Credit Crunches, Information Failures, and the Persistence of Pessimism*

Ludwig Straub **Robert Ulbricht**[†]

[†]Job Market Paper

January 9, 2013

Abstract

This paper examines how financial crises affect the ability of agents to learn about economic fundamentals, and how this in turn affects the transmission of financial shocks through the economy. To this end, we introduce a model where noise in the financial market drives business cycles. Agents endogenously learn about fundamentals from market prices, but financial constraints systematically destroy the informational capacity of prices in financial crises. This is because financially constrained agents stop responding to available information, reducing the efficiency of prices in aggregating information that is dispersed across the economy. As a result, times of financial crisis are marked by both endogenously increasing uncertainty and increasingly persistent pessimism, providing a powerful amplification mechanism for financial shocks. Importantly, this mechanism is inherently nonlinear. Whereas small or positive financial shocks have only a little influence on the economy, unusually adverse shocks virtually shut down market learning and result in *dis*-proportionately severe and persistent crashes—characterized by substantial losses in employment, output, and asset prices; and high levels of uncertainty, volatility, and risk premia.

Keywords: Credit crises, endogenous uncertainty, financial frictions, heterogeneous information, asymmetric and nonlinear business cycles.

JEL Classification: D83, E32, E44, G01.

*We would like to thank Matthias Doepke, Klaus M. Schmidt, and Mirko Wiederholt for extensive feedback and guidance throughout this project. We are also grateful to Klaus Adam, George-Marios Angeletos, Marco Bassetto, Markus K. Brunnermeier, Larry Christiano, Martin Eichenbaum, Christian Hellwig, Guido Lorenzoni, Alessandro Pavan, Giorgio Primiceri, Aleh Tsyvinski, and seminar participants at Munich and Northwestern Universities for many helpful comments and discussions. This research started while the second author was visiting Northwestern; he thanks the economics department for their hospitality. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged. *Email Addresses:* straub@mit.edu, robert.ulbricht@lrz.uni-muenchen.de.

1 Introduction

Observers of the recent financial crisis often emphasize the role of uncertainty for the transmission and amplification of financial shocks. Particularly, a widespread idea is that increasingly uncertain business conditions are a key factor for the persistence of the crisis. For instance, IMF chief economist Olivier Blanchard argues that: "(Financial) crises feed uncertainty. And uncertainty affects behavior, which feeds the crisis. Were a magic wand to remove uncertainty ... the crisis would largely go away."[1]

Understanding these ideas requires to think about uncertainty as being endogenous to the state of the economy. In this paper, we examine how financial distress affects the ability of agents to learn about economic fundamentals, and how this in turn affects the transmission of financial shocks through the economy. Specifically, we study a dynamic macroeconomic model where agents learn about economic fundamentals from market prices. The presence of financial frictions endogenously determines the efficiency of the pricing mechanism in aggregating available information and thereby governs uncertainty among agents.

The model highlights a novel mechanism that explains the characteristic persistence of financial crises. In contrast to other types of crises, *financially* constrained firms cannot step up their investments when they become more confident in their business outlooks. As a result, in states of financial distress, real business activity does *not* reflect actual business conditions, leaving market observers uncertain about the state of the economy. With high uncertainty feeding back into the financial market, this perpetuates financial distress and creates a persistent cycle of uncertainty and financial constraints. As the paper shows, this feedback loop has important implications for the behavior of financial markets and the behavior of the production sector in response to financial shocks.

Preview of the model The analysis is based on a stylized two-sector economy. In a production sector, entrepreneurs produce a single consumption good, using labor provided by workers as the only input factor; and in a financial sector, both workers and entrepreneurs trade an asset whose returns correlate with the economy's average productivity (the exogenous fundamental of the economy).

The model is built on two assumptions. First, fluctuations in the financial market affect the real sector via financial frictions. Specifically, we assume that entrepreneurs must borrow in order to pay their workers and that, as in Kiyotaki and Moore (1997), all debt must be secured with collateral. Having entrepreneurs use the financial asset as collateral, this creates a cap on hiring which tightens as asset prices tumble. Second, we assume that agents cannot observe the economy's average productivity directly. Instead they endogenously learn about productivity by observing the market-clearing prices in both sectors, which aggregate further exogenously available information that is dispersed across the economy.

[1] This quote is taken from a guest article written by Olivier Blanchard for The Economist, January 31, 2009.

Results The results in this paper are driven by the interaction of informational and financial frictions. This combination leads adverse financial shocks to systematically destroy the real economy's "informational capacities" (i.e., less can be learned from observing the production sector). The reason is that when entrepreneurs become financially constrained, they cease to respond to available information. Wages—or, any other production-based source of information[2]—therefore becomes less efficient in aggregating information that is dispersed across the production sector. This reduces overall learning and increases uncertainty among agents in times of financial distress.

The endogenous nature of the real economy's informational capacities gives rise to two *information*-based mechanisms that amplify adverse financial shocks.[3] To see this, note that when less can be learned from today's production sector, agents place more weight on other sources of information. In particular, agents' opinions are more affected by information contained in asset prices as they tumble. This amplifies pessimism among agents, tightens financial constraints even more, and creates a harmful feedback loop. At the same time, with little to learn from today's economy, agents are also affected more by prior information in forming their opinions. This causes pessimism on the financial market to become inherently persistent, thereby increases future financial distress, and in turn inhibits future learning. Metaphorically speaking, the financially constrained economy gets stuck in a "pessimism trap".

Importantly, both of these mechanisms are inextricably tied to the endogenous nature of uncertainty—were one to remove uncertainty during a financial crisis (using the "magic wand" imagined by Olivier Blanchard), the crisis would indeed largely go away. Moreover, while the bulk of the literature on financial frictions discusses amplification mechanisms that are symmetric,[4] the *information*-based amplification mechanisms just described are inherently asymmetric. That is, whereas adverse financial shocks inhibit learning and have an amplified and persistent impact on the economy, positive financial shocks improve learning and have a *de*-amplified and *non*-persistent impact.

More generally, we find that the more negative a financial shock, the greater are both the amplification and the persistence of the shock. In particular, while the amplification and persistence of small shocks is negligible, rare adverse ("tail") shocks virtually destroy the real economy's informational capacities and entail highly amplified and persistent crashes. This non-linearity (or convexity) provides a novel theoretical explanation for why during normal times the day-to-day fluctuations in the financial market appear to have only a small impact on the real economy, whereas

[2]In our model, the only statistic that is (directly) affected by entrepreneurs' production choices are wages. In a richer model, however, the efficiency in aggregating information would not only be reduced for wages, but for all observable prices (or quantities) that vary with the economy's production.

[3]To sharpen our results, we abstract from any risk-related mechanisms or any other mechanisms that directly translate increases in uncertainty into real effects. However, including any such mechanism only amplifies our findings, which we illustrate in Section 7.3 where we extend the model to study the effects of risk-aversion.

[4]This is, for instance, true for the seminal contributions by Kiyotaki and Moore (1997) and Bernanke, Gertler and Gilchrist (1999), and most of the literature thereafter. An important exception are Brunnermeier and Sannikov (2012), who consider an economy with a financial sector, in which shocks have similar nonlinear effects. See the literature review for details.

unusually adverse financial shocks propagate persistently throughout the whole economy.

Taken together, the aforementioned results imply that financial crises are characterized by (i) amplified and (ii) persistent losses in output, employment, and asset prices; (iii) as well as high uncertainty. In addition, the collapse in the informational capacities explains three further key characteristics of financial crises, namely: (iv) highly diverse views on the state and fate of the economy; (v) volatile asset prices; and (vi) large risk premia. Highly diverse views result from agents placing more weight on private sources of information—which are inherently diverse—when less can be learned from today's production sector. At the same time, as discussed above, agents also pay more attention to asset prices. As in any rational expectations equilibrium, this implies that asset prices become more exposed to noisy demand shocks within the financial market, increasing volatility. Lastly, when agents are risk-averse, risk premia on asset prices naturally rise as uncertainty increases.

Methodological contribution The information loss in our model is inextricably tied to endogenous learning governed by *nonlinear* laws of motion.[5] That is, agents learn about the fundamental not only through linear signals but also through nonlinear ones. More specifically, we show that, in virtue of financial frictions, real sector prices (in the presence of noisy demand) are informationally equivalent to observing a perturbed, *concave* function of the fundamental. The slope of the function is decreasing in the "constrainedness" of the economy. In a general theorem, we then prove that "well-behaved" concave signals generally result in higher uncertainty when the signal realizes in flatter regions. This theorem applies to a large class of information structures and holds independent of the specifics of our model.

One technical challenge in analyzing the dynamic properties of our model is that nonlinear Gaussian signal structures generally do not pair with conjugate prior distributions. To address this problem, we develop a *quasi*-Gaussian framework, departing from the assumption that the small additive noise terms included in the nonlinear signals are normally distributed. In particular, we construct the noise terms in such a way that the nonlinear signals behave *as if* they were linear normal signals with a *state-dependent* signal precision. Within this framework, our general theorem then maps every state of the economy to a unique signal precision, which is decreasing in the economy's constrainedness. Because, in the limit of signals becoming linear, the state-dependent signal precision becomes state-*in*dependent, quasi-Gaussian signals can be understood as a natural extension of the standard linear Gaussian framework to the case of nonlinear signals.

Related literature At a methodological level, the two building blocks of our model relate this paper to two strands of modern macroeconomics. First, there is a large literature on financial frictions that demonstrates how small shocks can get amplified through the financial system. In particular,

[5] Mertens (2011) and Hassan and Mertens (2011) also analyze a model with heterogeneous information and nonlinear laws of motion. However, in their paper, the information structure is such that nonlinearities are transformed away and agents actually update according to linear prices, avoiding endogenous uncertainty.

our formalization of credit constraints is based on Kiyotaki and Moore (1997).[6] We contribute to this literature by identifying a new *informational* role of financial frictions that complements the *constraining* role known from previous works. Throughout the paper, we highlight the consequences of this new informational mechanism by comparing our model to the counterfactual case where the constraining role of credit constraints remains intact, but all informational effects of frictions are shut down. As with the majority of the financial frictions literature, we find that in this counterfactual case, shocks *symmetrically* and *linearly* affect the economy. That is, both adverse and positive shocks are amplified through financial frictions in exactly the same way. In contrast, our findings that amplification is asymmetric and nonlinear is a novel feature of the *information*-based mechanism introduced in this paper.[7]

Second, our paper closely connects to an emerging literature on heterogeneous information in macroeconomics and finance (see, e.g., Morris and Shin, 2002 and Woodford, 2003).[8] From a methodological perspective, we contribute to this literature by showing how learning from nonlinear signals gives rise to endogenous uncertainty and by embedding nonlinear signals in a conjugate prior framework. From an applied perspective, our paper is similar to Angeletos, Lorenzoni and Pavan (2010) who also study how learning from the real sector affects the financial market, but do not consider how in the presence of financial constraints the financial market feeds back to the information aggregation and how that causes learning to collapse during financial crises. Perhaps closest to our model is a framework by La'O (2010), which also combines informational with financial frictions. However, because La'O resolves all dispersion of information at the time agents learn from financially constrained markets, posterior uncertainty in her model is completely determined by the exogenous amount of information available to the economy, ruling out the informational mechanism that drives our results.[9]

At a more applied level, our paper also relates to a recent literature following Bloom (2009) that puts forth the idea of uncertainty-driven business cycles resulting from *exogenous* uncertainty shocks.[10] Our approach relates to these works in two ways. First, we provide a microfoundation for

[6]More recent studies based on credit constraints include, e.g., Krishnamurthy (2003); Iacoviello (2005); Kiyotaki and Moore (2008); Gertler and Karadi (2011); and Gertler and Kiyotaki (2011). Important contributions that are based on other financial frictions include, e.g., Bernanke and Gertler (1989); Carlstrom and Fuerst (1997); and Bernanke, Gertler and Gilchrist (1996, 1999), which are based on Townsend's (1979) costly state verification approach; and Kurlat (2010); Bigio (2011, 2012); and Boissay, Collard and Smets (2012), which consider frictions originating in adverse selection. See Appendix B.2 for a discussion how the ideas developed in this paper can be applied to the costly-state-verification and adverse selection approaches.

[7]An important exception are Brunnermeier and Sannikov (2012), who consider an economy with a financial sector, in which shocks have similar nonlinear effects. While the findings of Brunnermeier and Sannikov are similar in spirit, their mechanism is, however, not. In particular, information in their model is perfect and uncertainty is constant over time.

[8]More recent contributions to the dispersed information literature with a macroeconmic focus include Adam (2007); Angeletos and Pavan (2004, 2007, 2009); Amato and Shin (2006); Morris and Shin (2006); Amador and Weill (2008); Lorenzoni (2009, 2010); Hellwig and Veldkamp (2009); Hassan and Mertens (2011); Goldstein, Ozdenoren and Yuan (2011); and Angeletos and La'O (2012*a*,*b*).

[9]We also differ from La'O (2010) in that we focus on business cycle dynamics, whereas La'O studies the (static) composition of output and price volatility in fundamental and noise shocks in a single period model.

[10]See also Sim (2008); Bachmann and Bayer (2009); and Bloom et al. (2012).

why uncertainty increases specifically during times of financial distress, which is also when empirical measures of uncertainty are highest. An important insight from our microfoundation is that the endogenous nature of uncertainty unleashes a powerful feedback loop, which is absent in business cycles that are driven by exogenous uncertainty shocks. Therefore, in contrast to Bloom, who finds that uncertainty shocks give rise to rapid drops and rebounds in economic activity, we find that high uncertainty goes along with amplified and persistent crises. Second, the literature on exogenous uncertainty shocks complements our findings in that it discusses a number of additional channels, absent in our model, by which increases in uncertainty may propagate through the economy. In particular, Christiano, Motto and Rostagno (2009) and Gilchrist, Sim and Zakrajšek (2010) illustrate how fluctuations in uncertainty are amplified through a combination of risk-aversion and financial frictions and have strong effects on the real sector.[11]

Finally, our finding that financial frictions destroy information relates to a small and closely related literature that studies endogenous fluctuations in uncertainty. Van Nieuwerburgh and Veldkamp (2006) explore the idea that learning about total factor productivity is slow in recessions when total business activity is low. The reason is that if output is perturbed by an additive noise term, then this noise term contributes relatively more to output when output is low, leading to higher uncertainty during recessions (see also, Veldkamp, 2005 and Ordoñez, 2010). The mechanism studied in this paper differs from the mechanisms in these papers in that the efficiency of learning is governed by the degree to which the economy is constrained rather than the level of output. As outlined above, this difference leads to a number of important implications for the transmission of financial shocks. Apart from this, our paper also differs in that it considers learning from price signals and provides an explicit foundation for why prices vary in their informational content in times of financial distress. This approach is shared with two related contributions by Yuan (2005) and Albagli (2011). However, both papers focus on one-shot financial market settings and do not include a real sector, preventing them from analyzing the transmission of financial shocks through the economy, which is at the core of our contribution.[12]

Outline The plan for the rest of the paper is as follows. The next section introduces the model economy. Section 3 examines how financial frictions affect the ability of agents to learn from market prices. Section 4 characterizes the full equilibrium. Section 5 then explores how shocks are transmitted through the economy. Section 6 illustrates our theoretical results with a numerical example. Section 7 points out some further empirical predictions, and Section 9 concludes.

[11] While we abstract from risk-related transmission mechanism of uncertainty in our model to sharpen our results, these mechanisms are clearly important and strongly amplify the role of the key mechanism identified in this paper. For a illustration, see Section 7.3 where we consider an extension of our model to the case where agents are risk averse.

[12] Another paper sharing the broad theme is Bachmann and Moscarini (2011), which looks at a mechanism that increases the cross-sectional dispersion of beliefs during crises, but in which posterior uncertainty remains constant. Also, there is a growing literature on rational inattention, which is based on the idea that learning is costly, effectively leading agents to endogenously pick their desired level of uncertainty (see, e.g., Sims, 2003, 2006; Maćkowiak and Wiederholt, 2009, 2010; and Woodford, 2009).

2 The model

Our model is based on two ingredients: financial frictions and endogenous learning. In the interest of analytical tractability, we make a number of simplifying assumptions. In particular, we focus on labor as a single input good, so that learning from the real sector takes the form of extracting information from the market clearing wage. Nonetheless, our analysis can be applied to any other price that varies with entrepreneurs' optimal production choices, and is meant to more generally capture the idea that agents learn about business conditions by observing the real sector. Also, we focus on a single, stylized financial friction to model spillovers from the financial market. However, while the model heavily rests on the constraining role of asset prices, it is irrelevant by which financial friction this is explained (see Appendix B.2).

Preferences and technologies Consider a discrete time, infinite horizon economy with a continuum 1 of risk-neutral, one-period lived agents.[13] A proportion m of each generation's agents are farmers, while the remaining $1-m$ are gatherers. With a slight abuse of notation, we use $\mathcal{F}$ and $\mathcal{G}$ to denote the set of farmers and gatherers at a given date, respectively. Both farmers and gatherers consume a single consumption good, a perishable fruit, which is produced using two distinct technologies. First, there is an entirely exogenous production unit. In reference to Lucas (1978) it is helpful to think about this unit as a "tree" (or asset), which bears a random number $\tilde{A}_t$ of fruits and comes in a total supply of 1, equally distributed across each generation. Second, farmers have access to a "field" which transforms labor input n_{it} into additional fruits. The production function for field work is given by

$$F(\tilde{A}_{it}, n_{it}) = \tilde{A}_{it} \log(n_{it}),$$

where $\tilde{A}_{it}$ is an idiosyncratic random productivity parameter of farmer $i \in \mathcal{F}$ at date t.

For simplicity, farmers are excluded from doing fieldwork themselves, but may employ gatherers for the purpose of cultivating their fields. Gatherer i's disutility of working is given by $v : \mathbb{R}_+ \to \mathbb{R}_+$, a twice differentiable, increasing, and strictly convex function, with $v'(0) = 0$ and $\lim_{n\to\infty} v'(n) = \infty$. Gatherer i thus wishes to maximize the quantity

$$\mathbb{E}\left\{\tilde{c}_{it} - v(n_{it}) | \mathcal{I}_{it}\right\}, \tag{1}$$

and farmer i wishes to maximize

$$\mathbb{E}\left\{\tilde{c}_{it} | \mathcal{I}_{it}\right\}, \tag{2}$$

where $\tilde{c}_{it}$ represents consumption of fruits, and $\mathbb{E}\left\{\cdot | \mathcal{I}_{it}\right\}$ is an expectations operator given information set $\mathcal{I}_{it}$.[14]

[13] Agents in our model are one-period lived to induce a common prior among agents at all times. This ensures that heterogeneously informed agents in our model do not run into Townsend's (1983) infinite regress problem, allowing us to derive all our results in an analytic fashion.

