this feedback effect is strong enough, the central bank can hit the zero lower bound and fail to achieve full employment, while still feeding over-optimism with excessive asset prices. At the heart of this mechanism is a time-inconsistency: even if the central bank would rather commit to never lean against the wind, it will always be optimal to do so if asset prices become high enough.

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. These findings narrowed the set of theories that can explain why buoyant credit markets and asset price booms predict financial crises, and put behavioral explanations at the forefront.[9] Direct evidence of such

[7] Previous literature showed that monetary policy can be a substitute instrument when traditional macroprudential tools are constrained (Caballero and Simsek 2020a; Farhi and Werning 2020).

[8] An 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 disastrous welfare consequences. If that is the case, my model suggests that an interest rate hike is warranted.

[9] Recent work refined our understanding of this predictability, and identified many other predictive factors. Baron and Xiong (2017) and Richter and Zimmermann (2021) examine bank equity returns and profitability. Greenwood and Hanson (2013) focus on a measure of credit quality, and find that credit booms are accompanied by a deterioration of the average quality of corporate issuers, and that a high share of risky loans forecasts negative corporate bond returns. López-Salido, Stein and Zakrajšek (2017) demonstrate predictable mean-reversion in credit spreads, and that elevated credit-market sentiment predicts a decline in economic activity in the following years. Kirti (2018) and Krishnamurthy and Muir (2020) use interactive regression specifications by combining credit growth with a proxy for sentiment, and find results consistent with the idea that the interplay between leverage and mispricing is central. Mian, Sufi and Verner (2017) show that household debt is also a good predictor of future economic slowdowns, an indication that systematic

biases comes from survey data: Bordalo et al. (2018) document the predictability of forecast errors for analysts' expectations regarding the Baa bond – Treasury credit spread. Finally, Jordà et al. (2015) and Greenwood et al. (2020) show that combining credit growth measures with asset price growth substantially increases the out-of-sample predictive power on a subsequent financial crisis. These facts motivate my analysis, where behavioral distortions in asset markets spill over the entire credit sector. 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."[10,11]

My paper integrates these lessons into the traditional literature on normative macrofinance.[12] My framework follows from earlier work characterizing generic inefficiencies created by incomplete markets, starting with Geanakoplos and Polemarchakis (1985) and Greenwald and Stiglitz (1986). In my model, markets are incomplete because contingent bonds are not available, and the amount of borrowing is limited by the expectation of the asset's future payoffs, a friction similar to Kiyotaki and Moore (1997).[13] Most of the recent normative literature, like Mendoza (2010), Bianchi (2011) and Jeanne and Korinek (2019), uses a collateral externality 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. A different strand of the literature has been preoccupied by aggregate demand (rather than pecuniary) externalities and the need for macroprudential policy.[14] A general treatment is

extrapolation errors are not specific to the finance sector.

[10] Recent theoretical work introduced extrapolative expectations into financial frictions models, and in particular showed how behavioral biases allow standard models to match the observed behavior of credit spreads before crises (Bordalo et al. 2018; Greenwood et al. 2019; Maxted 2020; Krishnamurthy and Li 2020; Bordalo, Gennaioli, Shleifer and Terry 2021; Camous and Van der Ghote 2021). Other papers started integrating behavioral distortions into business cycle analysis, eg. L'Huillier, Singh and Yoo (2021) and Bianchi, Ilut and Saijo (2021). Chodorow-Reich, Guren and McQuade (2021) study housing, where improvement in fundamentals triggers a boom-bust-rebound driven by over-optimism. All of these papers use the diagnostic expectations mechanism of Bordalo et al. (2018). Agents learn about the fundamentals by observing "dividends" but become over-optimistic. They thus do not feature endogenous sentiment, a feature that has different implications for policy as I show in this paper.

[11] 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 (2017a), 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 when biases depend on asset prices) but does not study financial crises or optimal macroprudential and monetary policies.

[12] To provide a rigorous welfare analysis under behavioral distortions, I rely on a recent literature of "behavioral public finance," with Gruber and Köszegi (2001), O'Donoghue and Rabin (2006) and Mullainathan, Schwartzstein and Congdon (2012). Farhi and Gabaix (2020) provide a general treatment of optimal taxation with behavioral agents, and I use their result and their concept of a "behavioral wedge" to characterize uninternalized welfare effects.

[13] In my model, assets never change hands in equilibrium since all borrowers are identical. This is in contrast with the notion of "fire sales," developed first in Shleifer and Vishny (1992), where liquidation does not necessarily allocate assets to the highest value users. Dávila and Korinek (2018) call these "distributive" externalities, where redistribution of wealth between agents with different marginal rates of substitution creates an inefficiency. This includes, for instance, the models in Caballero and Krishnamurthy (2003), Lorenzoni (2008) and Fanelli and Straub (2021). Gromb and Vayanos (2002) is an example featuring both distributive and collateral externalities. Dávila and Korinek (2018) show that distributive externalities can lead to under- as well as over-borrowing, whereas my model features no room for policy in the rational version. This allows me to compare my results to a simple benchmark where *laissez-faire* is optimal.

[14] Examples include Eggertsson and Krugman (2012), Korinek and Simsek (2016) and Guerrieri and Lorenzoni (2017).

developed in Farhi and Werning (2016).

The last section of the paper considers the use of monetary policy to "lean against the wind." The 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).[15] Galí (2014) adds to this argument by showing that, in a rational bubble setup, increasing interest rates actually enhances bubble growth.[16] I study the spillovers created by rule-based leaning against the wind: when agents take into account that the central bank might raise rates in the future to tame behavioral biases, this impacts the regular conduct of monetary policy to stimulate aggregate demand. To the best of my knowledge, Boissay, Collard, Galí and Manea (2021) is the only work looking at these issues. They study rule-based leaning against the wind in a New Keynesian environment augmented with endogenous financial crisis.[17]

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 prices.[18] They focus on the effects of over-optimism in booms and how it implies a need for monetary policy when leverage restrictions are constrained, while most of my results are arising from the presence of biases *inside* crises. Second, Dávila and Walther (2021) study an environment with general belief distortions during the boom, and characterize optimal leverage and monetary policies. However their setup does not include leverage constraints. I also contribute to this line of research by providing an alternative way of modeling general belief distortions that allows for endogenous sentiment (e.g. biases can depend on asset prices), while Dávila and Walther (2021) restrict their analysis to exogenous probability measure distortions.[19] Third, Caballero and Simsek (2020a) also feature behavioral elements in the form of heterogenous beliefs, and study

[15] This is also related to the large literature on monetary policy and financial stability, that abstract from bubbles or irrational expectations. See Bernanke and Gertler (2000), Woodford (2012) and Gourio, Kashyap and Sim (2018) among others. Smets (2014) provides a clear review of this literature.

[16] There is a theoretical literature studying financial frictions while allowing for rational bubbles *à la* Tirole (1985). Farhi and Tirole (2012a) add rational bubbles in a dynamic environment with financially constrained firms. Bubbles can help alleviate a shortage of collateral, similar in spirit to my model where over-optimism helps overcome under-investment issues. This idea was also already present in the corporate finance literature, see Stein (1996) and Baker, Stein and Wurgler (2003). Martin and Ventura (2016) show in a similar setup that there is an "optimal" bubble size that maximizes long-run output, providing a new motive for macroprudential policy to align the equilibrium bubble with the optimal one. Biljanovska, Gornicka and Vardoulakis (2019) explicitly study optimal policy in such a setting and, consistent with my paper, find that policy should lean against the bubble more aggressively to mitigate the pecuniary externalities from a deflating bubble when constraints bind. I differ from these papers, and at the same time sidestep the indeterminacy issue that comes with rational bubbles, by deviating from rational externalities.

[17] I discuss the relation between my results and these recent papers in more details in Section 6.

[18] In the model of Farhi and Werning (2020), wages are rigid and a Zero Lower Bound binds during a crisis. Macroprudential and monetary policy are thus needed in their rational benchmark. By contrast, my model does not leave any room for policy when agents are rational. I build on their insights, but also allow for more general departures from rationality, an investment margin, as well as incomplete information about sentiment, the possibility of ex-post intervention, and the dynamic effects of monetary policy.

[19] 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) 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.

monetary policy when macroprudential policy is constrained. I build on their results and also complement them by showing that the central bank can raise interest rates even when macroprudential tools are fully unconstrained, in order to preventively tame future extrapolation.

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, starting with Lorenzoni (2008). To isolate the effects of behavioral biases, it features a borrowing constraint that *does not* create externalities in a rational equilibrium.[20] I introduce behavioral biases in Section 2.2. I close this section by characterizing the decentralized equilibrium.

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:[21]

$$U^b = \mathbb{E}_1 \left[\ln(c_1) + \beta \ln(c_2) + \beta^2 c_3 \right] \tag{1}$$

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:

$$U^h = \mathbb{E}_1 \left[c_1^h + \beta c_2^h + \beta^2 c_3^h \right]. \tag{2}$$

Financial Assets: There are two financial assets in the economy: deposits and the risky asset. Financial intermediaries issue deposits d_t 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 $c(H)$. The equilibrium price of the risky asset at $t = 1$, by no-arbitrage, is thus $q_1 = c'(H)$. This asset pays

[20] All 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 Appendix C and the discussion below.

[21] This functional form is adopted for simplicity. It allows for tractability in the equilibrium expressions during a financial crisis. Online Appendix N 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.

stochastic dividends z_2 and z_3 in future periods, drawn from independent cumulative probability distributions F_2 and F_3. Only financial intermediaries have the necessary human capital to hold risky assets.[22]

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 its future payoff. We assume that the collateral constraint takes the specific form:

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

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.

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

Assumption 1. *The financial friction parameter is small enough such that financial crises are possible:*

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

If the discount factor is lower than ϕ, then an additional unit of the risky asset is worth less than the additional borrowing capacity it brings, implying that borrowing is unlimited. Assumption refassum:betaphi thus ensures that the collateral constraint is not always slack.

Constraints: Financial intermediaries' constraints for their optimization are then as follows:

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

$$c_2 + d_1(1 + r_1) + q_2 h_2 \leq d_2 + (z_2 + q_2)h_1 \tag{6}$$

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

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

where H is the quantity of the asset created, h_1 is the quantity intermediaries decide to keep at $t = 1$, 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 similar. Financial intermediaries have an endowment e_1 in the initial period.

In order for financial intermediairies to always be able to repay their debt, I make the following parametric assumption.

[22] Although this is a rather stark assumption, it is consistent with the evidence presented for example by He, Khang and Krishnamurthy (2010), documenting that toxic MBS were always on the balance sheet of financial intermediaries during the 2008 financial crisis. What ultimately matters for my paper is that the financial intermediaries are the marginal holders of these assets. Haddad and Muir (forthcoming) provide further evidence suggesting that intermediaries are responsible for a large fraction of risk premium variation in various asset classes.

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

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

This ensures that even in the worst-case scenario, financial intermediaries can repay their debt. This is satisfied as long as the financial friction parameter is small enough, and that the minimum payoff of the asset in the last period is strictly positive.

The budget constraints of households are given by:

$$c_1^h + d_1 \leq e_1^h \tag{10}$$

$$c_2^h + d_2 \leq e_2^h + d_1(1 + r_1) \tag{11}$$

$$c_3^h \leq e_3^h + d_2(1 + r_2) \tag{12}$$

where e_t^h denotes the endowment, in consumption goods, of households at period t.

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

$$n_2 = z_2 H - d_1(1 + r_1). \tag{13}$$

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

The risky asset can be understood as any asset used as collateral for short-term debt by financial intermediaries. A favoured interpretation is that H represents real estate held by the financial sector: the dividends are then simply rents coming from these operations. The cost $c(H)$ then has a simple construction interpretation.[23] Alternatively, one can picture the intermediaries as a firm/bank coalition, where H represent C&I loans or simply projects funded by the intermediaries. H may also represent Mortgage-Backed Securities, complex products widely used in repo markets. In this last case, the costs $c(H)$ should be interpreted as securitization costs (legal fees, or the wages of structured traders for example).

I have made a number of simplifying assumptions in order to focus on the intuition underlying the mechanisms. Lenders have linear utility so that the interest rate is exogenously set, and there is no need to worry about market clearing for savings. Financial intermediaries have linear utility in the last period in order to simplify their pricing kernel in crises, and to be able to derive closed-form solutions even in the presence of behavioral biases.

[23] Online Appendix M presents a simple example of a microfoundation where heterogeneous entrepreneurs finance the construction of real estate projects by borrowing from financial intermediaries.

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

1. Financial intermediaries lack commitment to repay in the final period ;
2. Financial intermediaries must take the decision of whether to default before observing the realization of z_3 ;
3. 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.[24] 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 will be constrained-efficient when expectations are rational (see Section 3.1).

Most 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:

$$d_2 \leq \phi H q_2. \tag{14}$$

This type of collateral constraint is used for example in the work of Bianchi (2011), Bianchi and Mendoza (2018) and Jeanne and Korinek (2019). It can be microfounded by assuming that avoiding repayment requires diverting ressources in the current period, and this is perfectly observed by lenders.[25] Ottonello et al. (2021) show that the quantitative predictions of both types of constraint are similar. Without taking a stance on which type of constraint is more realistic, I focus on the future payout constraint in equation (8) since it cleanly isolates the effects of behavioral biases, and discuss in the paper and in Appendix C the robustness of the results to this alternative formulation.

2.2 Beliefs

I allow for a general class of deviations from rationality: a behavioral bias enters the pricing equation of the risky asset as a location shifter on expected dividends. Specifically, I model it as a constant that depends on the information available at time t, $\mathcal{I}_t$. The bias shifts the whole distribution of dividends expected at $t+1$. I denote it by:

$$\Omega_{t+1}(\mathcal{I}_t) \equiv \Omega_{t+1} \tag{15}$$

[24] The 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 $\mathbb{E}_2[z_3]$.

[25] An alternative, more behavioral, interpretation of the contemporaneous price collateral constraint is the following. Even if defaulting by borrowers implies that lenders only recover a fraction ϕ of the collateral assets in the next period, these assets are typically complex (see the discussion above). Lenders are unable to use complicated models to price these assets when they lend to the bank: they use their best guess for what could be the value in the future, and so they use the current market value. This results in equation (14).

Agents at t thus expect the dividend realization in the next period to be $z_{t+1} + \Omega_{t+1}$ instead of z_{t+1}. In that respect, Ω_{t+1} exactly represents the *predictable component* at t of forecast errors realized at $t+1$. The bias enters the pricing equation at t in the following way:

$$q_t = \beta \mathbb{E}_t \left[\frac{\lambda_{t+1}(z_{t+1} + \Omega_{t+1})}{\lambda_t} \left(z_{t+1} + \Omega_{t+1} + q^r_{t+1}(z_{t+1} + \Omega_{t+1}) \right) \right] \tag{16}$$

where $q^r_{t+1}(z_{t+1} + \Omega_{t+1})$ is the price that would prevail, at $t+1$, in a rational environment where the state of the world (i.e. the dividend) realizes at $z_{t+1} + \Omega_{t+1}$. The r superscript makes it clear that this price will *not* necessarily occur even in the event that the realized dividend is indeed $z_{t+1} + \Omega_{t+1}$: if agents' future selves are also subject to behavioral biases, the price q_{t+1} will feature a term Ω_{t+2}. Importantly, I assume that agents fully neglect that other agents, and themselves, will be subject to behavioral biases in the future.[26]

Throughout the paper, I use a streamlined notation:

$$q_t = \beta \mathbb{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t} \left(z_{t+1} + \Omega_{t+1} + q^r_{t+1} \right) \right] \tag{17}$$

where the dependence of the stochastic discount factor λ_{t+1}/λ_t and of the price on the behavioral bias are kept implicit.

The bias can potentially depend on several variables (z_{t-i}, q_{t-i}, or sunspot shocks s_{t-i}).[27] This approach is particularly flexible for the subsequent welfare analysis, since it summarizes all possible distortions in a single quantity. Online Appendix P shows a formal correspondence between this approach and the commonly used framework in the literature where agents use a distorted probability measure regarding the exogenous state of the world.

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

Throughout the paper the bias Ω is kept general, which highlights the properties of sentiment that matter for welfare. It will be useful to flesh out specific examples to build intuition, however. In particular, 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.[28]

[26] Some of the results on optimal policy are robust to agents realizing that the market will be subject to sentiment in the future. I discuss this robustness in Section 4.1.

[27] In Section 6.3, I also entertain the possibility that current sentiment depends directly on the values of past sentiment, for example if the bias is slow-moving or mean-reverting. This possibility is not particularly insightful in a 3-period model, but becomes interesting once the dynamic build-up of sentiment is a concern.

[28] A particularly clear survey of this literature can be found in Barberis (2018).

Fundamental Extrapolation: This case captures models where investors extrapolate fundamentals, here 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) to explain long-term reversal and the value premium in the cross-section. This extrapolation can come from a variety of psychologically founded biases. Constraints 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) and Bordalo, Gennaioli, La Porta and Shleifer (2019) link extrapolative beliefs about fundamentals to the representativeness heuristic. In Rabin and Vayanos (2010), extrapolative beliefs stem from believing in the law of small numbers.

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 $\Omega_2 = \alpha(z_1 - z_0)$, 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 of Gennaioli and Shleifer (2018) is thus represented by fundamental realizations $z_1 > z_0$ (good news at $t = 1$) followed by $z_2 < z_1$ (disappointment).

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 models include papers by DeLong, Shleifer, Summers and Waldmann (1990), Hong and Stein (1999) and Barberis and Shleifer (2003). Recent research leverages the use of survey data to motivate price or return extrapolation, as in Cassella and Gulen (2018). 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. Bastianello and Fontanier (2021) propose a microfoundation for price extrapolation, where agents extract information from prices using a misspecified model of the world. 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 the price of the risky asset and the behavioral bias are thus determined *jointly*: q_t depends on Ω_{t+1}, which itself depends on q_t. Solving for the equilibrium thus requires solving a fixed-point problem between outcomes and beliefs.[29] Most importantly, agents' present and *future* beliefs now move with policies that influence asset prices (a potential channel for monetary policy, as studied in Section 6).[30]

Other Models: While the core of the paper focuses on these two cases, other behavioral models can be nested by the Ω formulation, such as inattention. I present and discuss several cases in Online Appendix J.

2.3 Equilibrium

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

Households: Households are passive throughout the three periods, and their only role is to pin down the rate of interest through their Euler equation:

$$\beta(1 + r_t) = 1 \tag{20}$$

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.

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

$$\lambda_2 = \frac{1}{c_2} = \mathbb{E}_2[\lambda_3] = 1 \tag{21}$$

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:

$$q_2 = \beta\mathbb{E}_2[z_3 + \Omega_3] \tag{22}$$

Crisis: In this case the collateral constraint is binding. The Lagrange multiplier on the collateral constraint, κ_2, is therefore given by:

$$\kappa_2 = \lambda_2 - 1 > 0 \tag{23}$$

[29] While it is entirely possible that agents only extrapolate *past* price changes, in the present setup it would not deliver insightful results, because of the three-period structure.

[30] Notice 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 Bastianello and Fontanier (2021) or Chahrour and Gaballo (2021). Subjective prices are then also distorted since agents form expectations in a consistent manner. Online Appendix J also discusses the case where the behavioral bias is distorting expected *prices*, while expectations of fundamentals stay rational. This is for example the case if agents are "internally rational" as in Adam and Marcet (2011). The externalities I am uncovering disappear in the benchmark case where the collateral constraint depends on the expectation of fundamentals, but the results are similar when it directly depends on equilibrium asset prices.

which directly quantifies the severity of the crisis: it encodes how far we are from the unconstrained equilibrium. The asset price is given by:

$$q_2 = \beta c_2 \mathbb{E}_2[z_3 + \Omega_3] + \phi(1 - c_2)\mathbb{E}_2[z_3 + \Omega_3] \tag{24}$$

where the last term is a *collateral premium*, illustrating that holding marginally more of the asset is valuable since it relaxes financial constraints.[31] Consumption, on the other hand, is directly coming from the budget constraint of financial intermediaries (6), since agents are against the collateral constraint:

$$c_2 = z_2 H - d_1(1 + r_1) + \phi H \mathbb{E}_2[z_3 + \Omega_3]. \tag{25}$$

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. There is another effect on equilibrium, however, when Ω_3 depends on q_2. When that is the case, the equilibrium must be determined through a fixed-point problem: consumption depends on sentiment, consumption determines asset prices, and asset prices determine sentiment. I now go over three benchmark cases in detail to build intuition: the Rational Expectations Equilibrium, exogenous Ω_3, and price-dependent $\Omega_3(q_2)$.

Rational Equilibrium at $t = 2$: In the event where $\Omega_3 = 0$, the equilibrium is determined by:

$$c_2 = n_2 + \phi H \mathbb{E}_2[z_3] \tag{26}$$

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

While the determination of the equilibrium is trivial in this case, a graphical representation helps fixing ideas, and most of all facilitates the distinction with the behavioral case. Figure 1 illustrates how a shock to net worth does not trigger any amplification: the fall in consumption is commensurate to the size of the shock to net worth, since as can be seen from equation (27), $dc_2/dn_2 = 1$.

Exogenous Sentiment at $t = 2$: In the case where Ω_3 is exogenously set, the budget constraint equation is still sufficient to obtain the consumption level in a crisis. It simply shifts consumption by a constant relative to the REE benchmark. The effect on asset prices is more severe: pessimism impacts the stochastic discount factor, c_2, and the expectation of future prospects. But this drop in asset prices still does *not* spill back to consumption. Figure 2 illustrates this equilibrium determination with a $\Omega_3 < 0$. Exogenous pessimism makes the pricing condition steeper, but consumption is

[31] Technically, the collateral premium comes from the dependence on h of the collateral constraint, $d_2 \leq \phi h \mathbb{E}_2[z_3 + \Omega_3]$. This constraint is associated with the Lagrange multiplier κ which, when multiplied by c_2, yields the $(1 - c_2)$.

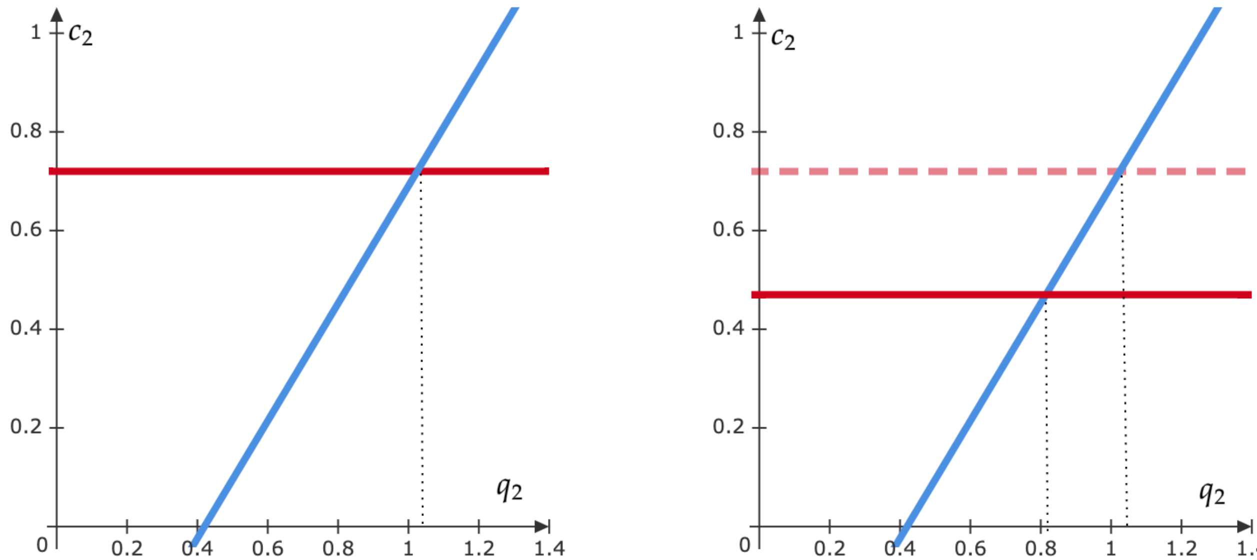

Figure 1: Graphical Illustration of REE Determination at $t = 2$. The red line represents the budget constraint equation (25), and the blue line represents the pricing equation (24). The right panel illustrates how the equilibrium shifts after an exogenous shock to net worth n_2.

pinned down independently. Specifically, the asset price is given by:

$$q_2 = \beta(n_2 + \phi H \mathbb{E}_1[z_3 + \Omega_3])\mathbb{E}_1[z_3 + \Omega_3] + \phi(1 - n_2 - \phi H \mathbb{E}_1[z_3 + \Omega_3])\mathbb{E}_1[z_3 + \Omega_3] \quad (28)$$

Endogenous Sentiment at $t = 2$: When the behavioral bias Ω_3 depends on equilibrium prices q_2, the budget constraint is not enough 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 3, and the left panel illustrates the rational benchmark for comparison.

Figure 3 shows the presence of a new feature that I call *belief amplification*.[32] 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.[33] The arrow on Figure 3 illustrates the further contraction in consumption levels c_2 due to this belief amplification.[34]

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

[33] The idea that irrational pessimism is a key aspect of financial crises, and works similarly to financial amplification, dates back to Fisher (1932), who laid out the theory of debt-deflation mechanism which is now part of a large number of financial crisis narratives. Fisher also mentions endogenous pessimism as a key feature: *"All of the down movements [...] have psychological effects. Already we have seen that shrinking net-worth leads to distress selling. But distress selling implies distress."*

[34] As can be seen in Figure 3, the equilibrium is ensured to be unique in the exogenous sentiment case. This is not immediate

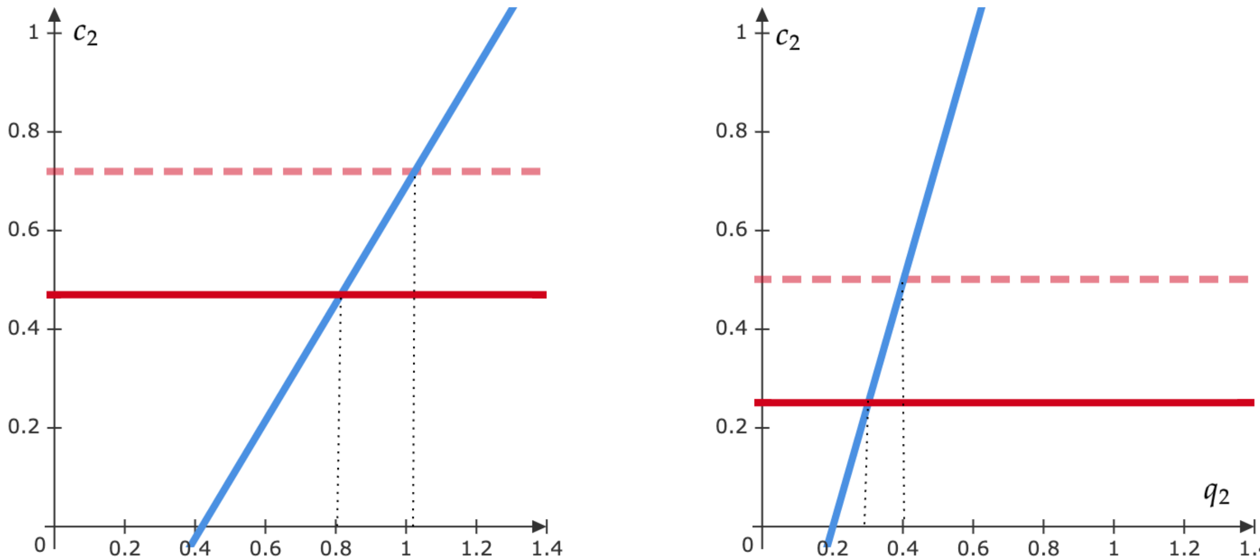

Figure 2: Graphical Illustration of Equilibrium Determination with exogenous sentiment at $t = 2$. The red line represents the budget constraint equation (25), and the blue line represents the pricing equation (24). The left panel illustrates how the equilibrium shifts after an exogenous shock to net worth n_2 when agents are rational. The right panel illustrates the same experiment but with an exogenous $\Omega_3 < 0$.

This belief amplification channel has important implications for welfare, and can also be understood analytically by studying the properties of prices in response to changes in net worth. The price *sensitivity* to changes in net worth at $t = 2$ can be expressed as:

$$\frac{dq_2}{dn_2} = \frac{(\beta - \phi)\mathbb{E}_2[z_3 + \Omega_3]}{1 - (\phi + (\beta - \phi)(2c_2 - n_2))\frac{d\Omega_3}{dq_2}}. \tag{29}$$

The derivative $d\Omega_3/dq_2$ magnifies the sensitivity of asset prices to changes in net worth, as long as $d\Omega_3/dq_2 > 0$ (which is the relevant case: agents become more optimistic when current asset prices are more elevated). This expression also makes clear that this is an amplification mechanism, since the denominator represents infinite rounds of feedback effects.[35] There is a second effect, working through the numerator. Here, over-pessimism ($\Omega_3 < 0$) *decreases* the sensitivity of asset prices to changes in net worth. It becomes a quantitative question whether this sensitivity is higher or lower than in a rational model. Nonetheless, this sensitivity has repercussions for the level of consumption *only* in the behavioral model where these price changes impact the collateral constraint.[36]

anymore for endogenous sentiment. In Online Appendix L I show how linear forms of price extrapolation 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.

[35] Simply because $1/(1 - x) = 1 + x + x^2 + ...$

[36] This is not the case anymore in the price-collateral constraint, as in Appendix C. There, these two countervailing effects alter the size of the pecuniary externality.

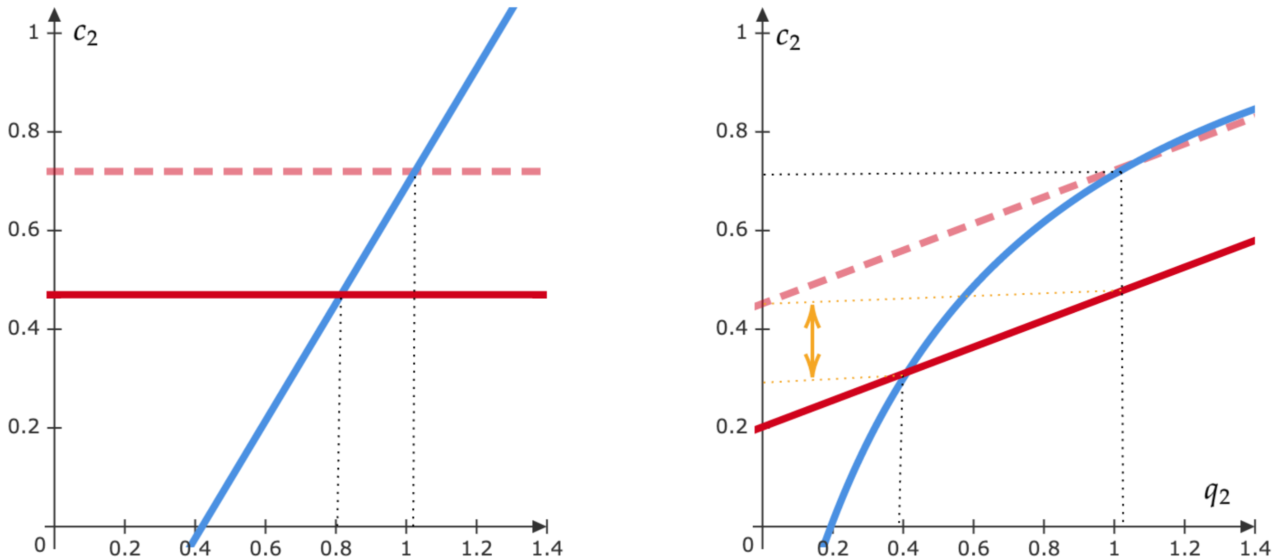

Figure 3: Graphical Illustration of Equilibrium Determination with exogenous sentiment at $t = 2$. The red line represents the budget constraint equation (25), and the blue line represents the pricing equation (24). The right panel illustrates how the equilibrium shifts after an exogenous shock to net worth n_2 when agents are rational. The left panel illustrates the same experiment but with an endogenous $\Omega_3(q_2) = \alpha(q_2 - q_1)$.

Crisis cutoff: I briefly characterize the occurrence of financial crises in this model. At the limiting state z_2 that delimitates the crisis region, the non-constrained Euler equation holds, the non-constrained pricing equation holds, and borrowing is at the limit. These conditions thus imply:

$$1 = z_2 H - d_1(1 + r_1) + \phi H \mathbb{E}_2(z_3 + \Omega_3) \tag{30}$$

which defines the crisis cutoff state:[37]

$$z^* = \frac{1 + d_1(1 + r_1) - \phi H \mathbb{E}_2(z_3 + \Omega_3)}{H} \tag{31}$$

where Ω_3 is the bias agents hold at this state of the world.[38] As can readily be seen from this expression, the cutoff is increasing with the level of outstanding debt (which will itself be increasing in initial optimism Ω_2) and with pessimism. Crucially, this cutoff is not used by private agents to compute the probability of a crisis to happen in the future. The objective probability of a crisis happening at time $t = 2$ is:

$$F_2\left(\frac{1 + d_1(1 + r_1) - \phi H \mathbb{E}_1(z_3 + \Omega_3)}{H}\right) \tag{32}$$

[37] I assume that the cutoff is unique. This is a natural assumption and is ensured as long as Ω_3 is increasing in z_2 and q_2 for example.

[38] In this expression Ω_3 is kept general. If Ω_3 depends on z_2 or q_2, this becomes a fixed-point equation implicitly defining the cutoff, as before. At the exact cutoff, we have $q_2 = \beta \mathbb{E}_2(z_3 + \Omega_3)$ since $c_2 = 1$ and hence the collateral premium disappears.

Instead, agents neglect their future bias Ω_3, but have a current bias Ω_2. Therefore they believe that the probability of a future crisis is:

$$F_2\left(\frac{1+d_1(1+r_1)-\phi H\mathbb{E}_1(z_3)}{H}-\Omega_2\right) \tag{33}$$

Both Ω_2 and Ω_3 contribute to the difference between the objective and perceived probability of a crisis happening. Notice also how d_1 is increasing in Ω_2 in equilibrium.[39]

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

$$1=\mathbb{E}_1\left[\frac{\lambda_2}{\lambda_1}\right] \tag{34}$$

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

$$q_1=c'(H)=\beta\mathbb{E}_1\left[\frac{\lambda_2}{\lambda_1}(z_2+\Omega_2+q_2^r)\right]. \tag{35}$$

Because consumption inside a crisis, c_2, depends directly 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 over-consumption at $t = 1$ relative to the rational benchmark, financed by borrowing (so a higher d_1). This gap between the expected c_2 and the realized one is driven as well by *future* sentiment. A predictable $\Omega_3 < 0$ at $t = 2$ leads to a collapse in the expectation of future payoffs, so a tighter collateral constraint, translating into even lower consumption than expected by behavioral agents.[40]

3 Welfare Analysis

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

1. The social planner takes uninternalized general equilibrium effects into account ;
2. The social planner computes expectations using rational expectations.

In this setup, the full information rational expectations hypothesis assumes that the planner uses the same probability distribution as agents, but recognizes that they are subject to a bias Ω_2, and

[39] These expressions highlight the role of sentiment as a "trigger" of crises, while models with only financial frictions have to rely on exogenous shocks. Stein (2021) argues that it is necessary to integrate the two approaches to understand if, quoting his title, "policy can tame the credit cycle".

[40] Even if agents do not have a behavioral bias Ω_2 at $t = 1$, the possible future presence of Ω_3 makes it easier to work with Ω-biases rather than with a distorted probability measure.

that they *will be* subject to a bias Ω_3 in the future. The way this bias will be determined, i.e. its dependence on future and past fundamentals and prices, is known to the Planner.[41] I adopt the notation $\mathbb{E}^{SP}$ to denote expectations formed according to this process. For instance, when agents are optimistic they expect their marginal utility of consumption at $t = 2$ to be lower than in reality, leading to the following relation:

$$1 = \mathbb{E}_1\left[\frac{\lambda_2}{\lambda_1}\right] < \mathbb{E}_1^{SP}\left[\frac{\lambda_2}{\lambda_1}\right] \tag{36}$$

Similarly, even if agents are correct about z_2, i.e. $\Omega_2 = 0$, but the social planner believes that in all states of the world at $t = 2$ there will be a bias $\Omega_3 < 0$, this inequality will hold.

3.1 Externalities

Before proceeding to the full-fledged welfare analysis, it is instructive to understand why, in the rational version of this model, there are no externalities.

The welfare function used at time $t = 2$, $\mathcal{W}_2(n_2, H; q_2, z_2)$, is given by:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]\right) + \beta^2\left(\mathbb{E}_2[z_3]H - \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}_2[z_3]H + n_2\right) & \text{otherwise} \end{cases} \tag{37}$$

The difference is that the social planner realizes that in financial crises ($z_2 \leq z^*$) this welfare function encodes general equilibrium effects through the price q_2, while a private agent only considers changes in d_1 and H, taking q_2 as given. This leads to pecuniary externalities in models with contemporaneous prices in the collateral constraint, but not here. Indeed, private agents have a first-order condition on borrowing such as:

$$u'(c_1) = (1 + r_1)\mathbb{E}_1\left[\frac{\partial \mathcal{W}_2}{\partial n_2}\right] \tag{38}$$

while the social planner has an extra-term corresponding to the pecuniary impact of private borrowing decisions:

$$u'(c_1) = (1 + r_1)\mathbb{E}_1^{SP}\left[\frac{\partial \mathcal{W}_2}{\partial n_2} + \frac{\partial \mathcal{W}_2}{\partial q_2}\frac{\partial q_2}{\partial n_2}\right] \tag{39}$$

and similarly for investment, since q_2 depends indirectly on H and n_2 (see Section 2.3 for the full description of equilibrium prices).

This change in equilibrium prices, however, has no impact on welfare. Consumption levels at $t = 2$, in the rational benchmark, are set independently of prices q_2. In other words, $\partial \mathcal{W}_2 / \partial q_2 = 0$ and the rational equilibrium is *constrained efficient*.

I now turn to the welfare analysis with behavioral biases. The equilibrium is constrained in-

[41] Section 5 explores the implications of incomplete information about Ω for the social planner.

efficient because of the difference in expectations between agents and the planner. Furthermore, *pecuniary* externalities can arise, in contrast to the rational benchmark.

3.2 Welfare 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.

3.2.1 Leverage

I start by analyzing how changes in debt d_1, fixing all others variables at $t = 1$, affect the welfare of individual agents.

Proposition 1 (Uninternalized Effects of Leverage)**.** *The uninternalized first-order impact on welfare when the level of short-term debt is marginally increased is given by:*

$$\mathcal{W}_d \quad = \quad \underbrace{\left(\mathbb{E}_1[\lambda_2] - \mathbb{E}_1^{SP}[\lambda_2]\right)}_{\mathcal{B}_d} \quad - \quad \underbrace{\mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \frac{dq_2}{dn_2}\right]}_{\mathcal{C}_d}. \tag{40}$$

Proof. All proofs are provided in Appendix A. □

The first term of equation (40) is the *behavioral wedge*. It is the difference between agents' perceived future marginal utility and what the social planner expects this marginal utility to be. When agents are over-optimistic, they expect their marginal utility to be, on average, lower than what a rational agent would expect. As a consequence, this difference is strictly negative in this case, meaning that a marginal increase in leverage has a negative first-order impact on welfare. This object is similar to the behavioral wedge of Farhi and Gabaix (2020), and is central to their analysis. The second term is a novel collateral externality that works through *future beliefs* and *future prices*. It is operative even though, as explained earlier, there is no collateral externality in the rational benchmark. I now explore these two terms in detail.

Behavioral Wedge: The strength of the behavioral wedge is driven by two separate forces (i) the difference in expected severity of crisis, and (ii) different expected probability 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 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. 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 $\mathcal{B}_d$ for short-term debt can be expressed as:*

$$\mathcal{B}_d \simeq \underbrace{-\boldsymbol{\Omega_2} H \mathbb{E}_1^{SP}\left[\boldsymbol{\lambda_2^2} \mathbb{1}_{\kappa_2>0}\right]}_{(i)} + \underbrace{\phi H \mathbb{E}_1^{SP}\left[\boldsymbol{\Omega_3 \lambda_2^2} \mathbb{1}_{\kappa_2>0}\right]}_{(ii)} \tag{41}$$

The first term quantifies the welfare losses from the 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, a recurring theme of this paper.

The second term quantifies welfare changes emanating from the predictable behavior of future deviations form rationality, Ω_3. Once again, predictable pessimism in the future is not enough to generate 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.[42]

Remark 2 (Sentiment and Financial Frictions). The welfare decomposition of equation (41) is always null when there is no possibility of crises in the future, as can be seen by the presence of the terms $\mathbb{1}_{\kappa_2>0}$ (remembering that κ_2 is the Lagrange multiplier on the collateral constraint, and is thus zero outside a crisis, since it is defined as an event where the collateral constraint binds). Of course, a model without any financial friction and behavioral deviations from rationality would generate *some* welfare losses, but these are higher-order than the terms presented above and thus negligible, assuming that the biases are quantitatively small. This feature can be seen as a direct application of the envelope theorem: in frictionless models, all agents are on their first-order condition for consumption. Accordingly, small perturbations to any parameter (including a perturbation of their expectations) do not have first-order welfare consequences. In this model, this does not apply since in a crisis agents are *not* on their first-order condition for consumption, and κ_2 quantifies the distance to the frictionless benchmark. One can also interpret terms like $\Omega_2 H \mathbb{E}^{SP}\left[\lambda_2^2 \mathbb{1}_{\kappa_2>0}\right]$ as the product of the mistake (Ω_2) by the cost of making the mistake (the expectation term). Without financial frictions the cost goes to 0, so mistakes are benign.

Remark 3 (Predictable Losses). An interesting case in point of equation (41) 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

[42] One might find it surprising that the difference in expected probability of crises does not enter this expression. This is because this term is negligible at the first-order: see Appendix A.2 for details.

period $t = 1$. This implies that is it not necessary for the social planner to know the current state of irrational exuberance to justified to act pre-emptively: knowing that agents will be pessimistic in bad states of the world is enough. Moreover, suggestive evidence supports the assumption that $\mathbb{E}^{SP}[\Omega_3 \lambda_2 \mathbb{1}_{\kappa_2 > 0}]$ is a significantly negative number. Figure 4 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) which computes an intermediary capital ratio.[43] The inverse of this capital ratio is proportional to λ_2 when agents have log-utility, as in this model. 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.[44,45] Figure 4 shows how Ω_3 is consistently negative in crises. In 2008, forecast errors Ω_3 crashed while marginal utility of intermediaries λ_2 spiked, suggesting sizeable welfare losses. A key point is that in events such as the dot-com bubble burst, pessimism was not accompanied by declines in financial health of intermediaries, and the theory I am developing suggests that these events are less of a concern for welfare.[46]

Collateral Externality: The second, and novel to the literature, term is a pecuniary externality that works through beliefs. This is the first paper, to the best of my knowledge, to identify that beliefs can give rise to such an externality. Let us examine in detail the terms composing this externality:

$$\mathcal{C}_d = -\mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \frac{dq_2}{dn_2}\right]. \tag{43}$$

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

[43] He et al. (2017) define the aggregate capital ratio η_t of the intermediary sector as the ratio of "aggregate value of market equity divided by aggregate market equity plus aggregate book debt of primary dealers:"

$$\eta_t = \frac{\sum_i \text{Market Equity}_{i,t}}{\sum_i \left(\text{Market Equity}_{i,t} + \text{Book Debt}_{i,t}\right)}. \tag{42}$$

Since only intermediaries can hold risky assets the capital ratio measures the wealth share of the intermediary sector. "Primary dealers" are designated by the NY Fed to serve as counterparties in the implementation of monetary policy. Most of them are large commercial banks.

[44] Forecast errors in the literature are traditionally defined as $FE_{t,t+1} = z_{t+1} - \mathbb{E}_t[z_{t+1}]$, so I multiply the time series by -1 to recover my definition of Ω_{t+1}.

[45] There is obviously no perfect measure of sentiment. I perform the same exercise for a variety of indicators that have been used in the literature to measure sentiment, and all of them depict the same variations in sentiment along the business cycle. These additional empirical results are presented in Online Appendix G

[46] In his famous "Irrational Exuberance" speech of 1996, Alan Greenspan alluded to this crucial interaction: "We as central bankers need not be concerned if a collapsing financial asset bubble does not threaten to impair the real economy [...]. But we should not underestimate or become complacent about the complexity of the interactions of asset markets and the economy. Thus, evaluating shifts in balance sheets generally, and in asset prices particularly, must be an integral part of the development of monetary policy."

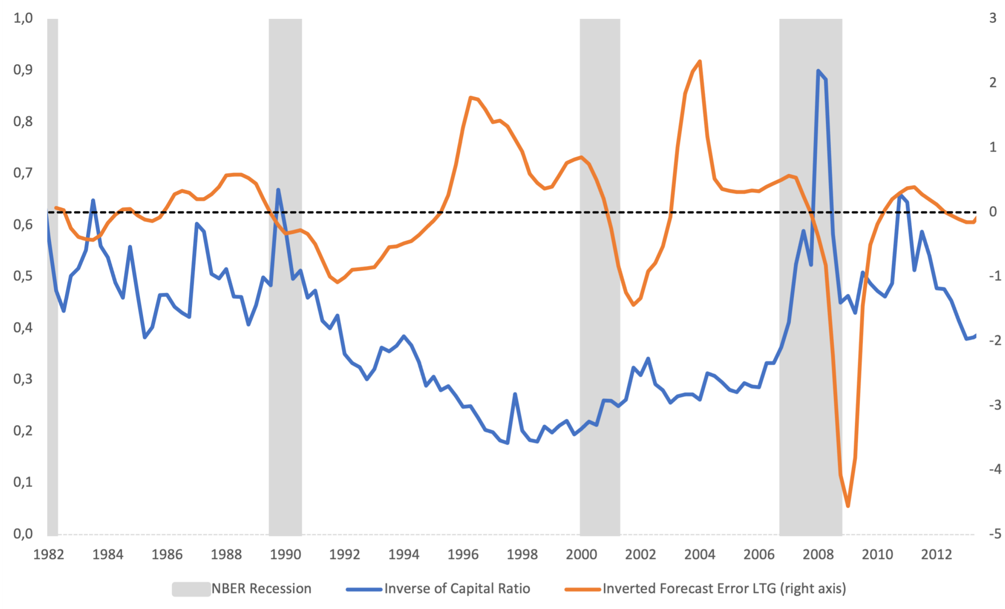

Figure 4: Time-series variation of proxies for λ_2 and Ω_3. For the financial health of intermediaries λ_2, I rely on He et al. (2017) which computes an intermediary capital ratio. The inverse of this capital ration is proportional to λ_2 when agents have log-utility, as in this model. For Ω_3, I use the forecast errors made by stock market analysts on the long-term growth of stocks, a measure of Bordalo et al. (2020)

prices in the future. 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 constrain 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 that $d\Omega_3/dq_2 \neq 0$. 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) does 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 negative: the private solution features excessive borrowing.