[14] Throughout, we differentiate between stochastic variables and their realizations by accentuating the stochastic version with a tilde ("~").

Field productivities $\{\tilde{A}_{it} : i \in \mathcal{F}\}$ are taken to be lognormally distributed, so that $\log(\tilde{A}_{it}) \equiv \tilde{\theta}_{it}$ has a normal distribution with mean $\tilde{\theta}_t$ and variance $1/\tau_\xi$, and where the average log productivity $\tilde{\theta}_t$ follows a first-order autoregressive process:

$$\tilde{\theta}_t = \rho\tilde{\theta}_{t-1} + \tilde{\epsilon}_t,$$

where $\tilde{\epsilon}_t$ is Gaussian noise with variance $1/\tau_\epsilon$. The dividend from fruit trees is assumed to be positively correlated with the average productivity and is given by

$$\log(\tilde{A}_t) = \gamma_0 + \gamma_1\tilde{\theta}_t + \tilde{u}_t,$$

where $(\gamma_0, \gamma_1) \in \mathbb{R} \times \mathbb{R}_+$ and $\tilde{u}_t$ is an independent (of $\tilde{\theta}_t$) random variable that possibly introduces additional noise to dividend payments.

Markets and credit constraints There are two types of markets operating at date t. First, a competitive labor market matches demand and supply for field work and determines the market clearing wage w_t. Second, a competitive stock market determines ownership of fruit trees and pins down an asset price q_t. Shares on trees are assumed to be perfectly divisible and entitle its owners to claim all $\tilde{A}_t$ fruits falling from the corresponding tree. In both markets, current period consumption serves as the unit of account. Furthermore, we simplify the analysis by ruling out margin trading and short selling of trees, effectively restricting asset holdings of agent i to $0 \leq x_{it} \leq 1$.[15]

We now describe the financial friction in our economy. Following Kiyotaki and Moore (1997), we assume that farmers lack the means to commit to paying their wage debt after production is sunk. As a consequence, gatherers refuse to do field work unless they are provided a security by farmers in exchange for their labor.[16] We assume that fruits harvested from fields are nontradeable, so that fields itself cannot be used as an security.[17] Instead, farmers may use trees as collateral to pay gatherers. Also, we simplify the problem of how to account the value of collateral by assuming that the asset market operates at least twice: A first time parallel to the labor market, ensuring that

[15] This specification is adopted from Albagli, Hellwig and Tsyvinski (2011) and in combination with our assumptions on noisy asset demand (see below), it keeps the law of motion of asset prices tractable within a conjugate prior framework. Further note, that these no-borrowing constraints are consistent with the lack of commitment power that we impose as a key friction on the labor market.

[16] Our assumption that farmers cannot commit to paying their wage bill is based on theoretical arguments developed by Hart and Moore (1994, 1998). In their 1994 paper, such commitment problem arises from the possibility to renegotiate wages at any point during the production process. Accordingly, if farmers are *indispensable* for reaping the benefits of field work (e.g., because fruits harvested on fields are nontransferable, as we assume in our setting), then the outside option of gatherers is reduced to the value of collateral, and farmers could renegotiate a smaller wage whenever the wage bill exceeds the value of collateral. Alternatively, if, as in their 1998 paper, farmers can "run away" after production is harvested, gatherers are likewise left with only the value of collateralized assets. In both cases, in anticipation of the moral hazard, gatherers will not accept an outstanding wage debt that exceeds the value of collateral.

[17] This assumption is again based on Hart and Moore's (1994) rationale for why farmers cannot commit to paying their wages in the first place: If farmers are indispensable for operating field production, then fields are naturally worthless to gatherers (see also, Footnote 16).

all information that is possibly aggregated by trading trees is already available when the value of collateral is determined; and a second time after production is realized and wages are paid, so that the value of collateral is based on the current market price q_t for trees. These assumptions jointly imply that the wage debt of a farmer i at date t is bounded from above by the market value of his asset holdings $x_{it}q_t$, constraining labor demand to satisfy

$$n_{it} \leq (q_t/w_t)\, x_{it}. \tag{3}$$

Information The average productivity takes the role of the "fundamental" in our economy. More generally, $\tilde{\theta}_t$ is meant to reflect the "profitability" of investments, comprising, e.g., technology shocks and aggregated business conditions. Agents base their expectations about $\tilde{\theta}_t$ on the information

$$\mathcal{I}_{it} = \{s_{it}\} \cup \{w_s, q_s\}_{s=1}^{t},$$

which, in addition to the publicly observable history of prices $\{w_s, q_s\}_{s=1}^{t}$, contains a private signal $\tilde{s}_{it}$, which reveals the true average productivity $\tilde{\theta}_t$ perturbed by some independent Gaussian noise $\tilde{\xi}_{it}$ with variance $1/\tau_\xi$:

$$\tilde{s}_{it} = \tilde{\theta}_t + \tilde{\xi}_{it}.$$

For simplicity, $\{\tilde{s}_{it}\}_{i\in\mathcal{F}}$ is assumed to be perfectly correlated with farmers' productivities $\{\tilde{\theta}_{it}\}_{i\in\mathcal{F}}$, so that by learning the realization of $\tilde{s}_{it}$ a farmer also learns the productivity of his field.

Furthermore, to prevent $\tilde{\theta}_t$ from being perfectly revealed by the market, prices w_t and q_t are perturbed by noise traders with stochastic asset demand $\Phi(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu))$ and labor demand $\Psi_{\theta_t,q_t}(\tilde{\omega}_t)$. Here $\tilde{\eta}_t$ and $\tilde{\omega}_t$ are independent Gaussian noise with variances $1/\tau_\eta$ and $1/\tau_\omega$, Φ is the cumulative standard normal distribution, μ is a constant which we conveniently set to offset the risk related components contained in $\tilde{q}_t$,[18] and $\Psi_{\theta_t,q_t} : \mathbb{R} \to \mathbb{R}$ is a function that transforms $\tilde{\omega}_t$ into a random variable $\tilde{\Psi}_t$ which may depend on the realizations of $\tilde{\theta}_t$ and $\tilde{q}_t$. Noisy asset demand $\Phi\left(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)\right)$ is divided between the two occurrences of the asset market in a fixed ratio of $1-m$ to m.

Note that to illustrate the precise conditions for which our main theorem holds, we deliberately keep Ψ_{θ_t,q_t} as general as possible for now, restricted only by the assumptions below. In addition, this generality, later grants us the freedom that is necessary to generalize the standard conjugate Gaussian framework and to extend it to the case of nonlinear learning.

[18] Although agents are risk-neutral in the model economy, the lognormal distribution of dividends implies that the asset price $\tilde{q}_t$ behaves as if agents were risk-seeking with respect to the fundamental $\tilde{\theta}_t$. By setting μ to $\gamma_1/2$ times the cross-sectional information dispersion (i.e., $\mu = \gamma_1/(2\tau_\xi)$), the risk-discount is exactly offset by the bias in noise traders' demand, yielding an asset price that behaves as if agents were risk-neutral with respect to $\tilde{\theta}_t$ and noise traders were unbiased. Also note that by transforming $\tilde{\eta}_t$ into $\Phi(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)) \in [0, 1]$, noise traders' demand matches the support of endogenous asset demand, ensuring the existence of a market clearing price. The unbiased version of this specification (with $\mu = 0$) is adopted from Albagli, Hellwig and Tsyvinski (2011) and keeps the law of motion of asset prices tractable within a conjugate prior framework (see also, Footnote 15).

Distributional assumptions As it will be seen, one key feature of this model is that the amount of information that is aggregated from the real sector through the labor market varies with the state of the economy. To define the conditions for which our characterization of learning holds, it is convenient to first introduce normalized versions of labor demand and supply, χ_s^d and χ_t^s, which describe them relative to the upper bound on labor as given by equation (3),[19]

$$\chi_t^d = \log\left(\frac{m\int_{\mathcal{F}} n_{it}\,\mathrm{d}i}{q_t/w_t}\right) \quad \text{and} \quad \chi_t^s = \log\left(\frac{(1-m)\int_{\mathcal{G}} n_{it}\,\mathrm{d}i}{q_t/w_t}\right);$$

and define

$$\chi_t^m = \log\left(\frac{\exp(\theta_t)/w_t}{q_t/w_t}\right) = \theta_t - \log(q_t),$$

which corresponds to the unconstrained relative demand of the median-productivity farmer. Intuitively, χ_t^d measures the fraction of farmers operating at their collateral constraint and provides a useful proxy for the constrainedness of the economy. Note that, even after observing w_t and q_t, $\tilde{\chi}_t^d|(w_t, q_t)$ is a nondegenerate random number that, via optimal labor demand, depends on $\tilde{\theta}_t$. In the next section, it will be seen that forming the posterior $\tilde{\chi}_t^d|\chi_t^s$ conveniently summarizes the information that can be extracted from the labor market.

To develop an intuition for the conditions under which our main results hold, it is helpful to define them in terms of the posterior $\tilde{\chi}_t^d|\chi_t^s$. Would there be no market noise, then labor demand would equal labor supply and we would have that $\tilde{\chi}_t^d|\chi_t^s = \chi_t^s$. Accordingly, the properties of the random variable $\tilde{\chi}_t^d|\chi_t^s$ reflect how the market noise $\tilde{\Psi}_t$ correlates with the state of the economy. In the following, we impose two restrictions on these properties. Importantly, even though $\tilde{\chi}_t^d|\chi_t^s$ arises endogenously, it is possible to map these restrictions on $\tilde{\chi}_t^d|\chi_t^s$ back into assumptions about the exogenous noise term $\tilde{\Psi}_t$ (for details, see Appendix B.3).[20]

With this in mind, we impose the following key restriction:[21]

Property 1. $\mathrm{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$ *is constant for all* $\chi_t^s \in \mathbb{R}$.

Property 1 ensures that the amount of information about the normalized labor demand $\tilde{\chi}_t^d$ that is contained in the market clearing wage $\tilde{w}_t$ is constant throughout all states of the economy. In Proposition 1, we show that this specification is equivalent to requiring that, in the absence of credit constraints, uncertainty about $\tilde{\theta}_t$ behaves exactly like in a standard economy where it is constant over time. This ensures that there is *no* time dependency of uncertainty inherent to the stochastic process $\tilde{\Psi}_t$, so that *any* variation of uncertainty will be the result of credit constraints.

Additionally, we shall also require the following regularity condition:

[19]Here we anticipate that in equilibrium all farmers will hold $x_{it} = 1$ assets at the time the labor market operates, so that the collateral constraint is given by $n_{it} \leq q_t/w_t$.

[20]The bottom line is that there exists a monotone transformation of $\tilde{\Psi}_t$ that directly enters the updating problem of agents which gives rise to $\tilde{\chi}_t^d|\chi_t^s$. By "backward-engineering" Bayes' law, any assumption on the posterior distribution, can therefore also be traced back to an assumption in terms of the signal structure defined by $\tilde{\Psi}_t$.

[21]Note that in Properties 1 and 2, the conditional distributions $\tilde{\chi}_t^d|\chi_t^s$ and $\tilde{\chi}_t^m|\chi_t^s$ are meant to denote posterior distributions that result from a flat prior.

Property 2. *It holds that*

(i) $\tilde{\chi}_t^d|\chi_t^s$ *satisfies the monotone likelihood ratio property (MLRP) with respect to* χ_t^s, **or**

(ii) $\tilde{\chi}_t^m|\chi_t^s$ *belongs to a location-scale family of distributions; i.e.,* $\tilde{\chi}_t^m|\chi_t^s = \alpha_1(\chi_t^s) + \alpha_2(\chi_t^s)\tilde{X}$ *where* $\tilde{X}$ *is a non-degenerate, square-integrable random variable with mean zero and* $\alpha_1 : supp(\tilde{\chi}_t^s) \to \mathbb{R}$ *increasing.*

This property states that observing a higher labor supply allows for the statistical inference that also the corresponding fundamental labor demand (net of market noise $\tilde{\Psi}_t$) is higher in the sense of stochastic ordering. This is the natural analogue to the case without market noise, where fundamental demand *exactly* equals supply. More specifically, Property 2 specifies two alternative ordering criterion. In case (i), we adopt the commonly used monotone likelihood ratio property.[22] In case (ii), we state an alternative distributional assumption which gives rise to a specific class of "location-scale" posteriors. Here, $\alpha_1(\chi_t^s)$ is the mean of the posterior and $(\alpha_2(\chi_t^s))^2$ is proportional to the posterior variance, where the ordering takes the form of assuming that α_1 is increasing. While it will be seen that tighter financial constraints imply an higher uncertainty in both cases, introducing the more specific location-scale setting allows us later to focus on a conjugate Gaussian framework for analyzing the dynamics of this economy.

Timing The timing of events within one period can be summarized as follows:

1. The random variables $\{\tilde{\epsilon}_t, \{\tilde{\xi}_{it} : i \in [0,1]\}\}$ are realized and agents learn the realizations of $\tilde{s}_{it}$.
2. Noise traders' demand and supply $\{\tilde{\eta}_t, \tilde{\omega}_t\}$ realize, the labor and asset market operate.
3. Field production takes place, farmers choose whether or not to pay their wage bill, and gatherers seize collaterals if farmers default on their wage debt.
4. The asset market operates again.
5. Fruits from trees are gathered and consumption takes place.
6. A new generation replaces the old one and period $t+1$ begins.

Equilibrium definition Because of the assumption that agents cannot trade on margin, farmers are prevented from sidestepping collateral constraints by buying additional trees. Accordingly, the only benefit of holding trees that is reflected in the market clearing price q_t is the expected dividend payoff. Moreover, a simple arbitrage argument then implies that trees are traded at the same price in both openings of the asset market at any date t. In appendix B.1, we show that without loss of generality we can treat the two asset markets as if all assets were traded in a single pooled market that operates parallel to the labor market, and where labor demand of all farmers

[22] Formally, MLRP states that $\chi_t^s < \hat{\chi}_t^s$ implies that $\text{Prob}(\tilde{\chi}_t^d|\chi_t^s)/\text{Prob}(\tilde{\chi}_t^d|\hat{\chi}_t^s)$ is decreasing in $\tilde{\chi}_t^d$.

is constrained to satisfy $n_{it} \leq q_t/w_t$. Accordingly, the information $\{\mathcal{I}_{it} : i \in [0,1]\}$ defined in the preceding paragraphs is the basis for all labor supply $\{n_{it} : i \in \mathcal{G}\}$, labor demand $\{n_{it} : i \in \mathcal{F}\}$, and asset demand choices $\{x_{it} : i \in [0,1]\}$ at date t. Given these considerations, a competitive rational expectations equilibrium is then defined in the usual manner.

Definition 1. Given a stochastic process of shocks $\{\tilde{\epsilon}_t, \tilde{\eta}_t, \tilde{\omega}_t, \{\tilde{\xi}_{it} : i \in [0,1]\}\}$, an equilibrium in this economy is a stochastic process of choices $\{\tilde{x}_{it}, \tilde{n}_{it} : i \in [0,1]\}$ and prices $\{\tilde{w}_t, \tilde{q}_t\}$, such that:

1. $\{\tilde{x}_{it}, \tilde{n}_{it} : i \in [0,1]\}$ maximize expected utility (1) and (2) given $\{\tilde{w}_t, \tilde{q}_t\}$ and $\{\tilde{s}_{it} : i \in [0,1]\}$;

2. markets clear, i.e.,

$$(1-m) \int_{\mathcal{G}} \tilde{n}_{it} \, \mathrm{d}i = m \int_{\mathcal{F}} \tilde{n}_{it} \, \mathrm{d}i + \Psi_{\tilde{\theta}_t, \tilde{q}_t}(\tilde{\omega}_t) \tag{4}$$

   and

$$\int_0^1 \tilde{x}_{it} \, \mathrm{d}i + \Phi(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)) = 1; \tag{5}$$

3. expectations in (1) and (2) are formed optimally given $\{\tilde{w}_t, \tilde{q}_t\}$ and $\{\tilde{s}_{it}\} = \{\tilde{\theta}_t + \tilde{\xi}_{it}\}$.

3 Learning with financial frictions

In this section, we explore the key mechanism of this paper. It will be seen how learning from the real sector breaks down when financial constraints are tight.

Agents learn from the real sector via the *endogenous* history of market prices $\{\tilde{w}_t, \tilde{q}_t\}$. Hereby, asset prices play a dual role. On the one hand, changes in $\tilde{q}_t$ tighten financial constraints and thereby affect the problem of extracting information from $\tilde{w}_t$. On the other hand, $\tilde{q}_t$ is also a source of information on its own. In the next section, it will be seen that this latter problem of extracting information from $\tilde{q}_t$ is standard. For now, we therefore focus on the novel problem of extracting information from $\tilde{w}_t$, by studying how an exogenously given asset price $\bar{q}$—without any informational content on its own—constrains farmers' choices and how this affects the information aggregation.

From the market clearing condition (4), we have that $\tilde{w}_t$ solves

$$(1-m)v'^{-1}(\tilde{w}_t) = n(\tilde{w}_t, \bar{q}, \tilde{\theta}_t) + \tilde{\Psi}_t,$$

where $v'^{-1}(\tilde{w}_t)$ is the optimal labor supply of a single gatherer, $\min\{A_{it}, \bar{q}\}/w_t$ is the optimal labor demand of a single farmer with productivity A_{it}, and

$$n(\tilde{w}_t, \bar{q}, \tilde{\theta}_t) = \frac{m}{\tilde{w}_t} \int_{-\infty}^{\infty} \min\{\exp(z), \bar{q}\} \, \mathrm{d}\Phi(\sqrt{\tau_\xi}\,(z - \tilde{\theta}_t)), \tag{6}$$

is the aggregated labor demand. Transforming noisy labor demand from an additive perturbation $\tilde{\Psi}_t$ to a multiplicative perturbation $\tilde{\psi}_t$ of farmers' labor demand,[23] the market clearing condition can be rewritten as

$$\tilde{\chi}_t^s = \tilde{\chi}_t^d + \tilde{\psi}_t, \tag{7}$$

where $\tilde{\chi}_t^s$ and $\tilde{\chi}_t^d$ are normalized labor supply and demand as defined in the previous section.

Conditional on any realization of $(\tilde{w}_t, \tilde{q}_t)$, $\tilde{\chi}_t^s$ is a publicly known number. Learning from the real sector is therefore equivalent to observing a signal $\tilde{\chi}_t^s$ which communicates the true value of $\tilde{\chi}_t^d$ perturbed by $\tilde{\psi}_t$. This amounts to a nonlinear signal structure. To see this, note that

$$\tilde{\chi}_t^d = H(\tilde{\theta}_t - \log(\bar{q})),$$

where $H : \mathbb{R} \to \mathbb{R}_-$ is defined by $H(x) = \log(n(1, 1, x))$. The key observation is that H is increasing and concave. Intuitively, aggregated labor demand is obviously increasing in the average productivity. However, as an increasing number of farmers is operating at their collateral constraint (i.e., as $(\theta_t - \log(\bar{q}))$ increases), fewer farmers respond to changes in their productivities. Aggregating thus implies that aggregated labor demand $\tilde{\chi}_t^d = H(\tilde{\theta}_t - \log(\bar{q}))$ is also less responsive to changes in the fundamental $\tilde{\theta}_t$. Hence the concavity of H.

3.1 Signal extraction *without* credit constraints

Before proceeding to our main theorem, it is insightful to first consider the limit case where all farmers are unconstrained. Formally, let $\bar{q} \to \infty$. Then $H' \to 1$ for all $\theta_t \in \mathbb{R}$, so that $\text{Var}\{\tilde{\theta}_t|\chi_t^s\} = \text{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$. Property 1 therefore exactly ensures that learning in the absence of credit constraints yields a posterior uncertainty that is constant over time.

Proposition 1. *Absent credit constraints,* $\text{Var}\{\tilde{\theta}_t|\chi_t^s\} = \text{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$*; i.e.,* $\text{Var}\{\tilde{\theta}_t|\chi_t^s\}$ *is constant if and only if Property 1 holds.*

Proof. In the appendix. □

By Proposition 1, uncertainty is constant in any unconstrained version of our economy. Any variations in uncertainty are therefore the exclusive result of learning in the presence of credit constraints.

3.2 Signal extraction *with* credit constraints

We now address how credit constraints affect learning from the real sector. We have already discussed that financial constraints give rise to a concave signal structure. The following theorem states that

[23]Formally, $\tilde{\psi}_t = \log(\tilde{\Psi}_t/n_t + 1)$; e.g., $\psi_t = 0.01$ refers to approximately an one percent amplification of fundamental labor demand $n_t = n(w_t, q_t, \theta_t)$.

independently from the specific properties of our model, such a signal structure always leads to signals with a precision that decreases as the signal realizes in flatter regions.[24]

Theorem 1. *Let $\tilde{s}$, $\tilde{\theta}$ and $\tilde{\epsilon}$ be three non-degenerate random variables, and let f be an increasing function defined on the convex hull of the support of $\tilde{\theta}$, such that $\tilde{s} = f(\tilde{\theta}) + \tilde{\epsilon}$ and $\tilde{\theta}$, $f(\tilde{\theta})$ square-integrable. Furthermore suppose that either $f(\tilde{\theta})|s$ satisfies the monotone likelihood ratio property with respect to s;* or *$\tilde{\theta}|s = \alpha_1(s) + \alpha_2(s)\tilde{X}$ for some non-degenerate, square-integrable random variable $\tilde{X}$ and some functions α_1, α_2 defined on the support of $\tilde{s}$ and α_1 increasing. Then*

1. $\text{Var}\{\tilde{\theta}|s\}$ *is increasing in s if f is concave and* $\text{Var}\{f(\tilde{\theta})|s\}$ *is nondecreasing in s,*

2. $\text{Var}\{\tilde{\theta}|s\}$ *is decreasing in s if f is convex and* $\text{Var}\{f(\tilde{\theta})|s\}$ *is nonincreasing in s.*

In both cases, the monotonicity of $\text{Var}\{\tilde{\theta}|s\}$ *is strict whenever f is strictly concave or convex.*

Proof. In the appendix. □

The intuition behind this theorem is quiet straightforward. When a signal $\tilde{s}$ realizes in flatter regions, then a "well-behaved" signal structure allows for the posterior belief that also the fundamental $\tilde{\theta}$ takes values for which f is flat.[25] But then, a Bayesian must also believe that the realization of $\tilde{s}$ is largely driven by noise $\tilde{\epsilon}$ rather than by the fundamental $\tilde{\theta}$. Hence the increase in posterior uncertainty.

Framed in terms of our model, this reasoning translates into wage signals that are largely driven by noisy demand fluctuations rather than fundamentals whenever the economy is in a constrained state. Formally, we have:

Proposition 2. $\text{Var}\{\tilde{\theta}_t|\chi_t^s\}$ *is strictly increasing in χ_t^s.*

Proof. In the appendix. □

The role of Proposition 2 in this paper cannot be overstated. It precisely tells us how and when uncertainty about $\tilde{\theta}_t$ is fluctuating in the economy. In particular, it establishes that uncertainty rises after negative shocks to $\tilde{q}_t$, which is at the core of all our results.

3.3 Quasi-Gaussian signal structure

Proposition 2 is inextricably tied to learning through the nonlinear function H. Unfortunately, learning from nonlinear Gaussian signals generally leads to posterior distributions that do not conjugate with the corresponding priors, making the dynamic analysis highly intractable. To address

[24]For readability, we state the theorem for increasing functions f only, but it is straightforward to generalize it to all monotonic f, leading to the opposite predictions for $\text{Var}\{\tilde{\theta}|s\}$ if f is decreasing.

[25]By "well-behaved", we refer to the assumption that either $f(\tilde{\theta}|s)$ are ordered according to the MLRP property—which then in Milgrom's (1981) language implies that "good news" for $\tilde{s}$ is also "good news" for $f(\tilde{\theta})$ and, hence, for $\tilde{\theta}$—or, alternatively, that the first moment of the posterior distribution of $\tilde{\theta}$ is increasing in s—similarly implying that good news for $\tilde{s}$ is good news for $\tilde{\theta}$.

this problem, we henceforth restrict ourselves to information structures that satisfy case (ii) of Property 2 with $\tilde{X}$ being a standard normal random variable. Under this assumption, beliefs of agents evolve *as if* agents observed a Gaussian signal with a *state-dependent* signal precision τ_υ. Using this approach of informationally equivalent Gaussian signals, we are able to embed the idea of time-varying uncertainty in a convenient framework with conjugate Gaussian priors.

Lemma 1. *Suppose Property 2(ii) holds with $\tilde{X}$ being standard normally distributed. Then, given a normal prior over $\tilde{\theta}_t$, observing $\tilde{\chi}_t^s = \chi_t^s$ is equivalent to observing a signal $\tilde{\theta}_t + \tilde{\upsilon}_t$ with realization $\alpha_1(\chi_t^s) + \log(\bar{q})$, where $\tilde{\upsilon}_t$ is Gaussian noise with variance $1/\tau_\upsilon(\theta_t + \upsilon_t - \log(\bar{q}))$ and $\tau_\upsilon : \mathbb{R} \to \mathbb{R}_+$ is strictly decreasing,* $\lim_{z\to-\infty} \tau_\upsilon(z) = 1/\mathrm{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$, *and* $\lim_{z\to\infty} \tau_\upsilon(z) = 0$.

Proof. In the appendix. □

The "quasi-Gaussian" signal structure that follows from Lemma 1 effectively decomposes the inference problem of agents into a straightforward interpretation of a Gaussian signal and a computational intensive calculation of the relevant signal precision. The benefit of this decomposition is that agents' beliefs can be computed straightforward within a conjugate prior framework. The computational intensive part, on the other hand, only has to be solved by the model analyst. However, applying the results from Proposition 2, we know that τ_υ is decreasing as the economy gets more constrained, enabling us to derive all of our results analytically without the need to solve the exact inference problem.

Figure 1 illustrates the properties of τ_υ that follow from Lemma 1: (i) τ_υ is decreasing as the economy gets more constrained (i.e., as $\alpha_1(\chi_t^s) = \theta_t + \upsilon_t - \log(\bar{q})$ increases, reflecting either tighter credit conditions, or an increased labor demand relative to existing credit conditions); (ii) τ_υ converges to $1/\mathrm{Var}\{\tilde{\chi}_t^d|\chi_t^s\} = const$ as the economy gets completely unconstrained (i.e, no information is lost when all farmers are unconstrained); and (iii) τ_υ converges towards zero, so that the signal gets completely uninformative, as the economy becomes fully constrained.

4 Equilibrium characterization

Equipped with the quasi-Gaussian representation, we are now ready to characterize the equilibrium. We proceed in two steps. First, we fix the information structure (by fixing τ_υ) and analyze the resulting signal extraction problem taking into account all signals. Because parts of agents' information is extracted from endogenous asset prices, this step involves finding the "usual" fixed point between a perceived law of motion of $\tilde{q}_t$ and its actual behavior. Second, allowing the information structure to vary with the state of the economy, we then establish the full informational equilibrium where information is simultaneously aggregated from labor and asset markets and where prices in both markets are consistent with the resulting beliefs.[26]

[26]Formally, agents post labor and asset demand schedules that are fully contingent on both prices. Agents therefore learn from observing prices in the labor and asset market simultaneously and the resulting beliefs have to be consistent

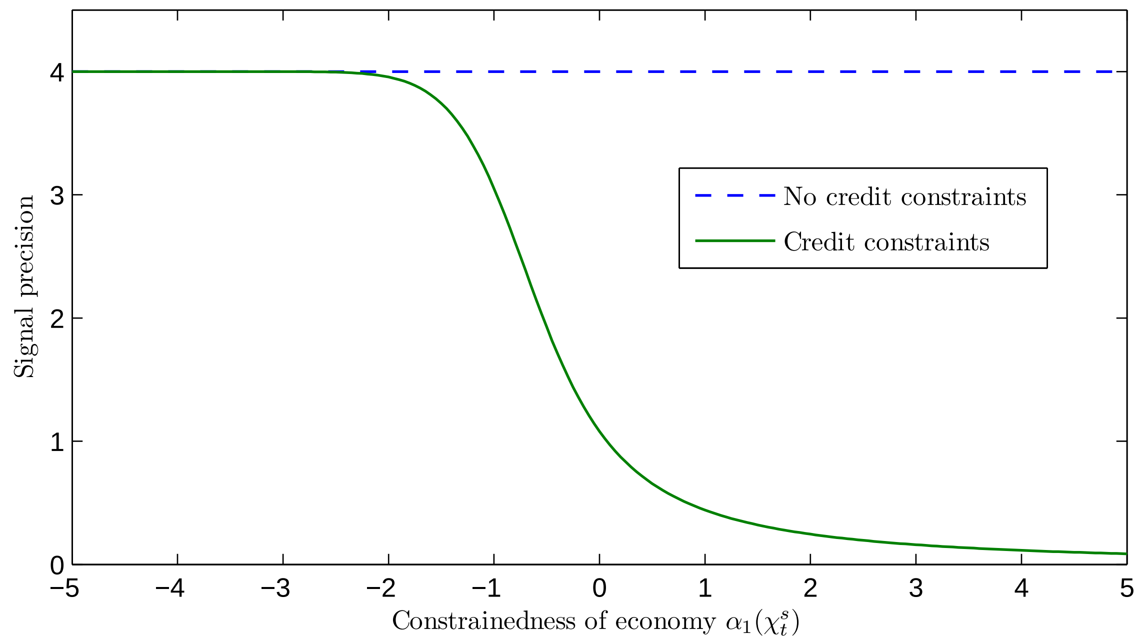

Figure 1: Endogenous signal precision.

4.1 Signal extraction from the financial market

From Lemma 1 it follows that we can focus on Normal posteriors. Accordingly, let $b_t \equiv \mathbb{E}\{\tilde{\theta}_t | \mathcal{I}_{it} \setminus s_{it}\}$ and $1/\pi_t \equiv \text{Var}\{\tilde{\theta}_t | \mathcal{I}_{it} \setminus s_{it}\}$ denote the posterior mean and variance that result from observing the *public* history of market prices up to date t. Further let $b_{it} \equiv \mathbb{E}\{\tilde{\theta}_t | \mathcal{I}_{it}\}$ and $1/\bar{\pi}_t \equiv \text{Var}\{\tilde{\theta}_t | \mathcal{I}_{it}\}$ denote the first two posterior moments given information set $\mathcal{I}_{it}$ (i.e., including the *private* signal $\tilde{s}_{it}$), where in anticipation of the results below, we drop the subscript i from the posterior precision $\bar{\pi}_t$. Then optimal asset demand is given by

$$x_{it} = \begin{cases} 1 & \text{if } \mathbb{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} > q_t, \\ [0,1] & \text{if } \mathbb{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} = q_t, \\ 0 & \text{if } \mathbb{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} < q_t, \end{cases}$$

where $\mathbb{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} = \exp(\gamma_0 + \gamma_1 b_{it} + \gamma_1^2/(2\bar{\pi}_t))$.[27] Aggregating, adding noise traders' demand, and rearranging yields the market clearing condition

$$\Phi(\sigma_t^{-1}(b_{mt}(q_t) - \bar{b}_t)) = \Phi(\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu)),$$

with the market clearing prices on both markets.

[27] Here we assume without loss of generality that $\tilde{u}_t$ is constant zero. Suppose it is not. Then by the independence of $\tilde{u}_t$ we have that $\mathbb{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} = \mathbb{E}\{\exp(\tilde{u}_t)\} \exp(\gamma_0 + \gamma_1 b_{it} + \gamma_1^2/(2\bar{\pi}_t)) = \exp(\gamma_0' + \gamma_1 b_{it} + \gamma_1^2/(2\bar{\pi}_t))$, where $\gamma_0' = \gamma_0 + \log(\mathbb{E}\{\exp(\tilde{u}_t)\})$ and, hence, any nonzero random noise $\tilde{u}_t$ can be absorbed by the constant γ_0.

where $b_{mt}(q_t)$ is the belief of the marginal trader with $\mathbb{E}\{\tilde{A}_t|b_{mt},\bar{\pi}_t\} = q_t$, $\bar{b}_t$ is the average belief $\int b_{it}\,\mathrm{d}i$, and $\sigma_t^2 = \mathrm{Var}\{b_{it}\}$ is the cross-sectional dispersion of beliefs. This pins down the marginal traders' belief $b_{mt} = \sigma_t\sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu) + \bar{b}_t$ and the equilibrium price. Given any conjectured law of motion for the market clearing price, this price also serves as an endogenous signal. In equilibrium the beliefs resulting from interpreting this signal have to give rise to optimal asset demands that yield an actual law of motion equal to the conjectured one. This fixed-point problem has a log-linear solution, which is established in the following lemma (for a detailed derivation, see, e.g., Hellwig (1980)).

Lemma 2. *Fix any τ_υ. Then there exists a unique log-linear asset price*

$$q_t = \exp\{\gamma_1(\bar{b}_t + \bar{\pi}_t^{-1}\tau_\xi\eta_t) + \gamma_0\} \tag{8}$$

which clears the asset market, where

$$\bar{b}_t = \bar{\pi}_t^{-1} \times \begin{bmatrix}\tau_\xi & \tau_\eta & \tau_\upsilon & \hat{\pi}_{t-1}\end{bmatrix} \times \begin{bmatrix}\theta_t \\ \theta_t + \eta_t \\ \theta_t + \upsilon_t \\ \rho b_{t-1}\end{bmatrix},$$

$$\bar{\pi}_t = \tau_\xi + \tau_\eta + \tau_\upsilon + \hat{\pi}_{t-1},$$

$$\hat{\pi}_{t-1} = \frac{\pi_{t-1}\tau_\epsilon}{\pi_{t-1} + \rho^2\tau_\epsilon},$$

and where (b_t, π_t) are given by $(\bar{b}_t, \bar{\pi}_t)$ after setting $\tau_\xi = 0$.

Proof. In the appendix. □

Given a public history summarized by b_{t-1} and π_{t-1}, and given a signal precision τ_υ, this lemma characterizes the unique log-linear relationship between the asset price at date t and the stochastic variables $\tilde{\theta}_t$, $\tilde{\eta}_t$, and $\tilde{\upsilon}_t$. The equilibrium price increases in all three variables, not just in the fundamental. The reason is that positive noisy demand realizations on the asset and labor market falsely suggest that the fundamental increased. The exact weight that is put on these sources of information, however, depends on the signal precision τ_υ. Intuitively, as the real sector aggregates less information, less weight is placed on real sector prices, and vice versa. We also see that posterior uncertainty $1/\bar{\pi}_t$ is increasing as the signal precision τ_υ of the labor market decreases.

4.2 General informational equilibrium

To characterize the full informational equilibrium that takes into account the mutual dependence of signal precision and asset price, define $g_q : \mathbb{R}_+^2 \times \mathbb{R}^4 \to \mathbb{R}$, such that $\log(q_t) = g_q(\tau_\upsilon, \Omega_t)$ as given by (8), and define $g_\tau : \mathbb{R}_+ \times \mathbb{R}^5 \to \mathbb{R}_+$, such that $\tau_\upsilon(\theta_t + \upsilon_t - \log(q_t)) = g_\tau(\log(q_t), \Omega_t)$ is the signal precision defined in Lemma 1. Here, $\Omega_t \equiv (\pi_{t-1}, b_{t-1}, \theta_t, \eta_t, \upsilon_t)$ is the vector of state variables in period t. Then the informational equilibrium can be computed pointwise for any state of the economy Ω_t, by solving the fixed-point problem[28]

$$g_\tau(\cdot, \Omega_t) - g_q^{-1}(\cdot, \Omega_t) = 0. \tag{9}$$

The following proposition establishes the existence of a solution to this problem for all possible states Ω_t and, hence, the existence of an informational equilibrium.