Finally, notice 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*:

Proposition 3 (Price Sensitivity With Sentiment)**.** *A change in net worth in period $t = 2$ impacts equi-*

librium asset prices as:

$$\frac{dq_2}{dn_2} = \frac{(\beta - \phi)\mathbb{E}_2[z_3 + \Omega_3]}{1 - (\phi + (\beta - \phi)(2c_2 - n_2))\frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}}} \tag{44}$$

Relative to a rational benchmark, where $\Omega_3 = 0$ and $d\Omega_3/dq_2 = 0$, sentiment creates two countervailing forces. First, over-pessimism ($\Omega_3 < 0$) makes the asset price less sensitive to changes in net worth, reducing the size of this sensitivity. Second, a positive change in net worth leads to a change in price through the stochastic discount factor c_2, which itself can lead to alleviating pessimism, supporting asset prices. This makes the price more sensitive to changes in net worth when $d\Omega_3/dq_2 > 0$.

Remark 4 (Collateral Externalities in the Literature). Dávila and Korinek (2018) use the term collateral externalities to externalities that apply when "agents are subject to a binding constraint that depends on aggregate variables". Here, the aggregate variable is the behavioral bias Ω_3, which is why I use this terminology. My paper thus shows that such externalities can still arise even if the contemporaneous price is not part of the collateral constraint, thus bridging the gap between these two commonly used models (see Ottonello et al. 2021). Appendix C further demonstrates that this effect is still present when contemporaneous prices are part of the collateral constraint: the two effects of belief and financial amplification now compound each other.

3.2.2 Investment

I perform the same type of welfare decomposition, but looking at a marginal increase in investment into the creation of collateral assets, keeping fixed d_1 and q_1. Results are similar to the borrowing case, but the sign can be ambiguous.

Proposition 4 (Uninternalized Effects of Investment)**.** *The uninternalized first-order impact on welfare when the level investment is marginally increased is given by:*

$$\mathcal{W}_H = \underbrace{\left(\beta\mathbb{E}_1^{SP}\left[\lambda_2(z_2 + q_2)\right] - \lambda_1 q_1\right)}_{\mathcal{B}_H} + \underbrace{\beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}}\left(\frac{dq_2}{dn_2}z_2 + \frac{dq_2}{dH}\right)\right]}_{\mathcal{C}_H} \tag{45}$$

composed of three distinct effects: a behavioral wedge $\mathcal{B}_H$, and collateral externality $\mathcal{C}_H$.

I again explore these two terms in turn.

Behavioral Wedge: Similar to the welfare costs of higher leverage, the behavioral wedge for investment is given by the difference between a rational valuation of the risky asset and private agents' valuation. This wedge is obviously negative when agents are over-optimistic, or when agents do no realize their future over-pessimism. As previously, we can approximate this behavioral wedge for small deviations from rationality, as in the following Proposition.

Proposition 5 (Behavioral Wedge Approximation for H)**.** *If Ω_2 and Ω_3 are small state-by-state, the behavioral wedge $\mathcal{B}_d$ for investment in the collateral asset can be expressed as:*

$$\mathcal{B}_H = \mathbb{E}_1^{SP}\left[\mathcal{B}_d(z_2)(z_2+q_2^r)\right] - \beta\Omega_2\mathbb{E}_1^{SP}\left[\lambda_2^r\left(1+(\beta-\phi)\right)Hz_3\mathbb{1}_{\kappa_2>0}\right] + \beta\mathbb{E}_1^{SP}\left[\Omega_3\lambda_2^r\frac{dq_2}{dz_3}\mathbb{1}_{\kappa_2>0}\right] \quad (46)$$

where $\mathcal{B}_d(z_2)$ is the behavioral wedge for leverage, from Proposition 2, for a realization z_2 of the dividend process at $t=2$:

$$\mathcal{B}_d(z_2) = \beta(\Omega_3(z_2)-\Omega_2)\lambda_2^2(z_2)\mathbb{1}_{\kappa_2(z_2)>0}. \quad (47)$$

This wedge is more complicated than in the borrowing case, $\mathcal{B}_H$, because the effects of belief distortions are impacting the stochastic discount factor as well as the expectation of future dividends and prices, which adds effects. The intuition, however, is similar: the wedge is negative when agents are irrationally exuberant ($\Omega_2>0$), when they will predictably be irrationally distressed ($\Omega_3<0$), and only if there is a probability of a financial crisis next period ($\mathbb{1}_{\kappa_2>0}$).

Collateral Externality: As with leverage, there is a collateral externality for investment even though the rational benchmark does not feature such an effect. It again works through the interaction of financial frictions, *future* asset prices, and sentiment:

$$\mathcal{C}_H = \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}}\left(\frac{dq_2}{dn_2}z_2+\frac{dq_2}{dH}\right)\right] \quad (48)$$

where dq_2/dH is the derivative of the asset price with respect to H, keeping net worth n_2 constant.

More surprisingly, this collateral externality is generally going in the opposite direction. Indeed, 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.[47] In turn, this ameliorates the borrowing capacity of the whole economy, thus improving welfare.

I already showed how the sensitivity of the price with respect to net worth was changed by sentiment. Similarly, how equilibrium prices move with the aggregate stock of collateral asset is changed by the behavioral wedges in a analogous way:

$$\frac{dq_2}{dH} = \frac{(\beta-\phi))z_2+\phi\mathbb{E}_2[z_3+\Omega_3])\mathbb{E}_2[z_3+\Omega_3]}{1-(\phi+(\beta-\phi)(c_2-\phi H\mathbb{E}_2[z_3+\Omega_3]))\frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}}} \quad (49)$$

where belief amplification, $d\Omega_3/dq_2$, is still at play.

Remark 5 (Exuberance and Under-investment). Unlike for the uninternalized welfare effects of increasing leverage, the behavioral wedge and the collateral externality for investment are going in *opposite* directions when agents are too optimistic at $t=1$ and when $d\Omega_3/dq_2>0$. Irrational exuberance

[47] Remember that we assumed positive dividends in all states of the world, $z_t>0$. In cases where dividends are negative this externality can become negative. This can for example be the case if collateral assets are *draining* liquidity in bad times.

leads agents to invest more than in the Rational Expectations equilibrium, which helps overcome the under-investment problem coming from financial frictions and the price-dependence of sentiment in a crisis. This is reminiscent of Martin and Ventura (2016), where the presence of bubbles alleviates financing frictions. 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.[48]

Remark 6 (Unambiguous Sign for Large Exuberance). The size of the collateral externality is bounded when Ω_2 increases, while the behavioral wedge is unboundedly negative when $\Omega_2 \to +\infty$. Hence the welfare impact of an additional unit of investment is unambiguously negative for large enough Ω_2.

3.2.3 Prices

In most models, like rational models or models with exogenous sentiment, the above two uninternalized effects are enough to characterize the efficiency of the equilibrium. Indeed, once allocations are set the equilibrium level of prices has no effect on welfare. To understand why, it is useful to take a step back and look at welfare maximization in a rational model. The social planner is maximizing a welfare function of the form:

$$\mathcal{W}\big(d_1, H; z_2, z_3\big).$$

As is obvious from inspecting any first-order conditions resulting from this maximization, it is enough for the social planner to impose a its specific desired allocation of (d_1, H). Price or quantity regulations, or any type of nudges are substitutable as long as the desired allocation is achieved. In the presence of exogenous sentiment, the problem takes a similar form:

$$\mathcal{W}\big(d_1, H; z_2, z_3 + \Omega_3\big).$$

The desired allocation can be different than under the rational counterfactual, but the practical implications are similar. The problem is different, however, in the presence of endogenous sentiment:

$$\mathcal{W}\big(d_1, H, \boldsymbol{q_1}; z_2, z_3 + \Omega_3\big).$$

A new state-variable now enters the optimal policy problem. The equilibrium level of asset prices today can enter the determination of *future* allocations, and thus the expected level of welfare. The following proposition illustrates this intuition, and shows that once again the interaction of endogenous sentiment and financial frictions is key.

[48] Farhi and Panageas (2004) empirically investigate whether sentiment corrects more inefficiencies than it causes, using a VAR methodology. They find that the negative effects of misallocation dominate.

Proposition 6 (Welfare Effects of Changing Asset Prices)**.** *The first-order impact on welfare when asset prices q_1 are marginally increased is non-zero and corresponds to a reversal externality:*

$$\mathcal{W}_q \quad = \quad \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_1}}\right] \tag{50}$$

This novel effect works through the interaction of financial frictions, *past* asset prices, and sentiment. The intuition for this term is as follows. When private agents marginally increase their investment in collateral assets, they push up the price of the asset today. This in turn might influence the formation of behavioral biases in the future, represented by the term $d\Omega_3/dq_1$.[49] 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 invest in risky assets they bid up prices which 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*.

3.3 Informational Advantage of the social planner

The theory developed above rests on the ability of the social planner to assess sentiment in real time, as well as the ability to forecast future movements in sentiment, while no one in the financial sector is taking contrarian positions. I briefly discuss the plausibility of this assumption, as well as the robustness of some of my results to this assumption.

First, is the regulator smarter than the market? Stein (2021) argues that even sophisticated arbitrageurs do not have the organizational structure that would allow them to bet aggressively against sentiment, consistent with the limits-to-arbitrage literature (Shleifer and Vishny 1997). Furthermore, non-financial entities appear successful in arbitraging macro conditions, suggesting that there is reliable predictive information about the extent of contemporaneous mispricing in the market.[50] Furthermore, survey data strongly supports the idea that forecast errors are reliably predictable ex-ante (Bordalo et al. 2018; Bordalo et al. 2019; De La O and Myers 2021).

Second, do all the aforementioned results disappear if the planner is subject to the same biases of the market? Not entirely. When agents know that everyone in the economy is subject to biases in the future, the impact of Ω_3 disappears from the behavioral wedges: agents take it into account when they make decisions. In other words, they realize that sentiment will hamper the borrowing

[49] This effect is also robust to a model where H is exogenously fixed: when agents increase their demand of the risk asset they bid up its price, creating welfare losses.

[50] Baker and Wurgler (2000) demonstrate that firms issue relatively more equity than debt just before periods of low market returns, suggesting firms time the market. Ma (2019) shows that net equity repurchases and net debt issuance both increase when expected excess returns on debt are particularly low, or when expected excess returns on equity are relatively high.

limit tomorrow, and thus take on less leverage today in consequence. The collateral and reversal externality, however, are still uninternalized. This is because they are essentially pecuniary externalities: to suppress them, agents would need to coordinate in order to reduce their leverage today. But with atomistic agents, this does not happen. This means that agents would realize that a high price today means excess pessimism tomorrow, but agents cannot act to reduce this price. Similarly, they realize that their leverage will impact prices tomorrow, and thus sentiment. The social planner still has a role to play in such an economy, as long as Ω_3, the behavioral bias during crises, directly depend on q_2 or q_1. Online Appendix O presents the details of the analysis with sophisticated agents and when the planner holds the same beliefs as private agents.

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 uninternalized effects from marginally altering leverage or investment, $\mathcal{W}_d$ and $\mathcal{W}_H$, into account. These objects are crucial to characterize optimal policy in this setting.[51]

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 Propositions 1, 4 and 6.

Proposition 7 (Second-Best Policy). *The social planner achieves the second-best by imposing:*

1. *A tax $\tau_d = -\mathcal{W}_d/\lambda_1$ on short-term borrowing ;*
2. *A tax $\tau_H = -\mathcal{W}_H/(\lambda_1 q_1^*)$ 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 7 is rather abstract, but makes three simple points. First, the calibration of macroprudential policy should be done by focusing on the key aspects of sentiment driving the uninternalized effects from the previous part, $\mathcal{W}_d$ and $\mathcal{W}_H$: (i) the current extent of sentiment Ω_2 ; (ii) the

[51] The 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. Appendix D allows for the simultaneous choice of ex-ante and ex-post policies. In particular, it shows that the possibility of intervention at $t = 2$ does not change the desirability of macroprudential interventions at $t = 1$.

future covariance of Ω_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. Finally, it makes it clear that without a positive probability of a crisis caused by binding collateral constraints, swings in exuberance should not impact the optimal conduct of macroprudential policy. My theory thus accounts for the sharply different aftermaths of the 2000 and 2008 bubble bursts. As noted by Blinder and Reis (2005), in 2000 the "biggest bubble" in history bursted without failing a single bank, which the authors take as a hint that preemptively taming asset bubbles or raising margin requirement might not be a worthwhile strategy. In the framework of the present paper, this is correct only when the asset affected by irrational exuberance is not used by financial intermediaries as collateral.

How can one interpret the results of Proposition 7 in terms of real-world policy? The optimal taxes on debt and investment correspond to the usual instruments in the macroprudential 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 7 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 4 and Online Appendix G), 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 take into account how it will influence the *future* realizations of Ω_3 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? The seminal work of 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 exogenous increase in sentiment Ω_2 increases their incentives to take on more debt, but agents simply cannot 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 (41): 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 8 (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 $\Omega_2 \geq 0$, 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.[52,53]

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 Propositions 1 and 4, 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, the leverage limit needs to be tightened or relaxed accordingly. For example, imagine a world where the planner believes that agents are not irrationally exuberant, $\Omega_2 = 0$, and that a financial crisis will not happen in the future. As demonstrated above, no leverage limit is required. If new information arrives, causing a sharp increase in q_1, this might not hold anymore. Even if Ω_2 stays at 0, the planner could fear that next period, these high prices will revert and lead to over-pessimism, causing a financial crisis. This would then create a need for preemptive leverage restrictions, even though the policy is enacted as a quantity restriction. Counter-cyclical restrictions are thus necessary, but because the likelihood of future over-pessimism fluctuates along the business cycles.

Remark 7 (Relation to the Literature). This paper is not the first to highlight possible disparities between price and quantity regulation in the optimal conduct of macroprudential policy. My results are complementary to those uncovered in recent papers. Clayton and Schaab (2020b) show that price regulation is superior in a multinational setting, since it forces national regulators to internalize the value of foreign banks, thus achieving the global optimum even if national authorities are acting non-cooperatively. Jeanne and Korinek (2020) demonstrate that quantity regulation has practical benefits when there is uncertainty about whether liquidity will be provided during crises in a targeted or untargeted form. Chen, Finocchiaro, Lindé and Walentin (2020) quantitatively compare the effect of different policies to curb household indebtedness when interest rates are low. Harper and Korinek (2021) show that quantity and price regulation have different distributional consequences: price regulation allows the social planner to determine the allocation of surplus between borrowers and lenders. In my paper, lenders have linear utility so distributional effects regarding lenders are irrelevant.

LTV Regulation: The second tax in Proposition 7 directly aims at regulating the *quantity* of risky investments. For this reason, this policy can be interpreted as loan-to-value (LTV) regulation, a widely used tool.[54] Importantly, the welfare analysis highlights again that the optimal LTV limit is time-varying, tracking the same behavioral biases as do leverage restrictions.

[52] Note that Proposition 8 looks at cases where the tax is fixed at an optimal level given some behavioral biases Ω_2 and Ω_3, and Ω_2 then exogenously moves. Section 5 looks at the optimal level of restrictions when the planner takes into account that Ω_2 is uncertain.

[53] Online Appendix F additionaly shows that a leverage limit is robust to the introduction of belief heterogeneity in the model, while a leverage tax becomes less efficient.

[54] According to Claessens (2014), LTV ratios are used by 55% of advanced economics.

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 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 3.2.2, 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.

The interpretation of investment regulation as LTV ratios squares naturally with the view that H represents real estate investment. In Online Appendix M I present a simple model in which financial intermediaries finance heterogenous construction entrepreneurs. Imposing a maximum LTV ratio prevents financial intermediaries from financing some entrepreneurs, thus limiting real estate investment. If one wishes to interpret H more broadly as investments made by firms, or C&I loans, this policy of restricting investment can simply be interpreted as "supervisory guidance": the financial stability authority nudges intermediaries towards reducing their financing of some activities or some sectors of the economy. Online Appendix M also offers an alternative interpretation where H are MBS: financial intermediaries pool mortgage loans to diversify the risk of idiosyncratic default. In this case, LTV regulation reduces the supply of mortgage loans.

Price Regulation: The third tax in Proposition 7 does not have a simple relation to the current macroprudential toolbox, however. This is because my model is the first to highlight 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 modelled 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. By altering discount rates, the central bank has a direct influence on equilibrium asset prices, and can thus complement the macroprudential toolbox. I provide an in-depth analysis of the use of monetary policy in my model in Section 6 by adding nominal rigidities, and explore its associated challenges.

I end this section with some specific cases in order to strengthen intuition.

4.3 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 (40), (45), and (50). At the rational expectations constrained-efficient equilibrium, behavioral wedges are zero, so the only left parts are the collateral externalities and the reversal externality:

$$\mathcal{C}_d = -\beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \frac{dq_2}{dn_2} \right] \tag{51}$$

$$\mathcal{C}_H = \beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \left(\frac{dq_2}{dn_2} z_2 + \frac{dq_2}{dH} \right) \right] \tag{52}$$

$$\mathcal{W}_q = \beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_1}} \right] \tag{53}$$

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

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 price of the asset tomorrow in a crisis, where agents are not on their Euler equation, has a first-order impact on welfare by aggravating financial crises. 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.[55]

4.4 Attainable Welfare Levels and Relative Sentiment

How does the chosen allocation, implemented following Proposition 7, compare to the allocation that a planner would choose in a rational world? Because the constrained efficiency concept allows the social planner to only choose the leverage and investment levels of financial intermediaries, the two will generally differ. Indeed, while the social planner can nudge agents with leverage taxes to counteract the effect of over-optimism, the planner cannot directly combat pessimism during crises. If the planner knows that Ω_3 will be negative during crises, welfare is maximized by imposing a leverage limit that is *lower* than in the rational counterfactual. This will result in a *lower* absolute welfare level in the behavioral case compared to the rational benchmark.

Of course, the opposite is also theoretically possible: if agents are always over-optimistic, the planner might impose a leverage limit at $t = 1$ but the overall welfare level could be above its rational counterpart: by being over-optimistic during crises, agents are effectively alleviating the pressure of financial frictions, bringing the economy closer to the first-best.

Finally, the sign of the behavioral wedge is a function of the *absolute* level of sentiment: what matters is whether agents are over-pessimistic, in the sense that they believe future payoffs to be lower than the objective distribution. The collateral and reversal externalities, on the other hand, are concerned with the *relative* levels of sentiment inside a crisis, embodied by the derivatives $d\Omega_3/dq_2$ and $d\Omega_3/dq_1$. Thus the planner is intervening to make agents *more optimistic*, irrespective of whether the financial crisis displays absolute levels of over-pessimism or over-optimism.

[55] Of course irrational exuberance is also costly, as it triggers more frequent credit crunches (see equation (31) and the discussion therein). 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. While the empirical literature on sentiment and financial crisis focused mainly on irrational optimism, excessive pessimism during bad times is also a robust feature of the data (see Bordalo et al. (2018) for an example on credit spreads forecasts). 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).

4.5 Third-Best Policy without Price Regulation

As mentioned earlier, the typical macroprudential toolbox only permits the use of capital structure regulation and LTV ratios. How should optimal policy be conducted by a regulator that acknowledges that prices are entering welfare but only has access to these limited instruments?

In this case, the social planner recognises that in equilibrium, the price will be changing with the level of investment, in order to stay on the $q_1 = c'(H)$ condition. We can thus write the uninternalized effects of marginally increasing H, taking into account how prices move, as:

$$\mathcal{W}_{H,q} = \left(\beta\mathbb{E}_1^{SP}\left[\lambda_2(z_2+q_2)\right] - \lambda_1 q_1\right) + \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H \boldsymbol{\frac{d\Omega_3}{dq_2}}\left(\frac{dq_2}{dn_2}z_2 + \frac{dq_2}{dH}\right)\right] + \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H \boldsymbol{\frac{d\Omega_3}{dq_1}} c''(H)\right] \quad (54)$$

where $c''(H) > 0$. The reversal externality thus enters investment welfare effect, adding a negative term (as long as sentiment in a crisis is indeed negatively related to past prices).

A direct consequence is thus that the third-best equilibrium features a *lower* level of investment than the second-best. In order to mitigate the effects of high asset prices for future sentiment, the social planner has to reduce investment levels. An equivalent, but perhaps more practical way to interpret this result, is that by adding a third instrument that directly controls asset prices the regulator can allow for a *higher* level of investment through relaxing LTV regulations, enhancing welfare and financial stability at the same time.

4.6 Summary

The theory I presented here unsurprisingly calls for more aggressive leverage regulation when sentiment is elevated. More surprisingly, the presence of predictable future sentiment has concrete implications for the conduct of macroprudential policy ex-ante. If future behavioral biases are negative (irrational pessimism) when financial intermediaries are distressed, leverage restrictions should be tightened. Furthermore, if future behavioral biases depend on the price of assets, there are novel collateral and reversal externalities that must be take into account, and can result in the need for an additional instrument.

Before diving into the analysis of monetary policy as a tool complementing the traditional macroprudential toolkit, it is useful to acknowledge that the social planner cannot have a infinitely precise estimate of the behavioral biases that agents are subject to. Indeed, Proposition 7 shows that to properly calibrate macroprudential policy, the planner needs to know the exact level of sentiment. The next section analyses how uncertainty alters the conduct of optimal policy.

5 Ω-Uncertainty

I so far assumed that the social planner had perfect information about the current state of exuberance Ω_2 and its state-contingent evolution Ω_3. A natural question of practical importance is whether these 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".

This aspect is of crucial importance regarding the practical implications of this paper. In a famous speech about asset price bubbles, Bernanke (2002) discussed the "identification problem" that naturally arises once the financial stability authority contemplates a proactive approach to bubbles. While recognizing that identifying a bubble is intrinsically difficult, this section shows that the widespread intuition that this uncertainty calls for *laissez-faire* is actually erroneous.

To this end, I leverage the prior equilibrium and welfare analysis. I here focus on the case where the planner is uncertain about the exact level of Ω_2, while Ω_3 is assumed to be certain and constant in the future.[56] Crucially, the results depend on whether or not the distribution of states of the world is common knowledge. 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 of z_2. I start with a modelling assumption.

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

This assumption implies 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. I add a natural assumption, made to rule out pathological cases:

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

I start briefly with the case where the planner knows $\bar{z}_2$. I then study the other extreme where the planner's prior over $\bar{z}_2$ is flat, and then finish with the intermediate and more general cases where the planner has some information about $\bar{z}_2$, but less than the private agents.

5.1 Full-Information Benchmark

Assume there that the level of $\bar{z}_2$ is common knowledge. The planner observes an equilibrium price q_1 in the initial period, and knows that agents are subject to behavioral biases. We can write the relation between prices, fundamentals, and sentiment conveniently as:

$$q_1 = g_q(\bar{z}_2 + \Omega_2) \tag{55}$$

[56] The analysis for varying or uncertain Ω_3 is presented in Online Appendix H. The results are identical.

Because of Assumption 4, this implies that the social planner can perfectly extract Ω_2 by inferring it from asset prices:

Proposition 9 (Full-Information Benchmark)**.** *If the mean of the dividend distribution $\bar{z}_2$ is common knowledge, the social planner extracts a behavioral bias of:*

$$\Omega_2 = g_q^{-1}(q_1) - \bar{z}_2 \tag{56}$$

and implements optimal policy according to Proposition 7 using this value for Ω_2. The planner's prior over Ω_2 is irrelevant in this case.

Finally, because of Assumption 4 the inferred bias is strictly increasing in asset prices, and as such optimal leverage and investment taxes are strictly increasing in asset prices.

5.2 Flat Prior over $\bar{z}_2$

I now investigate the polar case, by considering that the planner has a flat (improper) prior over $\bar{z}_2$. In this case the social planner's prior over sentiment matters. I assume that it is given by a uniform distribution:

$$w \sim \mathcal{U}\left[\bar{\Omega}_2 - \sigma_\Omega, \bar{\Omega}_2 + \sigma_\Omega\right] \tag{57}$$

where $\bar{\Omega}_2$ is the point estimate 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 sill use:

$$\Omega_2 + \bar{z}_2 = g_q^{-1}(q_1) \tag{58}$$

but since it has a flat prior over $\bar{z}_2$, the posterior distribution regarding sentiment stays the same as its prior, while the posterior mean is given by:

$$\bar{z}_2 = g_q^{-1}(q_1) - \bar{\Omega}_2 \tag{59}$$

In other words, observing the asset price does not change the point estimate of sentiment used by the planner, $\bar{\Omega}_2$. The uncertainty on sentiment, however, translates into the planner's objective function. Indeed, the planner now would use the same distribution as agents to optimally set short-term debt such that:

$$u'(c_1) = \frac{1}{2\sigma_\Omega}\int_0^\infty \left[\int_{-\sigma_\Omega}^{\sigma_\Omega} \frac{\partial \mathcal{W}_2}{\partial n_2}\left(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \omega_2\right) d\omega_2\right] f_2(z_2) dz_2 \tag{60}$$

This expression contains all of the intuition for how sentiment uncertainty can reinforce or weaken the need for preventive leverage tightening.[57] Once deducing the average behavioral error $\bar{\Omega}_2$, the

[57] Note that $\bar{\Omega}_2$ is an argument of $\mathcal{W}_2$ 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

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.

The presence of sentiment noise will thus affect the size of the expectational term according to a Jensen's type of argument. 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 $u'(c_1)$ and by diminishing expected 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 2.3, uncertainty about behavioral biases unambiguously calls for precautionary restrictions.

Proposition 10 (Ω_2-Uncertainty and Leverage Restrictions)**.** *If the social planner has a flat prior over $\bar{z}_2$ and believes that the behavioral bias at $t = 1$ can be expressed as $\bar{\Omega}_2 + \omega$, 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. It is strictly increasing as long as there exist a ω in $[-\sigma_\Omega, \sigma_\Omega]$ for which, if sentiment is $\bar{\Omega}_2 + \omega$, there is a positive probability of a crisis in the next period.*

The proof is rather involved (see Appendix A.11), but the intuition can be understood from studying a function of the following type:

$$g : \omega \to \int_0^{+\infty} \mathcal{W}_{n,2}(z - \omega) dF(z) \tag{61}$$

where $\mathcal{W}_{n,2}$ is the first derivative of $\mathcal{W}_2$ with respect to net worth. This integral measures expected welfare, when all states of the world are shifted by $-\omega$: a positive ω means that the distribution is shifted to the left, hence that the planner is using a distribution that is less optimistic than agents. The key is to notice that g is a *convex* function, as shown in Figure 5. Intuitively, sentiment uncertainty adds terms to the expectation computed by the planner, but the parts coming from intermediaries' optimism are more costly than the ones coming from pessimism.[58]

This analysis directly speaks to the debate about the potential costs of preventive action when identifying asset price bubbles is difficult. It is interesting to contrast this result with the prevailing view that uncertainty prevents the regulator from taking action. My model suggests that this is erroneous. The strong linearities associated with the interaction of sentiment and financial frictions make it attractive to tighten capital requirements in the face of uncertainty. Using the words of

add their bias $\bar{\Omega}_2$. The two are of course equivalent, but in the present Section this clarifies that the planner simply extracts the distribution used by private agents.

[58] The 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 H.

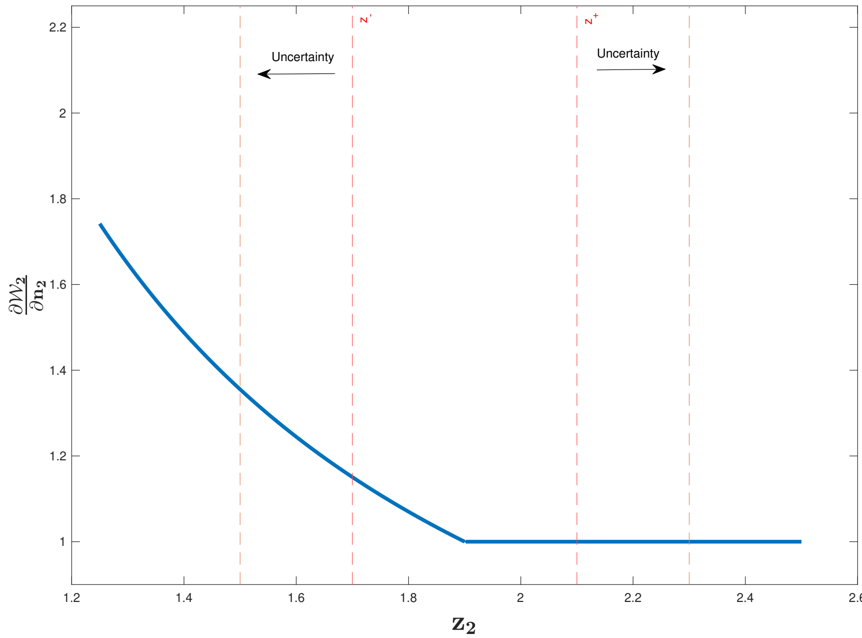

Figure 5: Non-linearities and Ω-uncertainty. This figure plots $\partial \mathcal{W}_2 / \partial n_2$ against the fundamental realization z_2. The partial derivative encodes all the uninternalized general equilibrium effects, in particular the variation of the price of the collateral asset when the net worth of financial intermediaries changes. The thick dotted lines correspond to the range of values of z_2 where the expectations is taken. The think dotted lines represent the widening of the range where the social planner computes expectations caused by the uncertainty on Ω_2. The discontinuity arises at z^*.

Yellen (2009), a "type 1" error is simply much less costly than a "type 2" error.[59]

Remark 8 (Uncertainty and Investment Regulation). Online Appendix H.3 studies the same uncertainty problem but for investment. Interestingly, the opposite result appears: 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 over-optimistic 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, this means that in times of heightened uncertainty, the regulator should tighten counter-cyclical capital buffers and marginally relax LTV ratios.

Remark 9 (Uncertainty and Behavioral Biases during Financial Crises). Online Appendix H.1 shows that uncertainty about future behavioral biases during crises, Ω_3, yield the same result. Furthermore, Online Appendix H.2 shows that the presence of belief amplification generally increases the precautionary motives for early intervention by adding convexity.

In this particular case with a flat prior over the fundamentals of the economy, optimal policy is

[59]Interestingly, the speech was title "A Minsky Meltdown: Lessons for Central Bankers." Janet Yellen mentioned the risk of collateral damage by stating: "There is also the harm that can result from "type 2 errors," when policymakers respond to asset price developments that, with the benefit of hindsight, turn out not to have been bubbles at all'.'

not directly time-varying. Because the planner has a flat prior over $\bar{z}_2$, all movements in asset prices are "absorbed" into the posterior estimation of the fundamental distribution. Unless the financial authority has a reason to believe that its confidence interval is expanding when asset prices are elevated, the optimal policy consists of a tight leverage limit, because of uncertainty, but one that is not evolving with asset prices. Proposition 10 thus suggests that the presence of sentiment calls for high, but unconditional, safeguards. As I show next, this premise is resting heavily on the assumption that the social planner has absolutely no information about fundamentals.

5.3 Time-Varying Optimal Policy and Ω-Uncertainty

I now depart from the stark hypothesis that the social planner has a flat prior over the distribution of future fundamentals. To make progress in a tractable way and still be able to distill some insights, I make the following assumptions for this section.

Assumption 5. *The social planner holds gaussian priors over $\bar{z}_2$ and Ω_2:*

$$\bar{z}_2 \sim \mathcal{N}\left(\mu_z, \sigma_z^2\right) \quad ; \quad \Omega_2 \sim \mathcal{N}\left(\bar{\Omega}_2, \sigma_\Omega^2\right) \tag{62}$$

Assumption 6. *The social planner is restricted to compute expectations over sentiment using a uniform distribution that minimizes the Kullback–Leibler divergence with its posterior.*

Assumption 5 is made for convenience: assuming normal priors allows for a tractable posterior expression. Assumption 6 allows for the use of my previous results, where a uniform distribution was used.

By observing a price q_1 the social planner now forms posterior beliefs about sentiment in the following way:

$$\Omega_2 \sim \mathcal{N}\left(\bar{\Omega}_2 + \frac{\sigma_\Omega^2}{\sigma_\Omega^2 + \sigma_z^2}\left[g_q^{-1}(q_1) - \bar{\Omega}_2 - \mu_z\right], \sigma_\Omega^2 \frac{1}{1 + \frac{\sigma_\Omega^2}{\sigma_z^2}}\right) \equiv \mathcal{N}\left(\bar{\Omega}(q_1), \Sigma_\Omega^2\right) \tag{63}$$

where the average level of sentiment extracted, $\bar{\Omega}(q_1)$, is increasing in q_1. Per assumption 6 the social planner uses the following uniform distribution to compute expectations:[60]

$$\Omega_2 \sim \mathcal{U}\left[\bar{\Omega}(q_1) - \sqrt{\frac{3}{2}}\Sigma_\Omega \quad , \quad \bar{\Omega}(q_1) + \sqrt{\frac{3}{2}}\Sigma_\Omega\right] \tag{64}$$

We can now directly apply Proposition 10. The planner still has a confidence interval for sentiment, but its point estimate for Ω_2 is varying with asset prices.

[60] The derivations for the minimization of the KL-divergence between the two distributions are presented in Appendix A.12.

Proposition 11 (Ω-Uncertainty and Time-Varying Policy)**.** *Under Assumptions 5 and 6, the social planner's optimal leverage tax is increasing in both equilibrium prices q_1 and sentiment uncertainty σ_Ω.*

This last proposition bridges the gap between the two polar cases studied above. Intuitively, the more certain the planner is about the objective distribution of future fundamentals, the less uncertainty it has over sentiment. Consequently, the less uncertainty there is about sentiment, the more the planner can adapt its leverage limits to observable conditions like asset prices. An interesting avenue for future research is to quantify these effects over the business cycle to understand how optimal macroprudential policy evolves from precautionary motives to more directly targeting sentiment.

Remark 10 (Relation to the Literature). These results might seem surprising in light of the famous Brainard (1967)'s "attenuation principle." The contradiction is only apparent, however.[61] Brainard (1967) presents a model where the central bank faces uncertainty over the parameter governing how its policy tools impact the level of aggregate activity. Intuitively, using uncertain tools introduce further volatility, which the policymaker dislikes. My results are obtained by assuming a whole different form of uncertainty: the planner faces no uncertainty over the impact of a leverage restriction on leverage in my model. In related work, Bahaj and Foulis (2017) show in a linear-quadratic and fully rational framework that the asymmetry in the objective function of the planner leads to a more active policy under uncertainty. Finally, Montamat and Roch (2021) look at macroprudential policy when the planner fears model misspecification. They analyze policy under robustness, *à la* Hansen and Sargent (2011): which policy is optimal under the "worst-case" scenario. In my framework, such a robustness exercise is immediate to carry out, since the worst-case scenario occurs when sentiment is equal to $\bar{\Omega}_2 + \sigma_\Omega$. It then follows immediately that leverage restrictions are increasing in σ_Ω.

6 Monetary Policy

I complete the study of optimal policy with the incorporation of monetary policy. 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 (2020a) and Farhi and Werning (2020) show that when traditional macroprudential policy is constrained, leaning against the wind with a monetary tightening is valuable. In both papers, 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 (thus assuming that the output gap can be perfectly closed).[62]

[61] In the words of Bahaj and Foulis (2017), "Brainard's results are sometimes misleadingly cited as a general rule that a policymaker should do less in the face of uncertainty."

[62] In contrast, Svensson (2017) finds that the costs of leaning against the wind exceed its benefits. Monetary tightening traditionally affects aggregate demand today, entailing a second-order welfare costs, but it is further assumed that this

In this section I first confirm this insight in my framework. But my model also features a different channel through which monetary policy can affect welfare. As developed in Section 4.1, by changing asset prices at $t = 1$ monetary policy is indirectly altering the formation of behavioral biases at $t = 2$ inside a financial crisis.[63] This makes monetary policy a *complementary* tool, rather than merely a *substitute* for existing tools. Finally, I extend the model to show that leaning against the wind can have perverse effects once one takes a more dynamic perspective: when agents expect the central bank to tighten when asset prices are rising, this weakens the stimulus power of an interest cut in normal times, creating a time-inconsistency problem. To the best of my knowledge, this paper is the first to highlight these potential dynamic costs of leaning-against-the-wind policies.

6.1 Nominal Rigidities

I start by introducing nominal 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 demand-determined at $t = 1$ by assuming wages are fully rigid.[64]

Concretely, households now have the following utility function:

$$U^h = \mathbb{E}_1 \left[\left(\ln(c_1^h) - \nu \frac{l_1^{1+\eta}}{(1+\eta)} \right) + \beta c_2^h + \beta^2 c_3^h \right]$$

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 at $t = 1$, 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* quantifies how far off are workers from their optimality condition:

$$\mu_1 = 1 - \nu c_1^h l_1^\eta \tag{65}$$

as can be seen from simply taking the first-order condition with respect to labor supply. 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 group at $t = 1$, in order to dismiss redistribution concerns.

will automatically translate into weaker aggregate demand if a recession ensues. This feature is entirely absent of my model: if anything, price extrapolation would translate into the opposite, since the size of the bust is positively related to the size of the boom.

[63] In addition, the presence of an asset creation margin implies that monetary policy can also have spillover effects on investment, a channel also not considered in the aforementioned papers.

[64] In 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 enough in crises.

6.2 Optimal Monetary Policy

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

$$\begin{aligned}\frac{d\mathcal{W}_1}{dr_1} &= \underbrace{\frac{dY_1}{dr_1}\mu_1}_{(i)} + \underbrace{\frac{dd_1}{dr_1}\mathcal{W}_d}_{(ii)} + \underbrace{\frac{dH}{dr_1}\mathcal{W}_H}_{(iii)} \\ &\quad + \underbrace{\frac{d\Omega_2}{dq_1}\frac{dq_1}{dr_1}\left(\frac{dd_1}{d\Omega_2}\mathcal{W}_d + \frac{dH}{d\Omega_2}\mathcal{W}_H\right)}_{(iv)} + \underbrace{\frac{dq_1}{dr_1}\beta\mathbb{E}_1\left[\kappa_2\phi H\frac{d\Omega_3}{dq_1}\right]}_{(v)} \end{aligned} \tag{66}$$

where $\mathcal{W}_d = \mathcal{B}_d + \mathcal{C}_d$, the sum of the behavioral wedge and the collateral externality for leverage, and $\mathcal{W}_H = \mathcal{B}_H + \mathcal{C}_H$, the sum of the behavioral wedge and the collateral externality for investment. The last term is proportional to $\mathcal{W}_q$, the reversal externality, (see Section 3.2 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 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 does not necessarily imply that leaning against the wind (by increasing r_1 above its value that achieves full employment) is 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 $\mathcal{W}_d = 0$, 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:

$$\frac{d\mathcal{W}_1}{dr_1} = \frac{dH}{dr_1}\mathcal{W}_H \tag{67}$$

because investment in unambiguously decreasing in the interest rate r_1, tightening is desirable only if $\mathcal{W}_H < 0$, i.e. if the uninternalized welfare effects of marginally increasing the creation of collateral assets is negative. As fully explained in Section 3.2.2, 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.[65]

[65] My framework also abstracts from other considerations that could argue against tightening in such a situation. For exam-

Notice from equation (66) that the ability of the central bank to improve financial stability largely depends on the reaction of beliefs to policy, a recurrent theme of this paper. 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 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. Monetary policy thus provides the planner with a supplementary instrument that can affect equilibrium prices, and not only real allocations, a desirable feature discussed in Section 4.1.

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 with the credit cycle (Dudley 2015; Caballero and Simsek 2020a ; Farhi and Werning 2020 ; Stein 2021). Inspecting Proposition 12, 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.

Proposition 13 (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:*

$$\frac{d\mathcal{W}_1}{dr_1} = \frac{dY_1}{dr_1}\mu_1 + \beta\mathbb{E}_1\left[\kappa_2\phi H \frac{d\Omega_3}{dq_1}\frac{dq_1}{dr_1}\right]. \tag{68}$$

This particular case calls for leaning against the wind in order to tame *current* asset prices, which will then tame *future* pessimism in a possible crisis – a new channel for monetary policy. Furthermore, such action does not require *any* information about contemporaneous biases. It is possible to be in a situation where a sharp increase in asset prices is entirely due to fundamentals, but the planner has an incentive to make prices deviate from their rational value today.

An interesting example is the housing boom of the 2000s. Fuelled by monetary easing, the years 2001 to 2003 witnessed sharp increases in house prices. As noted by many analysts and policymakers at the time, it was not obvious that these prices were above fundamentals. This idea is expressed by Kohn (2003): “in sum, the rise in housing prices and the increase in household investment in houses and consumer durables do not appear out of line with what might be expected

ple, 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 (forthcoming) 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 (2019) 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.

in the current environment." My model suggests that the worry 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, with adverse welfare consequences. In that case, an interest rate hike is recommended.[66]

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.

To conclude, while the literature on leaning against the wind found that monetary policy can be a substitute tool for leverage restrictions, my analysis highlights that monetary policy can actually be used as a *complement*, thanks to its ability to influence current asset prices. This channel is not dependent on the ability of the central bank to distinguish fundamental-driven movements from speculative bubbles.

6.3 Early vs. Late Tightening

The previous Section showed how a monetary tightening can improve welfare by reducing optimism today and indirectly alleviating future pessimism through its effect on asset prices. Because of the 3-period framework, I took as given the initial conditions of the economic system: given these initial conditions the central bank optimally tightens when $d\mathcal{W}_1/dr_1 > 0$. But asset price bubbles and leverage cycles form over long horizons, and the financial authority has arguably several occasions to act pre-emptively. This section asks under which conditions a preventive tightening should be triggered earlier or later along the credit boom phase.

To this end, I add an hypothetical time period, $t = 0$, where the central bank can decide to deviate from inflation targeting. I allow the central bank to deviate from inflation targeting only once, either at 0 or at $t = 1$. Agents are unaware that the central bank might deviate from its mandate.[67] At period 0, agents have some level of behavioral bias Ω_1. To focus on the timing and horizon issues raised by this question, I assume that Ω_t can also depend directly on past biases Ω_{t-i}.[68] Raising the nominal interest rate r_0 at $t = 0$ will impact welfare through different channels. It will change the equilibrium price q_0. This, in turn, will affect future sentiment through first a direct effect since Ω_2 and Ω_3 can depend on the evolution of past prices. But it can also work through current sentiment, if Ω_1 directly feeds into future sentiment Ω_2 and Ω_3.[69]

[66] Evidently, this policy problem is also plagued with uncertainty. In Online Appendix H.4, I show that the reversal externality part is also increasing with sentiment uncertainty when there is price extrapolation. Consequently, the incentive for the central banker to tighten interest rate after asset price soar is higher when there is uncertainty about Ω_2 and Ω_3.

[67] The next Section looks at the implications for monetary policy when agents anticipate that the central bank might lean against the wind in the future.

[68] This allows me to study the cases where sentiment is "sticky," slow-moving, or to the contrary subject to reversals, which will be important in what follows.

[69] In recent work, Bianchi et al. (2021) estimate a New Keynesian model with Diagnostic Expectations. To fit the data best, they estimate "memory weight" that imply that reference points are taken from 4 to 8 quarters in the past. Similarly, Bordalo et al. (2019) find a sluggishness of the reference point of 11 quarters to fit the behavior of forecast errors. Maxted

Denote by $d\Omega_t^e$ the first-order change in sentiment by tightening early in period 0, and $d\Omega_t^l$ if tightening late in period 1. Late tightening therefore leads to the following first-order changes to behavioral biases in period $t = 1$ and $t = 2$:

$$d\Omega_2^l = \frac{d\Omega_2}{dq_1}\frac{dq_1}{dr_1}dr_1 \tag{69}$$

$$d\Omega_3^l = \left[\frac{d\Omega_3}{d\Omega_2}\frac{d\Omega_2^l}{dr_1} + \frac{d\Omega_3}{dq_1}\frac{dq_1}{dr_1}\right]dr_1 \tag{70}$$

while early action yields also a change in behavioral biases at $t = 0$:

$$d\Omega_1^e = \frac{d\Omega_1}{dq_0}\frac{dq_0}{dr_0}dr_0 \tag{71}$$

$$d\Omega_2^e = \left[\frac{d\Omega_2}{d\Omega_1}\frac{d\Omega_1^e}{dr_1} + \frac{d\Omega_2}{dq_0}\frac{dq_0}{dr_0}\right]dr_0 \tag{72}$$

$$d\Omega_3^e = \left[\frac{d\Omega_3}{d\Omega_2}\frac{d\Omega_2^e}{dr_1} + \frac{d\Omega_3}{d\Omega_1}\frac{d\Omega_1^e}{dr_1} + \frac{d\Omega_3}{dq_0}\frac{dq_0}{dr_0}\right]dr_0 \tag{73}$$

I focus on the belief channel by assuming that $\mathcal{W}_H = 0$ through an optimal LTV ratio regulation, and that $dd_1/dr_1 = 0$ as discussed earlier. In both cases (late and early tightening), welfare effects are given by:

$$d\mathcal{W}_1 = \frac{dd_1}{d\Omega_2}\mathcal{W}_d d\Omega_2 + \beta\mathbb{E}_1\left[d\Omega_3\kappa_2\phi H\right] \tag{74}$$

To flesh out the difference forces that shape this trade-off, assume the following linear form of behavioral biases.

Assumption 7. *The behavioral bias in period t, Ω_{t+1} is a linear function of current and past prices, as well as past sentiment. Furthermore, the coefficients are constant over time:*

$$\Omega_{t+1} = \alpha_0 q_t + \alpha_1 q_{t-1} + \alpha_2 q_{t-2} + \gamma_0\Omega_t + \gamma_1\Omega_{t-1} \tag{75}$$

where the γ_i coefficients encode the dependence on past behavioral biases.[70] Next, for simplicity, assume also a linear formulation for the influence of interest rates on asset prices.

Assumption 8. *The first-order effect of a change in interest rates r_t on q_t is constant over time:*

$$\frac{dq_t}{dr_t} = \iota < 0 \tag{76}$$

This allows for a clear comparison of the welfare effects of tightening early or late in the cycle.

(2020) fits a parameter that governs the persistence of sentiment in a continuous-time model, and finds a half-life of sentiment of 5 years. Other behavioral finance models also incorporate slow-moving sentiment. For example, Adam et al. (2017a) proposes a model where agents gradually update their beliefs in the direction of past price growth observations, in the form of $\tilde{E}_t[P_{t+1}/P_t] = (1-g)\tilde{E}_{t-1}[P_t/P_{t-1}] + gP_{t-1}/P_{t-2}$ and estimate a persistence parameter of $1 - g = 0.9736$.