Proposition 3. *For all $\Omega_t \in \mathbb{R}_+ \times \mathbb{R}^4$, there exists a solution to the fixed-point problem* (9)*. The solution is unique for all Ω_t inside a set $\Xi \subset \mathbb{R}_+ \times \mathbb{R}^4$. In particular, it is unique for all Ω_t that satisfy*

$$(\gamma_1 - 1)(\theta_t + \upsilon_t) < M_1$$

or

$$\theta_t + \upsilon_t < M_2 + \gamma_1^{-1} \log(q_t|_{\tau_\upsilon \to 0})$$

where $M_1, M_2 \in \mathbb{R}$ are parameters defined by the primitives of the model. In contrast, if $\Omega_t \notin \Xi$, the economy is in a "sunspot" state where (9) *has multiple solutions.*

Proof. In the appendix. □

Proposition 3 implies that for any initial state Ω_0 the equilibrium dynamics of the model economy can be computed recursively by computing a solution to (9) given Ω_t and then using Lemma 2 to determine Ω_{t+1}. As long as $\{\Omega_s\}_{s=1}^t \in \Xi^t$, this pins down a unique equilibrium path. This is the case as long as the information aggregated in the labor and asset market is not too conflicting. In states where the labor market signals a realizations for $\tilde{\theta}_t$ that is sufficiently more optimistic than what is suggested by the asset market, the equilibrium gives rise to self-fulfilling sunspots: If agents coordinate on interpreting the conflicting information in an optimistic way, asset prices will be little constraining, so that observing the "good" labor market news will be sufficiently informative to justify an optimistic interpretation. On the other hand, if agents interpret the evidence in a pessimistic way, the resulting financial constraints will obscure the "good" news from the labor market and the pessimistic interpretation is indeed justified.

Because the findings in this paper generally carry over to sunspot regimes (but require a more subtle distinction between cases), we focus throughout most of this paper on paths where $\{\Omega_s\}_{s=1}^t \in \Xi^t$. A brief discussion of sunspot regimes can be found in Appendix B.4.

[28]For expositional reasons, we abstract from the non-generic case where g_q is flat throughout the main body of the paper. The case where g_q^{-1} does not exist is carefully treated in all formal proofs.

5 Transmission of shocks

We now explore how random shocks to the economy are propagated through the informational equilibrium. In the next two subsections, it will be seen how shocks are statically amplified or de-amplified depending on their size and composition. In Section 5.3, it will be seen that a similar distinction divides shocks in those that are dynamically persistent, and others that are non-persistent.

5.1 Static asymmetries: Amplifying and de-amplifying shocks

Consider the solution to the fixed point problem (9). We say that a shock is *amplified* through the endogenous information structure if and only if its absolute impact on q_t is larger than in the hypothetical benchmark in which τ_υ is fixed at the level it would attain in the absence of shocks. By that definition, which shocks are amplified and which shocks are de-amplified upon impact?

For answering this question, it is useful to define $\mathring{\tau}_\upsilon \equiv \tau_\upsilon(-\gamma_0)$ to be the precision of υ_t in the absence of shocks.[29] In the benchmark case, the market-clearing price then reads $\mathring{q}_t \equiv \exp(g_q(\mathring{\tau}_\upsilon, \Omega_t))$. Contrast this with the case, where τ_υ adjusts endogenously to the asset price. Then the endogeneity of the information structure unfolds a feedback loop. Intuitively, the benchmark price $\mathring{q}_t$ can be seen as the "initial" impact of the shock. In response to this change in $\tilde{q}_t$ and also in response to the shock itself (compare Lemma 1), τ_υ is now changing to $\tau_\upsilon^1 \equiv g_\tau(\log(\mathring{q}_t), \Omega_t)$. However, unless $\tau_\upsilon^1 = \mathring{\tau}_\upsilon$, agents who optimally form expectations respond to such changes in τ_υ by re-weighting the available sources of information, leading to an asset price q_t^1, and so on.

To answer which shocks are amplified, we need to compare the equilibrium of this feedback loop—i.e., the solution q_t^* to (9)—with $\mathring{q}_t$. Observe that the two directions of the feedback loop can be summarized by (i) the sign of $\partial g_q/\partial \tau_\upsilon$ (how do changes in τ_υ feed back to q_t), and (ii) the sign of $\{\tau_\upsilon^1 - \mathring{\tau}_\upsilon\}$ (after a single "cycle" through the feedback loop, how does τ_υ change in response to q_t^0).[30] Combining, we have four cases to consider, summarized by Figure 2. If g_q is increasing in τ_υ and $\tau_\upsilon^1 < \mathring{\tau}_\upsilon$, then g_q^{-1} intersects with g_τ to the left of $\log(\mathring{q}_t)$ and we have that $q_t^* < \mathring{q}_t$ (see Panel a). Likewise, if g_q is increasing in τ_υ and $\tau_\upsilon^1 > \mathring{\tau}_\upsilon$, then $q_t^* > \mathring{q}_t$ (see Panel b). A similar reasoning establishes the opposite relationship if g_q is decreasing in τ_υ (see Panels c and d).

So what determines the signs of $\partial g_q/\partial \tau_\upsilon$ and $\{\tau_\upsilon^1 - \mathring{\tau}_\upsilon\}$? From Lemma 2, we have that

$$\begin{aligned}\text{sign}\left\{\frac{\partial g_q}{\partial \tau_\upsilon}\right\} &= \text{sign}\left\{\frac{\partial}{\partial \tau_\upsilon}\left(\bar{\pi}_t^{-1}\begin{bmatrix}\tau_\xi + \tau_\eta & \tau_\upsilon & \hat{\pi}_{t-1}\end{bmatrix} \times \begin{bmatrix}\theta_t + \eta_t \\ \theta_t + \upsilon_t \\ \rho b_{t-1}\end{bmatrix}\right)\right\} \\ &= \text{sign}\left\{-(\theta_t + \eta_t) + \delta_1(\theta_t + \upsilon_t)\right\} ,\end{aligned}$$

where $\delta_1 : \mathbb{R} \to \mathbb{R}$ is defined by

[29] I.e., for $\theta_t = \eta_t = \upsilon_t = 0$ and $b_{t-1} = 0$, implying $\log(q_t) = \gamma_0$ and, hence, $\mathring{\tau}_\upsilon = \tau_\upsilon(-\gamma_0)$.

[30] Here we exploit that, as formally shown in the proof to Proposition 4, $\text{sign}\{\tau_\upsilon^1 - \mathring{\tau}_\upsilon\} = \text{sign}\{\tau_\upsilon^* - \mathring{\tau}_\upsilon\}$, allowing us to use $\text{sign}\{\tau_\upsilon^1 - \mathring{\tau}_\upsilon\}$ to determine the qualitative effects of the feedback loop on q_t^*.

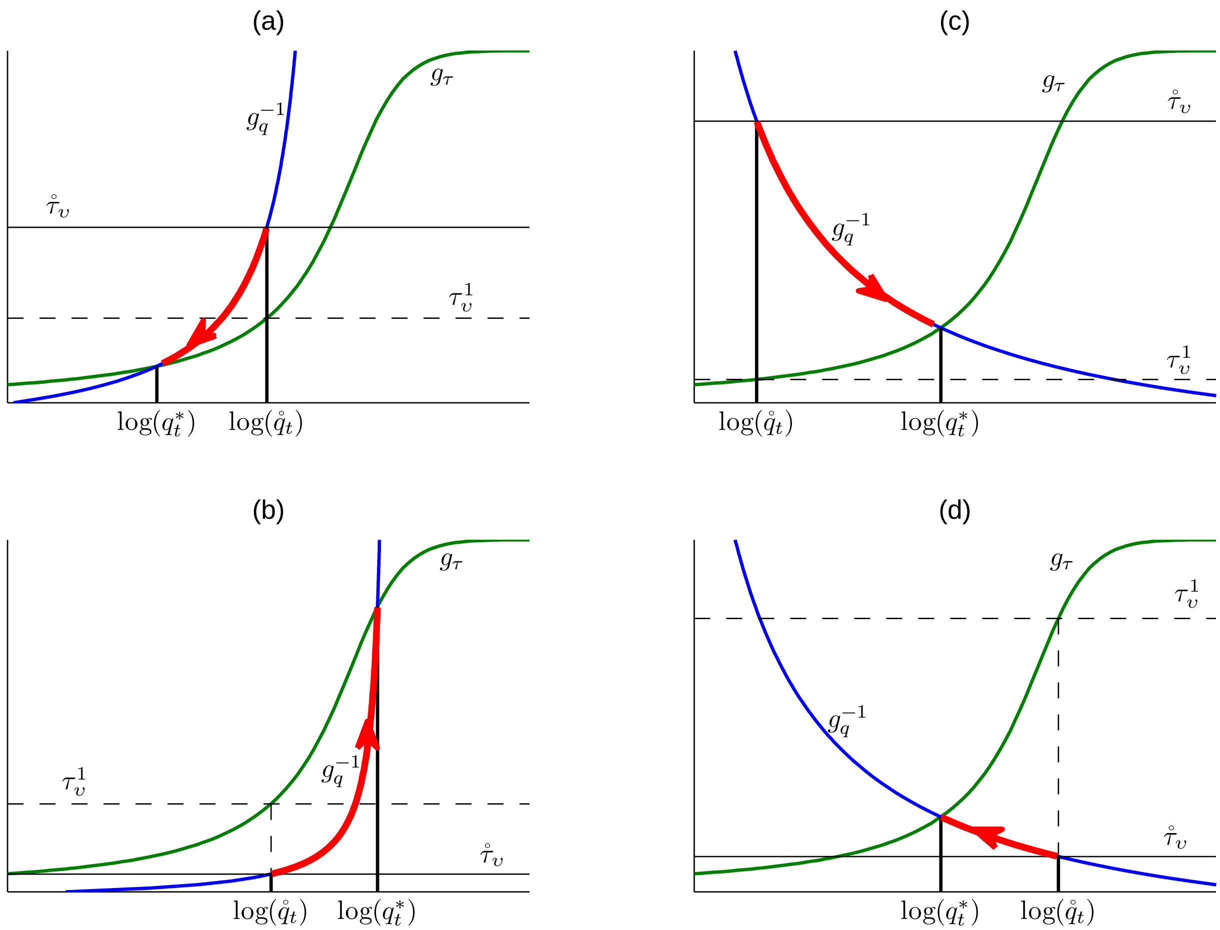

Figure 2: Location of equilibrium asset price relative to the counterfactual price.

$$\delta_1(x) = (\tau_\xi + \tau_\eta)^{-1}(-\check{b}_{t-1} + \check{\pi}_{t-1}x)$$

with $\check{b}_{t-1} = \rho b_{t-1}\hat{\pi}_{t-1}$ and $\check{\pi}_{t-1} = \tau_\xi + \tau_\eta + \hat{\pi}_{t-1}$. Here, $\delta_1^{-1}(\theta_t + \eta_t)$ is the residual asset price, which would obtain if there were no labor market signal. It includes all information that is inferred from *other* sources than w_t—i.e., the prior b_{t-1}, the idiosyncratic signals $\{s_{it}\}$, and news extracted from q_t itself. Weighting these sources against $\theta_t + \upsilon_t$—the information extracted from w_t—then pins down the equilibrium price. Intuitively, if $\theta_t + \upsilon_t$ is larger than $\delta_1^{-1}(\theta_t + \eta_t)$, then if agents increase the weight on w_t, they become more optimistic. As a result, g_q is increasing in τ_υ exactly if $\theta_t + \eta_t < \delta_1(\theta_t + \upsilon_t)$.

To determine the sign of $\{\tau_\upsilon^1 - \mathring{\tau}_\upsilon\}$, we apply the transformation τ_υ^{-1}. Since τ_υ is decreasing, the term is positive exactly if $\theta_t + \upsilon_t - \log(\mathring{q}_t) < -\gamma_0$, or if $\theta_t + \eta_t > \delta_2(\theta_t + \upsilon_t)$, where $\delta_2 : \mathbb{R} \to \mathbb{R}$ is defined by

$$\delta_2(x) = (\tau_\xi + \tau_\eta)^{-1}(-\check{b}_{t-1} + \gamma_1^{-1}(\check{\pi}_{t-1} + (1-\gamma_1)\mathring{\tau}_\upsilon)x).$$

Note that δ_2 is decreasing, if and only if $\gamma_1 > \check{\pi}_{t-1}/\mathring{\tau}_\upsilon + 1$. This reflects the case, where asset prices have a strong enough impact on τ_υ to compensate for any direct impact of $\theta_t + \upsilon_t$. By contrast, if γ_1 is sufficiently small, then any effect that a positive realization of $\tilde{\theta}_t + \tilde{\upsilon}_t$ has on q_t is dominated by additional labor demand. Thus the economy is effectively more constrained and generates less information as $\theta_t + \upsilon_t$ increases, which translates into a positively sloped δ_2.

We are now ready to address the key question in this section. Based on the affine functions δ_1 and δ_2, we can assign each state Ω_t one of the four cases depicted in Figure 2. If we also take into account whether $\mathring{q}_t$ is increased or decreased relative to the no-shock price γ_0, we can therefore determine whether endogenous information amplifies or de-amplifies the impact of shocks in period t. More specifically, when $\mathring{q}_t$ is increased compared to γ_0, then endogenous uncertainty amplifies the impact of Ω_t if $q_t^* > \mathring{q}_t$, and de-amplifies (or, possibly, reverses) the impact if $q_t^* < \mathring{q}_t$. The converse holds true if $\mathring{q}_t$ is decreased compared to the no-shock case. Comparing $\log(\mathring{q}_t)$ with γ_0, we find that $\log(\mathring{q}_t) > \gamma_0$ exactly if $\theta_t + \eta_t > \delta_3(\theta_t + \upsilon_t)$, where $\delta_3 : \mathbb{R} \to \mathbb{R}$ is defined by

$$\delta_3(x) = (\tau_\xi + \tau_\eta)^{-1}(-\check{b}_{t-1} - \mathring{\tau}_\upsilon x).$$

The state space Ω_t is thus divided into amplification and de-amplification regimes by lines δ_1, δ_2, and δ_3. The following proposition formalizes this result.

Proposition 4. *(a) Suppose that $\gamma_1 > 1$ and $\theta_t + \eta_t > \delta_3(\theta_t + \upsilon_t)$, that is, $\log(\mathring{q}_t) > \gamma_0$. Then $q_t^* > \mathring{q}_t$, so that the impact of Ω_t is amplified if and only if*

$$\delta_1(\theta_t + \upsilon_t) < \theta_t + \eta_t < \delta_2(\theta_t + \upsilon_t),$$

where the inequalities are reversed for $\gamma_1 < 1$.

(b) Suppose that $\gamma_1 > 1$ and $\theta_t + \eta_t < \delta_3(\theta_t + \upsilon_t)$, so that $\log(\mathring{q}_t) < \gamma_0$. Then $q_t^* < \mathring{q}_t$, so that the impact of Ω_t is de-amplified, if and only if

$$\delta_1(\theta_t + \upsilon_t) > \theta_t + \eta_t > \delta_2(\theta_t + \upsilon_t),$$

where the inequalities are reversed for $\gamma_1 < 1$.

Proof. In the appendix. □

Figure 3 illustrates the proposition. For all realizations of $\tilde{\theta}_t$, $\tilde{\eta}_t$, and $\tilde{\upsilon}_t$ that fall north-east of δ_3, the combined impact on $\mathring{q}_t$ is positive, so that $\log(\mathring{q}_t) > \gamma_0$. For these shocks, the impact of Ω_t compared to the exogenous uncertainty benchmark is de-amplified in region A and amplified in region B. Note that region A is split into two separate areas. The one to the north-west of B

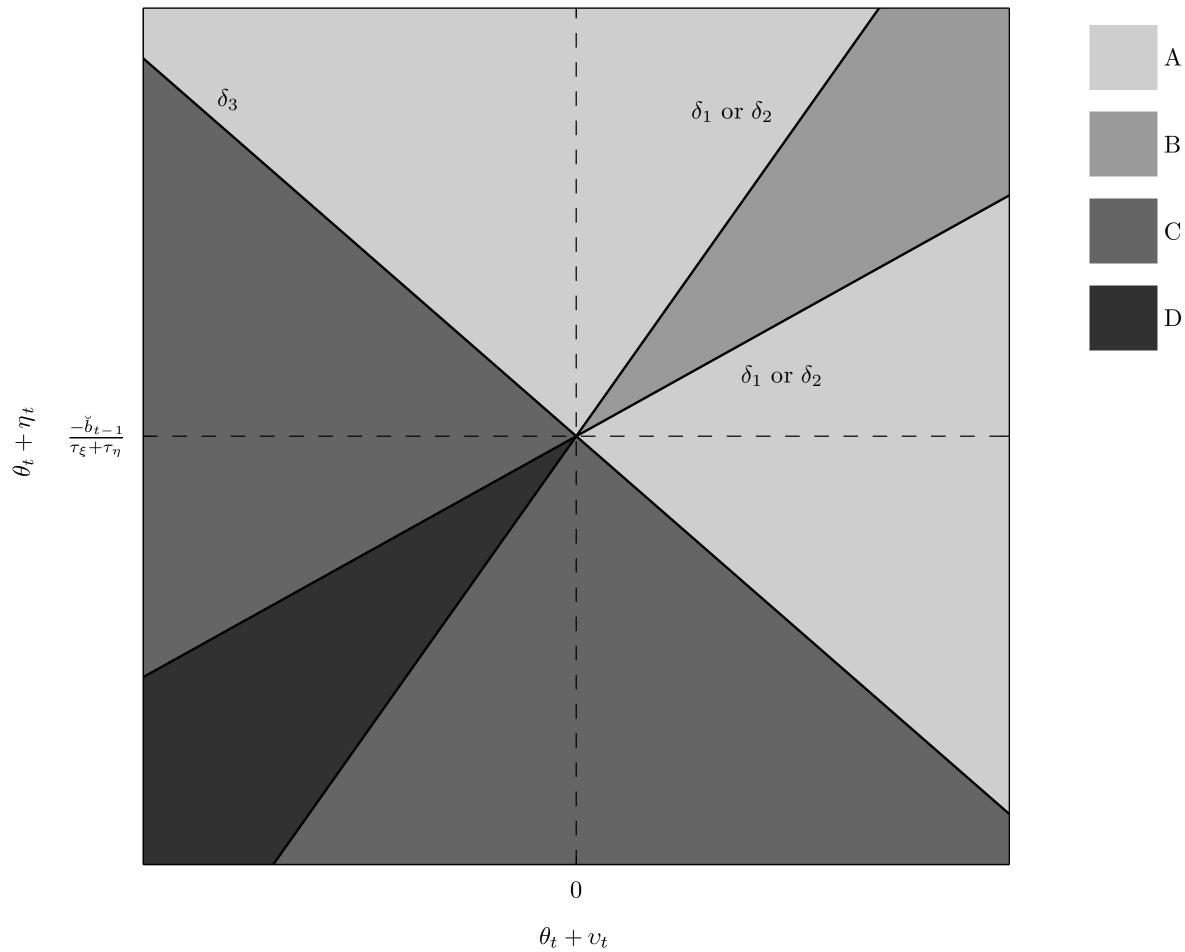

Figure 3: Impact asymmetries. *Note:* Shocks in regions A and B have an overall positive impact on $\mathring{q}_t$, shocks in regions C and D have an overall negative impact. Shocks in regions B and C are endogenously amplified, shocks in regions A and D are endogenously de-amplified.

corresponds to the case depicted in Panel (d) of Figure 2, and the one to the south-east corresponds to Panel (a). Region B corresponds to either Panel (b) or (c), depending on the value of γ_1. If $\gamma_1 > 1$, δ_2 is steeper than δ_1, and all Ω_t in the area bounded by these two lines are amplified by a decreasing g_q as depicted in Panel (c). If $\gamma_1 < 1$, then δ_2 has a smaller slope than δ_1, and Ω_t is amplified as depicted in Panel (b).