[70] Online Appendix J.4 gives a simple example of sticky beliefs, based on Bouchaud, Krueger, Landier and Thesmar (2019) that gives rise to a direct and linear dependence on past sentiment.

Proposition 14 (Early vs. Late Leaning Against the Wind)**.** *Under Assumptions 7 and 8, it is optimal to lean against the wind in period $t = 1$ rather than in period 0 if and only if:*

$$-\frac{dd_1}{d\Omega_2}\mathcal{W}_d\left(\alpha_0(1-\gamma_0)-\alpha_1\right) > \beta\mathbb{E}_1\left[\kappa_2\phi H\right]\left((\gamma_0\alpha_0+\alpha_1)(1-\gamma_0)-\gamma_1\alpha_0-\alpha_2\right) \tag{77}$$

Several insights can be gleaned from examining equation (77). First, a negative α_2 makes it more likely that an interest rate hike early in the cycle is beneficial. This is because it will support sentiment during a crisis: if behavioral biases depends negatively on asset prices in the distant past, lower prices in period 0 are beneficial. On the other hand, if α_2 is positive, meaning that high asset prices in the more distant past impact biases positively, the opposite is true: lowering asset prices in period 0 will be detrimental for the health of financial intermediaries in a crisis.

The role of α_1 is particularly interesting. This coefficient encodes how the most recent price impacts behavioral biases. A negative α_1, as in the simple price extrapolation example used throughout this paper, makes agents become optimistic, all else being equal, if the price last period was particularly low. The right-hand side of equation (77) captures the intuition contained in the earlier "reversal externality:" tightening in period $t = 1$ lowers the price of the asset q_1, which dampens over-pessimism in period $t = 2$ if $\alpha_1 > 0$. But the exact same effect is actually detrimental to welfare in the case of an early tightening, as can be seen form the left-hand side of equation (77). The intuition is that leaning against the wind in period $t = 0$ lowers the initial price of the asset, q_0, but since $\alpha_1 > 0$ this *exacerbates* future over-optimism in period $t = 1$, by lowering the reference point used by behavioral agents.[71] In this case, early leaning against the wind backfires: by lowering asset prices today, it is only kicking the can down the road, and encouraging irrational over-optimism later on. This effect is not purely hypothetical. In an experimental setting, Galí, Giusti and Noussair (2021) find that while increasing interest rates decreases the size of the bubble contemporaneously, the opposite effect is observed in the following period, when higher interest rates are associated with greater bubble growth. Relatedly, Galí and Gambetti (2015) estimate the response of stock prices to monetary policy shocks using a VAR methodology, and find cases where asset prices increase persistently after an exogenous tightening of monetary policy.

A similar intuition goes through for the influence of past biases (the γ_i coefficients). If γ_0 (dependence on the most recent behavioral bias) is positive, over-optimism today makes agents more optimistic next period. This implies that over-optimism in period $t = 1$ is less costly because it tames over-pessimism in a possible future financial crisis. When this is the case, tightening later in the cycle has ambiguous effects, since it forces agents to delever today, but also makes agents more pessimistic next period. There is thus a trade-off between making the financial system less fragile, and creating irrational over-pessimism in the future which can itself trigger a financial crisis.[72] It is possible that such undesired effects are responsible for the results uncovered in Schularick,

[71] This is similar to exogenously lowering the anchor q_0 that was used in the price extrapolation formula, $\Omega_2 = \alpha(q_1 - q_0)$.

[72] This is a possible microfoundation for the postulated reduced-form in Svensson (2017), where leaning against the wind translates into weaker aggregate demand after a crisis.

Ter Steege and Ward (2021). There the authors study empirically the effects of monetary policy on crisis risk. They show that discretionary leaning against the wind during credit and asset price booms is more likely to trigger crises than prevent them.

The effect of γ_1 (dependence on more distant behavioral biases), on the other hand, is unambiguous. It only enters the decision as an interest rate hike in period 0 will affect sentiment in a financial crisis, Ω_3. In this case, if sentiment is prone to reversals at this frequency, i.e. $\gamma_1 < 0$, it gives an argument for leaning against the wind early as exuberance in period 0 would create even stronger pessimism in the future, once a crisis hits.

In summary, even an excessively simple formulation for the temporal evolution of sentiment creates complex trade-offs for the central bank, once it recognizes its ability to influence the formation of behavioral biases. Further research is needed to understand over which horizon these biases are formed and how past outcomes enter their determination.[73]

Remark 11 (Low Interest Rates and Sowing the Seeds of the Next Crisis). The previous analysis highlighted how leaning against the wind early in the cycle can backfire by fuelling exuberance in the future. This insight relates to the debate surrounding the role of monetary policy in the formation of bubbles and subsequent financial crises. Proposition 14 was derived under the assumption that the central bank would perfectly target full employment, but it is easy to think of period 0 as the beginning of the credit cycle that led to the 2008 meltdown. In the early 2000s, low interest rates were needed to prevent a slump in aggregate demand. It also indirectly increased house prices. But if agents are price extrapolators, this monetary stimulus created irrational exuberance, expressed in equations (71) and (72), with an initial increase in q_0. It implies that stimulating the economy at $t = 0$ creates a motive for leaning against the wind in the next period. If agents anticipate this future tightening, however, the conventional stimulus of monetary policy can be severely impaired. This is the focus of the next section.

6.4 Dynamic Tradeoffs of Leaning against the Wind

I now turn to the indirect effects of preventive interest rate hikes for financial stability purposes on the regular conduct of monetary policy. To fully work out the consequences of such a deviation from classic inflation targeting, it is necessary to understand how the *anticipation* of such a policy would lead to a different set of initial conditions, and affect the economy during periods that are outside the simple framework used until now.[74]

To investigate these dynamic tradeoffs, I extend the mode as previously to add an initial period $t = 0$. For simplicity, I assume that financial intermediaries only enter the model at $t = 1$, and there

[73] In Online Appendix I I show how these results are still shaping the optimal policy response in an infinite-horizon version of the model.

[74] One interpretation could be that the previous analysis was made under the assumption that the interest rate hike was made unexpectedly. But after this first happened, agents now entertain the possibility that in future periods, the central bank will consider such a policy again. This is the focus of the present section.

is no more uncertainty. Households now have utility:

$$U^h = \mathbb{E}_0\left[\left(\ln(c_0^h) - \nu\frac{l_0^{1+\eta}}{(1+\eta)}\right) + \beta\left(\ln(c_1^h) - \nu\frac{l_1^{1+\eta}}{(1+\eta)}\right) + \beta^2 c_2^h + \beta^3 c_3^h\right]$$

Wages are still assumed to be fully sticky at $t = 0$ and $t = 1$, normalized to 1. In the Rational Expectations equilibrium, the monetary authority would perfectly stabilize the economy. This would lead to interest rates satisfying:

$$\beta^2(1+r_0^*)(1+r_1^*) = 1 \tag{78}$$

where r_1^* is such that the output gap is closed in the future: $\mu_1 = 0$.[75] Households, however, now expect the central bank to tighten in the event that asset prices are higher than some target price $\bar{q}$. Specifically, agents expect the future interest rate r_1^e to be determined by:

$$r_1^e = r_1^* + \rho(q_0 - \bar{q})^+ \tag{79}$$

with ρ a constant.[76] Furthermore, I illustrate the results with extrapolative expectations of the simple reduced-form:

$$\mathbb{E}_0[q_1] = q_1^r + \alpha(q_0 - q_{-1}) \tag{80}$$

where, again, q_{-1} is set exogenously. As in the baseline version of the model, a change in interest rates moves asset prices through the discount factor channel, and through the belief channel. But there is now a third effect: the a cut in rates increases asset prices today, which increases the expectation of future prices tomorrow, pushing current prices even more upward, which in itself creates an expectation of a future interest hike in period $t = 1$ through equation (79). The expectation of a recession next period engineered by the central bank creates a slump in aggregate demand today, thus calling for an even more aggressive cut in interest rates at 0.[77] This monetary transmission mechanism is presented schematically in Figure 6.

These multiple feedback effects lead the interest rate that closes the output gap to be lower with price extrapolation than in a rational world. Figure 7 represents graphically the equilibrium determination. In a rational world, the change in interest rates today has no impact on the expectation of

[75] In the event that the $t = 1$ REE equilibrium is not constrained efficient, then the interest rate at $t = 1$ could deviate from inflation targeting. This is the focus of Farhi and Werning (2020): an aggregate demand externality forces the social planner to try to reduce borrowing. The central bank would then target an interest rate such that, in my notation, $\mathcal{W}_1/dr_1 = 0$, by deviating from perfect inflation targeting. The intuition for the rest of this section would then be entirely similar, simply with a different expected interest rate r_1. This is also what would happen in my alternative model with the contemporaneous price in the collateral constraint. Note, however, that this is true only when macroprudential tools are constrained. If the planner can perfectly adapt counter-cyclical capital buffers and LTV ratios, the central bank in the REE benchmark of Farhi and Werning (2020) or in the alternative collateral constraint model would be perfectly targeting the output gap.

[76] The results in this section are also robust to an interest rule that is symmetric, i.e. when agents expect the central bank to lean against the wind, and to overheat the economy in order to stimulate the financial sector's sentiment.

[77] This is intuitively working like the now familiar forward guidance mechanism. Forward guidance consists of communicating such that agents expect future rates to be lower than normal, which is an expansionary force. Here, agents expect future rates to be higher than normal, which is a contractionary force.

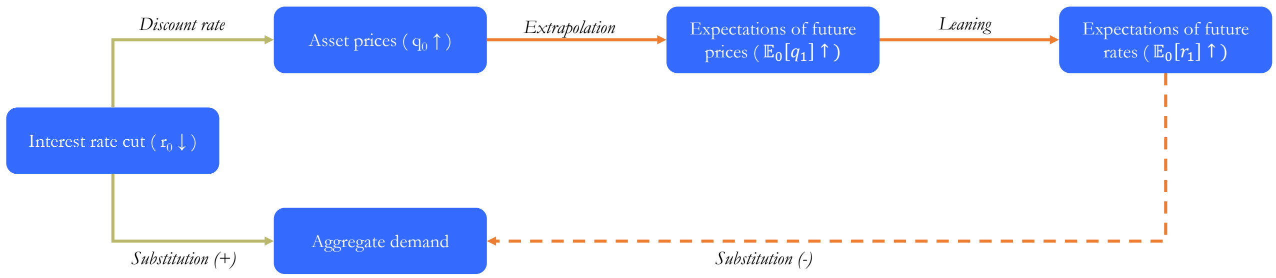

Figure 6: Monetary Transmission Mechanisms in the Model with Price Extrapolation.

future rates. This changes with extrapolation, creating a downward-sloping relation between current rates and tomorrow's rates. The equilibrium is determined at the intersection of this condition and the interest rate required to close the output gap, encoded in equation (79).

The following proposition provides a closed-form solution for the first-order approximation around the rational expectations equilibrium. The variables with stars are the rational values around which the approximation is taken.

Proposition 15 (Optimal Inflation Targeting at $t = 0$)**.** *The optimal interest rate at time* 0 *can be expressed as, in a first-order approximation around the rational benchmark* $\alpha \to 0$*:*

$$1 + r_0 = 1 + r_0^* - \frac{(1 + r_0^*)\rho\alpha(q_0^* - q_{-1})}{(1 + r_0^*)(1 + \rho q_0^*) + \rho q_0^*} \tag{81}$$

The properties of this expression are quite intuitive given the previous discussion. Two elements differ from the rational benchmark (where $\alpha = 0$). First, the term in the numerator quantifies how high asset prices today fuel extrapolation, which increases the expectations of future interest rates and thus forces the central bank to cut interest rates even more to close the output gap. Second, these effects are dampened by the terms proportional to ρ in the denominator. Indeed, when ρ is higher the central bank is expected to be very aggressive in its tightening next period, which decreases future asset prices and thus dampens exuberance today. As long as $r_0 \geq 0$, the central bank is still able to achieve perfect inflation targeting, but needs to be ready to cut rates more aggressively, and fuels irrational exuberance.

Trouble occurs when $r_0 < 0$, which is the case represented on Figure 8. With a zero lower bound constraint (ZLB), the central bank is unable to achieve full employment, because of endogenous expectations and the prospect of future leaning against the wind. If the ZLB is too severe, welfare losses can be so high that the planner would prefer to *not* lean against the wind in the future and suffer the welfare losses associated with higher exuberance. Nonetheless, the monetary authority runs into a time-consistency issue: as shown previously, it will always be optimal to tighten in period $t = 1$ if Ω_2 is high enough. These potential costs have thus to be factored in when the central bank contemplates leaning against the wind for the first time: it might involve a weakened power of conventional monetary policy in the future. It thus argues in favor of the view that monetary policy

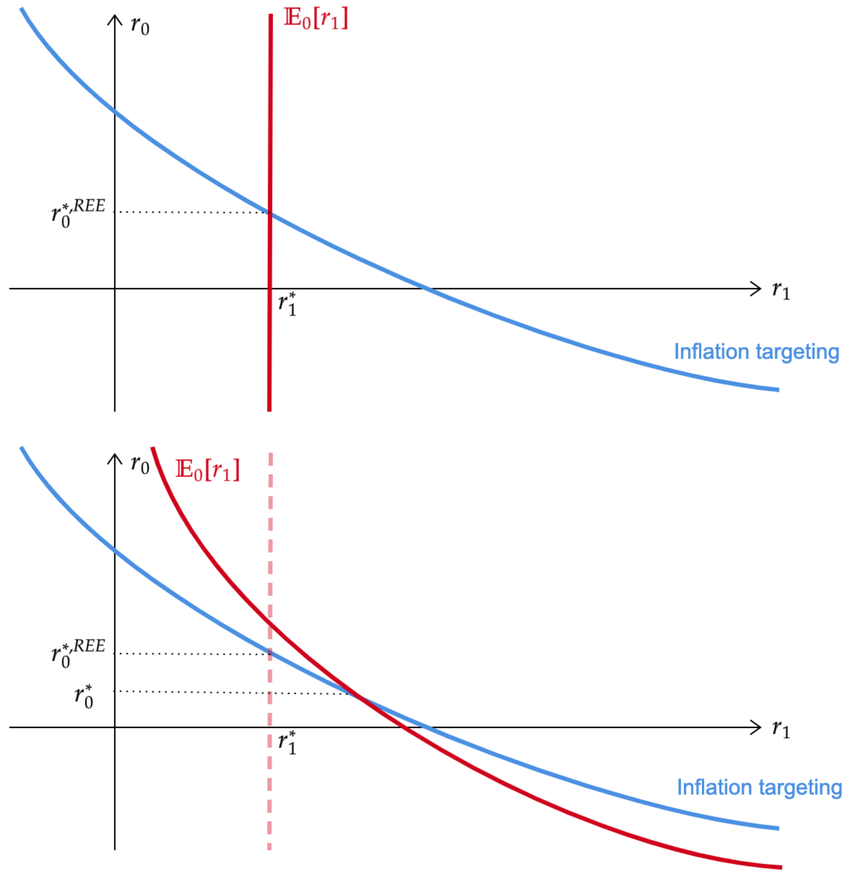

Figure 7: Equilibrium determination of the interest rate at $t = 0$ The red line represents the expectation of future interest rates by agents. The blue line represents the relation between current and future rates in order to close the output gap, as in equation (79). The top panel represent the rational case where expectations of future interest rates are independent of today's prices. The bottom panel features price extrapolation, and consequently the red line that represents expectations of future interest rates moves with current rates, through the impact on asset prices.

is not "the right tool for the job" (Bernanke 2002), albeit for a different and unexplored reason.[78] An immediate corollary of Proposition is that the central bank will have to react *more* to any demand shock, since its conventional stimulus power is weakened by expectations of future hikes through extrapolative beliefs. This is represented in Figure 9.

This result stands in contrast to the recent work of Boissay et al. (2021). In their paper, rule-based leaning against the wind is desirable: it amounts to providing households with an insurance against future aggregate shock. This helps smooth consumption, which reduces the incentives for households to accumulate capital. In their model, capital accumulation triggers financial crises, hence systematically leaning against the wind shields the economy against downturns. On the other hand, leaning against the wind in a discretionary way and late in the cycle triggers a fall

[78] Ueda and Valencia (2014) show that another time-consistency can arise when a central bank is in charge of price and financial stability: ex-post a crisis, the central bank has an incentive to inflate the debt away in order to reduce debt overhang problems.

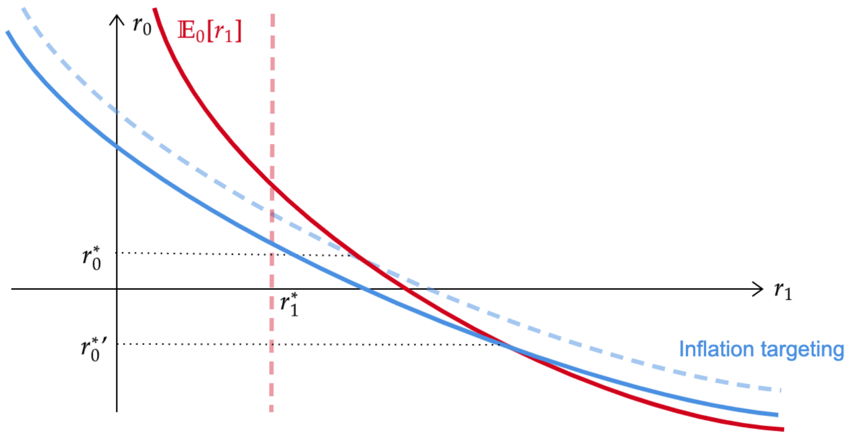

Figure 8: Central bank reaction and the Zero Lower Bound.

in aggregate demand that can itself cause a financial crisis.[79] These contrasting results highlight that the benefit and costs of leaning against the wind are dependent of the exact mechanism driving credit booms and busts. In my model, leaning against the wind works through the price of collateral assets and the expectation of agents, a feature orthogonal to the framework of Boissay et al. (2021).

This result is also separate from the recent work of Fanelli and Straub (2021). In their paper the central bank wants to lean against the wind of global capital flows, in order to dampen exchange rate movements. This is desirable because of a pecuniary externality, hence stabilizing the welfare of agents. The optimal policy in their case, however, calls for a smooth intervention, which leads to a time-inconsistency. The central bank promises future intervention, even though it will not be optimal anymore once the shock has passed. If the central bank lacks commitment, leaning against the wind is not optimal anymore. Proposition 15 highlights an entirely different form of time-inconsistency in my model: it can be optimal to commit to *never* lean against the wind because the anticipation of a future tightening has a negative impact on the conventional conduct of monetary policy today.

Remark 12 (Credit Booms at time 0). For simplicity, I assumed that financial intermediaries were only entering the model at $t = 1$, abstracting from the effects of loose monetary policy at $t = 0$ on credit. This is obviously an important part of conventional monetary policy, and this assumption was only made to focus on the key belief component that is at the heart of this paper. In a more complete version of this model the social planner would take into account how stimulating the economy at $t = 0$ creates a credit boom, such that financial intermediaries can already enter period $t = 1$ with high leverage. I leave this extension for future work.

Remark 13 (Other Mechanisms and Time-Inconsistency). The mechanism I just presented is only one

[79] The mechanism through which financial crises materialize in Boissay et al. (2021) comes from Boissay, Collard and Smets (2016). In their framework, firms are subject to idiosyncratic shocks, but the loan market is subject to financial frictions. The equilibrium then features multiple equilibria, and it is assumed that firms coordinate on the most efficient one. A small shock can make the good equilibrium disappear, and the loan market collapses to the lower-ranked equilibrium.

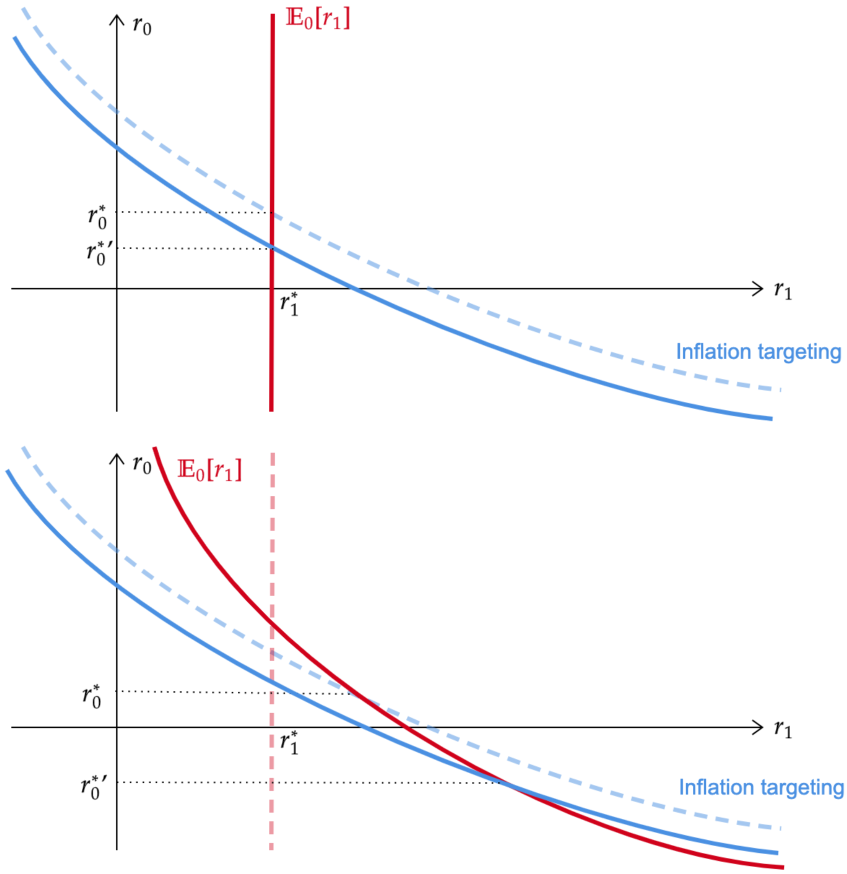

Figure 9: Central bank reaction after a negative demand shock. A negative demand shock exogenously changes the output gap relation, making it harder to reach perfect inflation targeting. In this specific example, this can be understood as a positive shock to β. The top panel represent the rational case where expectations of future interest rates are independent of today's prices. The bottom panel features price extrapolation, and consequently the red line that represents expectations of future interest rates moves with current rates, through the impact on asset prices.

way that this time-inconsistency might arise. Intuitively, only a few key features are needed for this result to emerge: (i) the central bank has some motive to lean against the wind when asset prices are higher than normal; (ii) private agents anticipate that there this is likely to occur in the future; and (iii) stimulating the economy today increases expectations of future asset prices. My model satisfies these conditions with only one unconventional assumption: agents' behavioral biases depend directly on recent prices.

7 Extensions and Robustness

The model presented above was deliberately stylized in order to flesh out the welfare implications of behavioral distortions in a model of financial crises. I discuss here the extensions presented in the Appendices. These various extensions show that the insights I uncovered for the conduct of optimal policy do not rely on the simplifying assumptions I made.

In the basic version of the model, households are only passively lending to financial intermediaries, and the only production is in the creation of collateral assets in period $t = 1$ by financial intermediaries. In Appendix B I extend the framework to allow for a real production sector in the intermediate period ($t = 2$): households supply labor to competitive firms, but a financial friction requires firms to borrow from financial intermediaries in order to cover a fraction of the wage bill in advance. When financial intermediaries are constrained (i.e. in a financial crisis), they cannot lend to the real sector the amount needed to obtain the desired level of output, which result in a fall in employment.[80] I show that the contraction in output is also driven by over-pessimism, extending the belief amplification mechanism to the real side of the economy. Finally, I extend the welfare analysis and show that adding production only adds a collateral externality term to the welfare objective of the planner, which is again proportional to the sensitivity of sentiment to changes in asset prices.

In Appendix C, I study an alternative formulation of the collateral constraint. There, contemporaneous prices are directly determining the borrowing capacity of financial intermediaries, creating pecuniary externalities even in the benchmark rational case. I nevertheless show that when agents are subject to behavioral biases, the analytical insights I uncovered are valid. In particular, the phenomenon of belief amplification is compounded by traditional financial amplification in the collateral and reversal externality expressions. The Ω-uncertainty implications are also preserved, highlighting the robustness of my results. The extensions with production, bailouts and monetary policy featuring the contemporaneous price in the collateral constraint are presented in Online Appendix E.

I allow for the simultaneous choice of ex-ante and ex-post policies in Appendix D. As in Jeanne and Korinek (2020), ex-ante regulation is still desirable even if ex-post liquidity injections are used. I study how moral hazard, due to the anticipation of future bailouts, is modified by the presence of behavioral distortions. Here again, whether sentiment is endogenous to asset prices or not matters. If sentiment is purely driven by exogenous shocks to fundamentals, moral hazard concerns are actually less acute than in a rational model. Indeed, when agents are too optimistic they expect financial crises to be less severe than in reality, which causes them to expect smaller bailouts than in reality. When sentiment comes from asset prices, however, anticipating bailouts will raise the attractiveness of holding financial assets since their price will be supported by the government in a crisis. This in turn exacerbates exuberance, and can backfire by pushing up leverage even more.

The baseline model also assumed that all agents hold the same beliefs, and thus that behavioral biases were homogeneously distributed in the population. The empirical literature finds widespread evidence of belief heterogeneity, however (Giglio, Maggiori, Stroebel and Utkus 2021; Mian and Sufi 2021; Meeuwis, Parker, Schoar and Simester 2021). Online Appendix F shows that my insights are preserved when I allow for a distribution of beliefs in the population around an average bias: only small modifications to the welfare decomposition of Proposition 1 are required.

[80] Chodorow-Reich (2014) shows that the reduction in firm borrowing from banks in the US can account for "between one-third and one-half of the employment decline at small and medium firms in the sample in the year following the Lehman bankruptcy." Similar results for Spain have been found by Bentolila, Jansen and Jiménez (2018).

Furthermore, I show that a leverage limit is robust to the introduction of heterogeneity, whereas an anonymous leverage tax looses its ability to fully restore the second-best.

Online Appendix G presents additional empirical evidence regarding the comovement of sentiment with the health of financial intermediaries. I use various measures of sentiment to highlight the robustness of this finding. Online Appendix H presents additional results regarding the impact of sentiment uncertainty on optimal policy. Online Appendix I presents a simple infinite-horizon version of the model to show that the welfare decomposition takes the same form and features the same forces identified in my baseline framework. Online Appendix J considers multiple psychological models of asset prices that have been proposed by the behavioral economics literature. I present the Ω-formulation that corresponds to these models, and highlight how their features imply different policy consequences.

8 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. The model features a collateral constraint and a general class of deviations from rationality. This allows me to isolate the properties of behavioral factors that matter for financial stability, and their interactions with financial frictions.

I derive a general welfare decomposition and use this breakdown to present several practical policy implications, some being natural and intuitive, others being more surprising. Naturally, over-optimism is a source of concern for the planner, and motivates stricter leverage restrictions, but only when there is a possibility of binding financial frictions in the future. More surprisingly, sentiment *inside* a crisis comoving with the health of the financial sector is a source of welfare loss that also calls for early intervention. Furthermore, the precise form of behavioral biases matters for welfare. Endogenous behavioral biases that develop through the observation of equilibrium prices or returns create novel externalities, even in models that do not feature any room for policy in their rational benchmark. Agents neglect that their actions impact current and future prices, which in turn impact sentiment inside a financial crisis. A practical implication is, therefore, that policymakers need additional instruments to control asset prices, since regulating only quantities becomes insufficient. While counter-cyclical capital buffers and LTV ratios are desirable, they need to be complemented with monetary policy.

I show that adding uncertainty about the precise extent of irrational exuberance in financial markets actually increases the incentives for the planner to act early by imposing restrictions in good times. This is due to a key non-linear interaction between sentiment and financial crises, which creates a role for precautionary restrictions. Counter-cyclical capital buffers thus need to be increased in times of heightened uncertainty. Finally, I show that monetary policy can play the role of such an additional instrument: under endogenous behavioral biases, leaning against the wind can be desirable even if capital buffers and LTV regulations are fully unconstrained. This effect is independent of the extent of irrational over-optimism: the central bank is concerned that high asset prices today might create extrapolation later on, and thus acts to temper the price boom. The systematic use of leaning against the wind, however, has costs. It can weaken the conventional stimulus power of interest rate cuts when agents expect leaning against the wind to happen in the future. This is due to a feedback loop between current prices, the expectations of future prices, and the expectations of future interest rates. It introduces a time-inconsistency, and can force the economy to hit the zero lower bound in normal times.

While the model can be extended along several dimensions, the results suggest a need for research on two specific dimensions. First, it is a recurrent theme of this paper that the specific form of deviations from rationality greatly matters for welfare. I showed which features of behavioral biases need to be quantified by future research. Furthermore when sentiment depends directly on asset prices, policy can influence outcomes by directly influencing beliefs. At the same time, it implies that allocations not only depend on past allocations, but also on past prices. On the other hand, if sentiment is driven by purely exogenous factors like fundamentals, irrational distress during crises is costly for welfare but policy will not be able to counteract it ex-post. While empirical research has convincingly demonstrated that overreaction, and thus optimism in good times and pessimism in bad times, is a feature of financial markets, we have less certainty about its drivers.[81] 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, I only scratched the surface of the dynamic tradeoffs faced by the central bank once leaning against the wind is part of the regulatory toolbox. In my model the small number of periods obfuscates the timing subtleties faced by the central bank. By stimulating the economy today, the monetary authority recognizes that it might trigger a credit boom and a surge in irrational exuberance, something it will have to fight in the future. 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. In addition, financial crises are often slow to develop even after substantial growth in credit and asset prices (Greenwood et al. 2020). Further empirical and theoretical research is needed to fully grasp the complex timing interactions between policy, crises, and behavioral biases.

[81] McCarthy and Hillenbrand (2021) propose a model with extrapolative beliefs on dividends and returns. They estimate that fundamental extrapolation explains 34% of movements in the S&P500 index, while return extrapolation accounts for 23%.

References

Adam, Klaus, Albert Marcet, and Johannes Beutel, "Stock price booms and expected capital gains," *American Economic Review*, 2017, *107* (8), 2352–2408.

— , — , **and** — , "Stock price booms and expected capital gains," *American Economic Review*, 2017, *107* (8), 2352–2408.

— **and** — , "Internal rationality, imperfect market knowledge and asset prices," *Journal of Economic Theory*, 2011, *146* (3), 1224–1252.

— **and Michael Woodford**, "Robustly optimal monetary policy in a new Keynesian model with housing," *Journal of Economic Theory*, 2021, *198*, 105352.

Allen, Franklin, Gadi Barlevy, and Douglas Gale, "Asset Price Booms and Macroeconomic Policy: A Risk-Shifting Approach," *American Economic Journal: Macroeconomics*, forthcoming.

Bahaj, Saleem and Angus Foulis, "Macroprudential Policy under Uncertainty," *International journal of central banking*, 2017, *13* (3), 119–154.

Baker, Malcolm and Jeffrey Wurgler, "The equity share in new issues and aggregate stock returns," *the Journal of Finance*, 2000, *55* (5), 2219–2257.

— **and** — , "Investor sentiment in the stock market," *Journal of Economic Perspectives*, 2007, *21* (2), 129–152.

— **, Jeremy C Stein, and Jeffrey Wurgler**, "When does the market matter? Stock prices and the investment of equity-dependent firms," *The Quarterly Journal of Economics*, 2003, *118* (3), 969–1005.

Barberis, Nicholas, "Psychology-based models of asset prices and trading volume," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 1, Elsevier, 2018, pp. 79–175.

— **and Andrei Shleifer**, "Style investing," *Journal of Financial Economics*, 2003, *68* (2), 161–199.

— , — **, and Robert Vishny**, "A model of investor sentiment," *Journal of Financial Economics*, 1998, *49* (3), 307–343.

— **, Robin Greenwood, Lawrence Jin, and Andrei Shleifer**, "Extrapolation and bubbles," *Journal of Financial Economics*, 2018, *129* (2), 203–227.

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

Basel Committee on Banking Supervision, *Basel III: A global regulatory framework for more resilient banks and banking systems* 2011.

Bastianello, Francesca and Paul Fontanier, "Partial Equilibrium Thinking in General Equilibrium," *Department of Economics, Harvard University*, 2021.

Bentolila, Samuel, Marcel Jansen, and Gabriel Jiménez, "When credit dries up: Job losses in the great recession," *Journal of the European Economic Association*, 2018, *16* (3), 650–695.

Bernanke, Ben S, "Asset-Price" Bubbles" and Monetary Policy," *Speech before the New York Chapter of the National Association for Business Economics*, 2002.

— **and Kenneth N Kuttner**, "What explains the stock market's reaction to Federal Reserve policy?," *The Journal of finance*, 2005, *60* (3), 1221–1257.

— **and Mark Gertler**, "Monetary policy and asset price volatility," Technical Report 2000.

Bianchi, Francesco, Cosmin Ilut, and Hikaru Saijo, "Diagnostic Business Cycles," *NBER Working Paper*, 2021.

Bianchi, Javier, "Overborrowing and systemic externalities in the business cycle," *American Economic Review*, 2011, *101* (7), 3400–3426.

— , “Efficient bailouts?,” *American Economic Review*, 2016, *106* (12), 3607–59.

— **and Enrique G Mendoza**, “Optimal time-consistent macroprudential policy,” *Journal of Political Economy*, 2018, *126* (2), 588–634.

Biljanovska, Nina, Lucyna Gornicka, and Alexandros Vardoulakis, “Optimal macroprudential policy and asset price bubbles,” *IMF Working Paper*, 2019.

Blinder, Alan S and Ricardo Reis, “Economic performance in the Greenspan Era: The evolution of events and ideas,” *The Federal Reserve Bank of Kansas City (ed.), The Greenspan Era: Lessons for the Future*, 2005.

Boissay, Frédéric, Fabrice Collard, and Frank Smets, “Booms and banking crises,” *Journal of Political Economy*, 2016, *124* (2), 489–538.

Boissay, Frederic, Fabrice Collard, Jordi Galí, and Cristina Manea, “Monetary policy and asset price bubbles: a laboratory experiment,” Technical Report 2021.

Bondt, Werner FM De and Richard H Thaler, “Financial decision-making in markets and firms: A behavioral perspective,” *Handbooks in operations research and management science*, 1995, *9*, 385–410.

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

— , — , — , and Stephen J Terry, “Real credit cycles,” *NBER Working Paper*, 2021.

— , — , Rafael La Porta, and Andrei Shleifer, “Diagnostic expectations and stock returns,” *The Journal of Finance*, 2019, *74* (6), 2839–2874.

— , — , — , and — , “Expectations of fundamentals and stock market puzzles,” *NBER Working Paper*, 2020.

Bordo, Michael D and Olivier Jeanne, “Monetary policy and asset prices: does ‘benign neglect’make sense?,” *International Finance*, 2002, *5* (2), 139–164.

Borio, Claudio, “Towards a macroprudential framework for financial supervision and regulation?,” *CESifo Economic Studies*, 2003, *49* (2), 181–215.

Borio, Claudio EV and Philip William Lowe, “Asset prices, financial and monetary stability: exploring the nexus,” *BIS working paper*, 2002.

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar, “Sticky expectations and the profitability anomaly,” *The Journal of Finance*, 2019, *74* (2), 639–674.

Brainard, William C, “Uncertainty and the Effectiveness of Policy,” *The American Economic Review*, 1967, *57* (2), 411–425.

Brueckner, Jan K, “Mortgage default with asymmetric information,” *The Journal of Real Estate Finance and Economics*, 2000, *20* (3), 251–274.

Brunnermeier, Markus K, Alp Simsek, and Wei Xiong, “A welfare criterion for models with distorted beliefs,” *The Quarterly Journal of Economics*, 2014, *129* (4), 1753–1797.

— **and Yuliy Sannikov**, “A macroeconomic model with a financial sector,” *American Economic Review*, 2014, *104* (2), 379–421.

Caballero, Ricardo J and Alp Simsek, “Prudential monetary policy,” *MIT Working Paper*, 2020.

— **and —** , “A risk-centric model of demand recessions and speculation,” *The Quarterly Journal of Economics*, 2020, *135* (3), 1493–1566.

— **and Arvind Krishnamurthy**, “Excessive dollar debt: Financial development and underinsurance,” *The Journal of Finance*, 2003, *58* (2), 867–893.

Caines, Colin and Fabian Winkler, “Asset price learning and optimal monetary policy,” *Federal Reserve Working Paper*, 2019.

Camous, Antoine and Alejandro Van der Ghote, "Financial Cycles under Diagnostic Beliefs," *Available at SSRN 3786680*, 2021.

Campbell, John Y and Joao F Cocco, "A model of mortgage default," *The Journal of Finance*, 2015, *70* (4), 1495–1554.

— **and Robert J Shiller**, "Stock prices, earnings, and expected dividends," *the Journal of Finance*, 1988, *43* (3), 661–676.

Cassella, Stefano and Huseyin Gulen, "Extrapolation bias and the predictability of stock returns by price-scaled variables," *The Review of Financial Studies*, 2018, *31* (11), 4345–4397.

Chahrour, Ryan and Gaetano Gaballo, "Learning from house prices: Amplification and business fluctuations," *The Review of Economic Studies*, 2021, *88* (4), 1720–1759.

Chen, Mr Jiaqian, Daria Finocchiaro, Jesper Lindé, and Karl Walentin, "The costs of macroprudential deleveraging in a liquidity trap," Technical Report 2020.

Chodorow-Reich, Gabriel, "The employment effects of credit market disruptions: Firm-level evidence from the 2008–9 financial crisis," *The Quarterly Journal of Economics*, 2014, *129* (1), 1–59.

— **, Adam M Guren, and Timothy J McQuade**, "The 2000s Housing Cycle With 2020 Hindsight: A Neo-Kindlebergerian View," *NBER Working Paper*, 2021.

Cingano, Federico, Francesco Manaresi, and Enrico Sette, "Does credit crunch investment down? New evidence on the real effects of the bank-lending channel," *The Review of Financial Studies*, 2016, *29* (10), 2737–2773.

Claessens, Mr Stijn, "An overview of macroprudential policy tools," *IMF Working Paper*, 2014.

Clayton, Christopher and Andreas Schaab, "Bail-Ins, Optimal Regulation, and Crisis Resolution," *Available at SSRN 3455684*, 2020.

— **and** — , "Multinational Banks and Financial Stability," *Available at SSRN 3512148*, 2020.

Coibion, Olivier and Yuriy Gorodnichenko, "What can survey forecasts tell us about information rigidities?," *Journal of Political Economy*, 2012, *120* (1), 116–159.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, "Investor psychology and security market under-and overreaction," *The Journal of Finance*, 1998, pp. 1839–1885.

Dávila, Eduardo and Ansgar Walther, "Prudential Policy with Distorted Beliefs," *Working Paper*, 2021.

— **and Anton Korinek**, "Pecuniary externalities in economies with financial frictions," *The Review of Economic Studies*, 2018, *85* (1), 352–395.

De La O, Ricardo and Sean Myers, "Subjective cash flow and discount rate expectations," *The Journal of Finance*, 2021, *76* (3), 1339–1387.

DeFusco, Anthony A, Charles G Nathanson, and Eric Zwick, "Speculative dynamics of prices and volume," Technical Report, National Bureau of Economic Research 2017.

Dell'Ariccia, Giovanni, Enrica Detragiache, and Raghuram Rajan, "The real effect of banking crises," *Journal of Financial Intermediation*, 2008, *17* (1), 89–112.

DeLong, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann, "Positive feedback investment strategies and destabilizing rational speculation," *the Journal of Finance*, 1990, *45* (2), 379–395.

Dewatripont, Mathias and Jean Tirole, "Liquidity regulation, bail-ins and bailouts," *Working Paper*, 2018.

Dudley, William, "Is the active use of macroprudential tools institutionally realistic?," *Federal Reserve Bank of New York Working Paper*, 2015.

Eggertsson, Gauti B and Paul Krugman, "Debt, deleveraging, and the liquidity trap: A Fisher-Minsky-Koo approach," *The Quarterly Journal of Economics*, 2012, *127* (3), 1469–1513.

Fanelli, Sebastián and Ludwig Straub, "A Theory of Foreign Exchange Interventions," *The Review of Economic Studies*, 03 2021.

Farhi, Emmanuel and Iván Werning, "A theory of macroprudential policies in the presence of nominal rigidities," *Econometrica*, 2016, *84* (5), 1645–1704.

— **and** — , "Taming a Minsky Cycle," Technical Report 2020.

— **and Jean Tirole**, "Bubbly liquidity," *The Review of Economic Studies*, 2012, *79* (2), 678–706.

— **and** — , "Collective moral hazard, maturity mismatch, and systemic bailouts," *American Economic Review*, 2012, *102* (1), 60–93.

— **and** — , "Deadly embrace: Sovereign and financial balance sheets doom loops," *The Review of Economic Studies*, 2018, *85* (3), 1781–1823.

— **and Stavros Panageas**, "The real effects of stock market mispricing at the aggregate: Theory and Empirical Evidence," *Available at SSRN 720462*, 2004.

— **and Xavier Gabaix**, "Optimal taxation with behavioral agents," *American Economic Review*, 2020, *110* (1), 298–336.

Fisher, Irving, *Booms and depressions: some first principles*, Adelphi Company New York, 1932.

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

Gabaix, Xavier, "Behavioral inattention," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 261–343.

Galí, Jordi, "Monetary policy and rational asset price bubbles," *American Economic Review*, 2014, *104* (3), 721–52.

— **and Luca Gambetti**, "The effects of monetary policy on stock market bubbles: Some evidence," *American Economic Journal: Macroeconomics*, 2015, *7* (1), 233–57.

— **, Giovanni Giusti, and Charles N Noussair**, "Monetary policy and asset price bubbles: a laboratory experiment," Technical Report 2021.

Gaspar, Vitor, Frank Smets, and David Vestin, "Inflation expectations, adaptive learning and optimal monetary policy," in "Handbook of monetary economics," Vol. 3, Elsevier, 2010, pp. 1055–1095.

Geanakoplos, John D and Heracles M Polemarchakis, "Existence, regularity, and constrained suboptimality of competitive allocations when the asset market is incomplete," *Cowles Foundation Discussion Papers*, 1985, *764.*

Geithner, Timothy F, *Stress test: Reflections on financial crises*, Broadway Books, 2014.

Gennaioli, Nicola and Andrei Shleifer, *A crisis of beliefs*, Princeton University Press, 2018.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, "Five facts about beliefs and portfolios," *American Economic Review*, 2021, *111* (5), 1481–1522.

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

Gorton, Gary, *Misunderstanding financial crises: Why we don't see them coming*, Oxford University Press, 2012.

Gourio, François, Anil K Kashyap, and Jae W Sim, "The trade offs in leaning against the wind," *IMF Economic Review*, 2018, *66* (1), 70–115.

Greenspan, Alan, "Economic Volatility," *Remarks by Chairman Alan Greenspan at a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming*, 2002.

Greenwald, Bruce C and Joseph E Stiglitz, "Externalities in economies with imperfect information and incomplete markets," *The quarterly journal of economics*, 1986, *101* (2), 229–264.

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

_ , _ , and Lawrence J Jin, "Reflexivity in credit markets," *NBER Working Paper*, 2019.

_ , _ , Andrei Shleifer, and Jakob Ahm Sørensen, "Predictable Financial Crises," Technical Report, Harvard Business School Working Paper 2020.

Gromb, Denis and Dimitri Vayanos, "Equilibrium and welfare in markets with financially constrained arbitrageurs," *Journal of Financial Economics*, 2002, *66* (2-3), 361–407.

Gruber, Jonathan and Botond Köszegi, "Is addiction "rational"? Theory and evidence," *The Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

Guerrieri, Veronica and Guido Lorenzoni, "Credit crises, precautionary savings, and the liquidity trap," *The Quarterly Journal of Economics*, 2017, *132* (3), 1427–1467.

Haddad, Valentin and Tyler Muir, "Do intermediaries matter for aggregate asset prices?," *The Journal of Finance*, forthcoming.

Hansen, Lars Peter and Thomas J Sargent, *Robustness*, Princeton university press, 2011.

Harper, Daniel and Anton Korinek, "Macroprudential Taxes, Ratios, and Rents: Equivalences and Not-Quite Equivalences," Technical Report 2021.

Hart, Oliver D, "On the optimality of equilibrium when the market structure is incomplete," *Journal of economic theory*, 1975, *11* (3), 418–443.

He, Zhiguo and Arvind Krishnamurthy, "A macroeconomic framework for quantifying systemic risk," *American Economic Journal: Macroeconomics*, 2019, *11* (4), 1–37.

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

_ , In Gu Khang, and Arvind Krishnamurthy, "Balance sheet adjustments during the 2008 crisis," *IMF Economic Review*, 2010, *58* (1), 118–156.

Herreño, Juan and Carlos Rondón-Moreno, "Overborrowing and Systemic Externalities in the Business Cycle Under Imperfect Information," *Available at SSRN 3597734*, 2020.

Hong, Harrison and Jeremy C Stein, "A unified theory of underreaction, momentum trading, and overreaction in asset markets," *The Journal of finance*, 1999, *54* (6), 2143–2184.

Jeanne, Olivier and Anton Korinek, "Managing credit booms and busts: A Pigouvian taxation approach," *Journal of Monetary Economics*, 2019, *107*, 2–17.

_ and _ , "Macroprudential regulation versus mopping up after the crash," *The Review of Economic Studies*, 2020, *87* (3), 1470–1497.

Jones, Bradley, "Asset bubbles: Re-thinking policy for the age of asset management," *IMF Working Paper*, 2015.

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

Khorrami, Paymon and Fernando Mendo, "Rational Sentiment and Financial Frictions," Technical Report, Working Paper 2021.

Kindleberger, Charles P, *Manias, panics and crashes: a history of financial crises*, Palgrave Macmillan, 1978.

Kirti, Divya, "Lending standards and output growth," *IMF Working paper*, 2018.

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

Kohn, Donald, "The Strength in Consumer Durables and Housing: Policy Stabilization or Problem in the Making?," *Remarks by Governor Donald L. Kohn*, 2003.

Korinek, Anton and Alp Simsek, "Liquidity trap and excessive leverage," *American Economic Review*, 2016, *106* (3), 699–738.

Krishnamurthy, Arvind and Tyler Muir, "How credit cycles across a financial crisis," *NBER Working Paper*, 2020.

__ and Wenhao Li, "Dissecting mechanisms of financial crises: Intermediation and sentiment," *NBER Working Paper*, 2020.

__ , Stefan Nagel, and Dmitry Orlov, "Sizing up repo," *The Journal of Finance*, 2014, *69* (6), 2381–2417.

L'Huillier, Jean-Paul, Sanjay R Singh, and Donghoon Yoo, "Incorporating Diagnostic Expectations into the New Keynesian Framework," *Available at SSRN 3910318*, 2021.