For realizations of $\tilde{\theta}_t$, $\tilde{\eta}_t$, and $\tilde{\upsilon}_t$ that fall south-west of δ_3, the combined impact has a negative effect on $\mathring{q}_t$, implying $\log(\mathring{q}_t) < \gamma_0$. In that case, realizations within region C are amplified—corresponding to the cases in Panel (a) in the south-east of region D, and Panel (d) in the north-west of region D. Realizations within region D are de-amplified—corresponding to Panel (b) if $\gamma_1 > 1$, and Panel (c) if $\gamma_1 < 1$.

Macroeconomists are often interested in the special case where the economy is hit by a single shock, shutting down all other stochastic channels through which the economy is impacted. Since δ_1 and δ_2 both have a finite slope, any state in which $|\eta_t|$ is sufficiently large compared to $|\theta_t|$ and $|\upsilon_t|$ is

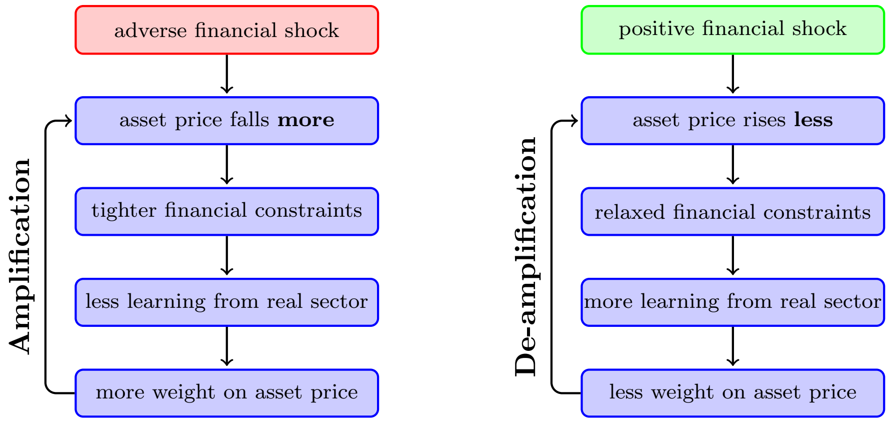

Figure 4: A schematic illustration of the feedback loop in the special case of an isolated financial shocks.

unambiguously amplified for $\eta_t < 0$ and de-amplified for $\eta_t > 0$. In particular, this adverse feedback loop applies to any financial "impulse" shocks; i.e., shocks along the vertical dashed axis through the origin of Figure 3. (See Figure 4 for a schematic illustration of the feedback loop induced by financial shocks.)

Corollary 1. *In the limit as $\theta_t \to 0$, $\upsilon_t \to 0$, and $b_{t-1} \to 0$, financial shocks are amplified if $\eta_t < 0$ and de-amplified if $\eta_t > 0$.*

Similarly, because b_{t-1} vertically shifts the origin of δ_1, δ_2, and δ_3, any prior pessimism is amplified and any prior optimism is de-amplified along an impulse response path.

Corollary 2. *In the limit as $\theta_t \to 0$, $\eta_t \to 0$, and $\upsilon_t \to 0$, prior beliefs are amplified if $b_{t-1} < 0$ and de-amplified if $b_{t-1} > 0$.*

The case where the economy is perturbed by a single shock on the labor market (i.e., shocks along the horizontal dashed axis through the origin of Figure 3) is less clear. This is because for large γ_1 the slope of δ_2 becomes negative (but never steeper than the slope of δ_3). Formally, this leads to the following result.

Corollary 3. *In the limit as $\theta_t \to 0$, $\eta_t \to 0$, and $b_{t-1} \to 0$, labor shocks are amplified if $\upsilon_t < 0$ and de-amplified if $\upsilon_t > 0$ if and only if $\gamma_1 < \breve{\pi}_{t-1}/\mathring{\tau}_\upsilon + 1$. Otherwise, the converse holds true.*

5.2 Static asymmetries: Non-proportionality in scale

Proposition 4 divides the state space into amplifying and de-amplifying regimes. It is silent, however, on how the *degree* of amplification or de-amplification changes within these regimes. We now address this question. In particular, we are interested in how the degree of amplification and de-amplification changes as shocks realize "further away" from the origin, $\mathcal{O}_t \equiv (0, -\breve{b}_{t-1}/(\tau_\xi + \tau_\eta))$, of Figure 3. The following proposition establishes that both the degree of amplification in amplification regimes and the degree of de-amplification in de-amplification regimes monotonically increases as shocks are scaled up relative to $\mathcal{O}_t$.

Proposition 5. *Consider any combination of shocks* $(\theta_t + \upsilon_t, \theta_t + \eta_t) \equiv \mathcal{S}_t + \mathcal{O}_t$. *Then scaling up these shocks to* $a\mathcal{S}_t + \mathcal{O}_t$, $a > 1$, *increases amplification in amplification regimes and decreases amplification in de-amplification regimes. Formally, that is,* $\log(q_t^*) - \log(\mathring{q}_t)$ *decreases in* a *if and only if* $sign\{\theta_t + \eta_t - \delta_1(\theta_t + \upsilon_t)\} = sign\{\theta_t + \eta_t - \delta_2(\theta_t + \upsilon_t)\}$.

Proof. In the appendix. □

Proposition 5 states that "scaling up" the combination of shocks that hit the economy at time t, implies that amplification or de-amplification of these shocks is both more pronounced. The reason is that absolute larger shocks lead to larger changes in uncertainty and hence to more pronounced amplification and de-amplification loops, respectively. Figure 5 illustrates these nonlinearities by plotting contours of the degree of amplification $\{\log(q_t^*) - \log(\mathring{q}_t)\} \times \text{sign}(\log(\mathring{q}_t))$. Negative contours (dashed lines) thus correspond to de-amplification regimes, positive contours (solid lines) correspond to amplification regimes. By Proposition 5, these contours are increasing towards the origin in de-amplification regimes, and are decreasing towards the origin in amplification regimes.

5.3 Dynamic asymmetries: Persistence and non-persistence of beliefs

We now consider the effect of changes in τ_υ on the persistence of shocks. Because our model abstracts from all intertemporal links other than the fundamental process of $\tilde{\theta}_t$, the only channel through which Ω_t may affect future periods other than through the autocorrelation of $\tilde{\theta}_t$, is through the persistence of public beliefs. For all $s > 0$, we define the persistence of belief b_t onto b_{t+s} as

$$\Lambda_{t,t+s} = \frac{\partial b_{t+s}}{\partial b_t}.$$

In the following, we focus on the specific signal structure underlying our model, but it is worth noting that the arguments generalize to arbitrary quasi-Gaussian signal structures (see the formal proof of Proposition 6). From Lemma 2, we have that

$$b_t = \pi_t^{-1} \times \begin{bmatrix} \tau_\eta & \tau_{\upsilon,t} & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \theta_t + \eta_t \\ \theta_t + \upsilon_t \\ \rho b_{t-1} \end{bmatrix}. \tag{10}$$

Recursively substituting and differentiating thus yields

$$\Lambda_{t,t+s} = \prod_{q=t+1}^{t+s} \lambda_q,$$

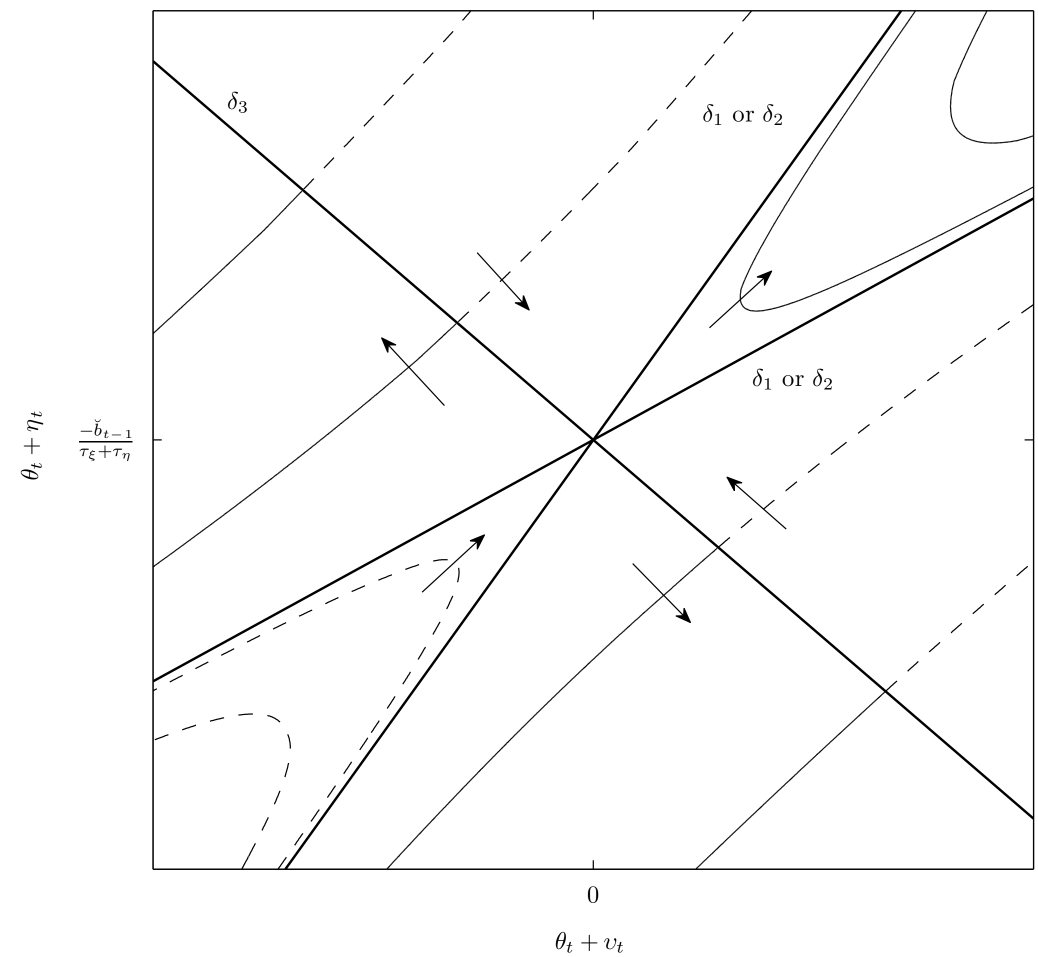

Figure 5: Amplification contours. *Note:* Positive contours (amplification regimes) are plotted solid, negative contours (de-amplification regimes) are dashed. Arrows point into the direction of increasing (more amplifying) contours.

where $\lambda_t \equiv \rho\hat{\pi}_{t-1}/\pi_t$. We are interested in the effect of an increase in uncertainty in period t on the persistence $\Lambda_{t-r,t+s}$ for $r, s \geq 0$. Suppose $\tau_{\upsilon,t}$ changes by a differential $\mathrm{d}\tau_{\upsilon,t}$. Consider the special case where $r = 0$. Then

$$\frac{\mathrm{d}\Lambda_{t,t+s}}{\mathrm{d}\tau_{\upsilon,t}} = \Lambda_{t,t+s} \times \sum_{q=t+1}^{t+s} \left(\frac{1}{\lambda_q} \frac{\partial \lambda_q}{\partial \pi_{q-1}} \prod_{p=t+1}^{q-1} \frac{\partial \pi_p}{\partial \pi_{p-1}} \right).$$

Here, the only term with a nontrivial sign is $\partial\lambda_q/\partial\pi_{q-1}$. However, because $\partial\pi_q/\partial\pi_{q-1} = \partial\hat{\pi}_{q-1}/\partial\pi_{q-1}$, we have

$$\frac{\partial \lambda_q}{\partial \pi_{q-1}} = \frac{\partial \hat{\pi}_{q-1}}{\partial \pi_{q-1}} \frac{\tau_\eta + \tau_{\upsilon,q}}{\pi_t^2} > 0,$$

establishing that $\mathrm{d}\Lambda_{t,t+s}/\mathrm{d}\tau_{\upsilon,t} > 0$. This reflects that more precise information in period t is unambiguously more relevant for forming *future* beliefs. Contrast this with the case where $s = 0$.

Then

$$\frac{\mathrm{d}\Lambda_{t-r,t}}{\mathrm{d}\tau_{\upsilon,t}} = \frac{\Lambda_{t-r,t}}{\lambda_t} \times \frac{\partial \lambda_t}{\partial \pi_t} < 0.$$

Now, an increase in period t's information unambiguously decreases the weight on *prior* information. So how do these two opposing effects add up in the general case where $r, s > 0$? The following proposition establishes that the decrease in weight on prior information always dominates all increases in future weights.

Proposition 6. *$\Lambda_{t-r,t+s}$ is decreasing in $\tau_{\upsilon,t}$ for all $r, s > 0$.*

Proof. In the appendix. □

Because τ_υ is decreased in financial crises, Proposition 6 implies that financial crises are inherently persistent. Intuitively, as the economy receives less news about the current state of the economy, more weight is put on prior information—which during a financial crisis is generally pessimistic. In turn, the asset market continues to constrain the real economy in future periods and, hence, continues to impede learning about $\tilde{\theta}_t$, throwing the economy into a "pessimism trap". In contrast, τ_υ increases during financial booms, making them inherently non-persistent. Moreover, as larger shocks have stronger effects on τ_υ, persistence and non-persistence are increasing when shocks are "scaled up" as in Proposition 5.

More precisely, from Lemma 2, we have that for $\theta_t = \upsilon_t = 0$,

$$\begin{aligned} b_{t+s} &= \pi_t^{-1} \times \begin{bmatrix} \tau_\eta & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \eta_t \\ \rho b_{t-1} \end{bmatrix} \times \Lambda_{t,t+s} \\ &= (\rho \hat{\pi}_{t-1})^{-1} \times \begin{bmatrix} \tau_\eta & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \eta_t \\ \rho b_{t-1} \end{bmatrix} \times \Lambda_{t-1,t+s}. \end{aligned}$$

Applying Proposition 6 then yields the formal result.

Corollary 4. *For $\theta_t = \upsilon_t = 0$, a financial shock η_t is persistent (compared to the fixed-τ_υ benchmark) if and only if $\eta_t < -\breve{b}_{t-1}/(\tau_\xi + \tau_\eta)$. Otherwise, η_t is non-persistent. Moreover, "scaling up" η_t as in Proposition 5 increases the persistence and non-persistence, respectively.*

5.4 Summary

The information-based feedback mechanism underlying our equilibrium creates two types of asymmetries. On the one hand, shocks are either amplified or de-amplified. On the other hand, shocks are also either persistent or non-persistent. In particular, our findings imply that adverse financial shocks are amplified and persistent, while positive financial shocks are de-amplified and non-persistent. Moreover, the underlying mechanisms are highly nonlinear, implying that "scaling up" a shock gives rise to more pronounced (de-)amplification and (non-)persistence, respectively. In consequence, the impact of small shocks is only little amplified and barely persistent, whereas rare adverse shocks virtually destroy the informational capacities of the real sector and thereby induce highly amplified and persistent crashes.

6 Illustration: Impulse responses to financial shocks

In this section, we illustrate our theoretical results using simulated impulse response paths to financial shocks. To highlight the *informational* role of credit constraints, we contrast the impulse responses with counterfactual paths where τ_υ is fixed at its steady state level, but credit constraints continue to *constrain* the economy. The only difference between our model and the counterfactual responses is that uncertainty is removed—just as if we were to use the "magic wand" imagined by Blanchard.[31]

6.1 Impulse responses to financial shocks

Consider the economy's response to a nonzero realization of noisy asset demand $\tilde{\eta}_t$ and, for simplicity, suppose that the economy is in its steady state prior to the arrival of the shock.[32] From (10),

$$b_{t+s} = (\tau_\eta/\pi_t)\, \Lambda_{t,t+s} \times \eta_t, \tag{11}$$

where $\{\pi_{t+s}\}$ are recursively defined by $\{\tau_{\upsilon,t+s}\}$ and $\{q_{t+s}\}$ solving (9), pinning down all other model variables. In particular, by Lemma 2:

Corollary 5. *If $\eta_t < 0$ ($>$), then for all $s \geq 0$, b_{t+s} and q_{t+s} are strictly smaller (larger) than their steady state levels in, both, the model and the counterfactual.*

Moreover, by Lemma 1:

Corollary 6. *If $\eta_t < 0$ ($>$), then for all $s \geq 0$, $\tau_{\upsilon,t+s}$ and π_{t+s} are strictly smaller (larger) than their steady state level in the model, but are constant in the counterfactual.*

Because the information-based amplification mechanisms crucially depend on the variability of τ_υ, it will be seen that this difference drives a wedge between the model and the counterfactual.

Spillovers to real sector To provide a simple closed form solution for the "real" variables of the model, consider the special case where $v(n_{it}) = n_{it}^2/2$ and $\alpha_1(\chi_t^s) = H^{-1}(\chi_t^s)$. This specification of α_1 ensures that wages will be unperturbed along the impulse response path; i.e., if $\upsilon_t = 0$, then $\psi_t = 0$, so that gatherers' labor supply equals farmers' labor demand n_t.[33] From (7), we then have that

[31] See the quote in the beginning of the introduction.

[32] As usual, we define the steady state as the situation where true productivity θ_t equals its unconditional expectation and there are no noisy perturbations ($\theta_t = \eta_t = \upsilon_t = 0$). Yet, agents are unaware of these realizations and beliefs are formed rationally. Prior expectations are undistorted ($b_{t-1} = 0$) and prior uncertainty is fixed at its stochastic steady state value given the corresponding steady state signal precision (for details see Hamilton, 1994, Ch. 13.5). To streamline the illustration in this section, we also focus on financial noise shocks. For noise that originates in the real sector (i.e., nonzero realizations of $\tilde{\upsilon}_t$), similar results hold for $\gamma_1 < \breve{\pi}_{\text{steady state}}/\mathring{\tau}_\upsilon + 1$ (see Corollary 3 for details).

[33] To see this, recall that by Lemma 1, $\tilde{\theta}_t + \tilde{\upsilon}_t = \alpha_1(\chi_t^s) + \log(q_t)$. For $\upsilon_t = 0$, $\alpha_1 = H^{-1}$ thus implies that $\chi_t^s = H(\theta_t - \log(q_t)) = \chi_t^d$. Hence, $\psi_t = 0$ by (7).

$$w_t = \frac{n_t}{1-m} = \left[\frac{q_t \exp(\chi_t^d)}{1-m}\right]^{1/2},$$

along the impulse response path. Because $\chi_t^d = H(\theta_t - \log(q_t))$ is increasing in $\theta_t - \log(q_t)$ with a slope smaller than unity, it follows that w_t and n_t are increasing in q_t. Moreover, field output is given by

$$y_t = m \int_{-\infty}^{\infty} \exp(z) \left(\min\{z, \log(q_t)\} - \log(w_t)\right) \mathrm{d}\Phi(\sqrt{\tau_\xi}\, z),$$

along the impulse response path. Substituting w_t, output y_t can be shown to be increasing in q_t, too. That is, tighter financial constraints spill over to the real sector, so that the "real" variables are decreased along the impulse response path, too:

Corollary 7. *If $\eta_t < 0$ ($>$), then for all $s \geq 0$, w_{t+s}, n_{t+s} and y_{t+s} are strictly smaller (larger) than their steady state levels in, both, the model and the counterfactual.*

Parametrization We set $m = \frac{1}{2}$, implying an equal mass of gatherers and farmers, and set $\rho = 0.98$, $\tau_\epsilon^{-0.5} = \frac{1}{3}$, and, $\tau_\xi^{-0.5} = 2$, corresponding to a persistent and predictable process for the average log-productivity $\tilde{\theta}_t$, with a strong cross-sectional dispersion of $\{\tilde{\theta}_{it}\}$. The standard deviations of market noise, $\tau_\omega^{-0.5}$ and $\tau_\eta^{-0.5}$, are set so that perturbations are high in the financial sector, $\tau_\eta^{-0.5} = 4$, and low in the real sector (i.e., the real sector is the predominant source of information to infer about business conditions). Because $\tau_\omega^{-0.5}$ only matters through $\mathrm{Var}\{\tilde{\chi}_t^d | \chi_t^s\}$, we directly set $\mathrm{Var}\{\tilde{\chi}_t^d | \chi_t^s\} = 4$, avoiding the need to specify $\Psi_{\tilde{\theta}_t, \tilde{q}_t}$. Finally, we use γ_0 and γ_1 to specify the fraction of firms that is constrained in the steady state and the relative amplitude of asset price fluctuations compared to productivity. We set γ_0, so that approximately 2.5 percent of firms are constrained in the steady state ($\gamma_0 = 4$). To emphasize the theoretical results, we pronounce the importance of financial fluctuations by setting γ_1 to 125, implying a relative amplitude that is about two to three times as high as its empirical counterpart.[34]

6.2 Amplification and persistence of financial crises

We first illustrate our results on the persistence of financial shocks. From Corollaries 1, 2, 4, and Proposition 5 it follows that:

Corollary 8. *For all $\eta_t \neq 0$ and $s \geq 0$, $\bar{b}_{t+s}$, q_{t+s}, w_{t+s}, n_{t+s}, and y_{t+s} are strictly smaller in the model economy than in the counterfactual. I.e., financial crises are more persistent than in the counterfactual, whereas financial booms are less persistent than in the counterfactual.*

In Figure 6, we plot the responses of the asset price (normalized to $\log(q_{t+s}/q_{\text{steady state}})$), employment n_{t+s}, output y_{t+s}, the fraction of constrained farmers in the economy (i.e., farmers

[34] With more conservative parameters choices for γ_1, we need unrealistically large financial shocks in order to see a notable amplification. In Section 7.3, we illustrate how small levels of risk aversion provide further amplification that substitutes for high values of γ_1, yielding similar responses for realistic values of γ_1.

with $n_{it+s} = q_{t+s}/w_{t+s}$), and the endogenous signal precision $\tau_{v,t+s}$ to an adverse (left column) and positive (right column) realization of $\tilde{\eta}_t$. In both columns, we consider a rare tail shock with a magnitude of 2.5 standard errors. With a quarterly interpretation of time, these shocks each correspond to events that occur roughly once every 40 years. As we emphasize in Section 6.4, because of the nonlinearity of our model, the large shock size considered here is crucial for the effects established in Corollary 8 to be economically significant. In all plots, the solid lines correspond to the responses in the model economy and the dashed lines correspond to the fixed-τ_v counterfactual.[35]

Consider first the case of financial crisis in the left column. Qualitatively, asset prices (row 1), employment (row 2), and output (row 3) all decline after a negative financial shock. The causal link between these responses are the financial constraints, implying that an increased fraction of farmers is financially constrained (row 4) after an initial decrease in the asset price, leading to the decline in employment and output. Note that—as known from the financial frictions literature—the contagion of the real sector is present in the counterfactual economy as well. We call this the *constraining* effect of credit constraints.

In addition to this constraining effect, our theoretical results suggest a novel *informational* effect of credit constraints. This effect results from the variations in the signal precision as seen in row 5. By fixing the precision in the counterfactual economy, the difference between our model and the counterfactual exactly amounts to this novel informational effect. As can be seen, this informational effect of credit constraints virtually shuts down the informational capacities of the real sector. This throws the economy in a "pessimism trap" that induces an amplified and highly persistent response to the considered shock. In contrast, removing uncertainty in the counterfactual, agents quickly learn about the noisy character of the crisis, and the crisis largely goes away as has been suggested by Olivier Blanchard.

6.3 De-amplification and non-persistence of financial booms

In contrast, after a positive financial shock (depicted in the right column of Figure 6), the model's response is less persistent than the counterfactual. This is because in financial booms, the real sector aggregates more information, so that agents learn faster about the bullish character of a non-fundamental boom (recall Proposition 6). Importantly, this implies that financial booms have

[35]Note that asset prices and output in the first and third row are truncated at ± 20 and 30, respectively, omitting the initial impact of the shock. The initial impact on $\log(q_t)$ amounts to -98 (49) in the model and -63 (63) in the counterfactual in the bust (boom) case. Note that the impact for the endogenous uncertainty model is higher than in the exogenous uncertainty benchmark, reflecting the initial amplification established in Corollary 1. Underlying the strong initial impact in, both, the model and the counterfactual is a dual role of financial shocks. First, a negative realization of noise traders' demand has a direct impact on the asset price in period t and, hence, also a direct impact on economic constraints of the real sector. Second, the perturbation of asset prices also plays an informational role, effecting the beliefs of agents in the economy. Because, the second effect persists over time, while the first effect only applies to period t, this leads to a "discontinuity" between initial impact and the response of the economy starting in period $t + 1$. Here, we choose to omit the initial impact to make the graphs more readable.

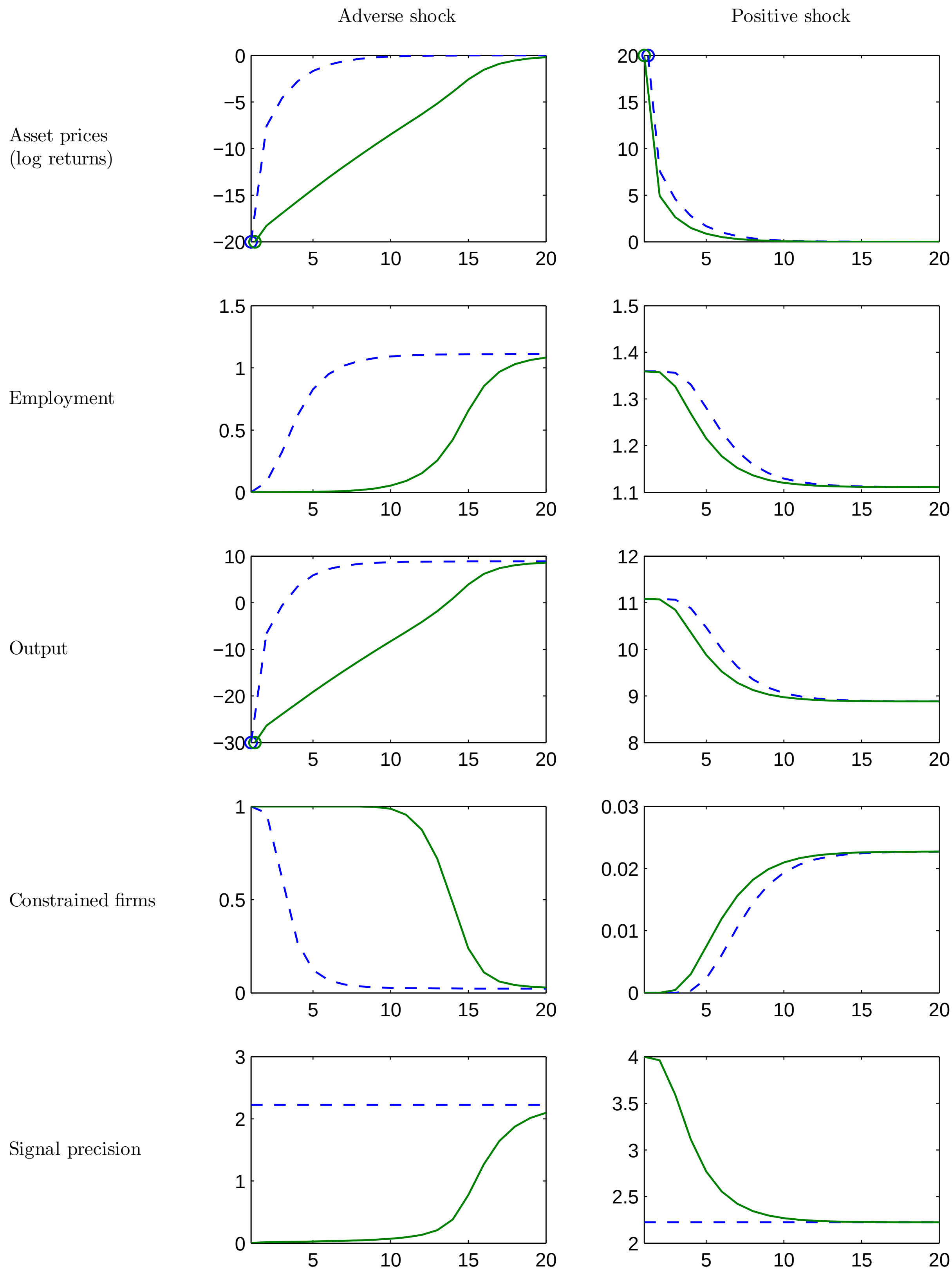

Figure 6: Impulse responses to financial shocks. *Notes:* Solid lines are model responses to shocks. Dashed lines are counterfactual responses in the exogenous uncertainty benchmark. Circles indicate that to retain readability the responses are truncated at the boundary of the graph.

necessarily smaller spillover effects on the real sector than financial crises.[36] (Note the small scale of the y-axis in the right column of Figure 6, when comparing the spillover effects to the left column).

6.4 Convexity of crises

In fact, by Proposition 5 and Corollary 4, the asymmetry between booms and busts generalizes to a non-linearity which, in particular, implies a general convexity of financial crises. The more negative a financial shock, the more information is destroyed and, hence, the more amplified and persistent is the resulting crisis. To illustrate this convexity, consider the half-life of average beliefs in our economy; i.e., the time s it takes until agents in the economy are half as pessimistic as they were at the time the shock hit the economy. In the counterfactual—as in any standard model—shocks are scale-independent, so that the half-life measure is independent of the shock size. That is, scaling up an initial shock leads to a proportional response along the whole response path, so that the *relative* realizations of beliefs across time remain unchanged. In contrast to this, the non-linearity of our model implies that larger shocks have an disproportionately severe effect, resulting in a convexity of financial crises.

Corollary 9. *Suppose the economy is in its steady state and let $T_{1/2}(\eta_t)$ denote the half-life of average beliefs in the economy along an impulse response path to a financial shock η_t; i.e., the time s it takes, such that $\bar{b}_{t+r} \leq \bar{b}_t/2$ for all $r \geq s$. Then it holds that $T_{1/2}$ is (weakly) decreasing in η_t.*

Note that because of the discrete nature of the half-life, it is necessarily locally constant almost everywhere, so that $T_{1/2}$ is only *weakly* decreasing. (Still, from our more general earlier results we know that crises are strictly more persistent as shocks get larger.)

In Figure 7, we plot the half-life in our model economy for negative financial shocks of different sizes. For our baseline simulations with a 2.5 standard error shock, the half-life is 9 quarters, while, in the conterfactual economy, any financial shock leads to a half-life of only 2 quarters.

From Proposition 5 it follows that small financial shocks, which are more frequent, have half-life periods that are similar to the exogenous uncertainty counterfactual, so that for small shocks our model behaves similar to models that only reflect the *constraining* role of credit constraints. However, in the event of a rare negative shock, the feedback loop and pessimism trap studied in Section 5 drive a significant wedge between the predictions of our model and those of the counterfactual.

[36]Nonlinearities in output and employment are responsible for parts of the asymmetry between the real variables during a financial boom compared to a financial bust. However, the responses for asset prices are necessarily perfectly symmetric in the counterfactual. This implies that the model's responses during a financial boom are bounded above by the responses of the counterfactual that are symmetric to the responses of the counterfactual during a crises. This then further implies that additionally to nonlinearities in output and employment, spillovers caused by the financial market are also necessarily less pronounced during a financial boom than during a bust.

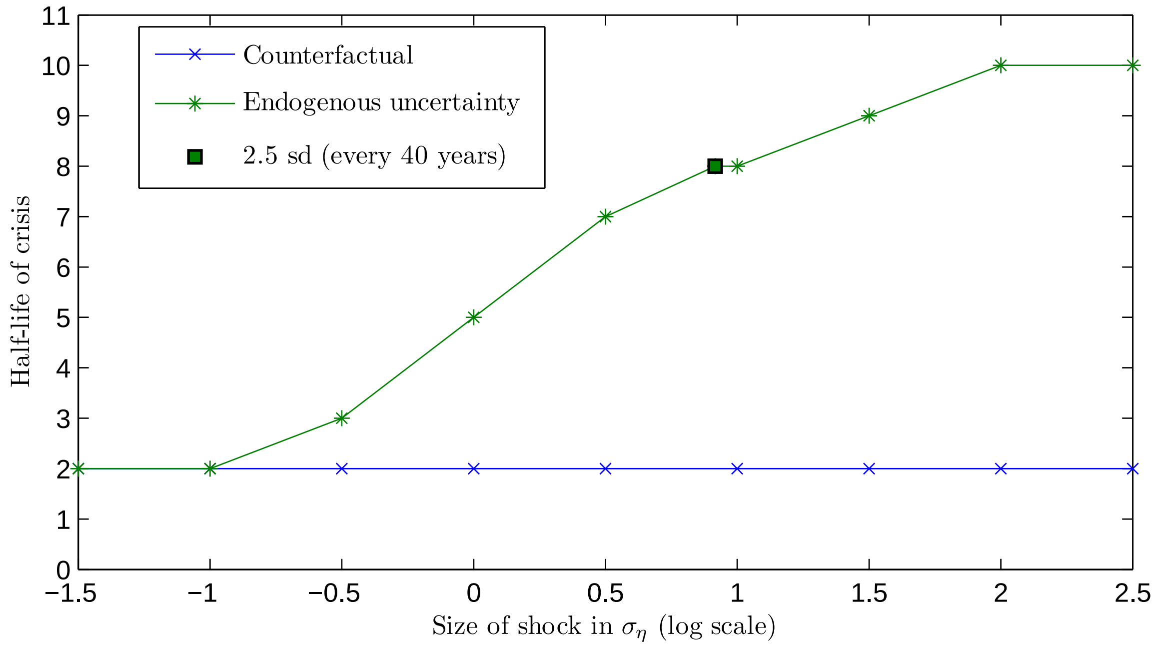

Figure 7: Half-life of adverse financial shocks.

7 Further empirical predictions

In this section, we point out some further implications of the endogeneity of the information structure. In particular, we show that common proxies of uncertainty, such as the dispersion of beliefs (as can be, for instance, measured by the survey of professional forecasters), risk premia and volatility of financial markets all increase in crisis times.

7.1 Dispersion of beliefs

One "measure of uncertainty" that is commonly used in the empirical literature is the diversity of beliefs in the economy. From Lemma 2 we know that the cross-section of beliefs is normally distributed around $\bar{b}_t$ with standard deviation $\sigma_t = \sqrt{\tau_\xi}/\bar{\pi}_t$. Here, an individual agent's uncertainty $1/\bar{\pi}_t = 1/(\pi_t + \tau_\xi)$ co-moves with the economy's (public) uncertainty $1/\pi_t$. In particular, when the economy is caught in a pessimism trap, agents increasingly refer to their own, private signals, creating a large dispersion in beliefs. On the other hand, during financial booms, public information becomes more valuable compared to private information, which reduces the dispersion of beliefs. In sum, opinions are *aligned* in booms and *dispersed* in crises.

Corollary 10. *If $\eta_t < 0$ (>), then for all $s \geq 0$, the cross-sectional dispersion of beliefs $\sigma_{t+s} = \sqrt{\tau_\xi}/(\pi_{t+s} + \tau_\xi)$ is strictly larger (smaller) than its steady state level in the model, but is constant in the counterfactual.*

7.2 Stochastic volatility

Another commonly used measure for uncertainty is the volatility of asset prices. During credit crises, high uncertainty induces volatile asset market behavior, as seen from the perspective of an outside observer. The reason is that, as the real sector becomes less informative, agents place more weight on signals from the financial market. This means that the asset price is more exposed to financial market noise and thus subject to larger conditional volatility.[37]

Corollary 11. $\mathrm{Var}\{\log(\tilde{q}_t)|\Omega_t \setminus \eta_t\}$ *is decreasing in* π_t*. Hence, if* $\eta_t < 0$ (>)*, then for all* $s \geq 0$*, the volatility* $\mathrm{Var}\{\log(\tilde{q}_t)|\Omega_t \setminus \eta_t\}$ *is strictly larger (smaller) than its steady state level in the model, but is constant in the counterfactual.*

7.3 Risk premium

Yet another natural consequence of high uncertainty in financial crises are large risk premia.