Lipsey, Richard G and Kelvin Lancaster, "The general theory of second best," *The Review of Economic Studies*, 1956, *24* (1), 11–32.

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

Lorenzoni, Guido, "Inefficient credit booms," *The Review of Economic Studies*, 2008, *75* (3), 809–833.

Ma, Yueran, "Nonfinancial firms as cross-market arbitrageurs," *The Journal of Finance*, 2019, *74* (6), 3041–3087.

Martin, Alberto and Jaume Ventura, "Managing credit bubbles," *Journal of the European Economic Association*, 2016, *14* (3), 753–789.

Martinez-Miera, David and Rafael Repullo, "Monetary policy, macroprudential policy, and financial stability," *Annual Review of Economics*, 2019, *11*, 809–832.

Maxted, Peter, "A Macro-Finance Model with Sentiment," *Working Paper*, 2020.

McCarthy, Odhrain and Sebastian Hillenbrand, "Heterogeneous Investors and Stock Market Fluctuations," *Available at SSRN 3944887*, 2021.

Meeuwis, Maarten, Jonathan A Parker, Antoinette Schoar, and Duncan I Simester, "Belief disagreement and portfolio choice," *Working Paper*, 2021.

Mendoza, Enrique G, "Sudden stops, financial crises, and leverage," *American Economic Review*, 2010, *100* (5), 1941–66.

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

__ and __ , "Credit Supply and Housing Speculation," *The Review of Financial Studies*, 03 2021.

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

Montamat, Giselle and Francisco Roch, "Robust Optimal Macroprudential Policy," *IMF Working Paper*, 2021.

Mullainathan, Sendhil, Joshua Schwartzstein, and William J Congdon, "A reduced-form approach to behavioral public finance," *Annual Review of Economics*, 2012, *4* (1), 511–540.

Odean, Terrance, "Volume, volatility, price, and profit when all traders are above average," *The journal of finance*, 1998, *53* (6), 1887–1934.

O'Donoghue, Ted and Matthew Rabin, "Optimal sin taxes," *Journal of Public Economics*, 2006, *90* (10-11), 1825–1849.

Ottonello, Pablo, Diego J Perez, and Paolo Varraso, "Are Collateral-Constraint Models Ready for Macroprudential Policy Design?," *Working Paper*, 2021.

Pflueger, Carolin, Emil Siriwardane, and Adi Sunderam, "Financial market risk perceptions and the macroeconomy," *The Quarterly Journal of Economics*, 2020, *135* (3), 1443–1491.

Rabin, Matthew and Dimitri Vayanos, "The gambler's and hot-hand fallacies: Theory and applications," *The Review of Economic Studies*, 2010, *77* (2), 730–778.

Richter, Björn and Kaspar Zimmermann, "The profit-credit cycle," *Working Paper*, 2021.

Rigobon, Roberto and Brian Sack, "The impact of monetary policy on asset prices," *Journal of Monetary Economics*, 2004, *51* (8), 1553–1575.

Rognlie, Matthew, Andrei Shleifer, and Alp Simsek, "Investment hangover and the great recession," *American Economic Journal: Macroeconomics*, 2018, *10* (2), 113–53.

Schmitt-Grohé, Stephanie and Martín Uribe, "Multiple equilibria in open economies with collateral constraints," *The Review of Economic Studies*, 2021, *88* (2), 969–1001.

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

_, Lucas Ter Steege, and Felix Ward, "Leaning against the wind and crisis risk," *American Economic Review: Insights*, 2021, *3* (2), 199–214.

Securities and Exchange Commission, *Primer: Money Market Funds and the Repo Market* 2021.

Shiller, Robert J, *Irrational exuberance*, Princeton university press, 2015.

Shleifer, Andrei and Robert W Vishny, "Liquidation values and debt capacity: A market equilibrium approach," *The Journal of Finance*, 1992, *47* (4), 1343–1366.

_ and _, "The Limits of Arbitrage," *The Journal of Finance*, 1997, *52* (1), 35–55.

Smets, Frank, "Financial stability and monetary policy: How closely interlinked?," *International Journal of Central Banking*, 2014.

Stein, Jeremy C, "Rational Capital Budgeting in an Irrational World," *Journal of Business*, 1996, *69* (4), 429–455.

_, "Can Policy Tame the Credit Cycle?," *IMF Economic Review*, 2021, *69* (1), 5–22.

Stiglitz, Joseph E, "The inefficiency of the stock market equilibrium," *The Review of Economic Studies*, 1982, *49* (2), 241–261.

Sufi, Amir and Alan M Taylor, "Financial crises: A survey," *NBER Working Paper*, 2021.

Svensson, Lars EO, "Cost-benefit analysis of leaning against the wind," *Journal of Monetary Economics*, 2017, *90*, 193–213.

Tirole, Jean, "Asset bubbles and overlapping generations," *Econometrica*, 1985, pp. 1499–1528.

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

Ueda, Kenichi and Fabian Valencia, "Central bank independence and macro-prudential regulation," *Economics Letters*, 2014, *125* (2), 327–330.

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

Winkler, Fabian, "The role of learning for asset prices and business cycles," *Journal of Monetary Economics*, 2020, *114*, 42–58.

Woodford, Michael, "Inflation targeting and financial stability," Technical Report, National Bureau of Economic Research 2012.

Yellen, Janet L, "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," Vol. 16 2009.

Appendices

A Proofs and Derivations

A.1 Proof of Proposition 1

At time $t = 2$, the welfare of financial intermediaries can be written as:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]\right) + \beta^2 \left(\mathbb{E}_2[z_3] H - \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}_2[z_3] H + n_2\right) & \text{otherwise} \end{cases} \tag{A.1}$$

with $n_2 = z_2 H - d_1(1 + r_1)$, while the Lagrangian corresponding to bankers' problem in period $t = 1$ is given by:

$$\mathcal{L}_{b,1} = \left[u(c_1) + \mathbb{E}_1[\mathcal{W}_2(n_2, H; q_2, z_2)]\right] - \lambda_1\left[c_1 + c(H) - d_1 - e_1\right] \tag{A.2}$$

the first-order condition on borrowing gives:

$$\frac{\partial \mathcal{L}_{b,1}}{\partial d_1} = \lambda_1 - \mathbb{E}_1\left[\lambda_2\right] \tag{A.3}$$

where λ_t is the Lagrange multiplier on the budget constraint at time t. The social planner maximizes the same function, but under its own expectations, and by also taking into account how a change in d_1 impacts asset prices in period 2. This leads to the following first-order condition:

$$\frac{\partial \mathcal{L}_{b,1}^{SP}}{\partial d_1} = \lambda_1 + \mathbb{E}_1^{SP}\left[\lambda_2\right] - \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2} \frac{\partial q_2}{\partial n_2}\right] \frac{dn_2}{dd_1} \tag{A.4}$$

where κ_2 is the Lagrange multiplier on the collateral constraint at $t = 2$. Hence simply by incorporating $\mathbb{E}_1[\lambda_2]$ we can express the total change in welfare as internalized plus uninternalized effects:

$$\frac{\partial \mathcal{L}_{b,1}^{SP}}{\partial d_1} = \underbrace{\lambda_1 - \mathbb{E}_1\left[\lambda_2\right]}_{\text{Internalized}} + \underbrace{\mathbb{E}_1\left[\lambda_2\right] - \beta \mathbb{E}_1^{SP}\left[\lambda_2\right] - \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2} \frac{\partial q_2}{\partial n_2}\right]}_{\text{Uninternalized}} \tag{A.5}$$

which proves Proposition 1. □

A.2 Proof of Proposition 2

I compute the difference between λ_2 expected by private agents and λ_2 expected by the Planner state by state z_2. When both expect a realization z_2 not to produce a financial crisis, marginal utilities are equalized to 1, so the difference disappears. For the rest there are two cases: either both marginal utilities correspond to binding collateral constraints, either one agent expect the friction to bind and the other not. The first case yields:

$$\frac{1}{c_2(z_2+\Omega_2,0)} - \frac{1}{c_2(z_2,\Omega_3)} = \frac{1}{(z_2+\Omega_2)H - d_1(1+r_1) + \phi H\mathbb{E}_2[z_3]} - \frac{1}{z_2 H - d_1(1+r_1) + \phi H\mathbb{E}_2[z_3+\Omega_3])} \tag{A.6}$$

I take the first-order approximation around the REE $\lambda_2 = 1/(z_2H - d_1(1+r_1) + \phi H\mathbb{E}_2[z_3]) = 1/c_2(z_2,0)$. It gives:

$$\frac{1}{(z_2+\Omega_2)H - d_1(1+r_1) + \phi H\mathbb{E}_2[z_3]} = \frac{1}{c_2(z_2,0)}\frac{1}{1+\frac{\Omega_2 H}{c_2(z_2,0)}} = \lambda_2\left(1 - \frac{\Omega_2 H}{c_2(z_2,0)}\right) \tag{A.7}$$

While the same algebra for the second part of equation (A.6) yields similarly:

$$\frac{1}{z_2H - d_1(1+r_1) + \phi H\mathbb{E}_2[z_3+\Omega_3]} = \frac{1}{c_2(z_2,0)}\frac{1}{1+\frac{\phi\Omega_3 H}{c_2(z_2,0)}} = \lambda_2\left(1 + \frac{\phi H\Omega_3}{c_2(z_2,0)}\right) \tag{A.8}$$

Taking the difference gives:

$$\frac{1}{c_2(z_2+\Omega_2,0)} - \frac{1}{c_2(z_2,\Omega_3)} = \lambda_2^2\left(H\Omega_2 - \phi H\Omega_3\right) \tag{A.9}$$

Lastly we need to consider the cases where the social planner and private agents disagree about the occurrence of a crisis for a given z_2. Without loss of generality, I assume that private agents are over-optimistic so for some range of states, $[z^* - dz, z^*]$ they expect to be at $c_2 = 1$, while the Planner expects the collateral constraint to be binding (where z^* is the crisis cutoff in the RE case). The size of the band is infinitesimal since, as can be seen in equations (32) and (33), the cutoff is only moving because of Ω_2 and Ω_3 which are small.

The difference, integrated on the band, can be expressed through a triangle approximation:

$$\int_{z^*-dz}^{z^*}\left(1 - \frac{1}{c_2(z_2,\Omega_3)}\right)\pi(z_2)dz_2 = \frac{dz\pi(z^*)}{2}\left(1 - \frac{1}{c_2(z^*-dz,\Omega_3)}\right) \tag{A.10}$$

Because the difference between $t = 1$ and $1/c_2(z^*-dz,\Omega_3^*)$, where Ω_3^* is the bias at the cutoff, is also infinitesimal, this term is negligible compared to the previous one.[82] It thus follows that, to the

[82] For completeness, its value can be approximated as:

$$\int_{z^*-dz}^{z^*}\left(1 - \frac{1}{c_2(z_2,\Omega_3)}\right)\pi(z_2)dz_2 \approx -(\Omega_2 - \phi H\Omega_3(z^*))\frac{(\Omega_2 - \phi H\Omega_3(z^*)) - \phi H\Omega_3(z^*)}{2}\pi(z^*)$$

Ω_2 enters this equation because it parametrizes the value of dz, i.e. the size of the band where agents do not expect a financial crisis but the planner does.

first order:

$$\mathcal{B}_d \simeq -\mathbf{\Omega_2} H\mathbb{E}^{SP}\left[\boldsymbol{\lambda}_\mathbf{2}^\mathbf{2}\mathbb{1}_{\kappa_2>0}\right] + \phi H\mathbb{E}^{SP}\left[\mathbf{\Omega_3}\boldsymbol{\lambda}_\mathbf{2}^\mathbf{2}\mathbb{1}_{\kappa_2>0}\right] \tag{A.11}$$

□

A.3 Proof of Proposition 3

The equilibrium pricing equation at $t = 2$ is given by:

$$q_2 = \beta c_2 \mathbb{E}_2[z_3 + \Omega_3] + \phi(1 - c_2)\mathbb{E}_2[z_3 + \Omega_3] \tag{A.12}$$

keeping in mind that $\Omega_3]$ can depend on q_2. Totally differentiating yields:

$$dq_2 = \beta dc_2 \mathbb{E}_2[z_3 + \Omega_3] + \beta c_2 d\Omega_3 - \phi dc_2 \mathbb{E}_2[z_3 + \Omega_3] + \phi(1 - c_2)d\Omega_3. \tag{A.13}$$

Then use the budget constraint, also totally differentiated, to get:

$$dc_2 = dn_2 + \phi H d\Omega_3. \tag{A.14}$$

since $c_2 = n_2 + \phi H\mathbb{E}[z_3 + \Omega_3]$. Combining these two conditions gives:

$$\begin{aligned} dq_2 = \beta(dn_2 + \phi H d\Omega_3)\mathbb{E}_2[z_3 + \Omega_3] + \beta c_2 d\Omega_3 \\ - \phi(dn_2 + \phi H d\Omega_3)\mathbb{E}_2[z_3 + \Omega_3] + \phi(1 - c_2)d\Omega_3. \end{aligned} \tag{A.15}$$

then notice that by assumption:

$$d\Omega_3 = \frac{d\Omega_3}{dq_2} dq_2. \tag{A.16}$$

Thus rearranging yields:

$$\begin{aligned} dq_2 \left(1 - \beta\phi H\mathbb{E}_2[z_3 + \Omega_3]\frac{d\Omega_3}{dq_2} - \beta c_2 \frac{d\Omega_3}{dq_2} + \phi^2 H\mathbb{E}_2[z_3 + \Omega_3]\frac{d\Omega_3}{dq_2} - \phi(1 - c_2)\frac{d\Omega_3}{dq_2}\right) \\ = \left(\beta\mathbb{E}_2[z_3 + \Omega_3] - \phi\mathbb{E}_2[z_3 + \Omega_3]\right) dn_2 \end{aligned} \tag{A.17}$$

Finally, notice that the factor on dq_2 can be simplified since:

$$\phi H\mathbb{E}_2[z_3 + \Omega_3] + \beta c_2 - \phi^2 H\mathbb{E}_2[z_3 + \Omega_3] + \phi(1 - c_2) = (\beta - \phi)(c_2 + \phi H\mathbb{E}_2[z_3 + \Omega_3]) + \phi. \tag{A.18}$$

and $c_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3] = 2c_2 - n_2$ through the budget constraint. This leads the price sensitivity to be equal to:

$$\frac{dq_2}{dn_2} = \frac{(\beta - \phi)\mathbb{E}_2[z_3 + \Omega_3]}{1 - (\phi + (\beta - \phi)(2c_2 - n_2)\,\frac{d\Omega_3}{dq_2}} \tag{A.19}$$

which is equation (29). □

A.4 Proof of Proposition 4

At time $t = 2$, the welfare of borrowers can be written as:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]\right) + \beta^2 \left(\mathbb{E}_2[z_3]H - \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}[z_3]H + n_2\right) & \text{otherwise} \end{cases} \tag{A.20}$$

while the Lagrangian corresponding to bankers' problem in period $t = 1$ is given by:

$$\mathcal{L}_{b,1} = \left[u(c_1) + \mathbb{E}_1[\mathcal{W}_2(n_2, H; q_2, z_2)]\right] - \lambda_1\left[c_1 + c(H) - d_1 - e_1\right] \tag{A.21}$$

the first-order condition on investment yields:

$$\frac{\partial \mathcal{L}_{b,1}}{\partial H} = -\lambda_1 c'(H) + \beta \mathbb{E}_1\left[\lambda_2(z_2 + \Omega_2)(z_2 + \Omega_2 + q_2^r(z_2 + \Omega_2))\right] \tag{A.22}$$

The social planner maximizes the same function, but under its own expectations, and by also taking into account how a change in d_1 impacts asset prices in period 1 and 2. This leads to the following first-order condition:

$$\frac{\partial \mathcal{L}_{b,1}^{SP}}{\partial H} = \beta \mathbb{E}_1^{SP}\left[\lambda_2(z_2 + q_2)\right] - \lambda_1 c'(H) + \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2}\left(\frac{\partial q_2}{\partial n_2} z_2 + \frac{\partial q_2}{\partial H}\right)\right] \tag{A.23}$$

Proposition 4 is then proved once we notice that:

$$\begin{aligned} \frac{\partial \mathcal{L}_{b,1}^{SP}}{\partial H} \quad = \quad & \underbrace{\beta \mathbb{E}_1\left[\lambda_2(z_2 + q_2)\right] - \lambda_1 q_1}_{\text{Internalized}} \quad + \\ & \underbrace{\beta \mathbb{E}_1^{SP}\left[\lambda_2(z_2 + q_2)\right] - \beta \mathbb{E}_1\left[\lambda_2(z_2 + q_2)\right] + \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2}\left(\frac{\partial q_2}{\partial n_2} z_2 + \frac{\partial q_2}{\partial H}\right)\right]}_{\text{Uninternalized}}. \end{aligned} \tag{A.24}$$

□

A.5 Behavioral Wedge for Investment

I use the same notation as for the proof of Proposition 2, presented in Appendix A.2. The behavioral wedge for investment can consequently be expressed state-by-state as:

$$\mathcal{B}_H(z_2) = [\lambda_2(0;\Omega_3)(z_2 + q_2(0;\Omega_3))] - [\lambda_2(\Omega_2;0)(z_2 + \Omega_2 + q_2(\Omega_2;0)] \tag{A.25}$$

As for leverage, it is sufficient to only look at states where the borrowing constraint binds both in the expectation of the social planner and of private agents. To the first-order, we can write:

$$\mathcal{B}_H(z_2) = (\lambda_2(0;\Omega_3) - \lambda_2(\Omega_2;0)(z_2 + q_2^r)) + \lambda_2^r \left(\Omega_3 \frac{dq_2}{d\Omega_3} - \Omega_2 \left(1 + \frac{dq_2}{d\Omega_2}\right)\right) \tag{A.26}$$

The part $\lambda_2(0;\Omega_3) - \lambda_2(\Omega_2;0)$ exactly corresponds to the behavioral wedge for leverage state-by-state, that we will denote by $\mathcal{B}_d(z_2)$ for conciseness. The behavioral wedge for investment can thus be expressed as:

$$\begin{aligned} \mathcal{B}_H \approx \beta \mathbb{E}_1^{SP}[\mathcal{B}_d(z_2)(z_2 + q_2^r)\mathbb{1}_{\kappa_2>0}] \\ - \beta\Omega_2\mathbb{E}_1^{SP}[\lambda_2(1 + (\beta - \phi)Hz_3)\mathbb{1}_{\kappa_2>0}] + \beta\mathbb{E}_1^{SP}\left[\Omega_3\lambda_2\frac{dq_2}{dz_3}\mathbb{1}_{\kappa_2>0}\right] \end{aligned} \tag{A.27}$$

where

$$\mathcal{B}_d(z_2) = (\Omega_3 - \Omega_2)\lambda_2^2 \tag{A.28}$$

□

A.6 Derivation of Equation (49)

I proceed as for the derivation of the price sensitivity to swings in sentiment, Proposition 3, as in Appendix (A.3). I start from the equilibrium condition that links the asset price at time $t = 2$ to consumption through the collateral constraint:

$$q_2 = \beta\left(n_2 + \phi H\mathbb{E}[z_3 + \Omega_3]\right)\mathbb{E}[z_3 + \Omega_3] + \phi(1 - n_2 - \phi H\mathbb{E}[z_3 + \Omega_3])\mathbb{E}[z_3 + \Omega_3] \tag{A.29}$$

I then differentiate with respect to H, acknowledging that q_2 and Ω_3 will be modified as a result:

$$\begin{aligned} dq_2 = \beta z_2\mathbb{E}[z_3 + \Omega_3] + \beta\phi dH\mathbb{E}[z_3 + \Omega_3]^2 + \beta\phi Hd\Omega_3\mathbb{E}[z_3 + \Omega_3] + \beta c_2 Hd\Omega_3 \\ + \beta\left(n_2 + \phi Hq_2\right)d\Omega_3 + \phi(1 - c_2)d\Omega_3 - \phi z_2\mathbb{E}[z_3 + \Omega_3] \\ - \phi^2 dH\mathbb{E}[z_3 + \Omega_3]^2 - \phi^2 H\mathbb{E}[z_3 + \Omega_3]d\Omega_3 \end{aligned} \tag{A.30}$$

Rearranging gives the desired result:

$$\frac{dq_2}{dH} = \frac{(\beta - \phi))z_2 + \phi \mathbb{E}_2[z_3 + \Omega_3])\mathbb{E}_2[z_3 + \Omega_3]}{1 - (\phi + (\beta - \phi)(c_2 - \phi H \mathbb{E}_2[z_3 + \Omega_3])) \frac{d\Omega_3}{dq_2}} \tag{A.31}$$

□

A.7 Proof of Proposition 6

The only variable that can be changed, at $t = 2$, by a change in q_1, is Ω_3 (remember that we are keeping everything else fixed at $t = 1$). Hence the welfare change is given by:

$$\frac{d\mathcal{W}_1}{q_1} = \beta \mathbb{E}_1^{SP}[\lambda_2 \phi H \frac{d\Omega_3}{dq_1} - \beta \phi H \frac{d\Omega_3}{dq_1}(1 + r_2)] \tag{A.32}$$

where once again the first part in the expectation corresponds to the change in consumption at $t = 2$ induced by the shift in the collateral limit, and the second part corresponds to the decrease in consumption at $t = 3$ since the amount that needs to be repaid is higher. That leads, using $\kappa_2 = \lambda_2 - 1$ and $\beta(1 + r_2) = 1$, to the reversal externality formulation:

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

□

A.8 Proof of Proposition 7

The proof of Proposition 7 is straightforward once the uninternalized effects of leverage and investment have been derived. By assumption, the planner can impose taxes or subsidies on leverage, on the creation of collaterals assets, and on the holdings of collateral assets, which are rebated or funded lump-sum. Denote these taxes/subsidies respectively by τ_d, τ_H and τ_q. The budget constraint can be written:

$$c_1 + c(H) + \tau_h H + q_1 h \leq e_1 + d_1(1 - \tau_d) + q_1 H + \tau_q h \tag{A.34}$$

where H is the amount invested and h is the amount kept on the balance sheet. Of course in equilibrium $h = H$.

The first-order conditions of private agents are given by:

$$\frac{\partial \mathcal{L}_{b,0}}{\partial d_1} = \lambda_1(1 - \tau_d) - \mathbb{E}_1[\lambda_2] = 0 \tag{A.35}$$

$$\frac{\partial \mathcal{L}_{b,0}}{\partial H} = c'(H) + \tau_H - q_1 = 0 \tag{A.36}$$

$$\frac{\partial \mathcal{L}_{b,0}}{\partial h} = \lambda_1 q_1 + \lambda_1 \tau_q - \mathbb{E}_1[\lambda_2(z_2 + \Omega_2 + q_2^r)] = 0 \tag{A.37}$$

The planner wants the agent to internalize the effects of leverage. This is simply done with a tax equal to:

$$\tau_d = -\frac{\mathcal{W}_d}{\lambda_1} \tag{A.38}$$

For investment, the planer wants to fix the level of investment at a level H such that:

$$c'(H) = \beta \mathbb{E}_1^{SP}\left[\lambda_2(z_2 + q_2)\right] + \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2}\left(\frac{\partial q_2}{\partial n_2} z_2 + \frac{\partial q_2}{\partial H}\right)\right] \tag{A.39}$$

and because

$$\beta \mathbb{E}_1^{SP}\left[\lambda_2(z_2 + q_2)\right] + \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2}\left(\frac{\partial q_2}{\partial n_2} z_2 + \frac{\partial q_2}{\partial H}\right)\right] = \beta \mathbb{E}_1[\lambda_2(z_2 + \Omega_2 + q_2^r)] - \mathcal{W}_H \tag{A.40}$$

the tax must simply be set equal to:

$$\tau_H = -\frac{\mathcal{W}_H}{\lambda_1} \tag{A.41}$$

Finally, denote by q_1^* the price at $t = 1$ such that the reversal externality is equal to 0. This is the price the planner wants to set. We thus simply want:

$$\lambda_1 q_1^* + \lambda_1 \tau_q - \mathbb{E}_1\left[\lambda_2(z_2 + \Omega_2 + q_2^r)\right] = 0 \tag{A.42}$$

so the tax should be set at:

$$\tau_q = \frac{\mathbb{E}_1\left[\lambda_2(z_2 + \Omega_2 + q_2^r)\right] - \lambda_1 q_1^*}{\lambda_1} \tag{A.43}$$

□

A.9 Proof of Proposition 8

I keep using the notation from the previous proof. Agents' private Euler equation when a tax is imposed on leverage is:

$$\lambda_1(1 - \tau_d) = \mathbb{E}_1[\lambda_2] \tag{A.44}$$

Since, in a crisis, $\lambda_2(d_1, z_2 + \Omega_2, H)$ is unambiguously decreasing in Ω_2, and because λ_1 is decreasing in d_1, leverage is increasing with Ω_2.

As long as $\Omega_2 > 0$, and there is a positive probability of a crisis, we have $\mathcal{W}_d < 0$. It directly implies, from equation (A.5), that this decreases welfare as evaluated from the planner.

However if the policy is put in place through a leverage limit, the allocation satisfies:

$$\lambda_1 = \max\left(\lambda_1^*, \mathbb{E}_1[\lambda_2]\right) \tag{A.45}$$

Since we assumed that $\mathcal{W}_d < 0$, this necessarily implies that $\lambda_1^* > \mathbb{E}_1[\lambda_2]$. In turn, this means that for a perturbation $d\omega < 0$ to initial exuberance:

$$\lambda_1^* > \mathbb{E}_1[\lambda_2(z_2 + \Omega_2)] > \mathbb{E}_1[\lambda_2(z_2 + \Omega_2 + d\omega)] \tag{A.46}$$

so leverage stays at the optimal level desired by the planner. Finally, regarding a downward movement to Ω_2, the assumption that $\mathcal{W}_d < 0$ implies that there is a non-zero gap between λ_1^* and $\mathbb{E}_1[\lambda_2]$, such that for a small enough $d\omega > 0$, it is also guaranteed that:

$$\lambda_1^* > \mathbb{E}_1[\lambda_2(z_2 + \Omega_2 - d\omega)] > \mathbb{E}_1[\lambda_2(z_2 + \Omega_2)] \tag{A.47}$$

hence guaranteed that allocations stay at the second-best. □

A.10 Proof of Proposition 9

This proposition is straightforward. Using Assumption 4, the function g_q is bijective. It allows the social planner to invert the price observation. Since $q_1 = g_q(\bar{z}_2 + \Omega_2)$ and $\bar{z}_2$ is known, the extent of sentiment at time $t = 1$ is exactly identified by:

$$\Omega_2 = g_q^{-1}(q_1) - \bar{z}_2 \tag{A.48}$$

□

A.11 Proof of Proposition 10

As explained in the main text, the social planner's optimality condition under the premises of Proposition 10 can be expressed as:

$$u'(c_1) = \frac{1}{2\sigma_\Omega} \int_0^\infty \left[\int_{-\sigma_\Omega}^{\sigma_\Omega} \frac{\partial \mathcal{W}_2}{\partial n_2} (d_1, H; z_2 - \bar{\Omega}_2 - \omega_2)\, d\omega_2 \right] f_2(z_2) dz_2. \tag{A.49}$$

Key to this proposition is the shape of $\partial \mathcal{W}_2 / \partial n_2$ with respect to z_2. First recall that:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(n_2 + \phi H \mathbb{E}_2[z_3]\right) + \beta^2 \left(\mathbb{E}^{SP}[z_3] H - \phi H \mathbb{E}_2[z_3]/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}^{SP}[z_3] H + n_2\right) & \text{otherwise} \end{cases} \tag{A.50}$$

so that the first derivative is equal to:

$$\frac{\partial \mathcal{W}_2}{\partial n_2} = \begin{cases} \beta \lambda_2 & \text{if } z_2 \geq z^* \\ \beta & \text{otherwise} \end{cases} \tag{A.51}$$

which is constant outside of a crisis, as expected. I use the following notation to simplify the exposition of the proof. First, the expectation over z_2 for a given w_2 is denoted by:

$$g(w_2) = \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \omega_2) f_2(z_2) dz_2 \tag{A.52}$$

while the integral taken over the uncertainty band is:

$$G(\sigma_\Omega) = \int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{g(w_2)}{2\sigma_\Omega} dw_2. \tag{A.53}$$

Given the continuity of $\partial \mathcal{W}_2 / \partial n_2$ (see equation A.50) we can differentiate with respect to σ_Ω:

$$\begin{aligned} G'(\sigma_\Omega) = &-\frac{1}{2\sigma_\Omega^2} \int_{-\sigma_\Omega}^{+\sigma_\Omega} \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; z_2 - \bar{\Omega}_2 - \omega_2) f_2(z_2) dz_2 dw_2 + \\ &\int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; z_2 - \bar{\Omega}_2 - \sigma_\Omega) f_2(z_2) dz_2 - \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; z_2 - \bar{\Omega}_2 + \sigma_\Omega) f_2(z_2) dz_2 \end{aligned} \tag{A.54}$$

which can be expressed in terms of the notation just defined above as:

$$G'(\sigma_\Omega) = -\frac{G(\sigma_\Omega)}{\sigma_\Omega} + \frac{1}{2\sigma_\Omega}(g(\sigma_\Omega) - g(-\sigma_\Omega)) \tag{A.55}$$

Before proceeding further, remember that the social planner optimally sets leverage such that:

$$u'(c_1) = G(\sigma_\Omega) \tag{A.56}$$

while the decentralized equilibrium is independent of σ_Ω. Thus, leverage restrictions are increasing in σ_Ω if and only if G is increasing in σ_Ω. This condition is then equivalent, using the derivative just computed, to:

$$\frac{g(\sigma_\Omega) - g(-\sigma_\Omega)}{2} > \int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{g(w_2)}{2\sigma_\Omega} dw_2. \tag{A.57}$$

Since $\partial \mathcal{W}_2 / \partial n_2$ is continuous in z and in ω_2, and since ω_2 is defined in the compact set $[-\sigma_\Omega, \sigma_\Omega]$, g is continuous (by continuity of parametric integrals) and Fubini's theorem implies that a sufficient condition for $G'(\sigma_\Omega) > 0$ is that:[83]

$$\frac{1}{2}\left(\frac{\partial \mathcal{W}_2}{\partial n_2}(z_2 + \sigma_\Omega) - \frac{\partial \mathcal{W}_2}{\partial n_2}(z_2 - \sigma_\Omega)\right) > \int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{\partial \mathcal{W}_2}{\partial n_2}(z_2 + \omega_2) \frac{d\omega_2}{2\sigma_\Omega} \quad \forall z_2 \in supp(f_2). \tag{A.58}$$

In other words, this condition requires that the average taken over a segment is below the average of the two extreme points of this same segment.

Next, notice that any convex function satisfies this requirement. For a convex function φ, Jensen's

[83] $\bar{\Omega}_2$ does not need to appear in this condition since this inequality is required to hold for all z_2 in the support of the definition, so equivalently for all $z_2 - \bar{\Omega}_2$ also in the support.

inequality yields:

$$\varphi(t\sigma_\Omega - (1-t)\sigma_\Omega) \leq t\varphi(\sigma_\Omega) + (1-t)\varphi(-\sigma_\Omega) \quad \forall t \in [0,1]. \tag{A.59}$$

Now integrate this inequality over t to get:

$$\int_0^1 \varphi(t\sigma_\Omega - (1-t)\sigma_\Omega)dt \leq \int_0^1 t\varphi(\sigma_\Omega)dt + \int_0^1 (1-t)\varphi(-\sigma_\Omega)dt. \tag{A.60}$$

A change of variable $t \to (x-\sigma_\Omega)/(2\sigma_\Omega)$ in the left-hand side thus yields:

$$\int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{\varphi(x)}{2\sigma_\Omega}dx \leq \frac{\varphi(\sigma_\Omega) - \varphi(-\sigma_\Omega)}{2} \tag{A.61}$$

which is exactly the relationship in equation (A.58).

We now have to prove that $\partial\mathcal{W}_2/\partial n_2$ is convex to end the proof of Proposition 10. Going back to equation (A.50), denote $\partial\mathcal{W}_2/\partial n_2$ by $\mathcal{W}_{2,n}$. Start with the derivative of marginal utility. We have:

$$\frac{d\lambda_2}{dz_2} = -\frac{H}{c_2^2} \tag{A.62}$$

and so:

$$\frac{d^2\lambda_2}{dz_2^2} = \frac{2}{c_2^3}H > 0 \tag{A.63}$$

Which concludes the proof.[84] □

Before moving to the next proof, notice that the convexity of marginal welfare was quite easy to prove. This is not the case anymore when the collateral constraint is of the form $\phi H q_2$. Indeed, when this is the case the marginal welfare function also features the price sensitivity, and its convexity is more involved to prove. It is possible to show that Proposition 10 still holds. See Proposition 19. The proof can be found in Online Appendix Q.6. Second, the convexity is also harder to prove when prices at $t=2$ impact sentiment Ω_3. See Online Appendix H.2 where it is shown that price extrapolation *amplifies* this convexity.

A.12 KL-Divergence

The Kullback–Leibler divergence between two distributions p and q is defined by the relative entropy:

$$KL(p,q) = \int_{-\infty}^{+\infty} p(x) \ln\left(\frac{p(x)}{q(x)}\right) dx. \tag{A.64}$$

[84] For the sake of brevity, Ω_3 is left out of the expressions as, by assumption, it is a constant. It thus only shifts the value of $\mathbb{E}_1[z_3]$ and that has no impact on the sign of these derivatives as long as $\mathbb{E}_1[z_3] + \Omega_3 > 0$, which we always assume to be the case.

Here we are interested in the KL-divergence between a Gaussian and a uniform random variables. I thus define:

$$q \sim \mathcal{U}\left[\Omega_2' - \omega; \Omega_2' + \omega\right] \tag{A.65}$$

$$p \sim \mathcal{N}\left(\Omega_2, \sigma^2\right) \tag{A.66}$$

and the objective is to find the Ω_2' and ω that minimize the KL-divergence (A.64).[85] Consequently, the objective is:

$$\min_{\Omega_2',\omega} \quad \int_{\Omega_2'-\omega}^{\Omega_2'+\omega} \frac{1}{2\omega} \ln\left(\frac{\sigma\sqrt{2\pi}}{2\omega} e^{\frac{(x-\Omega_2)^2}{2\sigma^2}}\right) dx \tag{A.67}$$

which conveniently leads to a simpler expression:

$$\min_{\Omega_2',\omega} \quad \int_{\Omega_2'-\omega}^{\Omega_2'+\omega} \frac{1}{2\omega} \left[\ln\left(\frac{\sigma\sqrt{2\pi}}{2\omega}\right) + \frac{(x-\Omega_2)^2}{2\sigma^2}\right] dx. \tag{A.68}$$

Integrating the two parts gives:

$$\min_{\Omega_2',\omega} \quad \left[\ln\left(\frac{\sigma\sqrt{2\pi}}{2\omega}\right) + \frac{1}{4\omega\sigma^2} \frac{(\Omega_2' - \Omega_2 + \omega)^3 - (\Omega_2' - \Omega_2 - \omega)^3}{3}\right]. \tag{A.69}$$

We can now easily minimize this expression by taking the first derivatives with respect to the average and spread of the targeted uniform distribution. Regarding the average, the uniform distribution is obviously centered on the same mean:

$$(\Omega_2' - \Omega_2 + w)^2 - (\Omega_2' - \Omega_2 - w)^2 = 0 \implies \Omega_2' = \Omega_2 \tag{A.70}$$

which in turn leads the minimization with respect to the spread of the uniform distribution to yield:

$$-\frac{1}{\omega} + \frac{2\omega}{3\sigma^2} = 0 \tag{A.71}$$

To conclude, the uniform distribution that minimizes the KL-divergence with a Gaussian distribution of parameters Ω_2 and σ^2 is:

$$q \quad \sim \quad \mathcal{U}\left[\Omega_2 - \sqrt{\frac{3}{2}}\sigma \quad ; \quad \Omega_2 + \sqrt{\frac{3}{2}}\sigma\right] \tag{A.72}$$

□

[85] Notice that because of the use of a uniform random variable, we can only compute $KL(p,q)$ and not $KL(p,q)$, since the Radon-N derivative of the

A.13 Proof of Proposition 11

This Proposition is straightforward once realizing, using equations (63) and (64), that $\bar{\Omega}_2(q_1)$ is increasing in q_1 since g_q^{-1} is an increasing function of its argument. This allows us to directly apply Proposition 10 with a band of width $\sqrt{3/2}\Sigma_\Omega$, given Assumption 6, and its associated computation in Appendix A.12. □

A.14 Proof of Proposition 12

The welfare function that the planner considers is given by:

$$\mathcal{W}_1 = \Phi^h \mathbb{E}_1^{SP}\Big(\ln\left[c_1^h - \nu \frac{l_1^{1+\eta}}{1+\eta}\right] + \beta c_2^h + \beta^2 c_3^h\Big) + \Phi^b \mathbb{E}_1^{SP}\big(\ln(c_1) + \beta \ln(c_2) + \beta^2 c_3\big) \tag{A.73}$$

where Φ^h and Φ^b are the Pareto weights attached to each group by the planner. In equilibrium, we have $Y_1 = l_1$ by assumption of linear production. We thus write utility of households at $t = 1$ as:

$$\mathcal{W}_1^h = \ln\left[c_1^h - \nu \frac{Y_1^{1+\eta}}{1+\eta}\right] + \beta c_2^h + \beta^2 c_3^h. \tag{A.74}$$

Households' welfare is affected by two effects: first, a change n r_1 changes the incentives for savings, forcing agents to substitute wealth across periods. Second, it changes output and thus consumption and labor supply levels. However, since households are on their Euler equation at $t = 1$, the first effect is exactly 0:

$$\frac{d\mathcal{W}_1^h}{dr_1} = \frac{Y_1}{dr_1}\lambda_1^h - \nu Y_1^\eta \frac{Y_1}{dr_1}\lambda_1^h + \underbrace{\frac{dc_1^h}{dr_1}\lambda_1^h - \beta \mathbb{E}_1 \frac{dc_1^h}{dr_1}}_{\text{Euler}=0}. \tag{A.75}$$

Next, the change in the interest rate have an impact on the borrowing of financial intermediaries. This is not zero as for households, because of the uninternalized effects explored in Section 3.2. It also has an impact on investment, which for the same reason is not zero in general. Finally, it has an impact on prices, which can spill over on sentiment. Because Pareto weights are chosen such as $\Phi^j = 1/\lambda_1^j$, we simply end up with:

$$\begin{aligned} \frac{d\mathcal{W}_1}{dr_1} \quad = \quad & \frac{dY_1}{dr_1}\mu_1 \quad + \frac{dd_1}{dr_1}\mathcal{W}_d \quad + \quad \frac{dH}{dr_1}\mathcal{W}_H \\ & + \quad \frac{d\Omega_2}{dq_1}\frac{dq_1}{dr_1}\left(\frac{dd_1}{d\Omega_2}\mathcal{W}_d + \frac{dH}{d\Omega_2}\mathcal{W}_H\right) \quad + \quad \mathbb{E}_1\left[\kappa_2 \phi H \frac{d\Omega_3}{dq_1}\frac{dq_1}{dr_1}\right] \end{aligned} \tag{A.76}$$

□

A.15 Proof of Proposition 13

This Proposition simply follows from the fact that optimal taxes are set such that $\mathcal{W}_d = \mathcal{W}_H = 0$, as per Proposition 7, so that:

$$\begin{aligned} \frac{d\mathcal{W}_1}{dr_1} &= \frac{dY_1}{dr_1}\mu_1 \quad + \frac{dd_1}{dr_1}\cdot 0 \quad + \quad \frac{dH}{dr_1}\cdot 0 \\ &\quad + \quad \frac{d\Omega_2}{dq_1}\frac{dq_1}{dr_1}\left(\frac{dd_1}{d\Omega_2}\cdot 0 + \frac{dH}{d\Omega_2}\cdot 0\right) \quad + \quad \mathbb{E}_1\left[\kappa_2\phi H\frac{d\Omega_3}{dq_1}\frac{dq_1}{dr_1}\right] \end{aligned} \tag{A.77}$$

□

A.16 Proof of Proposition 14

By assumption, $\mu_1 = 0$ where we take the first-order perturbation (perfectly closing the output gap). Since there is an optimal investment tax/subsidy, and that we take the benchmark case where $dd_1/dr_1 = 0$ (see the discussion on the core of the paper and Farhi and Werning 2020), Proposition 12 implies that:

$$d\mathcal{W}_1 = \frac{dd_1}{d\Omega_2}\mathcal{W}_d d\Omega_2 + \mathbb{E}_1\left[d\Omega_3\kappa_2\phi H\right] \tag{A.78}$$

We then need to plug in the first-order effects on sentiment. Substituting equations (69), (70), (72) and (73), we have the following welfare effects, respectively for late and early tightening:

$$d\mathcal{W}_1^l = \frac{dd_1}{d\Omega_2}\mathcal{W}_d\left(\alpha_0\iota dr_1\right) + \mathbb{E}_1\left[\left(\gamma_0\alpha_0 + \alpha_1\right)\iota dr_1\kappa_2\phi H\right] \tag{A.79}$$

$$d\mathcal{W}_1^e = \frac{dd_1}{d\Omega_2}\mathcal{W}_d\left(\gamma_0\alpha_0 + \alpha_1\right)\iota dr_0 + \mathbb{E}_1\left[\left(\gamma_0(\gamma_0\alpha_0 + \alpha_1) + \gamma_1\alpha_0 + \alpha_1\right)\iota dr_0\kappa_2\phi H\right]. \tag{A.80}$$

Since we are comparing the marginal benefits of the same tightening, $dr_1 = dr_0$. Rearranging and comparing the conditions for $d\mathcal{W}_1^l > d\mathcal{W}_1^e$ yields Proposition 14.

□

A.17 Proof of Proposition 15

Start with what households expect will occur at $t = 1$. Because utility is linear at $t = 2$ and there is no risk, household optimization implies:

$$\frac{1}{c_1^h} = \beta(1 + r_1(q_0)) \tag{A.81}$$

where the dependence of interest rates to asset prices is made explicit for clarity. Combining this expression with household optimization at $t = 0$ directly yields the following relation between contemporaneous interest rates and consumption, and future interest rates:

$$\frac{1}{c_0\beta(1+r_1(q_0))} = \beta(1+r_0) \tag{A.82}$$

which implies, under perfect inflation targeting (and so no output gap):

$$c_0 = 1 = \frac{1}{\beta^2(1+r_0)(1+r_1(q_0))} \tag{A.83}$$

This equilibrium relation makes clear that in the event of a higher price at $t = 0$, interest rates are expected to be higher and so the optimal r_0 to close the output gap decreases.

We can now proceed with the first-order perturbation to obtain the formulation in Proposition 15. The perturbation is made around the REE equilibrium, where the interest rate that closes the output gap at $t = 1$ is denoted by r_1^*, while the REE prices are denoted by q_0^* and q_1^*. The price of the asset initially is given by the pricing equation:

$$q_0 = \beta\left[\frac{c_0}{c_1}(q_1^r + \alpha(q_0 - q_{-1}))\right] \tag{A.84}$$

and using the Euler equation, this boils down to:

$$q_0 = \frac{1}{1+r_0}(q_1^e + \alpha(q_0 - q_{-1})) \tag{A.85}$$

But the price expected at $t = 1$ depends on the interest rate the bank will choose at $t = 1$. The pricing equation at $t = 1$ is given by:

$$q_1 = \frac{z_2 + q_2^r}{1+r_1} = \frac{z_2 + q_2^r}{1 + r_1^* + \rho(q_0 - \bar{q})^+} \tag{A.86}$$

which can be approximated as:

$$q_1 \approx q_1^*\left(1 - \rho\frac{(q_0 - \bar{q})^+}{1+r_1^*}\right) \tag{A.87}$$

Assume that prices at 0 are elevated (and simply check later once the equilibrium is solved that this is indeed the case). Plugging this back to the pricing equation at $t = 0$ leads to:

$$q_0 = \frac{1}{1+r_0}\left(q_1^*\left(1 - \rho\frac{(q_0 - \bar{q})^+}{1+r_1^*}\right) + \alpha(q_0 - q_{-1}\right) \tag{A.88}$$

which can be solved as:

$$q_0 = \frac{q_1^*\left(1 + \rho\frac{\bar{q}}{1+r_1^*}\right) - \alpha q_{-1}}{1 + r_0 + \rho q_0^* - \alpha} \tag{A.89}$$

This expression makes clear that the discount rate channel operates here: a fall in the interest rate r_0

boosts asset prices. Naturally, a fall in q_{-1} (the anchor) also boosts prices, with a coefficient of sensitivity α. Finally, all movements are amplified by belief amplification (the $-\alpha$ in the denominator). Similarly they are *dampened* by ρ: a bigger ρ creates a smaller price movement q_1.