Although agents in our model are risk-neutral with respect to $\exp(\tilde{\theta}_t)$, we can simulate any risk-attitude towards $\tilde{\theta}_t$ by shifting the realization of noisy asset demand shocks $\tilde{\eta}_t$ by a constant μ. In the baseline model, we use this feature to clear the model from any risk effects on the asset price, so that the price only reflects the first moment of agents' expectations (see also Footnote 18). Here, we apply this feature to obtain an equilibrium asset price which behaves as if agents exhibited risk preferences.

More specifically, suppose that we set μ to $\mu' + r/\tau_\xi$, where μ' denotes our baseline choice of μ that induces risk-neutral behavior, and $r > 0$. Then, the equilibrium asset price q_t reflects a risk-averse attitude towards $\tilde{\theta}_t$. In particular, one can show that in this case, all previous results hold exactly, except that q_t as defined by (8) is replaced by q_t^r, which relates to q_t as follows:

$$R_t^{-1} \equiv \frac{q_t^r}{q_t} = \exp\{-r\gamma_1\pi_t^{-1}\},$$

where R_t is the risk premium associated with each unit of the Lucas tree that is traded in t. Given this specification, Corollary 6 immediately implies that the risk premium is increased along the impulse response path.

Corollary 12. *Suppose* $r > 0$ *and* $\eta_t < 0$ (>)*. Then for all* $s \geq 0$*, the risk premium* R_t *is strictly larger (smaller) than its steady state level in the model, but is constant in the counterfactual.*

Figure 8 re-plots the economy's response to an adverse financial shock of 2.5 standard errors for $r = \frac{1}{2}$. In row 1, we plot asset prices (again, normalized as log returns) and the corresponding risk

[37] For technical reasons, we here focus on the conditional volatility of asset prices; i.e., the volatility of asset prices that is induced by noise originating in the financial market. Computing the unconditional volatility would require us to integrate out the distribution of noise in the labor market. While our quasi-Gaussian transformation is aimed at making the updating problem of agents tractable, there is no analytical solution available for computing the actual distribution of labor market noise (see also the discussion around Lemma 1).

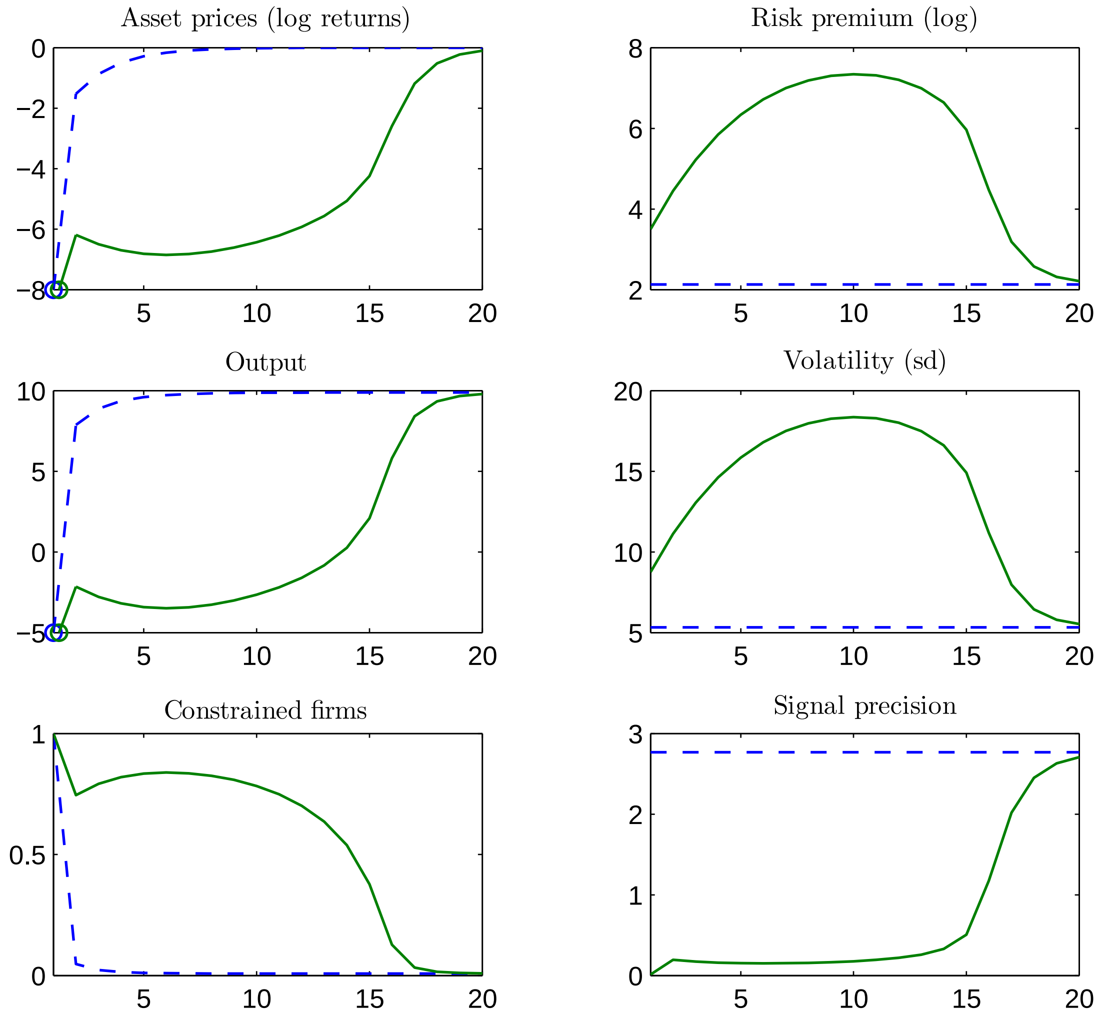

Figure 8: Volatility and risk premium response to an adverse financial shock.

premium (in logs). The response of the risk premium is hump-shaped. This is because the loss of information that results from tighter constraints on the real sector (see third row) slowly increases the posterior uncertainty. Accordingly, log assets returns in the first periods after impact mainly reflect the pessimism that is induced through learning. However, while in the case without risk aversion, prices monotonically increase as pessimism ebbs away, prices now further reflect the general uncertainty that characterizes a credit crisis in our model. Accordingly, as uncertainty increases along the crisis path, asset prices are more and more repressed by this increased uncertainty.

This additional downward pressure of prices tends to tighten financial constraints on the real sector even further and, therefore, introduces additional amplification. Note that in our numerical simulation this additional amplification is somewhat obscured by a change in the model parameters.[38]

[38]Because of the additional amplification due to risk-aversion, responses in the baseline parametrization lead to a crisis that spans more than a century. Here, we therefore reduce the value of γ_1 to 30, implying a relative amplitude of asset price fluctuations compared to productivity that is roughly in line with its empirical counterpart. We also set $\gamma_0 = 7$ in order to target again a fraction of 2.5 percent of farmers who are constrained in the steady state.

8 Concluding remarks

In this paper, we propose a novel mechanism that restricts the ability of agents to learn from the real sector during financial crises. Incorporating this idea into a dynamic macroeconomic model with a financial sector, we show that the transmission of financial shocks is inherently asymmetric and nonlinear. While fluctuations on the financial market have only little impact on the economy during "normal times", unusually adverse shocks destroy the informational capacities of the economy and therefore lead to disproportionately severe and persistent crises.

At a methodological level, we show that a combination of informational and financial frictions gives rise to a nonlinear signal structure, which explains why learning is less efficient in crisis times. Specifically, we establish that learning from "concave" signals leads to higher posterior uncertainty whenever the signal realizes in "flatter" regions. In a general theorem, this is shown to hold for a large class of information structures and to hold independently of the specific properties of our model. Equipped with these results, we then further demonstrate how learning from nonlinear signals can be incorporated within an analytically tractable conjugate Gaussian framework.

Going beyond our main results, our model also provides a number of further predictions that are in principle accessible to an empirical verification. In particular, we show that both uncertainty and also common empirical proxies of uncertainty (such as the dispersion of beliefs, the volatility of asset prices, and risk premia) increase during financial crises.

While these predictions are in line with conventional wisdom and stylized facts, a systematic empirical analysis on the causal links seems to be an important direction for future research. Another promising road along these lines is to directly examine how the informational capacity of the economy changes across different states of the world. In particular, applying the empirical methods recently developed by Coibion and Gorodnichenko (2012), a natural investigation suggested by our model is to examine the impact of credit conditions on the persistence of beliefs.

A Mathematical appendix

A.1 Proof of Proposition 1

As $\bar{q} \to \infty$, $x_{it} = \exp(\theta_{it})/w_t$ and hence $\tilde{\chi}_t^d = \log\{\int_{-\infty}^{\infty} \exp(z)\,\mathrm{d}\Phi(\sqrt{\tau_\xi}\,(z - \tilde{\theta}_t))\} - \log(\bar{q}) = \tilde{\theta}_t + 1/(2\tau_\xi) - \log(\bar{q})$. Therefore, $\mathrm{Var}\{\tilde{\theta}_t|w_t\} = \mathrm{Var}\{\tilde{\chi}_t^d|w_t\} = \mathrm{Var}\{\tilde{\chi}_t^d|\chi_t^s\}$, where the last equality follows since $\tilde{\chi}_t^s$ is a monotone transformation of $\tilde{w}_t$.

A.2 Proof of Theorem 1

We separate the proof into two steps. Lemma 3 establishes that $\mathrm{Var}\{\tilde{\theta}|s\} = \alpha_2(s)^2\,\mathrm{Var}\{\tilde{X}\}$ is increasing in s if we are in the case where $\tilde{\theta}|s = \alpha_1(s) + \alpha_2(s)\tilde{X}$ and where f is concave (the case where f is convex follows analogue). Lemma 5 establishes the corresponding results for the case where $f(\tilde{\theta})|s$ is ordered by the MLRP.

A.2.1 Location-scale distributions under concave transformations

Lemma 3. *Let $g : \mathbb{R} \to \mathbb{R}$ be a differentiable, (strictly) increasing, and (strictly) concave function and $\tilde{X}$ a square-integrable, non-degenerate random variable with mean zero over $\mathbb{R}$. Then, for any positive number $v > 0$, there exists a unique, differentiable function $\alpha_{21} : \mathbb{R} \to \mathbb{R}_{++}$ such that*

$$\mathrm{Var}\{g(\alpha_1 + \sqrt{\alpha_{21}(\alpha_1)}\tilde{X})\} = v \qquad \forall \alpha_1 \in \mathbb{R}.$$

Moreover, α_{21} is (strictly) increasing. When $g = H$, α_{21} has limits $\lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1) = (\mathrm{Var}\{\tilde{X}\})^{-1} v$, $\lim_{\alpha_1 \to \infty} \alpha_{21}(\alpha_1) = \infty$.

Proof. We only show the "strict" version of the lemma. Define $G(\alpha_1, \alpha_2) = \mathrm{Var}\{g(\alpha_1 + \sqrt{\alpha_2}\tilde{X}\} - v$. G is clearly differentiable in both arguments. We start by establishing that $G_{\alpha_1} < 0$ and $G_{\alpha_2} > 0$.

Suppose $\check{\alpha}_1 < \hat{\alpha}_1 \in \mathbb{R}$. Define $\check{g}(x) = g(\check{\alpha}_1 + x\sqrt{\alpha_2})$ and $\hat{g}(x)$ analogously. The function $\check{g}$ induces the density $F_{\check{g}(X)}(y) = F_X(\check{g}^{-1}(y))$ of $\check{g}(\tilde{X})$. Similarly we find $F_{\hat{g}(X)}(y)$. This means, $F_{\check{g}(X)}(y) = F_{\hat{g}(X)}(k(y))$, where it is straightforward to check that $k(y) \equiv \hat{g}(\check{g}^{-1}(y))$ is a differentiable contraction mapping with $k(y_0) = \mathbb{E}\{\hat{g}(X)\}$ for some y_0. In particular, k satisfies the condition in Lemma 4 and thus $G(\check{\alpha}_1, \alpha_2) > G(\hat{\alpha}_1, \alpha_2)$. This proves $G_{\alpha_1} < 0$.

For $G_{\alpha_2} > 0$, take $\check{\alpha}_2 < \hat{\alpha}_2 \in \mathbb{R}$ and define $\check{g}(x) = g(\alpha_1 + x\sqrt{\check{\alpha}_2})$ and $\hat{g}(x)$ analogously. Proceeding as before, we find that $F_{\hat{g}(X)}(y) = F_{\check{g}(X)}(k(y))$ with $k(y) = g(\alpha g^{-1}(y))$, $\alpha = \sqrt{\check{\alpha}_2/\hat{\alpha}_2} < 1$. Here, k need not be a contraction mapping in general. Still, for $y \leq y^*$ with $k(y^*) = y^*$, $k'(y) \leq \alpha < 1$ and $k(y) \leq y$ for $y > y^*$. Also, there exists a (unique) y_0 such that $k(y_0) = \mathbb{E}\{\check{g}(X)\} < g(0) = k(y^*)$. Using these two properties, the following inequalities hold,

$$|k(y) - k(y_0)| = |k(y) - k(y^*)| + |k(y^*) - k(y_0)| \leq |y - y^*| + \alpha|y^* - y_0| < |y - y_0|$$

if $y > y^*$, and

$$|k(y) - k(y_0)| \leq \alpha|y - y_0| \leq |y - y_0|$$

if $y \leq y^*$, ensuring that Lemma 4 is applicable. Consequently, $G(\alpha_1, \hat{\alpha}_2) > G(\alpha_1, \check{\alpha}_2)$ and $G_{\alpha_2} > 0$.

Now turn to the existence of α_{21}. For large α_2, $G(\alpha_{1,}\alpha_2)$ goes to infinity as can be seen by the inequality

$$\mathrm{Var}\{g(\alpha_1 + \sqrt{\alpha_2}\tilde{X}\} \geq \mathrm{Var}\{g'(\alpha_1)\sqrt{\alpha_2}\min\{\tilde{X}, 0\}\} = \alpha_2 \cdot const.$$

Also, $G(\alpha_1, 0) = -v < 0$, that is, by continuity of G and its monotonicity in α_2, a unique $\alpha_2 \equiv \alpha_{21}(\alpha_1) \in \mathbb{R}_{++}$ must exist, with $G(\alpha_1, \alpha_{21}(\alpha_1)) = 0$. By the implicit function theorem, α_{21} is continuous and differentiable with $\alpha_{21}'(\alpha_1) = -G_{\alpha_1}/G_{\alpha_2} > 0$. When $g = H$, it becomes almost linear for large negative values of α_1, i.e.

$$\begin{aligned} v = \lim_{\alpha_1 \to -\infty} \text{Var}\{H(\alpha_1 - \sqrt{\alpha_{21}(\alpha_1)}\tilde{X})\} &= \lim_{\alpha_1 \to -\infty} \text{Var}\{\alpha_1 - H(\alpha_1 - \sqrt{\alpha_{21}(\alpha_1)}\tilde{X})\} \\ &= \text{Var}\{\sqrt{\lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1)}\tilde{X}\} \\ &= \lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1)\text{Var}\tilde{X} \end{aligned}$$

and thus $\lim_{\alpha_1 \to -\infty} \alpha_{21}(\alpha_1) = (\text{Var}\tilde{X})^{-1}v$. Similarly, since $\lim_{\alpha_1 \to \infty} G(\alpha_1, \alpha_2) = -v$ for any level of α_2, $\lim_{\alpha_1 \to \infty} \alpha_{21}(\alpha_1) = \infty$. □

The main rationale behind the previous proof was to compare the variances of a random variable $\tilde{X}$ (in our case, this was $\breve{g}(\tilde{X})$ or $\hat{g}(\tilde{X})$) and its transform $k(\tilde{X})$ with k satisfying some regularity conditions. When these conditions are merely that k be a contraction mapping, the result is obvious and most textbooks (see for example,) mention it. In our case, however, a (much) more general result is needed since the second set of regularity conditions we derive in Lemma 3 clearly allow for functions with slopes larger than one.

Lemma 4. *Let $\tilde{X}, \tilde{Y}$ be two random variables over an interval $I \subseteq \mathbb{R}$ with cumulative distributions functions F_X, F_Y, and $k : I \to \mathbb{R}$ be an increasing function such that $F_Y(x) = F_X(k(x))$ and $|k(x) - \mathbb{E}\{\tilde{X}\})| \leq |x - x_0|$ for all $x \in I$ and a given $x_0 \in I$. Then,* $\text{Var}\{\tilde{X}\} \leq \text{Var}\{\tilde{Y}\}$. *If $|k(x) - \mathbb{E}\{\tilde{X}\}| < |x - x_0|$ somewhere in the support of $\tilde{X}$ the inequality is strict.*

Proof. We only show the "strict" version of the lemma. Without loss of generality we may assume $\mathbb{E}\{\tilde{X}\} = 0$, $I = \mathbb{R}$ (proof is analogous for any other interval) and $x_0 = 0$ ($\text{Var}\{\tilde{Y}\}$ is invariant under shifts $x \mapsto x + x_0$). First note, that we can restrict our attention to functions $k = k_+$ with $k_+(x) = x$ for all $x \leq 0$. This is the case since any k with $k(0) = 0$ can be split up into two functions, $k_-$and k_+, which are just the identity on the positive (negative) side of zero and k on the other. We recover the result for general k by constructing an intermediate random variable $\tilde{Y}_+$with cdf $F_{Y_+}(x) = F_X(k_+(x))$. The intermediate random variable $\tilde{Y}$ has then mean $\mathbb{E}\{\tilde{Y}\} \geq 0 = \mathbb{E}\{\tilde{X}\}$ and variance $\text{Var}\{\tilde{Y}\} \geq \text{Var}\{\tilde{X}\}$. Now, $F_Y(x) = F_X(k(x)) = F_X(k_+(k_-(x))) = F_{Y_+}(k_-(x))$, and thus $F_{-Y}(x) = F_{-Y_+}(-k_-(-x))$. The function $-k_-(-\cdot)$ is the identity on the negative side of zero and the mean of $-\tilde{Y}$ is also negative. We can use the "reduced" result shown below and get $\text{Var}\{\tilde{Y}\} = \text{Var}\{-\tilde{Y}\} \geq \text{Var}\{-\tilde{Y}_+\} = \text{Var}\{\tilde{Y}_+\} \geq \text{Var}\{\tilde{X}\}$.