Assume further that the central bank tightens when prices are higher than in the rational benchmark, so $\bar{q} = q_0^*$. The price deviation which then feeds into the interest rate rule becomes:

$$q_0 - q_0^* = \frac{q_1^*\left(1+\rho\frac{\bar{q}}{1+r_1^*}\right) - \alpha q_{-1}}{1 + r_0 + \rho q_0^* - \alpha} - q_0^* \tag{A.90}$$

$$\implies q_0 - q_0^* = \frac{q_1^*\left(1+\rho\frac{\bar{q}}{1+r_1^*}\right) - \alpha q_{-1} - q_0^* - r_0 q_0^* - \rho q_0^{2*} + \alpha q_0^*}{1 + r_0 + \rho q_0^* - \alpha} \tag{A.91}$$

$$\implies q_0 - q_0^* = \frac{q_1^* - (1+r_0)q_0^* + \alpha(q_0^* - q_{-1})}{1 + r_0 + \rho q_0^* - \alpha} \tag{A.92}$$

Going back to inflation targeting, we can write the condition for closing the output gap as:

$$1 + r_0 = \frac{1}{\beta^2(1 + r_1(q_0))} \tag{A.93}$$

Denote for simplicity the deviation from the REE as ϵ for the current interest rate. Algebra yields:

$$1 + r_0^* + \epsilon = \frac{1}{\beta^2}\frac{1}{1 + r_1^* + \rho(q_0 - q_0^*)} \tag{A.94}$$

$$\implies 1 + r_0^* + \epsilon = \frac{1}{\beta^2}\left(\frac{1}{1+r_1^*} - \rho\frac{(q_0 - q_0^*)}{1+r_1^*}\right) \tag{A.95}$$

$$\implies 1 + r_0^* + \epsilon = 1 + r_0^* - \frac{1}{\beta^2}\rho\frac{(q_0 - q_0^*)}{1+r_1^*} \tag{A.96}$$

$$\implies \epsilon = -\frac{1}{\beta^2}\rho\frac{(q_0 - q_0^*)}{1+r_1^*} \tag{A.97}$$

$$\implies \epsilon = -\frac{1}{\beta^2}\rho\frac{\frac{q_1^* - (1+r_0^*+\epsilon)q_0^* + \alpha(q_0^* - q_{-1})}{1+r_0^*+\epsilon+\rho q_0^* - \alpha}}{1+r_1^*} \tag{A.98}$$

$$\implies \epsilon = -(1+r_0^*)\rho\frac{q_1^* - (1+r_0^*+\epsilon)q_0^* + \alpha(q_0^* - q_{-1})}{1 + r_0^* + \epsilon + \rho q_0^* - \alpha} \tag{A.99}$$

$$\implies \epsilon \approx -(1+r_0^*)\rho\left[\frac{\alpha(q_0^* - q_{-1}) - \epsilon q_0^*}{1 + r_0^* + \rho q_0^*}\left(1 + \frac{\alpha - \epsilon}{1 + r_0^* + \rho q_0^*}\right)\right] \tag{A.100}$$

$$\implies \epsilon\left[1 + r_0^* + \rho q_0^* + (1+r_0^*)\rho q_0^* + \frac{\alpha(q_0^* - q_{-1})}{1 + r_0^* + \rho q_0^*}\right] \approx -(1+r_0^*)\rho\left[\alpha(q_0^* - q_{-1})\right] \tag{A.101}$$

$$\implies \epsilon\left[1 + r_0^* + \rho q_0^* + (1+r_0^*)\rho q_0^*\right] \approx -(1+r_0^*)\rho\left[\alpha(q_0^* - q_{-1})\right] \tag{A.102}$$

$$\implies \epsilon \approx -\frac{(1+r_0^*)\rho\left[\alpha(q_0^* - q_{-1})\right]}{1 + r_0^* + \rho q_0^* + (1+r_0^*)\rho q_0^*} \tag{A.103}$$

$$ \tag{A.104} $$

Which concludes the proof simply by noting that:

$$1 + r_0 = 1 + r_0^* + \epsilon = 1 + r_0^* - \frac{(1 + r_0^*)\rho\alpha(q_0^* - q_{-1})}{1 + r_0^* + \rho q_0^* + (1 + r_0^*)\rho q_0^*} \tag{A.105}$$

□

A.18 Proof of Proposition 22

At time $t = 2$, the welfare of financial intermediaries can now be written as:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(n_2 + b^* + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)]\right) + \beta^2 \left(\mathbb{E}[z_3]H - \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)]/\beta - b^*/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}[z_3]H + n_2\right) & \text{otherwise} \end{cases} \tag{A.106}$$

with the level of bailouts determined optimally in equilibrium. The private first-order condition on borrowing is unchanged since agents to not internalize their impact on b^* (atomistic agents):

$$\frac{\partial \mathcal{L}_{b,1}}{\partial d_1} = \lambda_1 - \mathbb{E}_1\left[\lambda_2(b^*)\right] \tag{A.107}$$

The social planner maximizes the same function, but under its own expectations, and by also taking into account how a change in d_1 impacts asset prices in period 2 *and* the level of bailouts. This leads to the following first-order condition:

$$\frac{\partial \mathcal{L}_{b,1}^{SP}}{\partial d_1} = \lambda_1 - \mathbb{E}_1^{SP}\left[\lambda_2(b^*)\right] - \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2}\frac{\partial q_2}{\partial n_2}\right] + \frac{db^*}{dn_2}\lambda_2 - \frac{db^*}{dn_2}g'(b^*) - \frac{db^*}{dn_2} \tag{A.108}$$

And the last part is equal to zero since bailouts are chosen optimally:

$$g'(b^*) = \lambda_2 - 1 \tag{A.109}$$

which proves Proposition 22. □

A.19 Proof of Proposition 23

The behavioral wedge is given by:

$$\mathcal{B}_d = \mathbb{E}_1\left[\lambda_2(b^*)\right] - \mathbb{E}_1^{SP}\left[\lambda_2(b^*)\right] \tag{A.110}$$

We can simply compare the two marginal utilities state-by-state. Agents believe that:

$$\lambda_2(b^*) = ((z_2 + \Omega_2)H + b^*(z_2 + \Omega_2, 0) - d_1(1 + r_1) + \phi H\mathbb{E}[z_3])^{-1} \tag{A.111}$$

While the planner believes that:

$$\lambda_2^{SP}(b^*) = (z_2 H + b^*(z_2, \Omega_3) - d_1(1 + r_1) + \phi H\mathbb{E}[z_3 + \Omega_3])^{-1} \tag{A.112}$$

Since bailouts are proportional to the severity of the crisis. Using equation (D.24) yields:

$$\frac{d\mathcal{B}_d}{d\xi} = -\mathbb{E}_1\left[\frac{db^*}{d\xi}\lambda_2^2(b^*)\right] + \mathbb{E}_1^{SP}\left[\frac{db^*}{d\xi}\lambda_2^2(b^*)\right] \tag{A.113}$$

$$\implies \frac{d\mathcal{B}_d}{d\xi} = -\mathbb{E}_1\left[(\lambda_2(b^*) - 1)\lambda_2^2(b^*)\right] + \mathbb{E}_1^{SP}\left[(\lambda_2(b^*) - 1)\lambda_2^2(b^*)\right] \tag{A.114}$$

Which is positive since the λ_2 are always greater or equal to 1.

□

B Real Production

B.1 A Simple Extension with Production

To incorporate a real side to the model, I allow households to supply labor at $t = 2$. Households have linear utility over consumption, and have a convex disutility for supplying labor in the intermediate period:

$$U^h = \mathbb{E}_1\left[c_1^h + \beta\left(c_2^h - \nu\frac{l_2^{1+\eta}}{(1+\eta)}\right) + \beta^2 c_3^h\right] \tag{B.1}$$

where l_2 is the amount of labor supplied by households at time $t = 2$.

There is a fringe of competitive firms of measure one, producing from the labor of households. Firms use a decreasing returns to scale technology from labor, with productivity A:

$$Y_2 = A l_2^{\alpha} \tag{B.2}$$

To bridge the gap between Main street and Wall street, I add a financial friction. Firms need to pay a fraction γ of wage bills in advance to workers, which requires them to borrow from financial intermediaries. In period 2, firms need to borrow $f_2 = \gamma w_2 l_2$ from financial intermediaries. We assume that the interest rate required by financial intermediaries to advance such funds depends on the size of the loan according to:

$$1 + r_f = \frac{\delta}{f_2} \tag{B.3}$$

This innocuous trick allows the model to say away from corner solutions.[86] The set of budget constraint is now given for households by:

$$c_1^h + d_1 \le e_1^h \tag{B.4}$$
$$c_2^h + d_1 \le e_2^h + w_2 l_2 + d_1(1+r_1) + \pi_2 \tag{B.5}$$
$$c_3^h \le e_3^h + d_2(1+r_2) \tag{B.6}$$

and financial intermediaries.

$$c_1 + c(H) \le d_1 + e_1 \tag{B.7}$$
$$c_2 + d_1(1+r_1) + f_2 + q_2 h \le d_2 + (z_2 + q_2)H \tag{B.8}$$
$$c_3 + d_2(1+r_2) \le z_3 h + f_2(1+r_f) \tag{B.9}$$

Household optimization then simply yields:

$$w_2 = \nu l_2^\eta \tag{B.10}$$

It is also assumed for simplicity that loans made to firms cannot be used as collateral.[87] The specific form assumed in (B.3) simplifies matter since funds allocated to firms verify the following identity:

$$\frac{f_2}{\delta} = \beta c_2 \tag{B.11}$$

so that bankers' consumption and funds allowed to firms are proportional. Intuitively, when collateral constraints are extremely tight, this forces financial intermediaries to cut back on consumption *and* their traditional intermediary activities in the same way.[88] Thus the amount of labor used for production verifies:

$$l_2 = \left(\frac{z_2 H - d_1(1+r_1) + \phi H \mathbb{E}_2[z_3 + \Omega_3]}{\gamma \nu \left(1 + \frac{1}{\beta\delta}\right)} \right)^{\frac{1}{1+\eta}} \tag{B.12}$$

[86] This also allows for belief application to survive. Remember that belief amplification comes from the two-way feedback effect between the stochastic discount factor and the price fo the risky asset. A corner solution with respect to the borrowing of real firms would break this link.

[87] A more complete formulation of the collateral constraint would be:

$$d_2 \le \phi H \mathbb{E}_2[z_3] + \psi f_2$$

whereby assuming that a fraction of the amount lent to firms can be recovered by depositors in the (non-equilibrium) possibility of default. I am here analyzing the limiting case where $\psi \to 0$. The general case complexifies matters without bringing any new intuition. Analytical derivations of the general case are thus relegated to Appendix B.3.

[88] Consumption is needed for the SDF to generate amplification: a risk-neutral valuation pricing kernel breaks the feedback loop between the price of the asset and marginal utility. But one could think of c_2 as dividends or compensation.

which translates into a production level at time $t = 2$ of:

$$Y_2 = A\left(\frac{z_2 H - d_1(1+r_1) + \phi H \mathbb{E}_2[z_3 + \Omega_3]}{\gamma\nu\left(1 + \frac{1}{\beta\delta}\right)}\right)^{\frac{\alpha}{1+\eta}} \tag{B.13}$$

A drop in expectations directly impacts output, as well as a fall in financial intermediaries' net worth $z_2 H - d_1(1+r_1)$. Hence, looking at $\mathbb{E}_2[z_3 + \Omega_3]$ inside a crisis is a sufficient statistics even in this extended model with real production. A liquidity drought spills over the real sector and propagates to employment and output, consistent with empirical evidence (see, e.g. Dell'Ariccia, Detragiache and Rajan 2008, Cingano, Manaresi and Sette 2016 or Bentolila et al. 2018).

B.2 Welfare Analysis with Real Production

The planner maximizes the following object:

$$\mathcal{W}_1 = \Phi^h \mathbb{E}_1^{SP}\left(c_1^h + \beta\left[c_2^h - \nu\frac{l_2^{1+\eta}}{1+\eta}\right] + \beta^2 c_3^h\right) + \Phi^b \mathbb{E}_1^{SP}\left(\ln(c_1) + \beta\ln(c_2) + \beta^2 c_3\right) \tag{B.14}$$

where Φ^h and Φ^b are the Pareto weights attached to each group by the planner. I denote by V_2^h and V_2^h the value functions of each group at time $t = 2$.

Leverage: We are interested in the derivatives of these value functions at time $t = 2$ with respect to the amount of short-term debt (or savings) chosen at time $t = 1$. Because funds allocated to firms (f_2) are chosen optimally without a constraint (see equation B.11), an infinitesimal change in f_2 does not have a first-order impact on the welfare of bankers:

$$\frac{dV_2^b}{dd_1} = \phi H(\lambda_2 - 1)\frac{d\Omega_3}{dq_2}\frac{dq_2}{dd_1} + \underbrace{\beta\frac{\delta}{f_2} - \lambda_2}_{=0}. \tag{B.15}$$

For households, however, there is a new term coming from the expansion of bank lending to firms in the real sector:

$$\frac{dV_2^h}{dd_1} = \phi H\underbrace{(\lambda_3^h - \lambda_2^h)}_{=0}\frac{d\Omega_3}{dq_2}\frac{dq_2}{dd_1} + \max\left(\underbrace{A\alpha\left(\frac{z_2 H - d_1(1+r_1) + \phi H \mathbb{E}[z_3 + \Omega_3]}{\gamma\nu\left(1 + \frac{1}{\delta}\right)}\right)^{\frac{\alpha}{1+\eta}-1} - \nu, 0}_{\to 0 \text{ when unconstrained}}\right)\frac{dc_2}{dd_1}. \tag{B.16}$$

To understand why this second term is 0 when firms are unconstrained, notice that when firms are able to perfectly maximize profits they hire an amount of labor corresponding to:

$$\alpha A l_2^{\alpha-1} = w \tag{B.17}$$

which iself implies, when combined with households first-order condition for labor/leisure:

$$\alpha A l^{\alpha-1-\eta} = \nu. \tag{B.18}$$

Similarly, the derivative dc_2/dd_1 is also 0 when financial intermediaries are unconstrained. To conclude, the planner's optimality condition for short-term debt is given by:

$$0 = \Phi^h \mathbb{E}_1^{SP}\left[(\nu - \alpha A l_2^{\alpha-1})\left(\phi H \frac{d\Omega_3}{dq_2}\frac{dq_2}{dd_1} - (1+r_1)\right)\right] + \Phi^b\left\{\mathbb{E}_1[\lambda_2] - \mathbb{E}_1^{SP}[\lambda_2] + \mathbb{E}_1^{SP}[\phi H \kappa_2 \frac{d\Omega_3}{dq_2}\frac{\partial q_2}{\partial d_1}]\right\} \tag{B.19}$$

where $\nu - \alpha A l_2^{\alpha-1}$ plays the role of a "capacity wedge": it measures how far firms are from their first-best production level. When this wedge is negative (there is underemployment, since $\alpha < 1$) a reduction in the leverage of financial intermediaries is beneficial for households, since it increases the production of real goods in a crisis.

Collateral Asset Investment: The same analysis applies to the externalities created by investing in H, keeping q_1 fixed. Similarly, a supplementary term appears because a marginal change in H will cause a marginal change in c_2, and thus a change in real output in a financial crisis. We thus have, following the same derivations as just above, that the planner's optimality condition for the creation of collateral assets is given by:

$$0 = \Phi^h \mathbb{E}_1^{SP}\left[(\nu - \alpha A l_2^{\alpha-1})\left(\beta\phi H \frac{d\Omega_3}{dq_2}\frac{dq_2}{dH} + z_2 + \phi q_2\right)\right] + \Phi^b\left\{\beta\mathbb{E}_1^{SP}[\lambda_2(z_2+q_2)] - \lambda_1 q_1 + \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H \frac{d\Omega_3}{dq_2}\left(\frac{\partial q_2}{\partial n_2}z_2 + \frac{dq_2}{dH}\right)\right]\right\} \tag{B.20}$$

Current Prices: The reversal externality, similar to the collateral externality, also enters in production. The welfare effects of changing marginally equilibrium prices q_1 are given by:

$$\mathcal{W}_q = \Phi^h \mathbb{E}_1^{SP}\left[(\nu - \alpha A l_2^{\alpha-1})\left(\beta\phi H \frac{\partial\Omega_3}{\partial q_1}\right)\right] + \Phi^b\left\{\beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H \frac{\partial\Omega_3}{\partial q_1}\right]\right\} \tag{B.21}$$

Summary: The welfare analysis is very similar to the case without production studied in the main paper. In particular, the forces at play are exactly the same. Production simply reinforces the need for the planner to intervene in financial markets. Indeed, the worsening of pessimism during crises has repercussions on the level of employment and output, inflating the size of welfare losses. The important lesson of this extension is that the features of behavioral biases that matter for welfare are entirely identical to what was identifies in the baseline welfare analysis.

B.3 Pledgeable Private Sector Loans

The previous section assumed that loans to the real sector (f_2) could not be used as collateral by financial intermediaries. Here, I look at the complete formulation of the collateral constraint, given by:

$$d_2 \leq \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)] + \psi f_2$$

whereby assuming that a fraction of the amount lent to firms can be recovered by depositors in the (non-equilibrium) possibility of default. The first-order condition for loans to real firms is now given by;

$$\lambda_2 = (1 + r_f) + \kappa_2 \psi \tag{B.22}$$

since lending to firms also expand the borrowing capacity of financial institutions *vis-à-vis* households. Since $\kappa_2 = \lambda_2 - 1$ as usual, this yields:

$$\lambda_2 = \frac{1 + r_f - \psi}{1 - \psi} \tag{B.23}$$

$$\implies \frac{1}{c_2} = \frac{\frac{\delta}{f_2} - \psi}{1 - \psi} \tag{B.24}$$

$$\implies \frac{1 - \psi}{c_2} = \frac{\delta}{f_2} - \psi \tag{B.25}$$

$$\implies f_2 = \frac{\delta c_2}{1 - \psi + \phi c_2} \tag{B.26}$$

where it is clear that the relation between c_2 and f_2 is not linear anymore. Using the budget constraint since financial intermediaries are constrained:

$$c_2 + f_2 = n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)] \tag{B.27}$$

$$\implies c_2 + \frac{\delta c_2}{1 - \psi + \phi c_2} = n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)]. \tag{B.28}$$

The fixed-point problem corresponding to belief amplification is now complexified by this additional non-linearity:

$$c_2 + \frac{\delta c_2}{1 - \psi + \phi c_2} = n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)] \tag{B.29}$$

$$q_2 = \beta c_2 \mathbb{E}_2[z_3 + \Omega_3(q_2)] + \phi(1 - c_2)\mathbb{E}_2[z_3 + \Omega_3(q_2)]. \tag{B.30}$$

As in Section 2.3, we can represent this equilibrium graphically. This is depicted in Figure 10. This modification clearly magnifies belief amplification by making the budget constraint a convex function instead of a linear one inside a crisis. The assumption made that $\psi \to 0$ in the previous section were thus conservative in terms of spillovers from the banking sector to real production in terms of welfare.

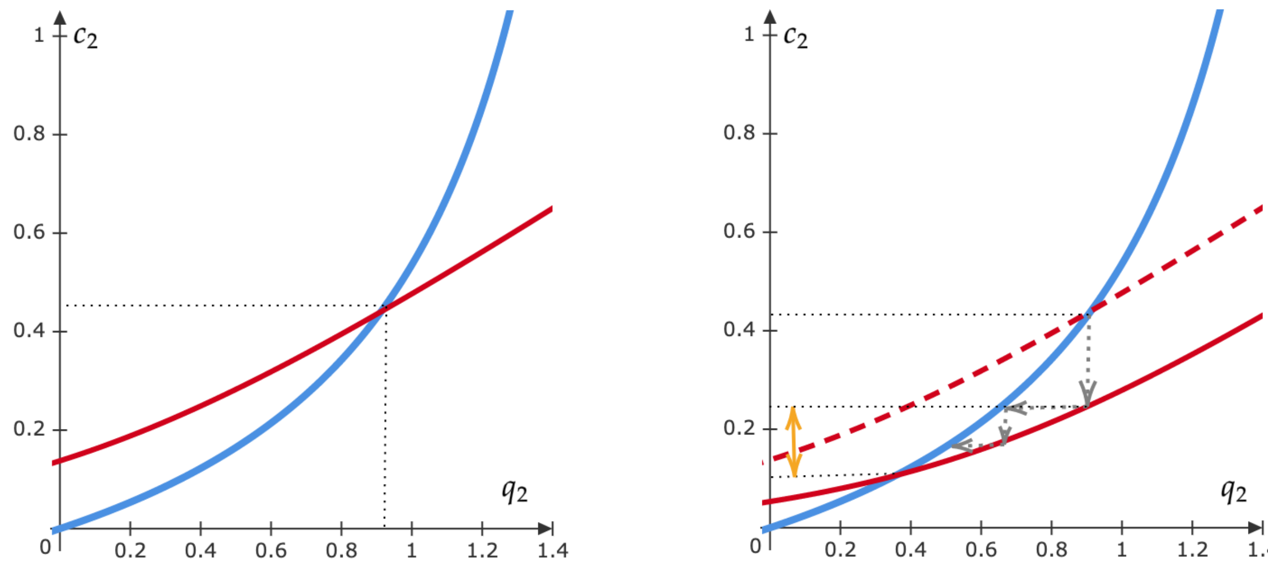

Figure 10: Graphical Illustration of Equilibrium Determination at $t = 2$ with pledgeable private sector loans. The red line represents the budget constraint, and the blue line represents the pricing condition. The right panel illustrates the phenomenon of belief amplification after a fall in net worth n_2. The arrows indicate the fixed-point problem that leads consumption to fall more than the size of the shock because of the tightening of the collateral constraint.

C Alternative Collateral Constraint with Current Prices

As mentioned in the paper, and as is well known in the financial frictions literature, the collateral constraint featuring $\mathbb{E}_2[z_]$ does not create any financial amplification, or any externality. This section shozs the robustness of my results when, instead, we consider a collateral constraint of the form:

$$d_2 \leq \phi H q_2. \tag{C.1}$$

This reliance on contemporaneous prices creates a feedback loop between the SDF and the tightness of the collateral constraint, which is at the heart of the financial amplification mechanism. This financial amplification is also why these models present inefficiencies: agents do not take into account that their leverage decision impact the price of the asset tomorrow, and hence the aggregate borrowing capacity of the financial sector. As shown for example by Ottonello et al. (2021), the quantitative predictions of the two models are very similar, making it hard to distinguish which type of friction is more likely to be relevant. My paper does take a stance on this debate, but rather shows that once endogenous sentiment is part of the picture, the gap between the two models is severely reduced.

The rest of this section follows the core of the paper and provide the same propositions and expressions, as well as intuitions, for this alternative collateral constraint. The proofs are relegated to Online Appendix Q.

C.1 Equilibrium

Financial Intermediaries at $t = 2$: When q_2 enters the collateral constraint, the asset price is given by:

$$q_2 = \beta c_2 \mathbb{E}_2[z_3 + \Omega_3] + \phi q_2(1 - c_2) \tag{C.2}$$

where the second term illustrates that holding marginally more of the asset is valuable since it relaxes financial constraints.[89] Financial amplification comes into play because the consumption level c_2 that prices the asset directly depends on the price of the asset through the collateral constraint (with $h = H$ in equilibrium):

$$c_2 = z_2 H - d_1(1 + r_1) + \phi H q_2. \tag{C.3}$$

A fall in the price of the risky asset tightens the budget constraint even more, thus leading the price to fall further as a result of stronger discounting, and so on. This *financial amplification* is represented on Figure 11.

As mentioned in the previous discussion the behavioral bias Ω_3 can also itself depend on q_2, for instance if agents extrapolate price changes. This adds the feature of *belief amplification* that compounds traditional financial amplification. Intuitively, a fall in the price of the risky asset creates endogenous pessimism, which leads the price of the asset to fall further. This tightens the borrowing constraints of financial intermediaries and in turns creates a further fall in the price that leads to more pessimism.

Asset price and consumption are determined in general equilibrium according to the fixed-point:

$$q_2 = \beta c_2(q_2) \mathbb{E}_2[z_3 + \Omega_3(q_2)] + \phi q_2(1 - c_2(q_2)). \tag{C.4}$$

I now illustrate the working of this fixed-point equation through the two examples of reduced-form behavioral biases laid out earlier. Here, and in the rest of the paper, it is useful to refer to n_2 as the net worth of the financial sector in period 2, i.e. $n_2 = z_2 H - d_1(1 + r_1)$. Consumption thus becomes $c_2 = n_2 + \phi H q_2$. If agents simply extrapolate fundamentals the fixed-point problem can be expressed as:

[89] This part of the expression complexifies the algebra, without bringing additional economic intuition. For this reason I present analytical examples that neglect this term, as in Jeanne and Korinek (2020). This is the case when the borrowing constraint takes the alternative form $d_2 \leq \phi q_2$, i.e. when the quantity of the risky asset does not enter the collateral constraint. A microfoundation of this constraint could be that lenders can only recover a fixed amount of the posted collateral.

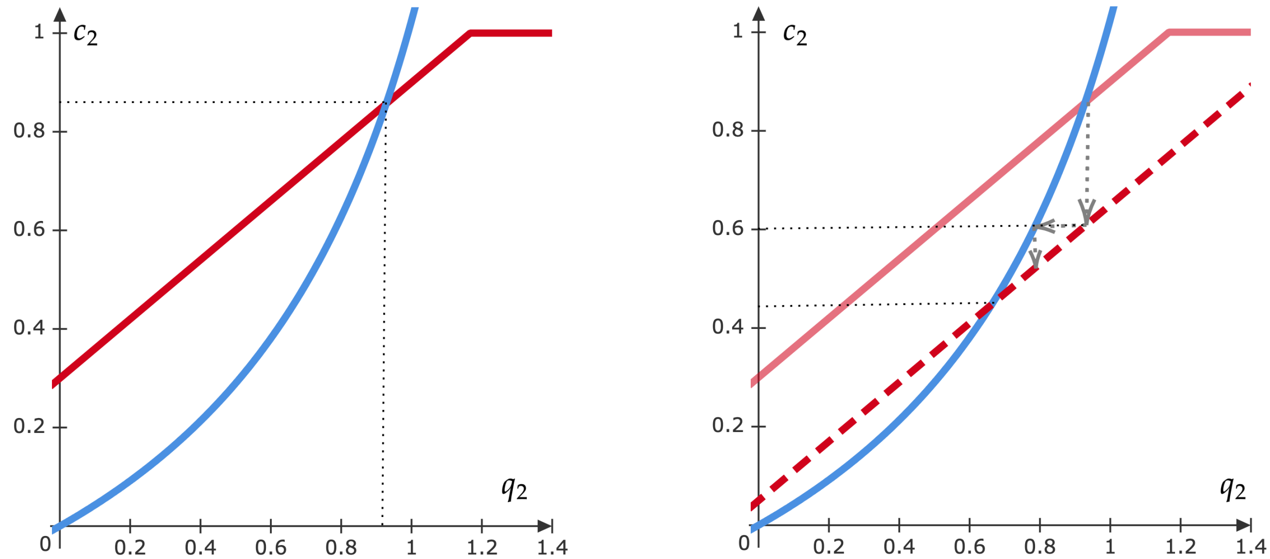

Figure 11: Graphical Illustration of Equilibrium Determination at $t = 2$. The red line represents the budget constraint equation (C.3), and the blue line represents the pricing equation (C.2). The right panel illustrates the phenomenon of financial amplification after a fall in net worth n_2. The arrows indicate the fixed-point problem that leads consumption to fall more than the size of the shock because of the tightening of the collateral constraint.

$$q_2 = \beta(n_2 + \phi H q_2)\mathbb{E}_2[z_3 + \alpha(z_2 - z_1)] + \phi q_2(1 - n_2 - \phi H q_2) \tag{C.5}$$

The first part of the right-hand side, embodying financial amplification through the pricing kernel, is linear in the price of the asset q_2. I thus present the results where the second part of the expression is negligible (full expressions are available in Online Appendix Q.1 for reference, but do not bring any additional intuition). The price is given by:

$$q_2 = \frac{\beta n_2\big(\mathbb{E}_2[z_3] + \alpha(z_2 - z_1)\big)}{1 - \phi H\big(\mathbb{E}_2[z_3] + \alpha(z_2 - z_1)\big)} \tag{C.6}$$

As can be readily seen from this expression, when agents are pessimistic during crises ($z_2 < z_1$) the asset price is lower with fundamental extrapolation. Interestingly the financial amplification channel (represented by the second negative term in the denominator) is also *weaker*, since a negative extrapolation term lowers the size of the feedback multiplier.

As in the core paper, a key object of interest in the welfare analysis is the price sensitivity to changes in net worth, $\partial q_2/\partial n_2$, as this object quantifies pecuniary externalities (as developed in Section C.2). In this case the sensitivity is:

$$\frac{\partial q_2}{\partial n_2} = \frac{\beta\big(\mathbb{E}_2[z_3] + \alpha(z_2 - z_1)\big)}{1 - \beta\phi H\big(\mathbb{E}_2[z_3] + \alpha(z_2 - z_1)\big)} \tag{C.7}$$

which, again, is weakened by pessimism during a crisis. In other words, it is harder to prop up the economy by injecting funds to financial intermediaries if entrenched pessimism is dragging down asset prices. This effect is represented schematically on Figure 12.

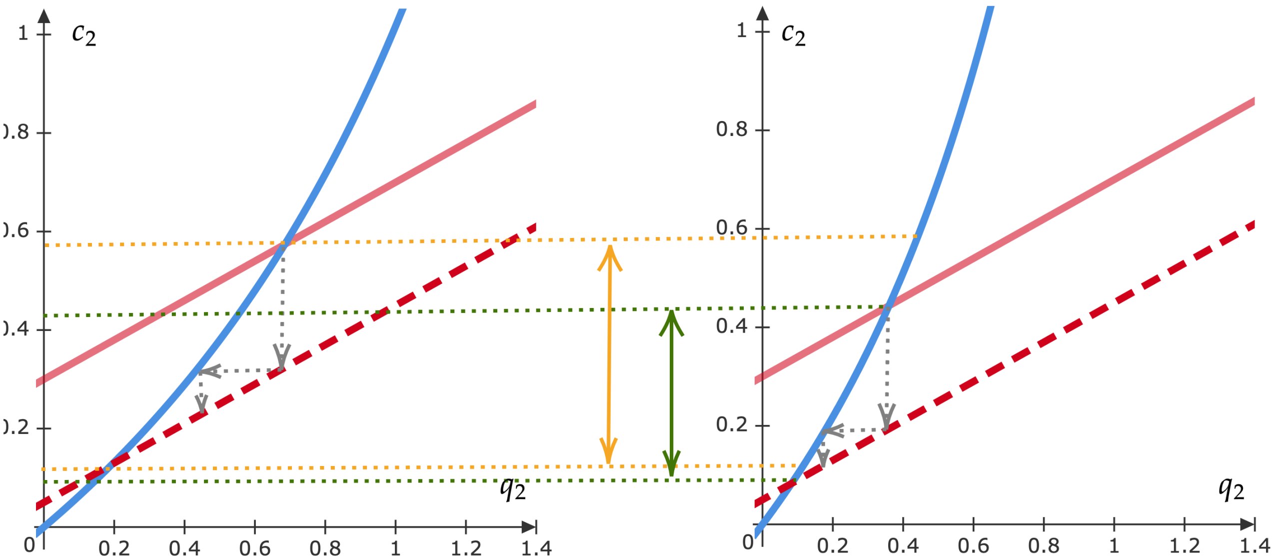

Figure 12: Graphical Illustration of Equilibrium Determination at $t = 2$ with Belief Amplification
The left panel illustrates the phenomenon of financial amplification after a fall in net worth n_2 in the rational case.The right panel adds an exogenous sentiment term corresponding to pessimism, $\Omega_3 < 0$. The yellows arrow indicates the size of financial amplification in the rational case, to be compared with the green arrow that shows the size of financial amplification once pessimism is entrenched into asset prices

Matters are different when agents extrapolate an endogenous object like the price q_2. In our price extrapolation formulation the pricing equation becomes (neglecting supplementary collateral terms as earlier):

$$q_2 = \beta(n_2 + \phi H q_2)\mathbb{E}_2[z_3 + \alpha(q_2 - q_1)]. \tag{C.8}$$

The pricing condition is now a quadratic equation, reflecting the multiplicative interaction of financial and belief amplifications. More interesting is the shape that the price sensitivity takes:

$$\frac{\partial q_2}{\partial n_2} = \frac{\beta\left(\mathbb{E}_2[z_3] + \alpha(q_2 - q_1)\right)}{1 - \beta\phi H \mathbb{E}_2[z_3] - \beta\alpha(c_2 + \phi H(q_2 - q_1))} \tag{C.9}$$

While the numerator is similar to the fundamental extrapolation case, whereby the price sensitivity is weakened by pessimism ($q_2 < q_1$), the denominator now has an extra term representing *belief amplification*. This new term compounds financial amplification and magnifies the sensitivity of asset prices – and thus of the borrowing capacity of the financial sector – to changes in net worth. Intuitively, injecting funds in this economy has powerful effects by relaxing collateral constraints and alleviating pessimism at the same time. These effects are illustrated on Figure 13. To ease exposition, parameters are chosen such that the equilibrium values are the same before the shock hits the net worth of financial intermediaries. By altering the shape of the pricing equation, belief amplification compounds financial amplification, leading shocks to have substantially larger effects.

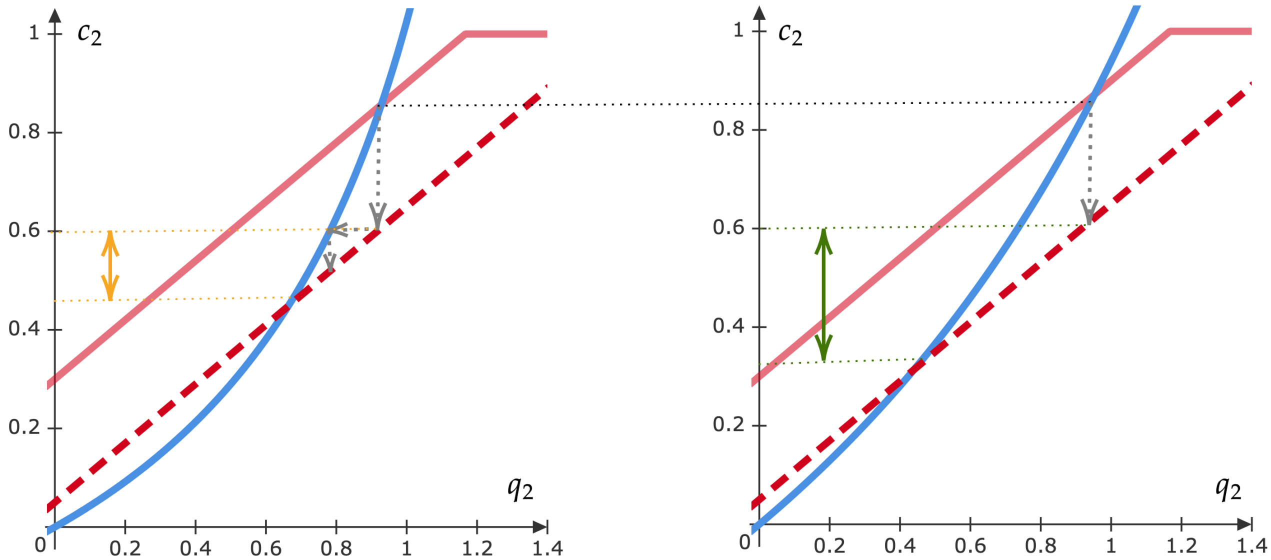

Figure 13: Graphical Illustration of Equilibrium Determination at $t = 2$ with Belief Amplification
The red line represents the budget constraint equation (C.3), and the blue line represents the pricing equation (C.2). The left panel illustrates the phenomenon of financial amplification after a fall in net worth n_2 in the rational case. The green arrows indicate the size of the amplification. The right panel adds price extrapolation of the form $\Omega_3 = \alpha(q_2 - q_1)$.

C.2 Welfare Analysis

C.2.1 Externalities

I start by listing the externalities now present in the rational version of this model, as this constitutes a policy benchmark. Two different, but related pecuniary externalities, usually require ex-ante correction to achieve constrained efficiency (Dávila and Korinek 2018). All externalities are working through the price of the asset used as collateral.

Private agents have a first-order condition on borrowing such as:

$$u'(c_1) = (1 + r_1)\mathbb{E}_1 \left[\frac{\partial \mathcal{W}_2}{\partial n_2} \right] \tag{C.10}$$

while the social planner has an extra-term corresponding to the pecuniary impact of private borrowing decisions:

$$u'(c_1) = (1 + r_1)\mathbb{E}_1^{SP} \left[\frac{\partial \mathcal{W}_2}{\partial n_2} + \frac{\partial \mathcal{W}_2}{\partial q_2} \frac{\partial q_2}{\partial n_2} \right] \tag{C.11}$$

and similarly for investment, since q_2 depends indirectly on H and n_2.

Borrowing Externality: First, agents are generically overborrowing. Atomistic financial intermediaries do not take into account that by increasing their leverage at $t = 1$, it subsequently lowers the price of the risky asset at $t = 2$ (through lowering financial intermediaries' net worth), which in turn hampers the aggregate financing capacity of the economy. This collateral externality is quantified by the following expression:

$$\mathcal{C}_D = -\mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{dq_2}{dn_2}\right] < 0 \tag{C.12}$$

which naturally features the sensitivity of the price with respect to net worth.

Investment Externality: Second, the same pecuniary externality pushes agents to generally underinvest in collateral assets. Similarly to the leverage externality, agents are not taking into account how a supplementary unit of collateral, by raising net worth next period, will ameliorate the borrowing capacity of the whole economy. We can similarly quantify this by:

$$\mathcal{C}_H = \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi \left(\frac{dq_2}{dn_2} z_2 + \frac{dq_2}{dH}\right)\right] \tag{C.13}$$

which is positive, as long as $z_2 \geq 0$ in all states of the world.[90]

Rational Benchmark: Traditional macroprudential policy, with perfectly rational agents, would offset these two pecuniary externalities using a tax on leverage and a subsidy on the creation of collateral assets (as shown by Dávila and Korinek 2018).[91] I now study welfare considering agents' departures from rationality.

C.2.2 Welfare Decomposition

Leverage We start by analyzing how changes in debt d_1 affect the welfare of individual agents.

Proposition 16 (Uninternalized Effects of Leverage)**.** *The uninternalized first-order impact on welfare when the level of short-term debt is marginally increased is given by:*

$$\mathcal{W}_d = \underbrace{\left(\mathbb{E}_1[\lambda_2] - \mathbb{E}_1^{SP}[\lambda_2]\right)}_{\mathcal{B}_d} - \underbrace{\mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{dq_2}{dn_2}\right]}_{\mathcal{C}_d}. \tag{C.14}$$

Behavioral Wedge: As in the core of the paper, an infinitesimal perturbation around the REE is enlightening (assuming Ω_2 and Ω_3 are small state-by-state):

Proposition 17 (Behavioral Wedge Approximation)**.** *If Ω_2 and Ω_3 are small state-by-state, the behavioral*

[90] Theoretically the externality can push towards under-investment when the asset drains liquidity in crisis times. In this paper I restricted the study to setups where $z_2 \geq 0$, which is the empirically relevant case for assets used in the repo market by financial intermediaries (like Mortgage-Backed Securities). My aim is not to claim that the benchmark should necessarily feature subsidies for holding collateral assets, but simply to highlight that, unlike for leverage, investment is not always associated with negative externalities.

[91] In effect, this could amount to a tax on consumption in this framework. I do not push this interpretation because this equivalence breaks down as soon as more margins of investment are introduced in the model.

wedge $\mathcal{B}_d$ for short-term debt can be expressed as:

$$\mathcal{B}_d \simeq \underbrace{-\boldsymbol{\Omega}_2 H\mathbb{E}^{SP}\left[\lambda_2^2\left(1+\phi\frac{dq_2}{dn_2}\right)\mathbb{1}_{\kappa_2>0}\right]}_{(i)} + \underbrace{\phi H\mathbb{E}^{SP}[\boldsymbol{\Omega}_3\lambda_2^2\frac{dq_2}{dz_3}\mathbb{1}_{\kappa_2>0}]}_{(ii)} \tag{C.15}$$

As can readily be seen from this expression, all the intuitions are preserved with this collateral constraint: the comovement of future sentiment with the health of the financial sector, and the necessary interaction with financial frictions. The new terms are simply coming from the fact that an error in the expectation of dividends directly spills over expected consumption, through the level of asset prices at $t = 2$.

Collateral Externality: The main difference with the pecuniary externality that would arise in a rational model is that the price sensitivity is different, because irrationality at $t = 2$, represented by Ω_3, influences equilibrium asset prices. The price sensitivity can be written generally, as in the next Proposition.

Proposition 18 (Price Sensitivity With Sentiment)**.** *A change in net worth in period 2 impacts equilibrium asset prices as:*

$$\frac{dq_2}{dn_2} = \frac{\beta\mathbb{E}_2\left[z_3+\boldsymbol{\Omega}_3\right]-\phi q_2}{1-\beta\phi H(\mathbb{E}_2\left[z_3+\boldsymbol{\Omega}_3\right])+2\phi^2 H q_2 - c_2\beta\frac{d\boldsymbol{\Omega}_3}{dq_2}} \tag{C.16}$$

Relative to a rational benchmark, where $\Omega_3 = 0$ and $d\Omega_3/dq_2 = 0$, sentiment creates two countervailing forces. First, entrenched pessimism makes the asset price less sensitive to changes in net worth, reducing the size of the pecuniary externality. Second, a change in net worth leads to a change in price because of financial amplification, which itself can lead to alleviating pessimism, supporting asset prices. This makes the price more sensitive to changes in net worth. Which of these two effects dominates is an empirical (to uncover the determinants of Ω) and quantitative question that lies outside the scope of this paper.[92] Nevertheless, even when the size of the pecuniary externality does not differ too much between a rational and a behavioral model, it can have drastic implications for welfare when models are calibrated according to the rational expectations hypothesis, an issue I explore in Online Appendix K.

Investment I perform the same kind of welfare decomposition, but looking at a marginal increase in investment into the creation of collateral assets. The price q_1 is kept fixed.

Proposition 19 (Uninternalized Effects of Investment)**.** *The uninternalized first-order impact on welfare*

[92] In Appendix Q.5, I present a Taylor expansion of the difference between the collateral externality evaluated by the rational planer, and the one that would be evaluated by a planner that respects private agents' beliefs.

when the level investment is marginally increased is given by:

$$\mathcal{W}_q = \underbrace{\left(\beta\mathbb{E}_1^{SP}\left[\lambda_2(z_2+q_2)\right]-\lambda_1 q_1\right)}_{\mathcal{B}_H} + \underbrace{\beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\left(\frac{dq_2}{dn_2}z_2+\frac{dq_2}{dH}\right)\right]}_{\mathcal{C}_H} \tag{C.17}$$

composed of two distinct effects: a behavioral wedge $\mathcal{B}_H$ and a collateral externality $\mathcal{C}_H$.

I explore these two terms in turn.

Behavioral Wedge: Similar to the welfare costs of higher leverage, the behavioral wedge for investment is given by the difference between a rational valuation of the risky asset, and private agents' valuation. Obviously this wedge is negative when agents are optimistic, or when agents do no realize their future pessimism. As previously, we can approximate this behavioral wedge for small deviations from rationality, as in the following Proposition.

Proposition 20 (Behavioral Wedge Approximation for H)**.** *If Ω_2 and Ω_3 are small state-by-state, the behavioral wedge $\mathcal{B}_d$ for investment in the collateral asset can be expressed as:*

$$\mathcal{B}_H = \beta\mathbb{E}_1^{SP}\left[\mathcal{B}_d(z_2)(z_2+q_2^r)\right] - \beta\Omega_2\mathbb{E}_1^{SP}\left[\lambda_2^r\left(1+\frac{dq_2}{dz_2}\right)\mathbb{1}_{\kappa_2>0}\right] + \beta\mathbb{E}_1^{SP}\left[\lambda_2^r\Omega_3\frac{dq_2}{dz_3}\mathbb{1}_{\kappa_2>0}\right] \tag{C.18}$$

where $\mathcal{B}_d(z_2)$ is the behavioral wedge for leverage, from Proposition 20, for a realization z_2 of the dividend process at $t=2$:

$$\mathcal{B}_d(z_2) = \Omega_2\lambda_2^2\left(H\Omega_2+\phi\frac{dq_2}{dn_2}\right)\mathbb{1}_{\kappa_2>0} - \phi H\Omega_3\lambda_2^2\frac{dq_2}{dz_3}\mathbb{1}_{\kappa_2>0}. \tag{C.19}$$

Collateral Externality: As explained in Appendix C.2.1, agents are not taking into account how a supplementary unit of collateral, by raising net worth next period, will ameliorate the borrowing capacity of the whole economy. I already showed how the sensitivity of the price with respect to net worth in this case was changed by sentiment. Similarly, how equilibrium prices move with the aggregate stock of collateral asset is changed by the behavioral wedges in a analogous way:

$$\frac{dq_2}{dH} = \frac{\beta\phi q_2\mathbb{E}_2\left[z_3+\mathbf{\Omega_3}\right]-\phi^2 q_2^2}{1-\beta\phi H(\mathbb{E}_2\left[z_3+\mathbf{\Omega_3}\right])+2\phi^2 Hq_2-\beta c_2\frac{d\mathbf{\Omega_3}}{dq_2}} \tag{C.20}$$

where the same two countervailing forces, from Ω_3 and $d\Omega_3/dq_2$, are still at play. As long as we restrict sentiment Ω_3 such that expected payoffs at $t=3$ are non-negative, this collateral externality has a positive sign, pushing towards under-investment in the decentralized equilibrium.

Welfare and Asset Prices The same reversal externality appears when changing asset price at $t=1$.

Proposition 21 (Welfare Effects of Changing Asset Prices)**.** *The first-order impact on welfare when asset prices q_1 are marginally increased is non-zero and corresponds to a reversal externality:*

$$\mathcal{R}_H = \beta\mathbb{E}_1^{SP}\left[\phi\kappa_2 H \frac{dq_2}{d\Omega_3}\frac{d\Omega_3}{dq_1}\right] \tag{C.21}$$

The intuition for this term is rather similar as in my baseline model and works as follows. When private agents marginally increase their investment in collateral assets, they push up the price of the asset today, in proportion of $c''(H)$, a positive term. This in turn might influence the formation of behavioral biases in the future, represented by the term $d\Omega_3/dq_1$. This change in sentiment at time 2 impacts the equilibrium price q_2, in proportion to $dq_2/d\Omega_3$, a positive quantity. This level change in asset prices impacts welfare if agents are against their borrowing constraint, since it directly alters the amount they can borrow. Finally, note that here this externality is plausibly sizeable: the sensitivity of prices to sentiment ($dq_2/d\Omega_3$) is magnified by the presence of belief amplification and financial amplification, as for the pecuniary externality (which is large enough to be a concern for policymakers and justify the installation of conventional macroprudential policies).

C.3 Optimal Policy

As in the baseline version of the paper, these two uninternalized welfare effects are enough to characterize optimal policy to achieve the second-best. Hence, Proposition 7 applies directly. The difference is simply that the comparison with the rational benchmark is less straightforward. Indeed, the rational benchmark now features a leverage tax and an investment subsidy (as explained in Appendix C.2.1). The need for a third instrument to control asset prices is again still valid because of the reversal externality.

Leverage Limit with $\phi M q_2$: Similarly, the insight that a leverage limit is more robust than a leverage tax to changes in contemporaneous behavioral biases still hold. A tempting, but erroneous, shortcut would be to then simply use the leverage limit recommended by a rational model, and be reassured that irrational exuberance would have no bite since agents would stay on the allocation desired bu the planner. In Online Appendix K, I show that the calibration of such models is highly dependent on the presence of sentiment in the model. This is because, to recover the size of financial frictions, a modeler typically calibrates the model to match the severity of financial crises and couples this with the Rational Expectations Hypothesis. When there are behavioral biases, however, the same severity of crises is achieved with *less strong* financial frictions, which in turn implies greater collateral externalities.

D Bailouts

While the previous analysis was made under the restriction of constrained efficiency, in reality financial crises are often addressed using direct liquidity injections. The possibility of ex-post "cleaning" is crucial to understand the policy debate around asset bubbles.[93] The so-called "Greenspan doctrine" states that it is preferable to clean, or "mop-up" once the crisis materialises, while the "'ex-ante leaning" camp argues for early intervention.[94] Farhi and Tirole (2012b) show that imperfectly targeted support to distressed institutions makes private leverage choices strategic complements, creating time-inconsistency and moral hazard problems. Jeanne and Korinek (2020) show that in a simplified environment with pecuniary externalities, the optimal policy mix involves both bailouts and ex-ante liquidity restrictions.[95]

This section investigates how bailouts, and their anticipation by agents, interact with irrational exuberance and distress concerns. The first question that naturally arises is whether the presence of behavioral biases changes the optimal policy mix between ex-ante and ex-post interventions. I then explore whether irrationality mitigates or amplify moral hazard problems.

D.1 A Stylized Model of Bailouts

The social planner can now directly inject liquidity into the financial system, by providing loans to financial institutions. Concretely, it transfers an amount b from households to financial intermediaries at time $t = 2$, and financial intermediaries reimburse households at $t = 3$ at the prevailing market risk-free rate. I assume that this transfer entails a quadratic cost $g(b)$, representing distortions arising from taxation or political economy concerns:

$$g(b) = \frac{b^2}{2\xi} \tag{D.22}$$

Outside of a financial crisis, there is no point in providing liquidity to financial intermediaries. Inside a crisis, welfare at $t = 2$ becomes:[96]

$$\mathcal{W}_2 = \ln\left(z_2 H - d_1(1+r_1) + b + \phi H \mathbb{E}_2[z_3 + \Omega_3]\right) \\ + \beta\left(\mathbb{E}^{SP}[z_3]H - \phi H \mathbb{E}_2[z_3 + \Omega_3]/\beta - b/\beta\right) - g(b). \tag{D.23}$$

[93] See Jones (2015) for a particularly clear summary of the "ex-post clean" and "ex-ante lean" paradigms.

[94] The "cleaning" camp was arguably dominant before the 2008 financial crisis. Early proponents of the "leaning" strategy include Bordo and Jeanne (2002) and Borio (2003).

[95] Other contributions include Bianchi (2016), Dewatripont and Tirole (2018), Farhi and Tirole (2018) and Clayton and Schaab (2020a).

[96] The welfare of households is irrelevant since the loan is made at the market rate, hence households stay on their Euler equation. Alternatively, g could represent the welfare costs borne by households if the loan make them deviate from their optimality conditions.

This leads to the following expression for the optimal bailout size:

$$b^*(d_1, H, z_2, \Omega_3) = \xi \left(\frac{\partial \mathcal{W}_2}{dn_2}(d_1, H, z_2, \Omega_3) - 1 \right) \tag{D.24}$$

where the partial derivative with respect to net worth can once again be expressed as:

$$\frac{\partial \mathcal{W}_2}{dn_2} = \kappa_2 \left(1 + \phi H \frac{d\Omega_3}{dq_2} \frac{da_2}{dn_2} \right). \tag{D.25}$$

Intuitively, a bailout is only desirable when $\kappa_2 > 0$: otherwise, there is no need to intervene to circumvent financial frictions that are not currently biting. The optimal bailout size is also increasing with κ_2: the more stringent frictions are, the more incentives to intervene and relax them. In particular, in the presence of excess pessimism $\Omega_3 < 0$, the financial crisis will be more severe and thus calling for stronger intervention. Furthermore, belief amplification also creates a new motive for ex-post intervention. By providing liquidity to distressed financial intermediaries, the social planner is indirectly supporting asset prices. This in turn can lessen pessimism and thus alleviate collateral constraints.[97]

D.2 Optimal Policy Mix

Does the possibility of bailouts in the future change the financial authority's incentive to impose leverage restrictions? Jeanne and Korinek (2020) show, in a somewhat related setup, that macroprudential policy is still desirable and can resolve any time-consistency problems that may arise from the use of ex-post liquidity provision. In this section I confirm that their results are still valid in the presence of behavioral factors. In other words, I show that the possible existence of irrational exuberance is not an argument in favor of the ex-post "cleaning" paradigm. As in Jeanne and Korinek (2020), this is intuitively because it is always optimal to use all second-best instruments in such settings with financial frictions, a general result originating in Lipsey and Lancaster (1956).

Proposition 22 (Uninternalized Welfare Effects with Bailouts)**.** *Under the presence of bailouts, the decomposition developed in Section 3.2 holds. The the uninternalized welfare effects of a marginal increase in leverage can be expressed as:*

$$\mathcal{W}_d = \underbrace{\beta \left(\mathbb{E}_1[\lambda_2(b^*)] - \mathbb{E}_1^{SP}[\lambda_2(b^*)] \right)}_{\mathcal{B}_d} - \underbrace{\beta \mathbb{E}_1^{SP} \left[\kappa_2(b^*) \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \frac{dq_2}{dn_2} \right]}_{\mathcal{C}_d}. \tag{D.26}$$

Equation D.26 makes the dependence of λ_2, marginal utility of financial intermediaries inside a

[97] Note that with this future-price collateral constraint, there is no reason to directly support asset prices through a TARP policy if sentiment is exogenous. When the collateral constraint features the current-price, it becomes valuable to directly target asset prices, but belief amplification is still reinforcing the incentives for intervention. See Online Appendix E.2 for the analysis of this case.

financial crisis, on the level of bailouts explicit.[98] One can readily see that, even in the presence of bailouts, the collateral externality is still present and uninternalized, thus calling for leverage restrictions.[99] Therefore, and naturally as in Jeanne and Korinek (2020), bailouts still need to be accompanied by ex-ante leverage restrictions to compensate for uninternalized welfare effects. The next part explores how the size of the optimal intervention changes because of moral hazard, following Farhi and Tirole (2012b).

D.3 Moral Hazard and Exuberance

The behavioral biases of agents furthermore interact with moral hazard concerns in a novel way.[100] This can be seen from inspecting the behavioral wedge of equation (D.26):

$$\mathcal{B}_{d,b^*} = \beta \mathbb{E}_1[\lambda_2(b^*)] - \beta \mathbb{E}_1^{SP}[\lambda_2(b^*)] \tag{D.27}$$

Marginal utility during a crisis depends on the level of bailouts b^*. But if agents recognize that bailouts will be determined optimally, according to equation (D.24), their expected bailout size state-by-state differs from the planner's. Indeed, b^* depends on the net worth of agents, but financial intermediaries believe that the asset will pay off $z_2 + \Omega_2$ instead of z_2 in each state. In other words, when agents are over-optimistic, they expect bailouts to be *smaller* than in reality, intuitively because they expect crises to be less severe than in reality. Hence, for a fixed $\Omega_2 > 0$, agents expect less aggressive bailouts than in reality: this directly reduces the behavioral wedge, which is the difference between expected marginal utilities between agents and the planner. Indeed, agents expect $\lambda_2(z_2 + \Omega_2, b^*(z_2 + \Omega_2, 0), 0)$, while the planner expected $\lambda_2(z_2, b^*(z_2, \Omega_3), \Omega_3)$. Moral hazard concerns are then attenuated by irrational optimism since $b^*(z_2 + \Omega_2, 0) < b^*(z_2, \Omega_3)$ and λ_2 is decreasing in b. This effect is further amplified by the fact that agents neglect the fact that the optimal bailout might be even larger since agents can be over-pessimistic in the future. This is summarized in the following proposition.

Proposition 23 (Moral hazard and Exogenous Biases)**.** *For a fixed* $\Omega_2 > 0$ *and fixed state-by-state* $\Omega_3 < 0$*, the behavioral wedge is negative and increasing in* ξ*.*

Matters, however, are more complicated when behavioral biases are endogenous, for example when Ω_2 depends on $(q_1 - q_0)$. In this case the expectation of future bailouts also raises the attractiveness of creating financial assets: their price will be supported by government's action in the

[98] One might seem surprising that the uninternalized welfare effect does not include a term $\partial b^* / \partial d_1$, that represents how increasing aggregate leverage changes the future size of bailouts. This is because bailouts are determined optimally in period $t = 2$, so the envelope theorem applies.

[99] This is assuming that bailouts are not effective enough to entirely prevent the occurrence of a financial crisis in the future. If bailouts are not costly at all, for example, the social planner will be able to provide enough liquidity in all states of the world such to achieve $\kappa_2 = 0$. Only under this extreme, and unrealistic case, are ex-ante restrictions undesirable.

[100] Dávila and Walther (2021) is the only work, to the best of my knowledge, that analyzes bailouts in an environment with distorted beliefs. They do not consider the moral hazard problems that arise from agents anticipating government intervention, neither do they study endogenous belief distortions, however.

intermediate period, lowering the risk premium. This pushes up the initial price of collateral assets, thereby fuelling irrational exuberance. This increase in Ω_2 leads financial intermediaries to augment their leverage. The initial equilibrium is then determined by multiple fixed-points between the value of bailouts, leverage, and sentiment, as can be seen from the following system:

$$u'(c_1) = -\mathbb{E}_1 \left[\frac{\partial \mathcal{W}_2}{\partial d_1} (d_1, b^*(d_1, H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)), H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)) \right] \tag{D.28}$$

$$\boldsymbol{q}_1 = \mathbb{E}_1 \left[\frac{\partial \mathcal{W}_2}{\partial H} (d_1, b^*(d_1, H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)), H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)) \right] \tag{D.29}$$

where bailouts b^* feed in the equilibrium price q_1 which then feeds into the Euler equation, and in turn changes the equilibrium value of bailouts, and so on. I represent this relation schematically in Figure 14.

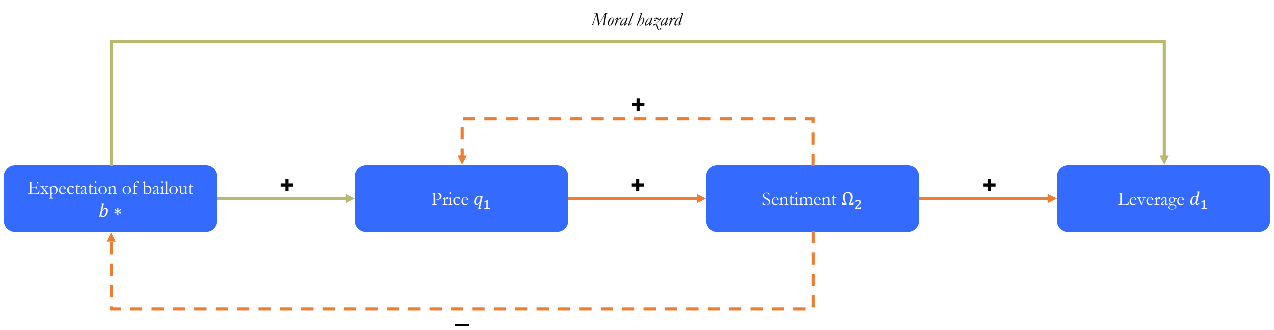

Figure 14: Impact of a bailout in the model with endogeneous sentiment.

I illustrate how this interaction between bailouts and sentiment depends on the determinants of Ω_2. Figure 15 presents the optimal leverage restriction that the planner needs to impose (in percentage of the decentralized equilibrium short-term debt) to attain the second-best, with and without bailouts, for different levels of initial sentiment. The left panel presents the case where Ω_2 is set exogenously. There, when optimism increases this reduces the value of the behavioral wedge and thus diminishes the size of optimal leverage reductions. The left panel then looks at the case where $\Omega_2 = \alpha(q_1 - q_0)$, and varird α. This time, even though optimism still weakens moral hazard concerns, it is compensated by the feedback effect that functions through asset prices.

Remark 14 (Timing of Announcement). The previous analysis rests on the idea that bailouts affect the price of the asset q_1, but what matters for beliefs is $q_1 - q_0$. This implicitly means that, at time 0, agents formed their expectations without taking future bailouts into account (or believing that bailouts are more costly than what they realize at $t = 1$). Hence, a corollary of this analysis is that announcing that bailouts will happen in case of a crisis must be done as early as possible if beliefs depend on price changes. Announcing bailouts at the last moment creates additional optimism in this case, right when the financial system is the most vulnerable.

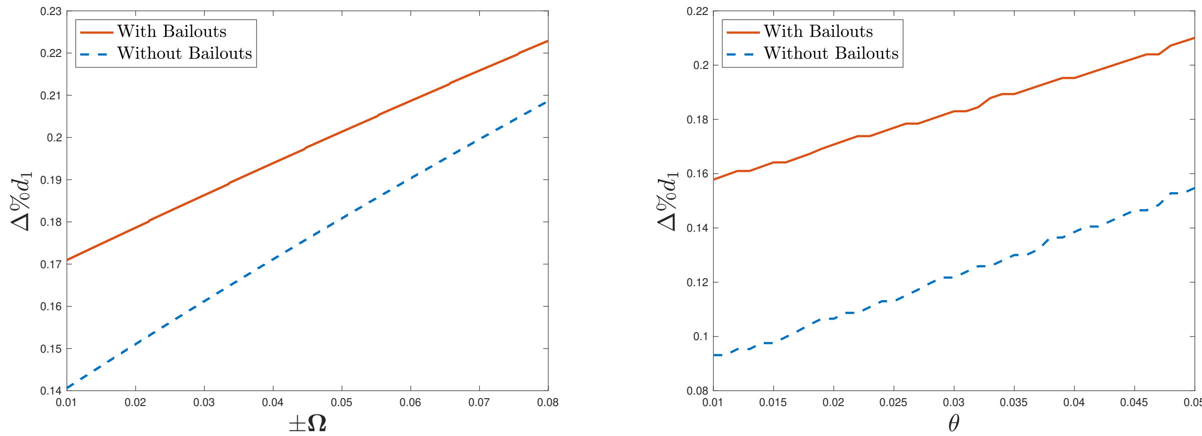

Figure 15: Supplementary Leverage Restrictions Required in the Exogenous Biases and and Endogenous Biases. In each panel the dotted lines plot the required decrease in leverage from the decentralized equilibrium to achieve the second-best in the absence of bailouts. Solid lines perform the same exercise but in the presence of bailouts in period $t = 2$. The behavioral bias in the left panel is of the fundamental extrapolation form, defined as $\Omega_{t+1} = \alpha(z_t - z_{t-1})$. z_0 and z_1 are chosen such that $\Omega_2 > 0$ to feature initial exuberance. The behavioral bias in the right panel is of the price extrapolation form, defined as $\Omega_{t+1} = \alpha(q_t - q_{t-1})$. q_0 is chosen such that $q_0 < q_1$ to feature initial exuberance.

E Extensions for the Current-Price Collateral Constraint

E.1 Real Production

E.1.1 A Simple Extension with Production

To incorporate a real side to the model, we allow households to supply labor at $t = 2$. Households have linear utility over consumption, and have a convex disutility for supplying labor in the intermediate period:

$$U^h = \mathbb{E}_1 \left[c_1^h + \beta \left(c_2^h - \nu \frac{l_2^{1+\eta}}{(1+\eta)} \right) + \beta^2 c_3^h \right] \tag{E.1}$$

where l_2 is the amount of labor supplied by households at time $t = 2$.

There is a fringe of competitive firms of measure one, producing from the labor of households. Firms use a decreasing returns to scale technology from labor, with productivity A:

$$Y_2 = A l_2^{\alpha} \tag{E.2}$$

To bridge the gap between Main street and Wall street, I add a financial friction. Firms need to pay a fraction γ of wage bills in advance to workers, which require them to borrow from financial intermediaries. In period 2, firms need to borrow $f_2 = \gamma w_2 l_2$ from financial intermediaries. We assume that the interest rate required by financial intermediaries to advance such funds depends on the size of the loan according to:

$$1 + r_f = \frac{\delta}{f_2} \tag{E.3}$$

This innocuous trick allows the model to say away from corner solutions and preserve financial amplification.[101] The set of budget constraint is now given by:

$$c_1^h + d_1 \leq e_1^h \tag{E.4}$$
$$c_2^h + d_1 \leq e_2^h + w_2 l_2 + d_1(1 + r_1) + \pi_2 \tag{E.5}$$
$$c_3^h \leq e_3^h + d_2(1 + r_2) \tag{E.6}$$

for households, and:

$$c_1 + c(H) \leq d_1 + e_1 \tag{E.7}$$
$$c_2 + d_1(1 + r_1) + f_2 + q_2 m \leq d_2 + (z_2 + q_2)H \tag{E.8}$$
$$c_3 + d_2(1 + r_2) \leq z_3 m + f_2(1 + r_f) \tag{E.9}$$

for financial intermediaries. Household optimization then simply yields:

$$w_2 = \nu l_2^{\eta} \tag{E.10}$$

It is also assumed for simplicity that loans made to firms cannot be used as collateral.[102] The specific form assumed in (E.3) simplifies matter since funds allocated to firms verify the following identity:

$$\frac{f_2}{\delta} = \beta c_2 \tag{E.11}$$

so that bankers' consumption and funds allowed to firms are proportional. Intuitively, when collateral constraints are extremely tight, this forces financial intermediaries to cut back on consumption *and* their traditional intermediary activities in the same way.[103] Thus the amount of labor used for production verifies:

$$l_2 = \left(\frac{z_2 H - d_1(1 + r_1) + \phi H q_2}{\gamma \nu \left(1 + \frac{1}{\beta\delta}\right)} \right)^{\frac{1}{1+\eta}} \tag{E.12}$$

[101] Remember that financial amplification comes from the two-way feedback effect between the Stochastic discount factor and the price fo the risky asset. A corner solution with respect to the borrowing of real firms would break this link.

[102] A more complete formulation of the collateral constraint would be:

$$d_2 \leq \phi H q_2 + \psi f_2$$

whereby assuming that a fraction of the amount lent to firms can be recovered by depositors in the (non-equilibrium) possibility of default. I am here analyzing the limiting case where $\psi \to 0$. The general case complexifies matters without bringing any new intuition. Analytical derivations of the general case are thus relegated to Appendix E.1.3.

[103] Consumption is needed for the SDF to generate financial amplification: a risk-neutral valuation pricing kernel breaks the feedback loop between the price of the asset and marginal utility. But one could think of c_2 as dividends or compensation.

which translates into a production level at time $t = 2$ of:

$$Y_2 = A\left(\frac{z_2 H - d_1(1+r_1) + \phi H q_2}{\gamma\nu\left(1+\frac{1}{\beta\delta}\right)}\right)^{\frac{\alpha}{1+\eta}} \tag{E.13}$$

Hence, a drop in the price of the risky asset q_2 directly impacts output, as well as a fall in financial intermediaries' net worth $z_2 H - d_1(1+r_1)$. Hence, looking at q_2 inside a crisis is a sufficient statistics even in this extended model with real production.

E.1.2 Welfare Analysis with Real Production

The planner maximizes:

$$\mathcal{W}_1 = \Phi^h \mathbb{E}_1^{SP}\left(c_1^h + \beta\left[c_2^h - \nu\frac{l_2^{1+\eta}}{1+\eta}\right] + \beta^2 c_3^h\right) + \Phi^b \mathbb{E}_1^{SP}\left(\ln(c_1) + \beta\ln(c_2) + \beta^2 c_3\right) \tag{E.14}$$

where Φ^h and Φ^b are the Pareto weights attached to each group by the planner. I denote by V_2^h and V_2^h the value functions of each group at time $t = 2$.

Leverage: We are interested in the derivatives of these value functions at time $t = 2$ with respect to the amount of short-term debt (or savings) chosen at time $t = 1$. Because funds allocated to firms (the f_2) chosen optimally without a constraint (see equation E.11), an infinitesimal change in f_2 will not have a first-order impact on the welfare of bankers:

$$\frac{dV_2^b}{dd_1} = \phi H(\lambda_2 - 1)\frac{dq_2}{dd_1} + \underbrace{\beta\frac{\delta}{f_2} - \lambda_2}_{=0}. \tag{E.15}$$

For households, however, there is a new term coming from the expansion of bank lending to firms in the real sector:

$$\frac{dV_2^h}{dd_1} = \phi H\underbrace{\left(\lambda_3^h - \lambda_2^h\right)}_{=0}\frac{dq_2}{dd_1} + \max\left(\underbrace{A\alpha\left(\frac{z_2 H - d_1(1+r_1) + \phi H q_2}{\gamma\nu\left(1+\frac{1}{\delta}\right)}\right)^{\frac{\alpha}{1+\eta}-1} - \nu, 0}_{\to 0 \text{ when unconstrained}}\right)\frac{dc_2}{dd_1}. \tag{E.16}$$

To understand why this second term is 0 when firms are unconstrained, notice that when firms are able to perfectly maximize profits they hire an amount of labor corresponding to:

$$\alpha A l_2^{\alpha-1} = w \tag{E.17}$$

which iself implies, when combined with households first-order condition for labor/leisure:

$$\alpha A l^{\alpha-1-\eta} = \nu. \tag{E.18}$$

Similarly, the derivative dc_2/dd_1 is also 0 when financial intermediaries are unconstrained. To conclude, the planner's optimality condition for short-term debt is given by:

$$0 = \Phi^h \mathbb{E}_1^{SP}\left[(\nu - \alpha A l_2^{\alpha-1})\left(\beta\phi H\frac{dq_2}{dd_1} - (1+r_1)\right)\right] + \\ \Phi^b\left\{\mathbb{E}_1[\lambda_2] - \mathbb{E}_1^{SP}[\lambda_2] - \mathbb{E}_1^{SP}[\phi H \kappa_2 \frac{\partial q_2}{\partial n_2}]\right\} \tag{E.19}$$

where $\nu - \alpha A l_2^{\alpha-1}$ plays the role of a "capacity wedge:" it measures how far firms are from their first-best production level. When this wedge is negative (there is underemployment, since $\alpha < 1$) a reduction in the leverage of financial intermediaries is beneficial for households, since it will increase the production of real goods in a crisis.

Collateral Asset Investment: The same analysis applies to the externalities created by investing in H, keeping q_1 fixed. Similarly, a supplementary term will appear because a marginal change in H will cause a marginal change in c_2, and thus a change in real output in a financial crisis. We thus have, following the same derivations as just above, that the planner's optimality condition for the creation of collateral assets is given by:

$$0 = \Phi^h \mathbb{E}_1^{SP}\left[(\nu - \alpha A l_2^{\alpha-1})\left(\beta\phi H\frac{dq_2}{dH} + z_2 + \phi q_2\right)\right] + \\ \Phi^b\left\{\lambda_1 q_1 - \beta\mathbb{E}_1^{SP}[\lambda_2(z_2+q_2)] - \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\left(\frac{\partial q_2}{\partial n_2}z_2 + \frac{dq_2}{dH}\right)\right]\right\} \tag{E.20}$$

Current Prices: The reversal externality, similar to the collateral externality, also enters in production. The welfare effects of changing marginally equilibrium prices q_1 are given by:

$$\mathcal{W}_q = \Phi^h \mathbb{E}_1^{SP}\left[(\nu - \alpha A l_2^{\alpha-1})\left(\beta\phi H\frac{\partial q_2}{\partial \Omega_3}\frac{\partial \Omega_3}{\partial q_1}\right)\right] + \Phi^b\left\{\beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\frac{\partial q_2}{\partial \Omega_3}\frac{\partial \Omega_3}{\partial q_1}\right]\right\} \tag{E.21}$$

E.1.3 Pledgeable Private Sector Loans

The previous section assumed that loans f_2 could not be used as collateral by financial intermediaries. Here, I look at the complete formulation of the collateral constraint, given by:

$$d_2 \le \phi H q_2 + \psi f_2$$

whereby assuming that a fraction of the amount lent to firms can be recovered by depositors in the (non-equilibrium) possibility of default. The first-order condition for loans to real firms is now given by;

$$\lambda_2 = (1 + r_f) + \kappa_2 \psi \tag{E.22}$$

since lending to firms also expand the borrowing capacity of financial institutions *vis-à-vis* households. Since $\kappa_2 = \lambda_2 - 1$ as usual, this yields:

$$\lambda_2 = \frac{1 + r_f - \psi}{1 - \psi} \tag{E.23}$$

$$\implies \frac{1}{c_2} = \frac{\frac{\delta}{f_2} - \psi}{1 - \psi} \tag{E.24}$$

$$\implies \frac{1 - \psi}{c_2} = \frac{\delta}{f_2} - \psi \tag{E.25}$$

$$\implies f_2 = \frac{\delta c_2}{1 - \psi + \phi c_2} \tag{E.26}$$

where it is clear that the relation between c_2 and f_2 is not linear anymore. Using the budget constraint since financial intermediaries are constrained:

$$c_2 + f_2 = n_2 + \phi H q_2 \tag{E.27}$$

$$\implies c_2 + \frac{\delta c_2}{1 - \psi + \phi c_2} = n_2 + \phi H q_2. \tag{E.28}$$

The fixed-point problem corresponding to financial amplification is now complexified by this additional non-linearity:

$$c_2 + \frac{\delta c_2}{1 - \psi + \phi c_2} = n_2 + \phi H q_2 \tag{E.29}$$

$$q_2 = \beta c_2 \mathbb{E}_1[z_3] + \phi q_2 (1 - c_2). \tag{E.30}$$

As in Section 2.3, we can represent this equilibrium graphically. This is depicted in Figure 16. This modification clearly amplifies financial amplification by making the budget constraint a convex function instead of a linear one inside a crisis. The assumption made that $\psi \to 0$ in the previous section were thus conservative in terms of spillovers from the banking sector to real production in terms of welfare.

E.2 Bailouts

Similarly to the baseline model in the main paper, the costs of bailouts are modeled in reduced-form as:

$$g(b) = \frac{b^2}{2\xi} \tag{E.31}$$

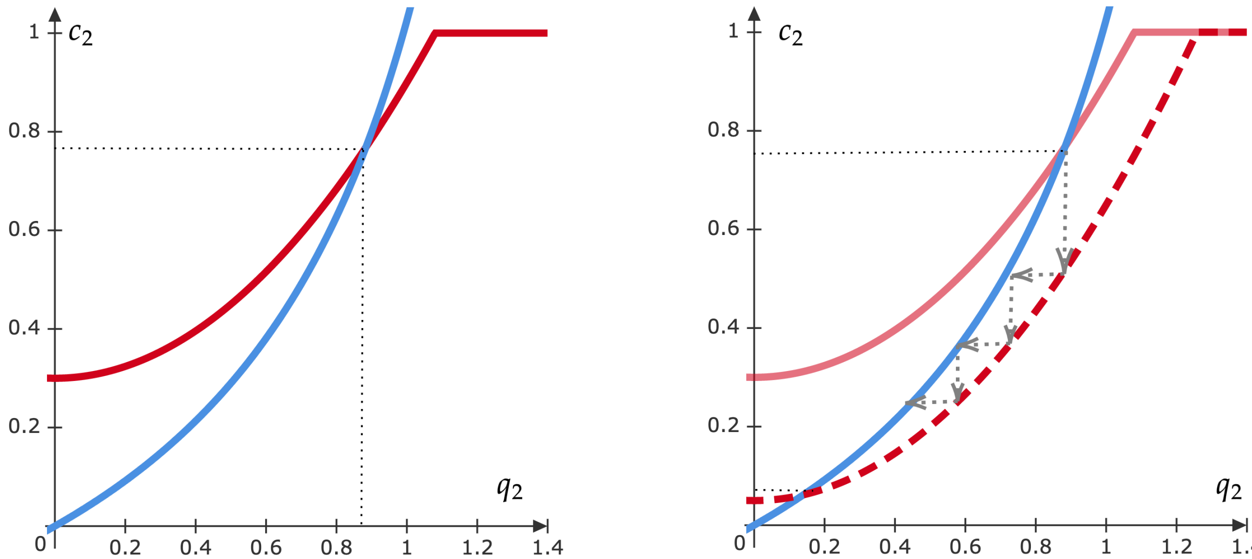

Figure 16: Graphical Illustration of Equilibrium Determination at $t = 2$ with pledgeable private sector loans. The red line represents the budget constraint, and the blue line represents the pricing condition. The right panel illustrates the phenomenon of financial amplification after a fall in net worth n_2. The arrows indicate the fixed-point problem that leads consumption to fall more than the size of the shock because of the tightening of the collateral constraint.

Outside of a financial crisis, there is no point in providing liquidity to financial intermediaries. In a crisis, welfare at $t = 2$ with this collateral constraint becomes:[104]

$$\mathcal{W}_2 = \ln\left(z_2 H - d_1(1+r_1) + b + \phi H q_2\right) + \beta\left(\mathbb{E}^{SP}[z_3]H - \phi H q_2/\beta - b/\beta\right) - g(b). \tag{E.32}$$

This leads to the following expression for the optimal bailout size:

$$b^*(d_1, H, z_2, \Omega_3) = \xi\left(\frac{\partial \mathcal{W}_2}{dn_2}(d_1, H, z_2, \Omega_3) - 1\right). \tag{E.33}$$

Intuitively, the optimal bailout size takes the same form as in the paper, but there will be an additional effect because of financial amplification. By providing liquidity to the financial sector, bailouts support asset prices and thus increase the borrowing capacity of the financial sector, an effect present even in the rational benchmark.

The behavioral wedge of equation takes the exact same form in both cases of collateral constraints:

$$\mathcal{B}_{d,b^*} = \mathbb{E}_1[\lambda_2(b^*)] - \mathbb{E}_1^{SP}[\lambda_2(b^*)] \tag{E.34}$$

Harginal utility during a crisis always depends on the level of bailouts b^*. But if agents recognize that bailouts will be determined optimally, according to equation (E.33), their expected bailout size state-by-state will differ from the planner's. The insight of the moral hazard consequences in

[104]The welfare of households is still irrelevant here since the loan is made at the market rate.

Appendix D are thus preserved here.

Similarly the insight about the interaction of endogenous sentiment and bailouts survives, since it only requires the *marginal welfare* functions to be impacted by price extrapolation and bailouts in the same way:

$$u'(c_1) = -\mathbb{E}_1 \left[\frac{\partial \mathcal{W}_2}{\partial d_1} \left(d_1, b^*(d_1, H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)), H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)\right) \right] \tag{E.35}$$

$$\boldsymbol{q}_1 = \mathbb{E}_1 \left[\frac{\partial \mathcal{W}_2}{\partial H} \left(d_1, b^*(d_1, H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)), H, z_2 + \Omega_2(\boldsymbol{q}_1 - q_0)\right) \right] \tag{E.36}$$

E.3 Monetary Policy

The only difference yields in the form of the reversal externality: changes in future sentiment now impact welfare *indirectly* by changing asset prices. This leads to the following expression for monetary policy:

$$\begin{aligned} \frac{d\mathcal{W}_1}{dr_1} = & \underbrace{\frac{dY_1}{dr_1}\mu_1}_{(i)} + \underbrace{\frac{dd_1}{dr_1}\mathcal{W}_d}_{(ii)} + \underbrace{\frac{dH}{dr_1}\mathcal{W}_H}_{(iii)} \\ & + \underbrace{\frac{d\Omega_2}{dq_1}\frac{dq_1}{dr_1}\left(\frac{dd_1}{d\Omega_2}\mathcal{W}_d + \frac{dH}{d\Omega_2}\mathcal{W}_H\right)}_{(iv)} + \underbrace{\beta\mathbb{E}_1\left[\frac{dq_2}{d\Omega_3}\frac{d\Omega_3}{dq_1}\frac{dq_1}{dr_1}\kappa_2\phi H\right]}_{(v)} \end{aligned} \tag{E.37}$$

F Heterogenous Beliefs

A fully general treatment of heterogeneity in beliefs inside the framework presented previously lies outside the scope of this paper. I thus focus on a stylized version of heterogeneity where all financial intermediaries are over-optimistic, but differ in their degree of over-optimism.[105] Financial intermediaries are indexed by $i \in [0, 1]$, and bank i holds a belief distortion of $\Omega_{2,i}$, with:

$$\Omega_{2,i} = \Omega_2 + \epsilon_2(2i - 1) \tag{F.1}$$

such that the most pessimistic bank is bank 0 with a bias of $\Omega_2 - \epsilon_2 > 0$, bank $1/2$ holds an average bias Ω_2 and bank $t = 1$ is the most optimistic with a bias of $\Omega_2 + \epsilon_2$. Put simply, financial intermediaries' beliefs are distributed uniformly around a value of Ω_2. Furthermore, I assume that this heterogeneity is common knowledge, and everyone agrees that there is no more heterogeneity in beliefs at time $t = 2$ onwards, and that the social planner can only impose a uniform tax or leverage

[105] Accordingly, the planner will use beliefs that are outside the convex combination of agents' beliefs. See Brunnermeier, Simsek and Xiong (2014) for an analysis of a welfare criterion with heterogeneous beliefs and when the planner does not take a stand on whose belief is correct.

limit (i.e., the planner imposes a uniform regulation). Last, I assume that the risky asset is in a fixed supply H to focus on leverage decisions. I start, as usual, by backward induction.

Financial Intermediaries at $t = 2$ Financial intermediaries enter the period with heterogeneous net worth $n_{2,i}$ (coming from heterogeneous leverage and heterogeneous holdings of the risky asset), and they hold homogeneous beliefs. Start with the following lemma:

Lemma 1. *In a crisis equilibrium, financial intermediaries have the same consumption level at $t = 2$, irrespective of the heterogeneity in net worth. This consumption level is given by:*

$$c_2 = \int_0^1 n_{2,i} di + \phi H q_2 \tag{F.2}$$

and the price of the risky asset in equilibrium is implicitly defined by:

$$q_2 = \beta \bar{c}_2 \mathbb{E}_2[z_3] + \phi q_2 (1 - \bar{c}_2) \tag{F.3}$$

Proof. An individual bank's optimality condition, in a crisis, for holding the risky asset is given by:

$$q_2 = \beta c_{2,i} \mathbb{E}_2[z_3] + \phi q_2 (1 - c_{2,i}) \tag{F.4}$$

which is a linear function of $c_{2,i}$, while all other variables are common to all agents. Thus, $c_{2,i} = c_{2,j}$ for all i and j in $[0, 1]$. Integrating over gives:

$$\int_0^1 c_{2,i} = \int_0^1 n_{2,i} + \phi q_2 \int_0^1 m_{2,i} \tag{F.5}$$

and by market clearing $m_{2,i} = H$, while $c_{2,i} = \bar{c}_2$ by what precedes. □

This lemma also implies that individual's holdings of the risky asset are given by:

$$m_{2,i} = \frac{\bar{c}_2 - n_{2,i}}{\phi q_2}. \tag{F.6}$$

Note that this means that financial intermediaries entering with higher net worth end up holding *less* of the assets. This is because they need to borrow less: indeed, the level of borrowing of bank i in equilibrium at $t = 2$ is $d_{2,i} = \phi m2, i q_2$. This level of consumption, however, is not what is expected by agents since they believe that the realization of the dividend z_2 will be higher on average. In other words, we have $\mathbb{E}_1^{SP}[\bar{\lambda}_2] > \mathbb{E}_{1,i}[\bar{\lambda}_2]$ for all i.

Welfare Analysis at $t = 1$: Taking this into account, the social planner first-order condition is given by:[106]

$$0 = \int_0^i \lambda_{1,i} - \mathbb{E}_1^{SP}\left[\bar{\lambda}_2\right] - \mathbb{E}_1^{SP}\left[\phi H \bar{\kappa}_2 \frac{\partial q_2}{\partial \bar{n}_2}\right]. \tag{F.7}$$

The utilitarian social planner can thus maximize welfare by imposing a uniform tax on leverage equal to:

$$\tau_d = \frac{\mathbb{E}^{SP}[\bar{\lambda}_2] - \int_0^i \mathbb{E}_{1,i}[\bar{\lambda}_2] + \mathbb{E}_1^{SP}\left[\phi H \bar{\kappa}_2 \frac{\partial q_2}{\partial \bar{n}_2}\right]}{\int_0^i \lambda_{1,i}} \tag{F.8}$$

which is, again, showing the robustness of the formulation in Proposition 1. And here again, intuitively, a leverage limit is robust to heterogeneity, whereas the tax is not. Since the planner's beliefs are outside the convex set of agents' beliefs, the required leverage is below the decentralized outcome for each financial intermediary, hence a leverage limit will be binding for every financial intermediary, and will bring back this margin to the second-best.

Impact of Heterogeneity on the Optimal Tax: A natural question that arises is whether heterogeneity in beliefs has a detrimental effect on the behavioral wedge and the collateral externality.

Bank i with beliefs $\Omega_{2,i}$ believes that the net worth of bank j in period $t = 2$ will be:

$$\mathbb{E}_{1,i}[n_{2,j}] = (z_2 + \Omega_{2,i})m_{1,j} - (1 + r_1)d_{1,j} \tag{F.9}$$

and so it believes that the aggregate net worth of the financial system will be:

$$\mathbb{E}_{1,i}[\bar{n}_2] = (z_2 + \Omega_{2,i})H - (1 + r_1)\int_0^1 d_{1,j}dj \tag{F.10}$$

but the $\int_0^1 d_{1,j}dj$ is correct since I assumed that belief disagreement were common knowledge. Hence the distribution of beliefs about aggregate net worth is uniformly distributed. It thus follows that the average belief about $\bar{n}_2$ is the same with and without heterogeneity, if $\int_0^1 d_{1,j}dj$ is kept constant.

In which direction goes aggregate leverage, $\int_0^1 d_{1,j}dj$? To understand what happens, consider the simplified case where there is no risk, and agents cannot trade their endowment of the risky asset at $t = 1$ (to prevent arbitrage). Then, (perceived) consumption smoothing implies that:

$$e_1 + d_{1,i} = (z_2 + \Omega_{2,i})H - (1 + r_1)d_{1,i} + \phi H q_2(\bar{c}_{2,i}) \tag{F.11}$$

which yields:

$$e_1 + (2 + r_1)d_{1,i} = (z_2 + \Omega_{2,i})H + \phi H q_2(\bar{c}_{2,i}). \tag{F.12}$$

[106] The aggregation made on the collateral externality part is made possible by the linearity of preferences at $t = 3$, also responsible for the fact that marginal utility is homogeneous at $t = 2$.

Aggregating over individuals, we get;

$$e_1 + (2 + r_1) \int_0^1 d_{1,i} di = (z_2 + \bar{\Omega}_2) H + \phi H \int_0^1 q_2(\bar{c}_{2,i}) di \tag{F.13}$$

and hence $\int_0^1 d_{1,i} di$ is implicitly defined by this relation since $\bar{c}_{2,i}$ is a function of $\int_0^1 d_{1,j} dj$. This is to be compared with the homogenous relation:

$$e_1 + (2 + r_1) d_{1,i} = (z_2 + \bar{\Omega}_2) H + \phi H q_2(\bar{c}_2). \tag{F.14}$$

Inspecting equations (F.13) and (F.14) shows that the behavior of aggregate leverage is determined by whether $\int_0^1 q_2$ is an increasing or decreasing function with respect to the heterogeneity of beliefs. The concavity of the price function (see Online Appendix Q.10) means that this is a decreasing function, implying that heterogeneity causes *lower* aggregate leverage (the slightly more optimistic financial intermediary takes on less additional leverage than what the pessimistic financial intermediary subtracts). Since the optimal leverage target of the planner is unchanged by the presence of heterogeneity, this heterogeneity reduces the gap between the aggregate decentralized solution and the planner's solution.

G Alternative Measures of Sentiment

I document the covariance between sentiment and financial intermediaries' health using six different measures that are common in the literature:

1. The HY indicator of Greenwood and Hanson (2013);
2. The GZ credit spreads Gilchrist and Zakrajšek (2012);
3. The LTG measure from Bordalo et al. (2020);
4. The PVS indicator of Pflueger, Siriwardane and Sunderam (2020);
5. The BW measure of sentiment of Baker and Wurgler (2007);
6. The $CAPE$ ratio of Campbell and Shiller (1988).

H Additional Results for Ω-Uncertainty

H.1 Ω_3-Uncertainty

This section extends the insights of Section 5 to the case where the uncertainty pertains to Ω_3. I start by studying the realization of only one state of the world, and complete the proof using the linearity of expectations.

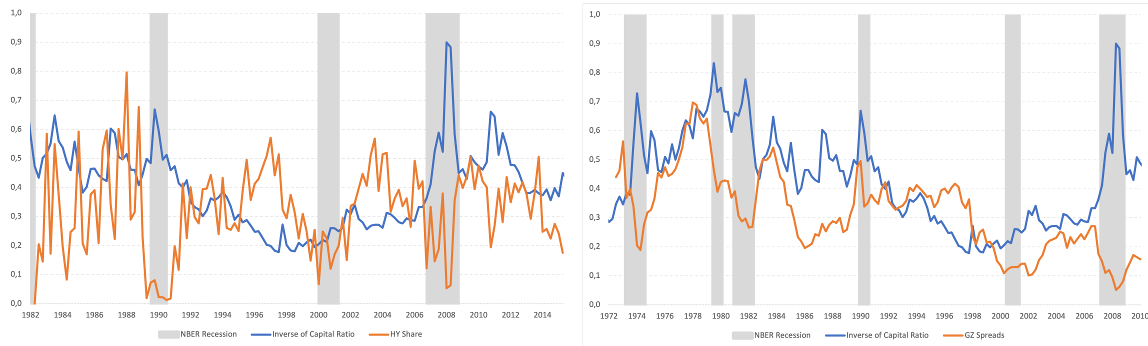

Figure 17: Time-series variation of λ_2 and credit-market proxies for Ω_3. For the financial health of intermediaries λ_2, I rely on He et al. (2017) which computes an intermediary capital ratio. The inverse of this capital ration is proportional to λ_2 when agents have log-utility, as in this model. For Ω_3, I use the High-Yield share of issuance measure of Greenwood and Hanson (2013) on the left panel and invert the credit-spread measure of Gilchrist and Zakrajšek (2012) on the right panel.(credit spreads are high when sentiment is low, and vice-versa).

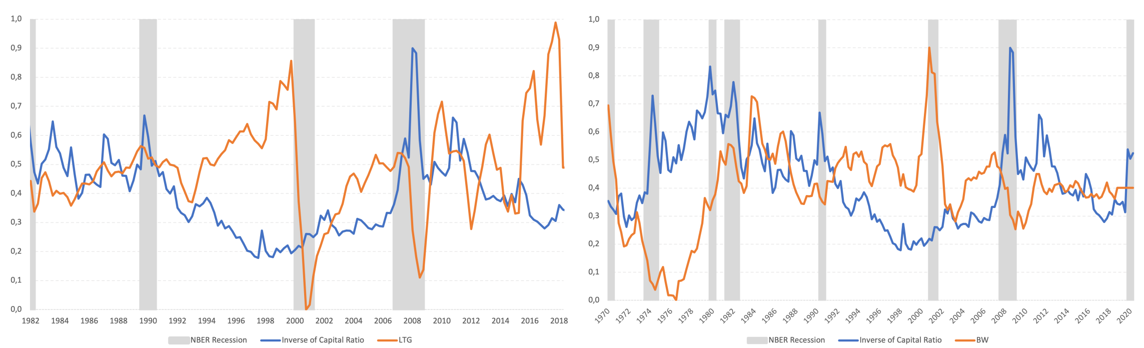

Figure 18: Time-series variation of λ_2 and stock-market sentiment for Ω_3. For the financial health of intermediaries λ_2, I rely on He et al. (2017) which computes an intermediary capital ratio. For Ω_3 on the left panel, I use the Long Term Growth (*LTG*) measure of Bordalo et al. (2020). This is directly constructed from survey data by aggregating stock market analysts' expectations. For the right panel, I use the Baker-Wurgler index of sentiment of Baker and Wurgler (2007).

I assume that for a given realization of z_2, the planner has a uniform distribution on sentiment during a crisis:

$$w_3 \sim \mathcal{U}\left[\bar{\Omega}_3 - \sigma_\Omega, \bar{\Omega}_3 + \sigma_\Omega\right] \tag{H.1}$$

The integral (denoted by L) used by the social planner to compute the marginal effect on welfare on increasing leverage becomes:

$$L = \frac{1}{2\sigma_\Omega} \int_{-\sigma_\Omega}^{\sigma_\Omega} \frac{\partial \mathcal{W}_2}{\partial n_2} \left(d_1, H; q_2, z_2, \bar{z}_3 - \bar{\Omega}_3 - \omega_3\right) d\omega_3 \tag{H.2}$$

Assume first that for all realisations of ω_3 the resulting equilibrium is a crisis one. This yields:

$$L = \frac{1}{2\sigma_\Omega} \int_{-\sigma_\Omega}^{\sigma_\Omega} \frac{1}{n_2 + \phi H(\bar{z}_3 - \bar{\Omega}_3 - \omega_3)} d\omega_3 \tag{H.3}$$

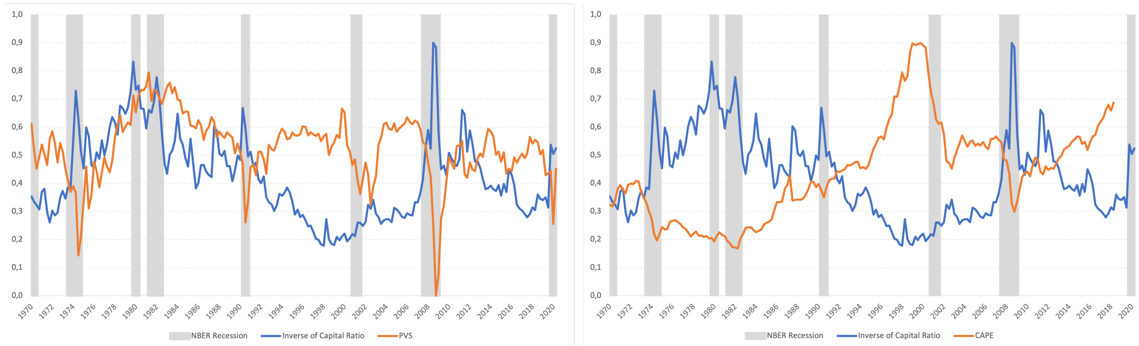

Figure 19: Time-series variation of λ_2 and additional proxies for Ω_3. For the financial health of intermediaries λ_2, I rely on He et al. (2017) which computes an intermediary capital ratio. For the left panel for Ω_3, I use the Price of Volative Stock (*PVS*) measure of Pflueger et al. (2020). For the right panel I use the the CAPE ratio of Campbell and Shiller (1988).

$$\implies L = -\frac{1}{(2\sigma_\Omega)\phi H}\left[\ln(n_2 + \phi H(\bar{z}_3 - \bar{\Omega}_3 - \omega_3))\right]_{-\sigma_\Omega}^{\sigma_\Omega} \tag{H.4}$$

$$\implies L = \frac{1}{(2\sigma_\Omega)\phi H}\ln\left(\frac{n_2 + \phi H(\bar{z}_3 - \bar{\Omega}_3 + \sigma_\Omega))}{n_2 + \phi H(\bar{z}_3 - \bar{\Omega}_3 - \sigma_\Omega))}\right) \tag{H.5}$$

This is a functions of the type:

$$f(x) = \frac{1}{x}\ln\left(\frac{K+x}{K-x}\right) \tag{H.6}$$

And we can show that this is increasing in x, for $x \in [0, K]$. Indeed, the derivative is given by:

$$f'(x) = \frac{(K^2 - x^2)\ln\left(\frac{K+x}{K-x}\right) + 2Kx}{x^2(K-x)(K+x)} \tag{H.7}$$

The denominator is clearly positive, but the denominator is indeterminate. Take the derivative of the denominator:

$$\frac{d}{dx}(K^2 - x^2)\ln\left(\frac{K+x}{K-x}\right) + 2Kx = 2x\ln\left(\frac{K+x}{K-x}\right) > 0 \tag{H.8}$$

The denominator is thus increasing and its limit in 0 is 0. Hence, f is increasing on $[0, K]$. Accordingly, L is increasing in σ_Ω.

Left now is the same calculation when for some parts of the uncertainty set, the economy is outside of a crisis. Following the same steps as before, this boils down to the study of, the time:

$$g(x) = \frac{1}{x}\ln\left(\frac{1}{K-x}\right) \tag{H.9}$$

Where the derivative is now:

$$g'(x) = \frac{\frac{x}{a-x} - \ln\left(\frac{1}{K-x}\right)}{x^2} \tag{H.10}$$

And the derivative of the numerator is:

$$\frac{d}{dx}\frac{x}{a-x} - \ln\left(\frac{1}{K-x}\right) = \frac{x}{(a-x)^2} > 0 \tag{H.11}$$

Since $g'(0^+) > 0$, g is increasing. Thus the same result applies. This concludes the proof by linearity of expectations: since this integral is increasing in σ_Ω, all components of the expectations over all future states of the world are increasing, and it then follows that the overall expectation is increasing in σ_Ω.

H.2 Amplification with Price Extrapolation

So far the exercise was done assuming that sentiment was constant state-by-state in period $t = 2$. Do the results change once we extend this to price-dependent biases? The answer lies in the shape of the marginal welfare functions once sentiment moves with prices inside a crisis:

$$\frac{d\mathcal{W}_2}{dn_2} = \lambda_2 + \kappa_2 \phi H \frac{d\Omega_3}{dq_2}\frac{dq_2}{dn_2} \tag{H.12}$$

The question is whether this added part, which is simply the collateral externality, is adding or retrenching convexity. With price extrapolation, we have:

$$\frac{d\Omega_3}{dq_2} = \alpha \tag{H.13}$$

So the only part left is the shape of $\frac{dq_2}{dn_2}$. Fortunately, we showed in Section Q.10 that this is also a convex function: see equations (Q.58) to (Q.63). Hence the marginal welfare function is more convex, amplifying the need for preventive restrictions in the face of uncertainty.

H.3 Ω-Uncertainty and Investment

So far, Proposition 10 was concerned about leverage restrictions. How is uncertainty changing the uninternalized effects of investment in H? Assume that sentiment is exogenous.[107] The first order condition becomes:

$$\lambda_1 c'(H) = \frac{1}{2\sigma_\Omega}\int_0^\infty \left[\int_{-\sigma_\Omega}^{\sigma_\Omega} \lambda_2(z_2 - \bar{\Omega}_2 - \omega_2)(z_2 - \bar{\Omega}_2 - \omega_2 + q_2(z_2 - \bar{\Omega}_2 - \omega_2))d\omega_2\right] f_2(z_2)dz_2. \tag{H.14}$$

Fortunately, it is now straightforward to sign the derivative of this function given the previous proofs. We know that $\lambda_2(z_2 - \bar{\Omega}_2 - \omega_2)$ is convex in ω_2. This is multiplied by a linear and positive function of ω_2 (the dividends), and then by the price realization at $t = 2$.

[107] I slightly abuse notations below by not writing Ω_3 for simplicity. This is harmless since we are fixing the first-order condition of private agents and simply study whether the first-order condition of the social planner is increasing or decreasing in σ_Ω.

The price at $t = 2$ is given by:

$$q_2 = \beta(n_2 + \phi M \mathbb{E}_2[z_3])\mathbb{E}_2[z_3] + \phi(1 - n_2 - \phi M \mathbb{E}_2[z_3])\mathbb{E}_2[z_3] \tag{H.15}$$

Which is clearly linear in ω_2 since net worth is linear in ω_2:

$$n_2 = (z_2 - \bar{\Omega}_2 - \omega_2)H - d_1(1 + r_1) \tag{H.16}$$

Hence this function is convex in ω_2, which implies that the right-hand side of the first-order condition is increasing in uncertainty. This time, however, this means that $c'(H)$ in equilibrium needs to be higher than in the decentralised equilibrium. Hence, uncertainty calls for increasing investment (or, in the case with large exuberance, less restrictions on investment). Intuitively, uncertainty increases the SDF that prices the asset, meaning that more consumption should be shifted to the future.

H.4 Ω-Uncertainty and Reversal Externality

How is sentiment uncertainty influencing the optimal conduct of monetary policy? The previous derivations can help us answer that question. The reversal externality that monetary policy explicitly targets is expressed as:

$$\mathcal{R}_q = \mathbb{E}_1 \left[\kappa_2 \phi H \frac{d\Omega_3}{dq_1} \frac{dq_1}{dr_1} \right] \tag{H.17}$$

The only unknown part of this expression, from the perspective of period $t = 1$, is the product $\kappa_2 d\Omega_3 / dq_1$. Fortunately, we just showed that κ_2 is a convex object with respect to sentiment uncertainty. It then depends on the shape of $d\Omega_3 / dq_1$ with respect to sentiment uncertainty. For instance, with price-extrapolation, $\kappa_2 d\Omega_3 / dq_1 = \kappa_2 \alpha$ and so this object is still convex.

Thus, sentiment uncertainty with linear price-extrapolation increases the incentive for the central bank to tighten interest rates when asset prices soar. To conclude, in times of heightened uncertainty about Ω_2 or Ω_3, with price extrapolation, the central planner should:

1. Tighten leverage limits;
2. Relax LTV ratios;
3. Increase the interest rate.

I Infinite-Horizon Model

This section provides a simple infinite-horizon version of the model. It shows how the insights derived in the main paper are not dependent on the 3-period structure assumed.

Financial intermediaries have a utility function given by:

$$U_t = \sum_{i\geq 0}^{+\infty} \beta^{t+i} \ln(c_{t+i}) \tag{I.1}$$

While households have again linear-utility throughout:

$$U_t^h = \sum_{i\geq 0}^{+\infty} \beta^{t+i} c_{t+i}^h \tag{I.2}$$

I assume that the stock of assets H is fixed and given. It can only be held by intermediaries. The budget constraint of financial intermediaries at t are:

$$c_t + d_{t-1}(1 + r_{t-1}) + q_t h \leq d_t + (z_t + q_t)H \tag{I.3}$$

$$d_t \leq \phi h \mathbb{E}_t[z_{t+1} + \Omega_{t+1}] \tag{I.4}$$

Where in equilibrium $h = H$. The first-order conditions, using the same notation for the Lagrange multipliers as in the core of the text, are thus given by:

$$\lambda_t = \frac{1}{c_t} \tag{I.5}$$

$$\lambda_t = \beta(1 + r_t)\mathbb{E}_t[\lambda_{t+1}] + \kappa_t \tag{I.6}$$

$$\lambda_t q_t = \beta \mathbb{E}_t[\lambda_{t+1}(z_{t+1} + \Omega_{t+1} + q_{t+1}^r)] + \phi \kappa_t \mathbb{E}_t[z_{t+1} + \Omega_{t+1}] \tag{I.7}$$

I assume that the planner can impose a tax on borrowing, or a tax on the holdings of the risky asset. Since H is fixed, this tax only purpose is to change the equilibrium price of the asset. Practically, this policy can be implemented through monetary policy, as explored in Section 6, with spillovers on inflation targeting. I focus on a simple asset tax here for simplicity.

One-Time Policy Intervention Start with the easiest case where the planner intervenes only once and commits to never intervene again afterwards. Thus the equilibrium is the laissez-faire one starting from $t + 1$. The planner chooses directly d_t and q_t at t, and takes as given the future values of d_{t+j} and q_{t+j} that will be freely determined in equilibrium.

The social planner maximises:

$$\mathcal{W}_t = \ln(c_t) + \beta \mathbb{E}_t[\mathcal{W}_{t+1}(d_t, q_t)] \tag{I.8}$$

The first-order conditions of the social planner are given by:

$$0 = \lambda_t - \beta \mathbb{E}_t[\lambda_{t+1}] - \sum_{j\geq 1}^{+\infty} \beta^{t+j} \mathbb{E}_t \left[\kappa_{t+j} \phi H \frac{d\Omega_{t+j}}{dq_{t+1}} \frac{dq_{t+1}}{dn_{t+1}} \right] \tag{I.9}$$

$$0 = \sum_{j \geq 0}^{+\infty} \beta^{t+j} \mathbb{E}_t \left[\kappa_{t+j} \phi H \frac{d\Omega_{t+j}}{dq_t} \right] \tag{I.10}$$

The social planner is thus trying to manipulate two things: (i) how future sentiment will be affected by future prices since a change in borrowing today impact prices tomorrow; and (ii) how future sentiment will be affected by current prices.[108]

Discussion of Implementability Constraints The above analysis allowed the Social Planner to directly choose the asset price at t. This simplifies the analysis but at the same time lacks concreteness. It is hard to imagine an infinite-horizon problem where the planner cannot realistically circumvent the market determination of asset prices at each t.

A full analysis of the problem where the social planner chooses short-term debt on behalf of private agents, and asset prices remain market-determined, is outside the scope of this paper. A few remarks can be made, nevertheless. Bianchi and Mendoza (2018) show that in a setup with a current-price collateral constraint, the optimal policy crucially depends on whether the planner has commitment or not. The intuition goes as follows: during a crisis, the planner would like to promise lower future consumption. This changes the stochastic discount factor, and thus props up asset prices, relaxing the borrowing constraint. This is however time-inconsistent: next period, it will be sub-optimal for the planner to implement this low level of consumption.

Such an effect would also arise here in the case of endogenous sentiment: the planner would like to prop up asset prices at t in order to prop up Ω_{t+1} and relax the collateral constraint (again with belief amplification replacing the traditional role of financial amplification). Note, however, that the problem would be vastly more complicated: the consumption that the planner would promise is not the one expected by private agents, since agents expect future consumption to depend on their biased estimate of future dividends. But to prop up asset prices like suggested by Bianchi and Mendoza (2018), it has to be that agents believe future consumption to be lower than under laissez-faire, since it is private agents' pricing condition that implements asset prices in equilibrium.

This also raises the more general question of policy in models where agents are behavioral. In my baseline setup, agents should be surprised that the planner is intervening: the model does not feature any externality from a rational perspective. There is thus the open question of what agents believe about future policy (an issue I briefly touched upon in Section 6.4), and whether agents should adapt in the face of recurrent intervention. These fascinating issues are left to future research.

[108] The derivates effect $d\Omega_{t+j}/dq_{t+1}$ are assumed to be taking into account the full effects on Ω_{t+j} for conciseness. For example for Ω_{t+2}, it implicitly factors in how prices at $t+1$ directly impact sentiment at $t+2$, but also how the change in Ω_{t+1} changes q_{t+2} and thus Ω_{t+2}. See Section 6.3 for an example on the 4-period model.

J Various Psychological Models of Asset Prices and Ω-Correspondence

J.1 Diagnostic Expectations

Diagnostic expectations are a psychologically founded model of belief formation in light of new data. It builds on the representativeness heuristic of Tversky and Kahneman (1983): agents overweight attributes of a class that are more frequent in that class than in a reference class. Bordalo et al. (2018) apply this logic to belief formation about aggregate economic condition. Specifically, assume that the state of the world follows an AR(1) process:

$$z_t = bz_{t-1} + \epsilon_t \tag{J.1}$$

with $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$. By taking as a reference point the state where there is no news, Bordalo et al. (2018) derive that the diagnostic distribution is also normal, with the same variance, but with mean:

$$\mathbb{E}_t^\theta[z_{t+1}] = \mathbb{E}_t^{SP}[z_{t+1}] + \theta\left(bz_t - b^2 z_{t-1}\right) \tag{J.2}$$

where θ is the parameter governing the representativeness bias. Diagnostic expectations are thus nested as:

$$\Omega_{t+1} = \theta\left(bz_t - b^2 z_{t-1}\right) \tag{J.3}$$

which is close to the reduced-form used in the core of the paper, $\Omega_{t+1} = \alpha\left(z_t - z_{t-1}\right)$. The difference is that, for diagnostic expectations, what matters is not the *per se* movements in z, but the unexpected component of these movements.

J.2 Internal Rationality

Adam and Marcet (2011) present a model where agents are not "externally rational:" they do not know the true stochastic process for payoff relevant variables beyond their control, i.e. prices in my setup. Adam, Marcet and Beutel (2017b) apply this idea in an asset pricing framework, giving rise to boom-bust cycles. Here I adapt their idea to my setup with some simplifying assumptions, and show in which circumstances the results change.

Agents are rational regarding the distribution of z_t, but they believe prices evolve according to:

$$q_{t+1} = q_t + \beta_{t+1} + \epsilon_{t+1} \tag{J.4}$$

with ϵ_{t+1} is a transitory shock and β_{t+1} is a persistent component evolving as:

$$\beta_{t+1} = \beta_t + \nu_{t+1}. \tag{J.5}$$

Furthermore, all innovations are jointly normal. Adam et al. (2017b) show that under some conditions, and when agents are using a steady-state precision, the filtering problem boils down to expectations evolving as:

$$\tilde{E}_t[q_{t+1}] = (1+g)(q_t - q_{t-1}) + (1-g)\tilde{E}_{t-1}[q_t] \tag{J.6}$$

where g is the equivalent of a Kalman gain, function of the variances of the noise terms. To make progress, I further assume that agents place a low conditional variance on this estimate, such that I can study the limiting case where this point estimate is believed to be certain (i.e. there is no risk for the price next period in agents' mind). I denote by $\tilde{q}_2$ this point estimate, such that agents' optimization yields:

$$q_1 = \beta \mathbb{E}_1 \left[\frac{\lambda_2}{\lambda_1}(z_2 + \tilde{q}_2) \right]. \tag{J.7}$$

Equation (J.7) can be rewritten using the correct price used by the planner q_2:

$$q_1 = \beta \mathbb{E}_1 \left[\frac{\lambda_2}{\lambda_1}(z_2 + q_2 + (\tilde{q}_2 - q_2)) \right]. \tag{J.8}$$

so an equivalent to the Ω_2 used throughout this paper is $\Omega_2^q = \tilde{q}_2 - q_2$: a bias on expected prices that is positive (exuberance) when the forecasted value if above the realized value, and vice-versa.

How does this impact the welfare analysis? It crucially depends on the form of the collateral constraint. If we stay in the benchmark case where the collateral constraint takes the form:

$$d_2 \leq \phi H \mathbb{E}_2[z_3] \tag{J.9}$$

then it is clear that since agents are correct about the distribution of fundamentals, they will make no mistake regarding their future net worth or the future borrowing capacity of the economy. Consequently, the only margin that is distorted if the investment margin: agents are too optimistic (pessimistic) regarding the payoffs of their investment, since they are too optimistic (pessimistic) regarding the resale value of the asset they are creating. Thus, only the behavioral wedge for investment is non-zero in this case.

Importantly, there are no externalities anymore. Indeed, decisions during the boom will impact time expectations of prices made at $t = 2$ but these expectations will not affect the tightness of collateral constraints.

This discussion makes clear that for externalities to survive in this case, it is necessary to have a collateral constraint that depends on prices (either current prices, or expected prices), whereas biases on fundamentals impact welfare in a "robust" way. When the collateral constraint takes the form:

$$d_2 \leq \phi H q_2 \tag{J.10}$$

then biases impact its tightness: when agents are over-pessimistic regarding future prices at $t = 3$, that impacts the equilibrium value of q_2.[109] In this case externalities survive. But notice that the sign of the key derivative for the reversal externality is clearly ambiguous:

$$\frac{d\Omega_3^q}{dq_1} = \frac{d\tilde{q}_3}{dq_1} = (1-g)\left(\frac{d\tilde{q}_2}{dq_1} - 1\right). \tag{J.11}$$

This is because sentiment is "sticky" with learning. If by reducing asset prices at $t = 1$, the planner makes future agents more pessimistic in a financial crisis, that hurts welfare.

J.3 Overconfidence

In an early behavioral finance survey, De Bondt and Thaler (1995) stated that "perhaps the most robust finding in the psychology of judgment is that people are overconfident." Overconfidence has been most widely used to explain large trading volume, by generating substantial disagreement between investors Odean 1998. Because this paper is about *aggregate* over-optimism or over-pessimism, I will focus in this section on the features of overconfidence that can generate momentum and reversals.[110] The interested reader can find an exploration of how heterogeneous beliefs among financial intermediaries impact the results in Appendix F.[111]

Financial institutions have a prior over the distribution of dividends in period $t = 2$:

$$z_2 \sim \mathcal{N}(\mu_0, \sigma_0^2) \tag{J.12}$$

and receive a signal $s = z_2 + \epsilon$ with:

$$\epsilon \sim \mathcal{N}(0, \sigma_s^2). \tag{J.13}$$

Overconfident financial intermediaries have a posterior of:

$$z_2 \sim \mathcal{N}\left(\mu_0 + \frac{\sigma_0^2}{\sigma_0^2 + \tilde{\sigma}_s^2}(s - \mu_0), \frac{\sigma_0^2}{1 + \frac{\sigma_0^2}{\tilde{\sigma}_s^2}}\right) \tag{J.14}$$

where $\tilde{\sigma}_s^2 < \sigma_s^2$, which means that overconfident agents believe that the signal has a higher precision than in reality. This directly implies that the bias, relative to the social planner valuation, is given by:

$$\Omega_2 = \frac{\sigma_s^2 - \tilde{\sigma}_s^2}{(\sigma_0^2 + \tilde{\sigma}_s^2)(\sigma_0^2 + \sigma_s^2)}\sigma_0(s - \mu_0) \tag{J.15}$$

so that agents become exuberant after positive news ($s > \mu_0$): $\Omega_2 > 0$.

Notice how the variance of the two distributions are different with overconfidence. As such, the

[109] I am here slightly abusing notation, since strictly speaking there is no price at $t = 3$. But claiming that there is no bias due to internal rationality in the crisis period would only come from the simplifying assumption that the horizon is finite.

[110] See e.g. Daniel, Hirshleifer and Subrahmanyam (1998).

[111] Caballero and Simsek (2020a) focus on prudential policies with financial speculation, but in an environment with aggregate demand – rather than pecuniary – externalities.

results in Propositions 2 and 5 are not directly applicable. But a higher $\tilde{\sigma}_s^2$ means that agents are using a narrower distribution than the social planner. This is reminiscent of the results presented in Section 5: intuitively, this will create an even larger gap between the two solutions since agents will neglect left-tail and right-tail events. As shown in Proposition 10, this is calling for tighter macroprudential regulation ex-ante.

J.4 Sticky Beliefs

While this paper is mostly concerned with investors that adjust their views too much in response to information, there is also widespread evidence of investors adjusting their beliefs *too little*. A recent example is the work of Bouchaud et al. (2019), where investors form expectations according to:

$$\tilde{\mathbb{E}}_1[z_2] = (1-\lambda)\mathbb{E}_1^r[z_2] + \lambda\tilde{\mathbb{E}}_0[z_2] \tag{J.16}$$

where $\mathbb{E}_1^r$ is the rational time 1 expectations about the future dividend. When $\lambda = 0$, expectations are fully rational. When $\lambda > 0$, expectations depend on past expectations. In terms of the notation of my paper, the bias can be expressed as:

$$\tilde{\mathbb{E}}_1[z_2] = \mathbb{E}_1^{SP}[z_2] + \lambda\left(\tilde{\mathbb{E}}_0[z_2] - \mathbb{E}_1^r[z_2]\right) \tag{J.17}$$

so that

$$\Omega_2 = \lambda\left(\tilde{\mathbb{E}}_0[z_2] - \mathbb{E}_1^r[z_2]\right). \tag{J.18}$$

Agents are thus over-optimistic in period t when the objective expected dividend is less than the expectation agents held in period $t-1$. Expanding this expression recursively yields:

$$\Omega_2 = \lambda\left(\mathbb{E}_0^r[z_2] - \mathbb{E}_1^r[z_2]\right) + \lambda\Omega_1. \tag{J.19}$$

which naturally gives rise to a formulation close to the one stipulated in Assumption 7.

Finally, note that this formulation does not necessarily imply pessimism during booms, and so calls for less aggressive macroprudential leverage limits. Indeed, agents are over-optimistic as long as $\tilde{\mathbb{E}}_0[z_2] > \mathbb{E}_1^r[z_2]$. It thus suffices that agents should revise their expectations down to create optimism. The three-period model is not suited to study this kind of dynamics, where a slowdown in growth for example creates irrational exuberance. But the unravelling of sentiment along such a cycle can be understood in the extended framework of Section 6.3. There, I showed that tightening later in the cycle has ambiguous effects since it also makes agents more pessimistic *during* a crisis.

J.5 Inattention

Gabaix (2019) argues that "much of behavioral economics may reflect a form of inattention." He proposes a theory of over- and under-reaction that rests on agents anchoring on a default autocorrelation parameter. Specifically, assume that the dividend process follows and AR(1) as in:

$$z_{t+1} = \rho z_t + (1-\rho)z_0 + \epsilon_{t+1} \tag{J.20}$$

Because agents have to deal with too many such processes, they may not fully perceive each autocorrelation, and instead use ρ_s to make forecasts, with:

$$\rho_s = m\rho + (1-m)\rho_d \tag{J.21}$$

where ρ_d is the average autocorrelation agents encounter. It is then straightforward to show that the bias used in this paper becomes:

$$\Omega_{t+1} = (\rho_s - \rho)(z_t - z_0). \tag{J.22}$$

Agents are thus overreacting when the autocorrelation parameter of the dividend process is less than the anchor value, ρ_d, since $\rho_s - \rho = (1-m)(\rho_d - \rho)$. When this is the case, agents make forecasts thinking that the dividend process is more persistent than in reality, thus putting too much weight on recent data and not enough on the unconditional mean of the process. The opposite happens when $\rho > \rho_d$.

K The Mistakes of Rational Calibration

In Appendix C.2.2, I showed that the collateral externality of the behavioral model differ from the rational counterfactual: sentiment creates two countervailing forces. First, entrenched pessimism makes the asset price less sensitive to changes in net worth, reducing the size of the pecuniary externality. Second, a change in net worth leads to a change in price because of financial amplification, which itself can lead to alleviating pessimism, supporting asset prices. This makes the price more sensitive to changes in net worth.

While it is entirely possible, given the presence of these two countervailing forces, that the introduction of sentiment in this model does not tremendously change the size of the pecuniary externalities, it can still imply large policy differences if the modeler uses the rational expectations hypothesis during a calibration. To understand this, notice that the pecuniary externality is a structural object:

$$\beta \mathbb{E}_1^{SP} \left[\phi \kappa_2 \frac{dq_2}{dn_2} \right] \tag{K.1}$$

and hence is not something that can be measured directly from the data. The pecuniary externality corresponds to a counterfactual exercise, that asks the question "by how much would the price of the collateral asset change if all financial intermediaries were to reduce their leverage exogenously before the crisis happens?" Quantitatively answering this question thus requires a calibration determining the value of each parameter, such as the strength of financial frictions ϕ, that controls the pecuniary externality.

One strategy used in the quantitative macroprudential literature, starting with the seminal work of Bianchi (2011) or more recently by Herreño and Rondón-Moreno (2020), calibrates the financial friction parameter ϕ combining (i) the rational expectation hypothesis, and (ii) a targeted moment on the probability or severity of financial crises.

To illustrate how behavioral forces might hinder this inference, I use a simplified version of the model where the collateral term is ignored in the pricing equation, and without risk.[112] I also assume that the stock of collateral assets H is exogenously fixed to streamline the exposition. Assume that we are aiming at calibrating our model such that a crisis provokes a price drop of X%. We can work through the rational equilibrium conditions to link the targeted moment X to the collateral parameter ϕ as:

$$\frac{1}{X} = 1 + \frac{Hz_3}{2 - \phi Hz_3} \tag{K.2}$$

which directly implies that a smaller ϕ (more stringent financial frictions) is needed to match larger asset price crashes. But a smaller ϕ directly implies a weaker sensitivity of the price with respect to net worth in period $t = 2$:

$$\frac{dq_2}{dn_2} = \frac{z_3}{1 - \phi Hz_3} \tag{K.3}$$

Intuitively, if financial frictions become extremely stringent, the borrowing capacity of the economy is at zero in period $t = 2$, and a change in net worth does not change this fact. Hence pecuniary externalities disappear when $\phi \to 0$. Calibrating the rational model to match more severe crises therefore automatically reduces the quantitative size of the inefficiencies.

In a behavioral model, however, parts of asset price crashes are attributable to swings in sentiment, and not only to binding collateral constraints.[113] This intuitively allows the calibration to match the same severity of crisis X% but with a higher value for the parameter ϕ, implicitly giving pecuniary externalities a greater weight.

I graphically illustrate these calibration issues in the case where sentiment is given by $\Omega_{t+1} = \alpha(q_t - q_{t-1})$, and I set q_0 such that there is initially irrational exuberance ($\Omega_2 > 0$). The left panel of Figure 20 presents the calibration step, and should be read from the y-axis to the x-axis. A modeler selects the severity of crisis observed in the data X and infer the value of ϕ. As we intuited earlier, for a given X the value of ϕ is greater in the extrapolative model. The right panel of Figure 20 then constructs the size of pecuniary externalities, by plugging the inferred value of X, read from the x-axis to the y-axis.

The parameters are deliberately chosen to feature small differences in the size of the pecuniary

[112]Note that, once again, the Ω-formulation allows me to flexibly work with behavioral biases in a riskless environment. Were one to decide to use a distorted probability measure instead, the task would prove to be more delicate.

[113]Swings in sentiment are also needed to match other moments which are defining features of financial crises: typically the behavior of credit spreads before crashes. Rational models with financial frictions, like Brunnermeier and Sannikov (2014) and He and Krishnamurthy (2019), cannot simultaneously generate elevated probability of crisis with decreasing credit spreads, a robust feature of the data (see Schularick and Taylor 2012 or López-Salido et al. 2017)

externality for a fixed ϕ. This exercise shows that these slight discrepancies might hide large differences when calibrated to the same moments. As can be seen from Figure 20, calibrating the model to $X = 77\%$ leads the rational model to estimate a pecuniary externality more than three times weaker than in the extrapolative model.

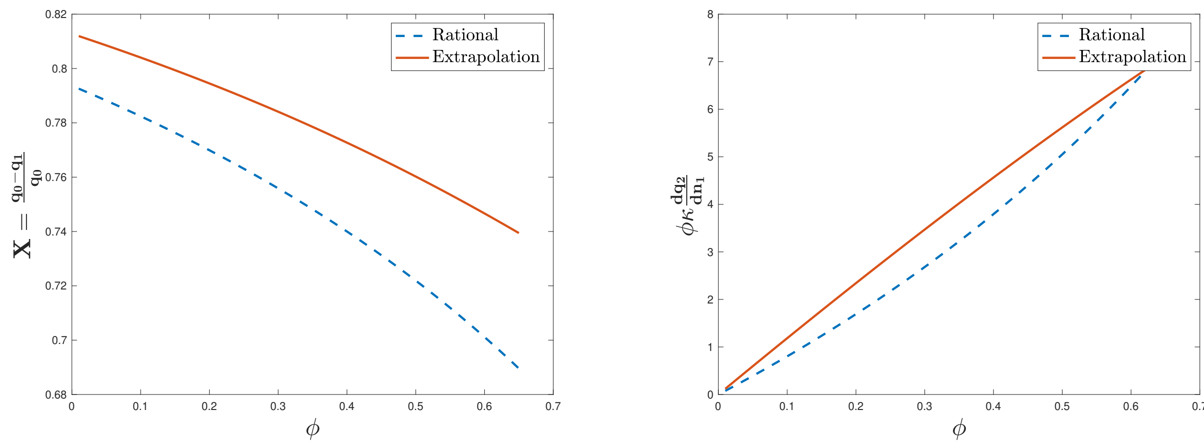

Figure 20: Calibration and Size of Pecuniary Externalities in the Rational and Extrapolative Cases. The behavioral bias is of the price extrapolation form, defined as $\Omega_{t+1} = \alpha(q_t - q_{t-1})$. q_0 is chosen such that $q_0 < q_1$ to feature initial exuberance.

L Multiple Equilibria

The analysis in the main paper as made under the assumption that the equilibrium was unique at $t = 2$ (see footnote 34). When sentiment is exogenous, the uniqueness of the equilibrium is straightforward to prove. It stems from the two equilibrium conditions:

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

$$c_2 = z_2 H - d_1(1 + r_1) + \phi H \mathbb{E}_2[z_3 + \Omega_3] \tag{L.2}$$

The second condition (coming from the budget constraint) directly pins down the consumption in equilibrium. This in turn directly pins down the asset price, and the equilibrium is unique.

This shows that multiple equilibria can arise only when sentiment depends on asset prices. This creates a feedback effect between prices and consumption, which can be strong enough to generate multiple equilibria. This is reminiscent of the literature on current-price collateral constraints: it is well known that financial amplification can lead to a multiplicity (see Jeanne and Korinek 2019; Schmitt-Grohé and Uribe 2021). Endogenous beliefs reintroduce this two-way feedback effect even in the future-price collateral constraint.[114]

[114] Khorrami and Mendo (2021) explore in general how this two-way feedback creates self-fulfilling fluctuations.

With endogenous biases, the system of equation becomes:

$$q_2 = \beta c_2 \mathbb{E}_2[z_3 + \Omega_3(q_2)] + \phi(1 - c_2)\mathbb{E}_2[z_3 + \Omega_3(q_2)] \tag{L.3}$$

$$c_2 = z_2 H - d_1(1 + r_1) + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)] \tag{L.4}$$

which makes it clear that, as long as Ω_3 is strictly increasing in q_2, different equilibrium levels of asset prices result in different equilibrium levels of consumption. The asset price determination is given by:

$$q_2 = \beta\left(n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)]\right)\mathbb{E}_2[z_3 + \Omega_3(q_2)] \\ + \phi(1 - (n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2)]))\mathbb{E}_2[z_3 + \Omega_3(q_2)] \tag{L.5}$$

Depending on the shape of $\Omega_3(q_2)$, an arbitrary number of equilibria are possible. I illustrate the problem with a linear function:

$$\Omega_3(q_2) = \alpha q_2 + \chi \tag{L.6}$$

The price condition is now:

$$q_2 = \beta\left(n_2 + \phi H \mathbb{E}_2[z_3 + \alpha q_2 + \chi]\right)\mathbb{E}_2[z_3 + \alpha q_2 + \chi] \\ + \phi(1 - (n_2 + \phi H \mathbb{E}_2[z_3 + \alpha q_2 + \chi]))\mathbb{E}_2[z_3 + \alpha q_2 + \chi] \tag{L.7}$$

This is a quadratic equation, hence will have at most two solutions. That means, however, that only one of them will be stable: since the consumption equation is linear in q_2, dc_2/dq_2 as computed along the pricing equation is necessarily below the slope of the budget constraint on one of the two equilibria. Figure 21 illustrates this instability. We can thus consider the case of unique equilibrium when sentiment is linear in prices. How more complicated forms of biases interact with frictions to create multiple equilibria is left for future work.

M Investment Microfoundations and LTV regulation

M.1 *H* as Housing

This section provides a concrete and simple example of microfoundations for the investment function, that highlights how LTV regulation impacts the model in practice.

There is a continuum of entrepreneurs, who are looking for funds to finance the construction of houses. Entrepreneurs are denoted by $j \in [0, \infty]$. Entrepreneurs are identical on all dimensions, expect the cost of their project. In particular, all entrepreneurs have the same net worth A, and their project is yielding the same stochastic payoffs Z_t in periods $t = 2$ and $t = 3$. An entrepreneur j must invest a total of I_j to complete its housing project. Entrepreneur j thus wants to raise $I_j - A$ of

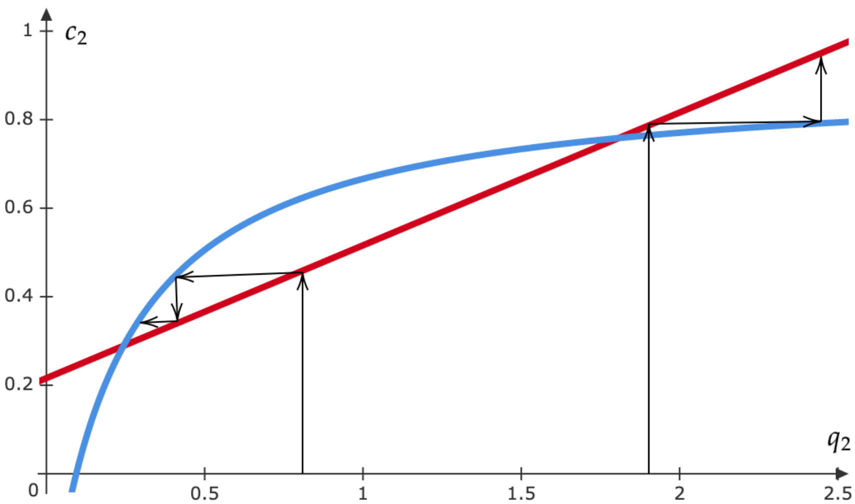

Figure 21: Graphical Illustration of Equilibrium Determination at $t = 2$ with sentiment linear in asset prices. The red line represents the budget constraint, and the blue line represents the pricing condition. The black arrows represent a tâtonnement process that starts at a given price. This price yields a certain level of sentiment and thus of consumption, which then gives rise to a different price, and so on. The right equilibrium is unstable, as the tâtonnement diverges to infinity on the right of the equilibrium, or converges to the stable equilibrium if the starting point is on the left.

outside funds, which financial intermediaries can provide.

At $t = 1$, once they obtained the funds, entrepreneurs can shirk to get private benefits of B next period. When entrepreneurs shirk, the housing project yields no payoff. Entrepreneurs are risk neutral and have no time discounting, and will thus exert effort only when their payoffs z'_t from the project are such that:

$$E_1[z'_2 + z'_3] \geq B \tag{M.1}$$

How the aggregate payoff Z_t is decomposed between z_t and z'_t is irrelevant here: the particular information and contracting frictions will give rise to an equilibrium z_t for the financial intermediaries, which are the payoffs are the risky asset that are used throughout the paper. The important take-away of this microfoundation is that the payoff z_t from project j does *not* depend on the amount $I_j - A$ and so does not depend on j. Payoffs of an individual project are thus fixed irrespective of the aggregate level of H.

Obviously, because of this specific structure, financial intermediaries will start by financing projects with low j since it requires a lower investment amount, but pays the same payoff. The cost of investing into H projects for the financial intermediary is thus:

$$c(H) = \int_0^H (I_j - A)dj \tag{M.2}$$

which is strictly convex in H as long as I_H is strictly increasing in H.

How is LTV regulation entering this problem? The marginal entrepreneur financed by intermediaries is borrowing $I_H - A$, for a total value of investment of I_H. The loan-to-value ratio is thus simply:

$$LTV_H = \frac{I_H - A}{I_H} \tag{M.3}$$

which is strictly increasing in H again, as long as I_H is strictly increasing in H. Therefore, by restricting LTV ratios to be below a certain amount, the regulator will forbid the financing of project by entrepreneurs above a limit $\bar{H}$. Setting an LTV regulation will directly control for the level of H in equilibrium.

Finally, note that I took the example of housing construction to make the model palatable. But a similar interpretation can be given about other types of activities financed by financial intermediaries, such as C&I loans. In this case, the policy instrument would not be LTV ratio regulation but rather "supervisory guidance:" the regulator would nudge intermediaries towards reducing their activities, therefore controlling H exactly like in the housing example.

M.2 H as Mortgage Loans

The collateral assets held by financial intermediaries can be interpreted as mortgage-backed securities, henceforth MBS. Collateralized mortgage obligations (CMO) and MBS still account for roughly 30% of the collateral assets used in repo markets (Securities and Exchange Commission 2021). During the 2007-2008 financial crisis, around 50% of Securities Lenders repo agreements were collateralized by agency securities (Krishnamurthy, Nagel and Orlov 2014). In this section I provide a simple model that microfounds this view, and show how LTV regulations are useful instruments when it comes to regulating the quantity of MBS held by banks, and connect the behavioral bias Ω to behavioral biases directly on house prices.

Setup I make several simplifying assumptions in order to adapt these micro-foundations to the baseline model presented in the paper, which I discuss at the end of this section. I draw on Brueckner (2000) standard model of mortgage default. Mortgage borrowers have a default cost of C, and a repayment of Z in the next period. If a mortgage borrower defaults on its loan, the financial intermediary seizes the house. House prices P next period are distributed according to a density function $F(P)$.

The mortgage borrower optimally defaults when:

$$C < B - P \tag{M.4}$$

since $P - B$ is housing equity. The expected payoff from the mortgage contract is thus:

$$z = \int_0^{B-C} Pf(P)dP + \int_{B-C}^{+\infty} Bf(P)dP. \tag{M.5}$$

The point of MBS is to pool many mortgage. contracts together. Consider for example the case where default costs are heterogenous (and unobserved by banks *ex ante*) and distributed uniformly in $[\underline{C}, \bar{C}]$. For a given price P, assuming that there is enough heterogeneity such that there are defaults and non-defaults for any P in the support of $f(P)$, the average payoff of a mortgage contract is thus:[115]

$$z(P) = \int_{\underline{C}}^{B-P} P \frac{dC}{\bar{C} - \underline{C}} + \int_{B-P}^{\bar{C}} B \frac{dC}{\bar{C} - \underline{C}} \tag{M.6}$$

which is simply equivalent to:

$$z(P) = \frac{P(B - P - \underline{C}) + B(\bar{C} - B + P)}{\bar{C} - \underline{C}} \tag{M.7}$$

$$\implies z(P) = \frac{B\bar{C} - P\underline{C} - (B - P)^2}{\bar{C} - \underline{C}}. \tag{M.8}$$

Because a MBS pools many different mortgages, this is the exact payoff of an MBS for given realization of P (by the law of large numbers).[116] Although not immediately obvious, this payoff is unambiguously increasing in house prices:

$$\frac{dz(P)}{dP} = \frac{2(B - P) - \underline{C}}{\bar{C} - \underline{C}} > 0. \tag{M.9}$$

Behavioral Bias Consider now the case where a financial intermediary has a behavioral bias, and believes that house prices will be $P + \omega$ instead of P. Assume, to stay within our assumptions, that the bias is such that there are still defaults as well as non-defaults expected in the pool:

$$\underline{C} < B - P - \omega < \bar{C} \iff B - P - \bar{C} < \omega < B - P - \underline{C} \tag{M.10}$$

The payoff of the MBS for a given price realization P becomes:

$$z(P + \omega) = z(P) + \frac{2\omega(B - P) - \omega^2}{\bar{C} - \underline{C}} \tag{M.11}$$

As can be seen from inspecting this equation, there is no directly relation between ω and Ω. Indeed, the size of the behavioral bias on the payoff depends on P, the underlying stochastic variable. The implicit correspondence, for Ω to be constant, is that ω varies with P and needs to verify:

$$\omega(P) = (B - P) - \sqrt{(B - P)^2 - \Omega(\bar{C} - \underline{C})} \tag{M.12}$$

But note that, to the first-order in the bias:

$$w(P) \approx \Omega \frac{\bar{C} - \underline{C}}{B - P} \tag{M.13}$$

[115] Specifically, for any P such that $f(P) \neq 0$, we have $\underline{C} < B - P < \bar{C}$.

[116] In other words, MBS fully diversify the risk associated with stochastic default costs.

which means that when agents are over-optimistic, there are more optimistic regarding the left-tail of the distribution than over the right-tail.

Discussion This model of mortgage loans was deliberately stylized in order to fit my baseline framework of the core paper. In particular, I kept the payoffs of the loan (and of the housing project in the previous section) constant even when H varies. In general, the risk premium asked by the intermediary, as well as the payments specified in a mortgage contract or when funding entrepreneurs, should depend on the characteristics of the borrower. A more general treatment of these issues, for example following the model of mortgage contracts developed by Campbell and Cocco (2015), is an interesting question left for future research. Second, when collateral assets are loans, like in the MBS case, their payoff profile is generally flat in good times. This implies that behavioral distortions will have different impacts depending on whether they apply to the left-tail or the right-tail of the distribution.[117]

N General Intertemporal Elasticity of Substitution

The paper made two assumptions on the utility function form of financial intermediaries: (i) log-utility in the first two periods, and (ii) linear utility in the last period. These assumptions were made for tractability, and to avoid over-complicating expressions without bringing any new intuition. In this section, I show that a model with a more general intertemporal elasticity of substitution (henceforth IES) delivers the exact same insights.

The utility function of banks is now given by:

$$U^b = \mathbb{E}_1 \left[\frac{c_1^{1-\sigma}}{1-\sigma} + \beta \frac{c_2^{1-\sigma}}{1-\sigma} + \beta^2 \frac{c_3^{1-\sigma}}{1-\sigma} \right] \tag{N.1}$$

where σ is the inverse of the IES. The equilibrium is now characterized by the Lagrange multiplier on the collateral constraint, κ, expressed as:

$$\kappa = \lambda_2 - \mathbb{E}_2[\lambda_3] \tag{N.2}$$

where the marginal utility is now given by:

$$\lambda_t = c_t^{-\sigma}. \tag{N.3}$$

The pricing equation at $t = 2$ is thus now slightly more complicated than before:

$$q_2 = \beta \mathbb{E}_2 \left[\frac{\lambda_3}{\lambda_2} (z_3 + \Omega_3) \right] + \phi \left(1 - \mathbb{E}_2 \left[\frac{\lambda_3}{\lambda_2} \right] \right) \mathbb{E}_2 \left[(z_3 + \Omega_3) \right] \tag{N.4}$$

[117] See Dávila and Walther (2021) for a related exploration of this issue.

However, it should be clear by now that the uninternalized welfare effects take exactly the same form I presented in Proposition 1 and 4. Why? The welfare of intermediaries at time $t = 2$ during a crisis can be written as:

$$\mathcal{W}_2 = \beta u\left(n_2 + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]\right) + \beta^2 u\left(\mathbb{E}_2[z_3]H - \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]/\beta\right) \tag{N.5}$$

with u the CRRA utility function and $n_2 = z_2 H - d_1(1 + r_1)$, while the Lagrangian corresponding to bankers' problem in period $t = 1$ is given by:

$$\mathcal{L}_{b,1} = \left[u(c_1) + \mathbb{E}_1[\mathcal{W}_2(n_2, H; q_2, z_2)]\right] - \lambda_1\left[c_1 + c(H) - d_1 - e_1\right] \tag{N.6}$$

the first-order condition on borrowing still gives:

$$\frac{\partial \mathcal{L}_{b,1}}{\partial d_1} = \lambda_1 - \mathbb{E}_1\left[\lambda_2\right] \tag{N.7}$$

where λ_t is the Lagrange multiplier on the budget constraint at time t. The social planner maximizes the same function, but under its own expectations, and by also taking into account how a change in d_1 impacts asset prices in period 2. This leads to the following first-order condition:

$$\frac{\partial \mathcal{L}_{b,1}^{SP}}{\partial d_1} = \lambda_1 + \mathbb{E}_1^{SP}\left[\lambda_2\right] - \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial \Omega_3}{\partial q_2} \frac{\partial q_2}{\partial n_2}\right] \frac{dn_2}{dd_1} \tag{N.8}$$

where the only difference is now that $\kappa_2 = \lambda_2 - \mathbb{E}_2[\lambda_3]$ instead of $\lambda_2 - 1$. Obviously, the same algebra ensures that Proposition 4 is in the same way still valid.

It is less obvious to sign the derivative $\partial q_2/\partial n_2$ in this general case. But inside a financial crisis, this sensitivity is unambiguously positive. Indeed, we have:

$$\frac{dc_2}{dn_2} = 1 + \phi H \frac{d\Omega_3}{dq_2}\frac{dq_2}{dn_2} \implies d\lambda_2 = -\sigma\left(1 + \phi H \frac{d\Omega_3}{dq_2} dq_2\right)\lambda_2^{\frac{\sigma+1}{\sigma}} \tag{N.9}$$

and

$$\frac{dc_3}{dn_2} = -\phi H \frac{d\Omega_3}{dq_2}\frac{dq_2}{dn_2}(1 + r_2) \implies d\lambda_3 = \sigma\left(\phi H \frac{d\Omega_3}{dq_2} dq_2 (1 + r_2)\right)\lambda_3^{\frac{\sigma+1}{\sigma}} \tag{N.10}$$

which implies that $d\lambda_3$ is of the sign as dq_2. In other words, the IES value (whether it is above or below 1) is irrelevant inside a crisis, because the amount of borrowing is fixed by the collateral constraint. In the case of an exogenous behavioral bias, the price sensitivity can be written:

$$dq_2 = \beta\sigma dc_2 c_2^{\sigma-1}\mathbb{E}_2[\lambda_3(z_3 + \Omega_3)] - \phi\sigma dc_2 \mathbb{E}_2[\lambda_3]\mathbb{E}_2[(z_3 + \Omega_3)] \tag{N.11}$$

which can be simplified as:

$$dq_2 = \beta\sigma dc_2 c_2^{\sigma-1}\Big((\beta - \phi)\mathbb{E}_2[\lambda_3(z_3 + \Omega_3)] + \phi Cov(\lambda_3, z_3)\Big) \tag{N.12}$$

dc_2 is obviously positive when the change is in net worth. Because of Assumption 1, the first term in the parentheses is positive. The second term, however, is negative.[118] While we can entertain the possibility that the covariance is strongly negative, this is not robust to changes in the micro-foundations of the collateral constraint. Indeed, if we assume that agents can default after observing the realization in z_3, the collateral constraint becomes of the form $d_2 \leq \phi H \min z_3$ and in this case the price sensitivity is:

$$dq_2 = \beta\sigma dc_2 c_2^{\sigma-1}\Big(\beta\mathbb{E}_2[\lambda_3(z_3+\Omega_3)] - \phi\mathbb{E}_2[\lambda_3](\min z_3 + \Omega_3)\Big) \tag{N.13}$$

which is unambiguously positive with Assumption 1. in this section, I thus only study the natural case where $dq_2/dn_2 > 0$.[119]

This calculation was made with a fixed Ω_3, but is still valid with an endogenous bias. Indeed, movements in Ω_3 only amplify this price sensitivity:

$$\begin{aligned} dq_2 = \beta\sigma dc_2 c_2^{\sigma-1}\Big((\beta-\phi)\mathbb{E}_2[\lambda_3(z_3+\Omega_3)] + \phi Cov(\lambda_3, z_3)\Big) \\ + \Big(\mathbb{E}_2[\frac{d\lambda_3}{\lambda_2}(z_3+\Omega_3)] - \phi\mathbb{E}_2[\frac{d\lambda_3}{\lambda_2}[\mathbb{E}_2[(z_3+\Omega_3)]]\Big) \\ + d\Omega_3\Big(\beta\mathbb{E}_2[\frac{\lambda_3}{\lambda_2}] + \phi(1-\mathbb{E}_2[\frac{\lambda_3}{\lambda_2}])\Big) \end{aligned} \tag{N.14}$$

where dc_2 also incorporates how the price in q_2 impact Ω_3 and thus the borrowing capacity. There is also a term (the second line) expressing how a change in sentiment brought by a change in asset prices impact future marginal utility, λ_3. Under the same condition as before, this term is also positive (similarly, it is only needed that ϕ is small enough, and this condition disappears under the alternative collateral formulation involving the minimum payoff).

Using the same welfare function:

$$\mathcal{W}_2 = \beta u\left(n_2 + \phi H\mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]\right) + \beta^2 u\left(\mathbb{E}_2[z_3]H - \phi H\mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)]/\beta\right), \tag{N.15}$$

the general formulation in Proposition 4 is also still valid:

$$\mathcal{W}_H = \left(\beta\mathbb{E}_1^{SP}\left[\lambda_2(z_2+q_2)\right] - \lambda_1 q_1\right) + \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\frac{d\Omega_3}{dq_2}\left(\frac{dq_2}{dn_2}z_2 + \frac{dq_2}{dH}\right)\right] \tag{N.16}$$

Here again, however, the sign of dq_2/dH is harder to determine without the linearity of utility in the ultimate period, since movements in H have effects on the future marginal utility. A first thing to notice is that even if for some levels of IES, dq_2/dH becomes negative, that is still unlikely to overturn the result that the collateral externality pushes towards under-investment. Indeed, as I

[118] This could not happen in the linear utility at time $t=3$, since then λ_3 was a constant.

[119] Dávila and Korinek (2018) also assume that the price of capital assets is increasing in the net worth of the financial sector.

just showed the first term of the collateral externality dq_2/dn_2 is positive. So dq_2/dH needs to be strongly negative to compensate for this effect. In other words, the linearity of utility at $t = 3$ or the log-utility at $t = 2$ are not directly responsible for this result: it is the assumption that $z_2 > 0$ (see Dávila and Korinek 2018 for examples where over-investment arises because dividends are negative in bad states of the world).

But in general, for the same reason dq_2/dn_2 is positive, this derivative will also be positive. Intuitively, dq_2/dH measures how an expansion of the borrowing capacity of financial intermediaries impact the equilibrium asset price. If dq_2/dn_2 is positive, we should expect the same thing for dq_2/dH: an increase in the borrowing capacity is similar to an increase in net worth during a financial crisis. Indeed, consider the case with an exogenous bias for intuition first (remember that these derivatives are keeping the net worth constant):

$$\frac{d\lambda_2}{dH} = -\sigma\phi\mathbb{E}_2[z_3 + \Omega_3]\lambda_2^{\frac{\sigma+1}{\sigma}} > 0 \tag{N.17}$$

$$\frac{d\lambda_3}{dH} = \sigma\phi\mathbb{E}_2[z_3 + \Omega_3](1 + r_2)\lambda_3^{\frac{\sigma+1}{\sigma}} < 0 \tag{N.18}$$

so that it is clear that the stochastic discount factor (λ_3/λ_2) is increasing in H. The price sensitivity can be expressed as always as:

$$dq_2 = \beta\mathbb{E}_2\left[d\frac{\lambda_3}{\lambda_2}(z_3 + \Omega_3)\right] - \phi\mathbb{E}_2\left[d\frac{\lambda_3}{\lambda_2}\right]\mathbb{E}_2\left[(z_3 + \Omega_3)\right] \tag{N.19}$$

which again will be negative only in the case where the covariance is strongly negative:

$$dq_2 = (\beta - \phi)\mathbb{E}_2\left[d\frac{\lambda_3}{\lambda_2}(z_3 + \Omega_3)\right] + \phi Cov(d\frac{\lambda_3}{\lambda_2}, z_3) \tag{N.20}$$

And, once again, this is not robust to alternative collateral constraints like $d_2 \leq \phi H \min[z_3 + \Omega_3]$. Lastly, this goes through with endogenous sentiment (as previously for net worth):

$$dq_2 = (\beta - \phi)\mathbb{E}_2\left[d\frac{\lambda_3}{\lambda_2}(z_3 + \Omega_3)\right] + \phi Cov(d\frac{\lambda_3}{\lambda_2}, z_3) + d\Omega_3\left((\beta - \phi)\mathbb{E}_2\left[\frac{\lambda_3}{\lambda_2}\right] + \phi\right) \tag{N.21}$$

where $d\frac{\lambda_3}{\lambda_2}$ now also incorporates how changes in sentiment affect the SDF. Using $d\Omega_3 = \frac{d\Omega_3}{dq_2}dq_2$, we see that the sign of dq_2 is unchanged, movements in sentiment are simply amplifying the previous sensitivity.

To conclude, the model with a general CRRA utility function across all three periods deliver the same uninternalized welfare effects as in the baseline case. This generality comes at the cost of greater complexity, without bringing anymore intuition. Derivatives are harder to express, and are of the opposite sign as in the baseline case only in extreme situations, that are not robust to small changes in the micro-foundations of the collateral constraint. Importantly, whether the IES is above or below 1 is not the driving force behind the sign of these derivatives inside financial crises. This is because inside a financial crises, there is no ambiguity that additional wealth will be allocated to current consumption rather than future consumption, independent of σ.

O Sophisticated Agents and Optimal Policy

This section provides results in the case where agents are sophisticated, and the social planner is subject to the same biases as private agents. Specifically, private agents now realize that their future selves will have a behavioral bias Ω_3, but are unaware that they are biased today. The planner holds the same beliefs.

Under these conditions, private agents and the planner are effectively maximizing the same welfare function (inside a crisis for brevity):

$$\mathcal{W}_2 = \beta \ln \left((z_2 + \Omega_2) H - d_1(1 + r_1) + \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)] \right) \\ + \beta^2 \left(\mathbb{E}_2[z_3] H - \phi H \mathbb{E}_2[z_3 + \Omega_3(q_2, q_1)] / \beta \right) \quad \text{(O.1)}$$

This expression conceals the intuition for sophisticated agents. First, agents believe that dividends are going to be at a level of $z_2 + \Omega_2$, thus are biased. They also take into account that their borrowing capacity is going to be affected by Ω_3, and that this future bias can depend on asset prices $\Omega_3(q_2, q_1)$. But they also know that the payoff of the asset itself is going to be z_3 and not $z_3 + \Omega_3$, hence the unbiased expectation in last-period consumption. The amount of debt they need to repay at $t = 3$, however, depends on the bias since it corresponds to the maximum amount permitted by the collateral constraint at $t = 2$, hence the $z_3 + \Omega_3(q_2, q_1)$ in the last position of this expression.

Since private agents and the planner are maximizing the same function, there are no behavioral wedges anymore.[120] Nevertheless, there are still uninternalized welfare effects working through sentiment and prices, exactly like pecuniary externalities:

$$\mathcal{W}_d = -\mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \frac{dq_2}{dn_2} \right] \quad \text{(O.2)}$$

$$\mathcal{W}_H = \beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_2}} \left(\frac{dq_2}{dn_2} z_2 + \frac{dq_2}{dH} \right) \right] \quad \text{(O.3)}$$

$$\mathcal{W}_q = \beta \mathbb{E}_1^{SP} \left[\kappa_2 \phi H \frac{\boldsymbol{d\Omega_3}}{\boldsymbol{dq_1}} \right] \quad \text{(O.4)}$$

While agents are aware that their leverage today will have a negative impact on asset prices which itself can aggravate pessimism, they cannot take it into account in their maximization program since private agents are infinitesimal. The same applies to their level of investment and the equilibrium

[120] Recall that behavioral wedges were quantifying the difference in expectations in Propositions 1 and 4, which are now 0.

price at $t = 1$.

This specific case exemplifies the robustness of my results. Even in a framework that does not feature any externality in a rational benchmark, even by having sophisticated agents, and even by having the planner sharing the same beliefs, these three externalities survive. Thus, even in this extreme case should the planner have an additional instrument in order to tame asset prices. Note that in this case, not only are equilibrium prices *unbiased*, they can even be lower than in a rational counterfactual. Indeed, agents factor in their expectation that they will be over-pessimistic in a crisis, which reduces q_1 since it worsens the severity of future crises. Finally note that, since there is no behavioral wedge, this case unambiguously calls for investment *subsidies*, not restrictions.

P State-Space Representation of Belief Distortions

The paper works with an analytically convenient formulation for belief distortions, where behavioral biases are represented as a shifter in the pricing equation:

$$q_1 = \mathbb{E}_1 \left[\frac{\lambda_2}{\lambda_1} \left(z_2 + \Omega_2 + q_2^r \right) \right] \tag{P.1}$$

where Ω_2 is a constant. This definition might seem unconventional, since a large part of the behavioral literature instead works with distortions in *probability density*: agents use a density $\tilde{\pi}$ instead of the objective density π.[121] This section presents the formal correspondence between the two concepts. To this end, the relationship between endogenous objects and the state of world z_2 is made explicit. In particular, the marginal utility of the financial sector at $t = 2$ is expressed as $\lambda_2(z; \Omega(z))$ to highlight how its equilibrium value depends on the realization of the state of the world, and of the future behavioral bias at $t = 2$, which itself depends on the state of the world.

P.1 Correspondence at $t = 2$

From $\tilde{\pi}$ to Ω: Behavioral agents set their expectations using a distorted probability density function $\tilde{\pi}$:

$$\frac{1}{\lambda_2} \int z_3 \tilde{\pi}_3(z_3) dz_3 \tag{P.2}$$

which implies that we can simply set Ω_3 such that:

$$\Omega_3 = \int (\tilde{\pi}(z_3) - \pi(z_3)) z_3 dz_3 \tag{P.3}$$

This does not mean, however, that all results are the same whether we are using an Ω or a distorted density. With a distorted density, the expressions for the first-order approximations for example do not hold, and similarly the expressions for the collateral and reversal externalities. This is the

[121] As in Bordalo et al. (2018), Caballero and Simsek (2020b) or Dávila and Walther (2021) among others.

reason why working with this Ω-formulation is convenient for welfare analysis.

From Ω to $\tilde{\pi}$: Since we are summarizing an entire function with a single scalar, there are an infinite number of ways to proceed. A convenient approach for exposition is to define z_3^m as the median of the stochastic process, and distort the density with a constant factor above the median:

$$\tilde{\pi}_3(z_3) = \begin{cases} \tilde{X}_1 \pi_3(z_3) & \text{if } z_3 \geq z_3^m \\ (2 - \tilde{X}_1)\pi_3(z_3) & \text{if } z_3 < z_3^m \end{cases} \tag{P.4}$$

where $\tilde{X}_1$ is defined as:

$$\tilde{X}_1 = \frac{\int_{z_3^m}^{\infty} \pi_3(z_3) z_3 dz_3 - \int_0^{z_3^m} \pi_3(z_3) z_3 dz_3 + \Omega_3}{\int_{z_3^m}^{\infty} \pi_3(z_3) z_3 dz_3 - \int_0^{z_3^m} \pi_3(z_3) z_3 dz_3} \tag{P.5}$$

which is greater than 1 as long as $\Omega_3 > 0$, which means that agents use a distorted density that exaggerates the probability of being above the median z_3^m.

P.2 Correspondence at $t = 1$

We now have to take into account the influence of the dependence of the SDF with respect to the state of the world.

From $\tilde{\pi}$ to Ω: A rational agent, recognizing the future biases of agents, would set its expectation of the discounted payoff of the asset as:

$$\int \left(\lambda_2(z_2; \Omega_3(z_2)) z_2 + z_3 + \Omega_3(z_2)\right) \pi_2(z_2) dz_2 \tag{P.6}$$

where Ω_3 can possibly be defined with distorted probabilities as shown above. As such, if behavioral agents discount the exact same payoffs state-by-state by use a distorted density $\tilde{\pi}_2$, we can implicitly define Ω_2 as[122]:

$$\int \left(\lambda_2(z_2 + \Omega_2; 0)(z_2 + \Omega_2) + z_3\right) \pi_2(z_2) dz_2 = \int \left(\lambda_2(z_2; \Omega_3(z_2)) z_2 + z_3 + \Omega_3(z_2)\right) \tilde{\pi}_2(z_2) dz_2 \tag{P.7}$$

From Ω to $\tilde{\pi}$: Given Ω_2, behavioral agents set their expectations according to:

$$\int \left(\lambda_2(z_2 + \Omega_2; 0)(z_2 + \Omega_2) + z_3\right) \pi_2(z_2) dz_2 \tag{P.8}$$

where π_2 is the objective probability density function of dividends at $t = 2$. This expression makes clear that agents do not take into account that their future selves might be subject to behavioral

[122] Ω_2 is uniquely defined as long as the discounted payoff of the asset is increasing in optimism, a natural condition we shall assume.

biases, represented by the 0 in $\lambda_2(z_2+\Omega_2;0)$. I similarly use the median z_2^m to construct the distorted probability measure correspondence:

$$\tilde{\pi}_3(z_2) = \begin{cases} \tilde{X}_2\pi_3(z_2) & \text{if } z_2 \geq z_2^m \\ (2-\tilde{X}_2)\pi_3(z_2) & \text{if } z_2 < z_2^m \end{cases} \tag{P.9}$$

where $\tilde{X}_2$ is defined as:

$$\tilde{X}_2 = \frac{\int_0^\infty (\lambda_2(z_2+\Omega_2;0)(z_2+\Omega_2)+z_3)\pi_2(z_2)dz_2 - 2\int_0^{z_2^m}(\lambda_2(z_2;\Omega_3(z_2))z_2+z_3+\Omega_3(z_2))\pi_2(z_2)dz_2}{\int_{z_2^m}^\infty(\lambda_2(z_2;\Omega_3(z_2))z_2+z_3+\Omega_3(z_2))\pi_2(z_2)dz_2 - \int_0^{z_2^m}(\lambda_2(z_2;\Omega_3(z_2))z_2+z_3+\Omega_3(z_2))\pi_2(z_2)dz_2} \tag{P.10}$$

This exercise highlights two advantages of the Ω notation. First, it makes the handling of endogenous variables (and their dependence, in expectations, to sentiment) more convenient. Second, it can span a larger set of behavioral biases than the distorted density: while the range of values attainable by the distorted density is restricted by the support of the exogenous process, there is no such constraint when adding directly a location shifter to the dividend distribution.

Q Proofs and Derivations for the Current-Price Collateral Constraint

Q.1 Full expressions for Appendix C.1

This part provides the expressions of Section C.1 once the collateral pricing part is taken into account.

Fundamental Extrapolation: When $\Omega_3 = \alpha(z_2 - z_1)$ the price is determined by the quadratic equation:

$$q_2 = \beta(n_2+\phi H q_2)\big(\mathbb{E}_2[z_3]+\alpha(z_2-z_1)\big) + \phi q_2(1-n_2-\phi H q_2) \tag{Q.1}$$

leading to the full expression:

$$q_2 = \frac{-(1-\beta\phi(\mathbb{E}_2[z_3]+\alpha(z_2-z_1))-\phi(1-n_2))}{2H\phi^2} + \frac{\sqrt{(1-\beta\phi(\mathbb{E}_2[z_3]+\alpha(z_2-z_1))-\phi(1-n_2))^2+4H\phi^2\beta n_2}}{2H\phi^2} \tag{Q.2}$$

the sensitivity of the price with respect to changes in net worth is given by:

$$\frac{dq_2}{dn_2} = \frac{\beta(\mathbb{E}_2[z_3] + \alpha(z_2 - z_1)) - \phi q_2}{1 - \beta\phi H(\mathbb{E}_2[z_3] + \alpha(z_2 - z_1)) - \phi(1 - n_2 - \phi^2 H q_2)} \tag{Q.3}$$

□

Price/Return Extrapolation: When $\Omega_3 = \alpha(q_2 - q_1)$ the quadratic equation becomes:

$$q_2 = \beta(n_2 + \phi H q_2)\big(\mathbb{E}_2[z_3] + \alpha(q_2 - q_1)\big) + \phi q_2(1 - n_2 - \phi H q_2) \tag{Q.4}$$

leading to the full expression:

$$q_2 = -\frac{(1 - \beta\phi(\mathbb{E}_2[z_3] - \alpha q_1) - \phi(1 - n_2))}{2(H\phi^2 - \beta H\phi\alpha)}$$

$$+ \frac{\sqrt{(1 - \beta\phi(\mathbb{E}_2[z_3] - \alpha q_1) - \phi(1 - n_2))^2 + 4\beta n_2 H(\mathbb{E}_2[z_3] - \alpha q_1)(\phi^2 H\phi\alpha)}}{2(H\phi^2 - \beta H\phi\alpha)} \tag{Q.5}$$

and the price sensitivity:

$$\frac{dq_2}{dn_2} = \frac{\beta(\mathbb{E}_2[z_3] - \alpha q_1) - \phi q_2}{1 - \beta\phi H(\mathbb{E}_2[z_3] - \alpha q_1) - \phi(1 - n_2 - \phi^2 H q_2) - \beta(n_2 + \phi H q_2)\alpha} \tag{Q.6}$$

□

Q.2 Proof of Proposition 16

At time $t = 2$, the welfare of borrowers can be written as:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(z_2 H - d_1(1 + r_1) + \phi H q_2\right) + \beta^2\left(\mathbb{E}[z_3]H - \phi H q_2/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta\mathbb{E}[z_3]H + z_2 H - d_1(1 + r_1)\right) & \text{otherwise} \end{cases} \tag{Q.7}$$

while the Lagrangian corresponding to bankers' problem in period $t = 1$ is given by:

$$\mathcal{L}_{b,1} = \left[u(c_1) + \mathbb{E}_1[\mathcal{W}_2(n_2, H; q_2, z_2)]\right] - \lambda_1\left[c_1 + c(H) - d_1 - e_1\right] \tag{Q.8}$$

the first-order condition on borrowing gives:

$$\frac{\partial\mathcal{L}_{b,1}}{\partial d_1} = \lambda_1 - \mathbb{E}_1\left[\lambda_2\right] \tag{Q.9}$$

The social planner maximizes the same function, but under its own expectations, and by also taking into account how a change in d_1 impacts asset prices in period 2. This leads to the following first-order condition:

$$\frac{\partial \mathcal{L}^{SP}_{b,1}}{\partial d_1} = \lambda_1 - \mathbb{E}^{SP}_1[\lambda_2] - \mathbb{E}^{SP}_1[\phi H \kappa_2 \frac{\partial q_2}{\partial n_2}] \tag{Q.10}$$

Hence simply by incorporating $\mathbb{E}_1[\lambda_2]$ we can express the total change in welfare as internalized plus uninternalized effects:

$$\frac{\partial \mathcal{L}^{SP}_{b,1}}{\partial d_1} = \underbrace{\lambda_1 - \mathbb{E}_1[\lambda_2]}_{\text{Internalized}} + \underbrace{\mathbb{E}_1[\lambda_2] - \mathbb{E}^{SP}_1[\lambda_2] - \mathbb{E}^{SP}_1[\phi H \kappa_2 \frac{\partial q_2}{\partial n_2}]}_{\text{Uninternalized}} \tag{Q.11}$$

which proves Proposition 16. □

Q.3 Proof of Proposition 17

I compute the difference between λ_2 expected by private agents and λ_2 expected by the Planner state by state z_2. When both expect a realization z_2 not to produce a financial crisis, marginal utilities are equalized to 1, so the difference disappears. For the rest there are two cases: either both marginal utilities correspond to binding collateral constraints, either one agent expect the friction to bind and the other not. The first case yields:

$$\frac{1}{c_2(z_2+\Omega_2,0)} - \frac{1}{c_2(z_2,\Omega_3)} = \\ \frac{1}{(z_2+\Omega_2)H - d_1(1+r_1) + \phi H q_2(z_2+\Omega_2;0)} - \frac{1}{z_2 H - d_1(1+r_1) + \phi H q_2(z_2;\Omega_3)} \tag{Q.12}$$

I take the first-order approximation around the REE $\lambda_2 = 1/(z_2 H - d_1(1+r_1) + \phi H q_2(z_2;0)) = 1/c_2(z_2,0)$. It gives:

$$\begin{aligned}\frac{1}{(z_2+\Omega_2)H - d_1(1+r_1) + \phi H q_2(z_2+\Omega_2;0)} &= \frac{1}{c_2(z_2,0)} \frac{1}{1 + \frac{\Omega_2 H + \phi \Omega_2 \frac{dq_2}{n_2}}{c_2(z_2,0)}} \\ &= \lambda_2 \left(1 - \frac{\Omega_2 H + \phi \Omega_2 \frac{dq_2}{n_2}}{c_2(z_2,0)}\right)\end{aligned} \tag{Q.13}$$

While the same algebra for the second part of equation (Q.12) yields similarly:

$$\frac{1}{z_2 H - d_1(1+r_1) + \phi H q_2(z_2;\Omega_3)} = \frac{1}{c_2(z_2,0)} \frac{1}{1 + \frac{\phi \Omega_3 H \frac{dq_2}{z_3}}{c_2(z_2,0)}} = \lambda_2 \left(1 + \frac{\phi H \Omega_3 \frac{dq_2}{z_3}}{c_2(z_2,0)}\right) \tag{Q.14}$$

Taking the difference gives:

$$\frac{1}{c_2(z_2+\Omega_2,0)} - \frac{1}{c_2(z_2,\Omega_3)} = \lambda_2^2 \left(H\Omega_2 + \phi \frac{dq_2}{dn_2}\Omega_2 - \phi H \frac{dq_2}{dz_3}\Omega_3 \right) \tag{Q.15}$$

Lastly we need to consider the cases where the social planner and private agents disagree about the occurrence of a crisis for a given z_2. Without loss of generality, I assume that private agents are over-optimistic so for some range of states, $[z^* - dz, z^*]$ they expect to be at $c_2 = 1$, while the Planner expects the collateral constraint to be binding (where z^* is the crisis cutoff in the RE case). The size of the band is infinitesimal since, as can be seen in equations (32) and (33), the cutoff is only moving because of Ω_2 and Ω_3 which are small.

The difference, integrated on the band, can be expressed through a triangle approximation:

$$\int_{z^*-dz}^{z^*} \left(1 - \frac{1}{c_2(z_2,\Omega_3)}\right) \pi(z_2) dz_2 = \frac{dz\pi(z^*)}{2} \left(1 - \frac{1}{c_2(z^*-dz,\Omega_3)}\right) \tag{Q.16}$$

Because the difference between $t = 1$ and $1/c_2(z^* - dz, \Omega_3^*)$, where Ω_3^* is the bias at the cutoff, is also infinitesimal, this term is negligible compared to the previous one.[123]It thus follows that, to the first order:

$$\mathcal{B}_d = \Omega_2 \mathbb{E}_1^{SP} \left[\lambda_2^2 \left(H\Omega_2 + \phi \frac{dq_2}{dn_2} \right) \right] - \phi H \mathbb{E}_1^{SP} \left[\Omega_3 \lambda_2^2 \frac{dq_2}{dz_3} \right] \tag{Q.17}$$

□

Q.4 Proof of Proposition 18

The asset price is determined in a crisis equilibrium by:

$$q_2 = \beta \left(n_2 + \phi H q_2\right) \mathbb{E}[z_3 + \Omega_3] + \phi q_2 (1 - n_2 - \phi H q_2) \tag{Q.18}$$

A total differential yields:

$$\begin{aligned} dq_2 = \beta dn_2 \mathbb{E}[z_3+\Omega_3] + \beta\phi H dq_2 \mathbb{E}[z_3+\Omega_3] + \beta\left(n_2+\phi H q_2\right) d\Omega_3 \\ + \phi dq_2(1-n_2-\phi H q_2) + \phi q_2(-dn_2 - \phi H dq_2) \end{aligned} \tag{Q.19}$$

Because the variation in Ω_3 can only come from q_2 by assumption, rearranging gives:

[123] For completeness, its value can be approximated as:

$$\int_{z^*-dz}^{z^*} \left(1 - \frac{1}{c_2(z_2,\Omega_3)}\right) \pi(z_2) dz_2 \approx -(\Omega_2 - \phi\Omega_3(z^*)) \frac{(\Omega_2 - \phi\Omega_3(z^*))\left(1+\phi\frac{dq_2}{dn_2}\right) - \phi\Omega_3(z^*)}{2} \pi(z^*)$$

Ω_2 enters this equation because it parametrizes the value of dz, i.e. the size of the band where agents do not expect a financial crisis but the planner does.

$$\frac{dq_2}{dn_2} = \frac{\beta \mathbb{E}[z_3 + \Omega_3] - \phi q_2}{1 - \beta\phi H \mathbb{E}[z_3 + \Omega_3] + 2\phi^2 H q_2 - \beta c_2 \frac{d\Omega_3}{dq_2}} \tag{Q.20}$$

□

Q.5 Collateral Externality Perturbation

We can perform a perturbation analysis around the REE equilibrium to gain intuition about how the collateral externality is changed by sentiment. Let us elaborate on the difference with a social planner that would entirely respect the beliefs of private agents. I develop the Taylor expansion of the difference between the two expectations, where $(dq_1/dn_2)^e$ is the price sensitivity in the rational world, and similarly κ_2^e for the Lagrange multiplier (both are defined state-by-state).

$$\begin{aligned}\mathbb{E}_1^{SP}\left[\phi\kappa_2 \frac{dq_2}{dn_2}\right] - \mathbb{E}_1\left[\phi\kappa_2 \frac{dq_2}{dn_2}\right] &= -\mathbb{E}_1\left[\left(\frac{dq_1}{dn_2}\right)^e \mathcal{B}_d\right] \\ &+ \mathbb{E}_1\left[\kappa_2^e \frac{\mathbf{\Omega_3}}{1 - \phi H(\mathbb{E}_2[z_3]) + \phi^2 H q_2 - c_2 \frac{d\mathbf{\Omega_3}}{dq_2}}\left(1 + \frac{z_3 - \phi q_2}{1 - \phi H(\mathbb{E}_2[z_3]) + \phi^2 H q_2 - c_2 \frac{d\mathbf{\Omega_3}}{dq_2}}\right)\right]\end{aligned} \tag{Q.21}$$

This expression shows that the difference is increasing when the behavioral wedge $\mathcal{B}_d$ becomes more negative. Future pessimism reduces this difference, while a higher sensitivity of future sentiment with respect to asset prices increases it. Finally, note that Ω_2 is not part of the second term of this expansion. This is because, in my specific modeling framework with log utility, the price sensitivity in the rational benchmark is constant with respect to net worth.

□

Q.6 Proof of Proposition 19

At time $t = 2$, the welfare of borrowers can be written as:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(z_2 H - d_1(1 + r_1) + \phi H q_2\right) + \beta^2 \left(\mathbb{E}[z_3]H - \phi H q_2/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}[z_3]H + z_2 H - d_1(1 + r_1)\right) & \text{otherwise} \end{cases} \tag{Q.22}$$

while the Lagrangian corresponding to bankers' problem in period $t = 1$ is given by:

$$\mathcal{L}_{b,1} = \left[u(c_1) + \mathbb{E}_1[\mathcal{W}_2(n_2, H; q_2, z_2)]\right] - \lambda_1\left[c_1 + c(H) - d_1 - e_1\right] \tag{Q.23}$$

the first-order condition on investment yields:

$$\frac{\partial \mathcal{L}_{b,1}}{\partial H} = \lambda_1 c'(H) - \mathbb{E}_1\left[\lambda_2(z_2 + \Omega_2 + q_2(z_2 + \Omega_2))\right] \tag{Q.24}$$

The social planner maximizes the same function, but under its own expectations, and by also taking into account how a change in d_1 impacts asset prices in period 2 (recall that $q1$ is fixed by assumption). This leads to the following first-order condition:

$$\frac{\partial \mathcal{L}^{SP}_{b,1}}{\partial H} = \lambda_1 c'(H) - \beta \mathbb{E}^{SP}_1 [\lambda_2(z_2 + q_2)] - \beta \mathbb{E}^{SP}_1 [\kappa_2 \phi H \left(\frac{\partial q_2}{\partial n_2} z_2 + \frac{\partial q_2}{\partial H} \right)] - \beta \mathbb{E}^{SP}_1 [\kappa_2 \phi H \frac{dq_2}{dq_1} \frac{dq_1}{dH}] \tag{Q.25}$$

where the last part quantifies how a change in price today impacts the aggregate borrowing capacity of the financial sector. In most models, this term is zero since $dq_2/dq_2 = 0$: there is no reason a change in price today should directly change the price tomorrow. But in the case where sentiment Ω_3, which enters the determination of prices at period 2, depends on past prices, this derivative is not zero anymore.

Proposition 19 is then proved once we notice that $q_1 = c'(H)$ in equilibrium, so that $dq_1/dH = c''(H)$, while:

$$\frac{dq_2}{dq_1} = \frac{dq_2}{d\Omega_3} \frac{d\Omega_3}{dq_1} \tag{Q.26}$$

which yields the final formula for the uninternalized effects of marginally increasing investment:

$$\begin{aligned} \frac{\partial \mathcal{L}^{SP}_{b,1}}{\partial H} &= \underbrace{\lambda_1 q_1 - \beta \mathbb{E}_1 [\lambda_2(z_2 + q_2)]}_{\text{Internalized}} + \\ &\underbrace{\beta \mathbb{E}_1 [\lambda_2(z_2 + q_2)] - \beta \mathbb{E}^{SP}_1 [\lambda_2(z_2 + q_2)] - \beta \mathbb{E}^{SP}_1 \left[\kappa_2 \phi H \left(\frac{\partial q_2}{\partial n_2} z_2 + \frac{dq_2}{dH} \right) \right] - \beta \mathbb{E}^{SP}_1 \left[\kappa_2 \phi H \frac{\partial q_2}{\partial \Omega_3} \frac{\partial \Omega_3}{\partial q_1} c''(H)\right]}_{\text{Uninternalized}}. \end{aligned} \tag{Q.27}$$

□

Q.7 Proof of Proposition 20

I use the same notation as for the proof of Proposition 17, presented in Online Appendix Q.3. The behavioral wedge for investment can consequently be expressed state-by-state as:

$$\mathcal{B}_H(z_2) = [\lambda_2(0; \Omega_3)(z_2 + q_2(0; \Omega_3))] - [\lambda_2(\Omega_2; 0)(z_2 + \Omega_2 + q_2(\Omega_2; 0)] \tag{Q.28}$$

As for leverage, it is sufficient to only look at states where the borrowing constraint binds both in the expectation of the social planner and of private agents. To the first-order, we can write:

$$\mathcal{B}_H(z_2) = (\lambda_2(0; \Omega_3) - \lambda_2(\Omega_2; 0)(z_2 + q^r_2)) + \lambda^r_2 \left(\Omega_3 \frac{dq_2}{dz_3} - \Omega_2 \left(1 + \frac{dq_2}{dz_2} \right) \right) \tag{Q.29}$$

The part $\lambda_2(0;\Omega_3) - \lambda_2(\Omega_2;0)$ exactly corresponds to the behavioral wedge for leverage state-by-state, that we will denote by $\mathcal{B}_d(z_2)$ for conciseness. The behavioral wedge for investment can thus be expressed as:

$$\mathcal{B}_H(z_2) = \mathbb{E}_1^{SP}\left[\mathcal{B}_d(z_2)(z_2+q_2^r)\right] - \Omega_2\mathbb{E}_1^{SP}\left[\lambda_2^r\left(1+\frac{dq_2}{dz_2}\right)\right] + \mathbb{E}_1^{SP}\left[\lambda_2^r\Omega_3\frac{dq_2}{dz_3}\right] \tag{Q.30}$$

where

$$\mathcal{B}_d(z_2) = \Omega_2\lambda_2^2\left(H\Omega_2 + \phi\frac{dq_2}{dn_2}\right) - \phi H\Omega_3\lambda_2^2\frac{dq_2}{dz_3}. \tag{Q.31}$$

□

Q.8 Derivation of Equation C.20

I proceed as for the derivation of the price sensitivity to swings in sentiment, Proposition 18, as in Online Appendix Q.4. I start from the equilibrium condition that links the asset price at time $t=2$ to consumption through the collateral constraint:

$$q_2 = \beta\left(n_2 + \phi H q_2\right)\mathbb{E}[z_3+\Omega_3] + \phi q_2(1-n_2-\phi H q_2) \tag{Q.32}$$

I then differentiate with respect to H, acknowledging that q_2 and Ω_3 will be modified as a result:

$$\begin{aligned} dq_2 = \beta\phi q_2 dH\mathbb{E}[z_3+\Omega_3] + \beta\phi H dq_2\mathbb{E}[z_3+\Omega_3] + \beta\left(n_2+\phi H q_2\right)d\Omega_3 \\ + \phi dq_2(1-n_2-\phi H q_2) + \phi q_2(-\phi q_2 dH - \phi H dq_2) \end{aligned} \tag{Q.33}$$

Rearranging gives the desired result:

$$\frac{dq_2}{dH} = \frac{\beta\phi q_2\mathbb{E}_2\left[z_3+\mathbf{\Omega_3}\right] - \phi^2 q_2^2}{1-\beta\phi H(\mathbb{E}_2\left[z_3+\mathbf{\Omega_3}\right]) + 2\phi^2 H q_2 - \beta c_2\frac{d\mathbf{\Omega_3}}{dq_2}}. \tag{Q.34}$$

□

Q.9 Proof of Proposition 21

Using equation (Q.22), the derivative of total welfare with respect to changing asset prices at $t=1$ is:

$$\frac{\partial\mathcal{W}_1}{\partial q_1} = \beta\mathbb{E}_1^{SP}\left[\kappa_2\phi H\frac{dq_2}{dq_1}\right] \tag{Q.35}$$

In most models, this term is zero since $dq_2/dq_1 = 0$: there is no reason a change in price today should directly change the price tomorrow. But in the case where sentiment Ω_3, which enters the determination of prices at period 2, depends on past prices, this derivative is not zero anymore.

Proposition 21 is then proved once we notice that

$$\frac{dq_2}{dq_1} = \frac{dq_2}{d\Omega_3}\frac{d\Omega_3}{dq_1} \tag{Q.36}$$

which yields the final formula for the welfare effects of marginally changing asset prices:

$$\mathcal{W}_q = \beta \mathbb{E}_1^{SP}\left[\kappa_2 \phi H \frac{\partial q_2}{\partial \Omega_3}\frac{\partial \Omega_3}{\partial q_1}\right] \tag{Q.37}$$

□

Q.10 Proof of Proposition 10 for the Current-Price Collateral Constraint

As explained in the main text, the social planner's optimality condition under the premises of Proposition 10 can be expressed as:

$$u'(c_1) = \frac{1}{2\sigma_\Omega}\int_0^\infty \left[\int_{-\sigma_\Omega}^{\sigma_\Omega} \frac{\partial \mathcal{W}_2}{\partial n_2}\left(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \omega_2\right) d\omega_2\right] f_2(z_2) dz_2. \tag{Q.38}$$

Key to this proposition is the shape of $\partial \mathcal{W}_2/\partial n_2$ with respect to z_2. First recall that:

$$\mathcal{W}_2 = \begin{cases} \beta \ln\left(z_2 H - d_1(1+r_1) + \phi H q_2\right) + \beta^2\left(\mathbb{E}[z_3]H - \phi H q_2/\beta\right) & \text{if } z_2 \geq z^* \\ \beta\left(\beta \mathbb{E}[z_3]H + z_2 H - d_1(1+r_1)\right) & \text{otherwise} \end{cases} \tag{Q.39}$$

so that the first derivative is equal to:

$$\frac{\partial \mathcal{W}_2}{\partial n_2} = \begin{cases} \beta\left(1 + \phi H \frac{dq_2}{dn_2}\right)\lambda_2 - \beta \phi H \frac{dq_2}{dn_2} & \text{if } z_2 \geq z^* \\ \beta & \text{otherwise} \end{cases} \tag{Q.40}$$

which is constant outside of a crisis, as expected. More important is the behavior of this derivative inside of crises, which can be rewritten:

$$\frac{\partial \mathcal{W}_2}{\partial n_2} = \frac{\beta\left(1 + \phi H \frac{dq_2}{dn_2}\right)}{z_2 H - d_1(1+r_1) + \phi H q_2} - \beta \phi H \frac{dq_2}{dn_2} \quad \text{if } z_2 \geq z^* \tag{Q.41}$$

or for convenience:

$$\frac{\partial \mathcal{W}_2}{\partial n_2} = \lambda_2 + \phi H(\lambda_2 - 1)\frac{dq_2}{dn_2} \quad \text{if } z_2 \geq z^*. \tag{Q.42}$$

I use the following notation to simplify the exposition of the proof. First, the expectation over z_2 for a given w_2 is denoted by:

$$g(w_2) = \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \omega_2) f_2(z_2) dz_2 \tag{Q.43}$$

while the integral taken over the uncertainty band is:

$$G(\sigma_\Omega) = \int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{g(w_2)}{2\sigma_\Omega} dw_2. \tag{Q.44}$$

Given the continuity of $\partial\mathcal{W}_2/\partial n_2$ (see equation Q.39) we can differentiate with respect to σ_Ω:

$$\begin{aligned} G'(\sigma_\Omega) = -\frac{1}{2\sigma_\Omega^2} \int_{-\sigma_\Omega}^{+\sigma_\Omega} \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \omega_2) f_2(z_2) dz_2 dw_2 + \\ \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; q_2, z_2 - \bar{\Omega}_2 - \sigma_\Omega) f_2(z_2) dz_2 - \int_0^{+\infty} \frac{\partial \mathcal{W}_2}{\partial n_2}(d_1, H; q_2, z_2 - \bar{\Omega}_2 + \sigma_\Omega) f_2(z_2) dz_2 \end{aligned} \tag{Q.45}$$

which can be expressed in terms of the notation just defined above as:

$$G'(\sigma_\Omega) = -\frac{G(\sigma_\Omega)}{\sigma_\Omega} + \frac{1}{2\sigma_\Omega}(g(\sigma_\Omega) - g(-\sigma_\Omega)) \tag{Q.46}$$

Before proceeding further, remember that the social planner optimally sets leverage such that:

$$u'(c_1) = G(\sigma_\Omega) \tag{Q.47}$$

while the decentralized equilibrium is independent of σ_Ω. Thus, leverage restrictions will be increasing in σ_Ω if and only if G is increasing in σ_Ω. This condition is then equivalent, using the derivative just computed, to:

$$\frac{g(\sigma_\Omega) - g(-\sigma_\Omega)}{2} > \int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{g(w_2)}{2\sigma_\Omega} dw_2. \tag{Q.48}$$

Since $\partial\mathcal{W}_2/\partial n_2$ is continuous in z and in ω_2, and since ω_2 is defined in the compact set $[-\sigma_\Omega, \sigma_\Omega]$, g is continuous (by continuity of parametric integrals) and Fubini's theorem implies that a sufficient condition for $G'(\sigma_\Omega) > 0$ is that[124]:

$$\frac{1}{2}\left(\frac{\partial \mathcal{W}_2}{\partial n_2}(z_2 + \sigma_\Omega) - \frac{\partial \mathcal{W}_2}{\partial n_2}(z_2 - \sigma_\Omega)\right) > \int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{\partial \mathcal{W}_2}{\partial n_2}(z_2 + \omega_2) \frac{d\omega_2}{2\sigma_\Omega} \quad \forall z_2 \in supp(f_2). \tag{Q.49}$$

In other words, this condition requires that the average taken over a segment is below the average of the two extreme points of this same segment.

Next, notice that any convex function satisfies this requirement. For a convex function φ, Jensen's inequality yields:

$$\varphi(t\sigma_\Omega - (1-t)\sigma_\Omega) \leq t\varphi(\sigma_\Omega) + (1-t)\varphi(-\sigma_\Omega) \quad \forall t \in [0,1]. \tag{Q.50}$$

[124] $\bar{\Omega}_2$ does not need to appear in this condition since this inequality is required to hold for all z_2 in the support of the definition, so equivalently for all $z_2 - \bar{\Omega}_2$ also in the support.

Now integrate this inequality over t to get:

$$\int_0^1 \varphi(t\sigma_\Omega - (1-t)\sigma_\Omega)dt \leq \int_0^1 t\varphi(\sigma_\Omega)dt + \int_0^1 (1-t)\varphi(-\sigma_\Omega)dt. \tag{Q.51}$$

A change of variable $t \to (x - \sigma_\Omega)/(2\sigma_\Omega)$ in the left-hand side thus yields:

$$\int_{-\sigma_\Omega}^{+\sigma_\Omega} \frac{\varphi(x)}{2\sigma_\Omega}dx \leq \frac{\varphi(\sigma_\Omega) - \varphi(-\sigma_\Omega)}{2} \tag{Q.52}$$

which is exactly the relationship in equation (Q.49).

We now have to prove that $\partial\mathcal{W}_2/\partial n_2$ is convex to end the proof of Proposition 10. Going back to equation (Q.39), denote $\partial\mathcal{W}_2/\partial n_2$ by $\mathcal{W}_{2,n}$. Start with the derivative of marginal utility. We have:

$$\frac{d\lambda_2}{dz_2} = -\frac{H + \phi H \frac{dq_2}{dz_2}}{c_2^2} \tag{Q.53}$$

and so:

$$\frac{d^2\lambda_2}{dz_2^2} = -\frac{2}{c_2^3}\left(H + \phi H \frac{dq_2}{dz_2}\right) \tag{Q.54}$$

and notice that we will have the following for the derivative of consumption given the log-utility assumption:

$$\frac{dc_2}{dz_2} = H + \phi H \frac{dq_2}{dz_2} = -\frac{d\lambda_2}{dz_2}c_2^2 \tag{Q.55}$$

Turning now to the whole marginal welfare derivative:[125]

$$\frac{d\mathcal{W}_{2,n}}{dz_2} = \frac{d\lambda_2}{dz_2} - \phi H \frac{dc_2}{dz_2}\frac{dq_2}{dn_2} + \phi H(1-c_2)\frac{d^2 q_2}{dz_2 dn_2}. \tag{Q.56}$$

Finally we can inspect the second derivative of $\mathcal{W}_{2,n}$ to sign it. Basic algebra yields:

$$\begin{aligned}\frac{d^2\mathcal{W}_{2,n}}{dz_2^2} = &-\left[-\frac{2}{c_2^3}\left(H + \phi H\frac{dq_2}{dz_2}\right) + \phi H \frac{d^2q_2}{dn_2 dz_2}\right]\frac{dc_2}{dz_2} \\ &- \left[\frac{1}{c_2^2} + \phi H \frac{dq_2}{dn_2}\right]\phi H \frac{d^2 q_2}{dn_2 dz_2} - \left[H + \phi H \frac{dq_2}{dz_n}\right]\phi H \frac{d^2 q_2}{dn_2 dz_2} \\ &+ \phi H(1-c_2)\frac{d^3 q_2}{dn_2 d^2 z_2}\end{aligned} \tag{Q.57}$$

Inspecting the signs of the different terms, one can notice that $c_2 > 0$, $dq_2/dn_2 > 0$, $dc_2/dz_2 > 0$, and $1 - c_2 > 0$ since we are in a financial crisis (or equivalently, $\kappa_2 > 0$). This directly implies that a sufficient condition (but far form necessary) for this second derivative to be positive (and hence

[125] The behavior of λ_2 would be enough to characterize how leverage should move with sentiment uncertainty if we were to look at an infinitesimal agent. But the social planner takes pecuniary externalities into account, so we need to also compute the derivatives of the pecuniary externalities.

the function of interest convex) is that both $d^2q_2/(dn_2dz_2) < 0$ and $d^3q_2/(dn_2d^2z_2) > 0$. We can now conclude this proof by computing these two objects. First the second derivative of the price in a financial crisis:[126]

$$\frac{dq_2^2}{dn_2dz_2} \propto -\phi\frac{dq_2}{dn_2}\left((1-\beta\phi H\mathbb{E}_1[z_3]+2\phi^2Hq_2)+2\phi H(\beta\mathbb{E}_1[z_3]-\phi q_2)\right). \tag{Q.58}$$

and both terms inside the large parentheses are positive: they correspond to the denominator and numerator of the sensitivity dq_2/dn_2. The coefficient of proportionality is positive since it corresponds to the square of the denominator of this exact same sensitivity. Hence we unambiguously have:

$$\frac{dq_2^2}{dn_2dz_2} < 0 \tag{Q.59}$$

Now we are left with the the third derivative of the price function inside a financial crisis. Rather than directly compute its expression – which is extremely involved – a possible short-cut is instead to write the price sensitivity of the form:

$$\frac{dq_2}{dn_2} = b(z_2) = \frac{\beta z_3 - \phi q_2}{D(z_2)} > 0. \tag{Q.60}$$

This short notation implies that deriving this expression yields, using the computation just above:

$$b'(z_2) = -\phi\frac{dq_2}{dz_2}\frac{D(z_2)+2\phi H(\beta z_3-\phi q_2)}{D^2(z_2)} < 0 \tag{Q.61}$$

which can be rewritten for convenience as:

$$b'(z_2) = -\frac{\phi\frac{dq_2}{dz_2}}{D} - \phi\frac{dq_2}{dz_2}\frac{2\phi Hb(z_2))}{D(z_2)} < 0 \tag{Q.62}$$

since $(\beta z_3 - \phi q_2)/D^2 = b/D$. This eases the computation of the next derivative, yielding:

$$b''(z_2) = -\frac{\phi\frac{d^2q_2}{dz_2^2}}{D(z_2)} + \phi\frac{\frac{dq_2}{dz_2}D'(z_2)}{D^2(z_2)} - \phi\frac{\phi\frac{d^2q_2}{dz_2^2}}{D(z_2)}2\phi Hb(z_2) + \phi\frac{\phi\frac{dq_2}{dz_2}}{D(z_2)}2\phi Hb(z_2)D'(z_2) - \phi\frac{\phi\frac{dq_2}{dz_2}}{D(z_2)}2\phi Hb'(z_2) \tag{Q.63}$$

and all terms are positive, since $dq_2/dz_2 > 0$, $d^2q_2/dz_2^2 < 0$, $D'(z_2) > 0$ and $b'(z_2) < 0$.[127]

Summing up, $d^2q_2/(dn_2dz_2) < 0$ and $d^3q_2/(dn_2d^2z_2) > 0$. This implies that $d^2\mathcal{W}_{2,n}/dz_2^2$ is a convex function. This leads to G' being positive, thus the right-hand side of equation (Q.38)

[126] For the sake of brevity, Ω_3 is left ou of the expression as, by assumption, it is a constant. It thus only shifts the value of $\mathbb{E}_1[z_3]$ and that has no impact on the sign of these derivaties as long as $\mathbb{E}_1[z_3] + \Omega_3 > 0$, which we always assume to be the case.

[127] Indeed, $D'(z_2) \propto dq/dn > 0$.

to be increasing in σ_Ω. Since the optimality condition of private agents is independent of σ_Ω, this equivalently means that optimal leverage restrictions are increasing in σ_Ω. This concludes the proof. $\square$