Second, it is sufficient to prove the result for functions k_+ which are not only the identity for negative values of x but also above a certain (possibly large) level $M > 0$. This is without loss of generality by a standard limit argument.

Third, even simpler functions may be used to accomplish the proof, namely functions $k_{z,c}$ of the form

$$k_{z,c}(x) = \begin{cases} x & x > z + c \\ z & z < x \leq z + c \\ x & x \leq z \end{cases} \qquad z > 0, c > 0.$$

These functions only slightly differ from the identity but still they are very effective in that they are the building blocks of more general functions k_+. More precisely, any $k_+$with an upper bound of $M > 0$ can be decomposed as follows,

$$k_+ = \lim_{n\to\infty} k_{k_+(Nh),Nh-k_+(Nh)} \circ \ldots \circ k_{k_+(2h),2h-k_+(2h)} \circ k_{k_+(h),h-k_+(h)}, \tag{12}$$

where $h \equiv M/N$, $N = 2^n$, and where the limit holds under the sup-norm. Hence fix $z > 0$ and consider the random variable $\tilde{Y}_c$ given by cdf $F_{Y_c}(x) = F_X(k_{z,c}(x))$. We would like to show that $g(c) \equiv \mathrm{Var}\{\tilde{Y}_c\} - \mathrm{Var}\{\tilde{X}\}$ is strictly larger than zero for $c > 0$. Clearly $g(0) = 0$. Straightforward computation shows that

$$g'(c) = 2(z + c)(\Delta - \Delta^2) + 2\Delta \int_z^{z+c} x \,\mathrm{d}F_X(x) \geq 0, \tag{13}$$

where $\Delta = F_X(z + c) - F_X(z) \geq 0$. Given our assumption that $|k(x) - \mathbb{E}\{\tilde{X}\}| < |x - x_0|$ somewhere in the support of $\tilde{X}$, the approximation in (12) will always involve functions $k_{z,c}$ such that $\Delta > 0$ and (13) is strict.[39] Therefore, $\mathrm{Var}\{\tilde{X}\} < \mathrm{Var}\{\tilde{Y}\}$. $\square$

A.2.2 MLRP distributions under concave transformations

Lemma 5. *Let $I \subseteq \mathbb{R}$ be a nonempty (and possibly unbounded) real interval, let X_1, X_2 be two random variables over I which exhibit the monotone likelihood ratio property, i.e. f_2/f_1 is increasing with f_i being the (possibly degenerate) density of X_i, and let g be an increasing continuous function defined on I.*

1. $\mathrm{Var}\{g(X_1)\} \leq \mathrm{Var}\{g(X_2)\}$ *if g is convex and* $\mathrm{Var}\{X_2\} \leq \mathrm{Var}\{X_1\} < \infty$.

2. $\mathrm{Var}\{g(X_1)\} \geq \mathrm{Var}\{g(X_2)\}$ *if g is concave and* $\mathrm{Var}\{X_1\} \leq \mathrm{Var}\{X_2\} < \infty$.

In both cases, the inequality is strict whenever f_2/f_1 is strictly increasing and g is strictly convex or concave somewhere in the support of f_2.

Proof. We restrict ourselves to $I = [a, b]$ and distributions of X_1, X_2 with finite support. It is straightforward to generalize the result first to continuous distributions and then to arbitrary, possibly unbounded intervals. Moreover, it is sufficient to prove the result solely for increasing

[39] Because $\mathrm{Var}\{\tilde{Y}\}$ is monotonic in k—the closer k is to the identity, the closer $\mathrm{Var}\{\tilde{Y}\}$ is to $\mathrm{Var}\{\tilde{X}\}$—the "strictness" does not vanish in the limit of (12). It rather becomes larger with every increase in n.

convex functions g since all other cases can be reduced to this case by flipping g either vertically or horizontally. We will prove the result for functions g of the form

$$g_{z,c}(x) = x + (c-1)(x-z)\, 1_{\{x \geq z\}} \quad \forall z \in \mathbb{R}, c \geq 1.$$

The general case follows by iteration,

$$g(x) = \lim_{N\to\infty} \left(g_{g(z_N),g'(z_N)/g'(z_{N-1})} \circ \ldots \circ g_{g(z_1),g'(z_1)/g'(z_0)}\right) \left(g'(z_0)(x-z_0) + g(z_0)\right),$$

where $z_i = a + i(b-a)/N$. To have MLRP well-defined for pairs of discrete distributions, assume X_1 and X_2 share a common support $\{x_1 < x_2 < \ldots < x_N\}$, $N \in \mathbb{N}$, and assign probability weights (p_i) and (q_i) to the respective nodes. MLRP then translates to

$$0 \leq \frac{q_1}{p_1} \leq \frac{q_2}{p_2} \leq \ldots \leq \frac{q_N}{p_N} \leq \infty.$$

Now define $H : \mathbb{R}^N \to \mathbb{R}$ by

$$H(y_1, \ldots, y_N) = \sum (q_i - p_i) y_i^2 - \left(\sum q_i y_i\right)^2 + \left(\sum p_i y_i\right)^2.$$

Note that $H((x_i)) = \text{Var}\{X_2\} - \text{Var}\{X_1\} \geq 0$. By regarding the first two derivatives with respect to c we will now show that

$$h(c) \equiv H((g_{z,(c+1)}(x_i))_i) \geq 0 \quad \forall z \in \mathbb{R}, c \geq 0.$$

Assume $z \in [x_{j-1}, x_j)$. Then, $h(c)$ is a quadratic polynomial in $c \geq 0$ with the following two derivatives at $c = 0$,

$$\begin{aligned} h'(0) &= 2\sum_{i=j}^{N} q_i(x_i - z)(x_i - \mu_2) - 2\sum_{i=j}^{N} p_i(x_i - z)(x_i - \mu_1) \\ &\equiv 2B_z - 2A_z \end{aligned} \tag{14}$$

and

$$\begin{aligned} h''(0) = 2 &\left[\sum_{i=j}^{N} q_i(x_i - z)^2 - \left(\sum_{i=j}^{N} q_i(x_i - z)\right)^2\right] - 2\left[\sum_{i=j}^{N} p_i(x_i - z)^2 - \left(\sum_{i=j}^{N} p_i(x_i - z)\right)^2\right] \\ \equiv 2&D_z - 2C_z. \end{aligned} \tag{15}$$

If we can show that both expressions are nonnegative, we are done since $h(0) = \text{Var}\{X_2\} - \text{Var}\{X_1\} \geq 0$ by assumption.

1st derivative. First note that B_z and A_z are continuous and piecewise linear in $z \in [x_1, x_N]$ with $B_{x_1} - A_{x_1} \geq 0$ and $B_{x_N} - A_{x_N} = 0$. We will show that $B_z - A_z$ is single-peaked or quasi-concave on $[x_1, x_N]$, meaning it is impossible for $B_z - A_z$ to first drop below 0 and then rise again to 0 at the right boundary x_N. More precisely, we claim that if the derivative $B'_z - A'_z$ is nonpositive at some point z, it stays nonpositive thereafter, preventing u-shaped behavior.

Suppose $z \in [x_{j-1}, x_j)$ and $B'_z \leq A'_z$, i.e.

$$-B'_z = \sum_{i \geq j} q_i(x_i - \mu_2) \geq \sum_{i \geq j} p_i(x_i - \mu_1) = -A'_z. \tag{16}$$

It is sufficient to show that $B'_z \leq A'_z$ also in the next interval $[x_j, x_{j+1})$. The rest follows by iteration. We distinguish between three cases. First, assume $q_j(x_j - \mu_2) \leq p_j(x_j - \mu_1)$. Then, by omitting the terms with $i = j$ on both sides of (16) we only increase the inequality and trivially get that $-B'_z \geq -A'_z$ for all $z \in [x_j, x_{j+1})$. Note that we are automatically in the first case whenever $\mu_2 \geq x_j \geq \mu_1$. Second, assume $q_j(x_j - \mu_2) > p_j(x_j - \mu_1)$ and $x_j > \mu_2 \geq \mu_1$. Thus, using MLRP, we have for all $i > j$,

$$\frac{q_i}{p_i} \geq \frac{q_j}{p_j} \geq \frac{x_j - \mu_1}{x_j - \mu_2} > \frac{x_i - \mu_1}{x_i - \mu_2},$$

where the last inequality follows since $x \mapsto (x - \mu_1)/(x - \mu_2)$ is decreasing when $\mu_2 > \mu_1$ (which follows from MLRP). By $x_i > x_j > \mu_2$ this is easily rearranged to $q_i(x_i - \mu_2) > p_i(x_i - \mu_1)$ for all $i > j$. After summing over all $i > j$ we obtain for $z \in [x_j, x_{j+1})$,

$$-B'_z = \sum_{i > j} q_i(x_i - \mu_2) > \sum_{i > j} p_i(x_i - \mu_1) = -A'_z.$$

Third, assume $q_j(x_j - \mu_2) > p_j(x_j - \mu_1)$ and $\mu_2 \geq \mu_1 > x_j$. In a fashion similar to before we find that for all $i < j$

$$\frac{q_i}{p_i} \leq \frac{q_j}{p_j} \leq \frac{\mu_1 - x_j}{\mu_2 - x_j} < \frac{\mu_1 - x_i}{\mu_2 - x_i},$$

and so $q_i(\mu_2 - x_i) < p_i(\mu_1 - x_i)$ for all $i < j$. Using $\sum_{i=1}^N q_i(x_i - \mu_2) = 0$ we see that for $z \in [x_{j-1}, x_j)$,

$$-B'_z = \sum_{i \geq j} q_i(x_i - \mu_2) = \sum_{i < j} q_i(\mu_2 - x_i) < \sum_{i < j} p_i(\mu_1 - x_i) = -A'_z,$$

contradicting our assumption. The third case is therefore not possible given the assumption. This concludes the proof that $h'(0) \geq 0$ for all possible z.

2nd derivative. Again, note that C_z and D_z are both continuous in $z \in [x_1, x_N]$. However, other than before, these two functions are no longer piecewise linear but piecewise quadratic and therefore require a more subtle treatment. First, note that $D_z - C_z$ is quasi-concave on each sub-interval $[x_{j-1}, x_j]$. Suppose this did not hold. This means, the constant second derivative $D''_z - C''_z$ must be positive,

$$D''_z/2 = F_2(x_{j-1})(1 - F_2(x_{j-1})) > F_1(x_{j-1})(1 - F_1(x_{j-1})) = C''_z/2, \tag{17}$$

where F_i denotes the cumulative distribution function of X_i. At the same time, the first derivatives $D'_z - C'_z$ at the left and right boundaries must be negative and positive, respectively. Let us regard the right boundary. A positive first derivative implies,

$$\lim_{z \nearrow x_j} D'_z/2 = -F_2(x_{j-1}) \sum_{i \geq j} q_i(x_i - x_j) > -F_1(x_{j-1}) \sum_{i \geq j} p_i(x_i - x_j) = \lim_{z \nearrow x_j} C'_z/2. \tag{18}$$

It is a well-known feature of MLRP that the corresponding conditional distributions $(X_i - x_j)|(X_i \geq x_j)$ satisfy MLRP again. Necessarily, the conditional means must be ordered again,[40]

$$(1 - F_2(x_{j-1}))^{-1} \sum_{i \geq j} q_i(x_i - x_j) \geq (1 - F_1(x_{j-1}))^{-1} \sum_{i \geq j} p_i(x_i - x_j).$$

Now, we multiply this nonnegative inequality with (17) to obtain,

$$F_2(x_{j-1}) \sum_{i \geq j} q_i(x_i - x_j) \geq F_1(x_{j-1}) \sum_{i \geq j} p_i(x_i - x_j),$$

contradicting (18) and establishing quasi-concavity of $D_z - C_z$ on each interval $[x_{j-1}, x_j)$.

The quasi-concavity on the intervals allows us to restrict our attention to z's that lie on the interval boundaries, i.e. it is sufficient to prove $D_j \geq C_j$ where, with slight abuse of notation, we write D_j for D_{x_j} and similarly for C_j. Still, $D_1 \geq C_1$ and $D_N = C_N$.

We will now show the claim $D_j \geq C_j$ by induction over N. It is trivial for $N = 2$. In the following, suppose it holds for $N - 1$ points in the support of X_1 and X_2. To show it for N points, assume the contrary is true, namely there exists a j such that $C_j > D_j$ while at the same time $\text{Var}\{X_2\} \geq \text{Var}\{X_1\}$. We will try to increase $\text{Var}\{X_1\}$ relative to $\text{Var}\{X_2\}$ as much as possible but eventually see that it is not possible that the former exceeds the latter.

Define $\delta = x_2 - x_1$ and rewrite

$$C_1 = \text{Var}\{X_1\} = \delta^2(p_1 - p_1^2) + 2\delta p_1 s_1 + C_1^0,$$

where $s_1 = \sum_{i>3} p_i(x_i - x_2)$. C_1^0 is the variance of X_1 when we collapse x_1 and x_2 by reducing their distance δ to zero. Similar results emerge for $D_1 = \text{Var}\{X_2\}$, s_2 and D_1^0. By the induction hypothesis we must have $C_1^0 > D_1^0$, otherwise D_j would have to be larger than C_j. Note that s_1 and s_2 are independent of the actual levels of p_1 and p_2. From Lemma 6 we can infer that $p_1(1-p_1) \geq q_1(1-q_1)$ (otherwise we already have our contradiction $D_j \geq C_j$), thus for C_1 to be possibly larger than D_1, we would certainly need $s_2 > s_1$.

Now, let us study how these expressions change if we shift mass from x_2 to x_1. Since p_2 does not

[40]Note that $1 - F_2(x_{j-1})$ is precisely $\sum_{i \geq j} q_i$ and the same holds for F_1.

enter C_1 directly,[41] a mass increase towards x_1 by one percent yields

$$p_1 \frac{\partial C_1}{\partial p_1} = C_1 - C_1^0 - \delta^2 p_1^2 \geq 0 \tag{19}$$
$$q_1 \frac{\partial D_1}{\partial q_1} = D_1 - D_1^0 - \delta^2 q_1^2. \tag{20}$$

Here, the positivity of the bottom derivative relies on the fact that $q_1 \leq 1/2$ (which is true if $p_1 - p_1^2 \geq q_1 - q_1^2$ and $p_1 \geq q_1$) and therefore $\delta^2 q_1(1-2q_1) \geq 0$. In virtue of the positivity, we increase q_1 to make D_1 as large as possible without violating MLRP, i.e. we set

$$q_1 = p_1 \chi \equiv p_1 \frac{1 - \sum_{i>2} q_i}{1 - \sum_{i>2} p_i}.$$

The ratio χ here guarantees that $p_1/q_1 = p_2/q_2$. In other words, we used the positivity of (19) to reduce the two degrees of freedom of the two mass shifts to one. Precisely by percentage shifts up or down, we can now control how much weight *both* distributions lay on x_1, thereby keeping our equated MLRP condition $p_1/q_1 = p_2/q_2$ intact. The overall effect of these remaining percentage shifts on $D_1 - C_1$ is a comparison of (19) and (20).

First, assume $p_1 \partial C_1/\partial p_1 \geq q_1 \partial D_1/\partial q_1$. Then, adding the two inequalities $C_1^0 > D_1^0$ and $\delta^2 p_1^2 \geq \delta^2 q_1^2$ to this yields $C_1 > D_1$, a contradiction to our assumption $\text{Var}\{X_2\} \geq \text{Var}\{X_1\}$. Now, suppose $p_1 \partial C_1/\partial p_1 < q_1 \partial D_1/\partial q_1$. This means, we increase $D_1 - C_1$ by moving more weight from x_2 to x_1 until there is nothing left at x_2, i.e. $p_2 = q_1 = 0$. This is equivalent to omitting x_2, so, again, we know by the induction hypothesis that $D_1 \geq C_1$, another contradiction. In sum, we have shown that $D_j \geq C_j$ for all j whenever $\text{Var}\{X_2\} \geq \text{Var}\{X_1\}$, completing the proof. □

Lemma 6. *Let $(p_i)_{1\leq i\leq N}$, $(q_i)_{1\leq i\leq N}$, $(C_j)_{2\leq j\leq N}$, and $(D_j)_{2\leq j\leq N}$ be specified as above. If $p_1 - p_1^2 \leq q_1 - q_1^2$ then $D_j \geq C_j$ for all $j \geq 2$ and $N \geq 2$.*

Proof. Note that, since $p_1 \geq q_1$ by MLRP, p_1 must be larger than 1/2. In particular, for any $\ell \geq 1$,

$$F_1(x_\ell)(1 - F_1(x_\ell)) \leq F_2(x_\ell)(1 - F_2(x_\ell)). \tag{21}$$

If $F_2(x_\ell) \geq 1/2$, this immediately holds since $F_1(x_\ell) \geq F_2(x_\ell)$ by MLRP so both $F_2(x_\ell)$ and $F_2(x_\ell)$ are in the decreasing region of $x \mapsto x(1-x)$. Now suppose $F_2(x_\ell) < 1/2$. We know that $1/2 \geq F_2(x_\ell) \geq q_1 \geq 1 - p_1$ and all three are in the increasing region, i.e. $F_2(x_\ell)(1 - F_2(x_\ell)) \geq p_1(1-p_1) \geq F_1(x_\ell)(1-F_1(x_\ell))$, where the last inequality follows from the fact that $F_1(x_\ell) \geq p_1 \geq 1/2$, so both are in the decreasing region. This establishes (21).

The proof itself works by induction over N. The result is immediate if $N = 2$. Assume it holds for distributions with a support of $N-1$ points. We define $\delta = x_N - x_{N-1}$ and rewrite

$$C_j = \delta^2 p_N(1 - p_N) + 2\delta p_N r_1 + C_j^0,$$

[41] Notice that C_1^0 only the depends on the sum $p_1 + p_2$ and is invariant under mass shifts from one to the other.

where $r_1 = \left(\sum_{j<i<N} p_i(x_{N-1} - x_i) + (x_{N-1} - x_j)F_1(x_j)\right) \geq 0$. C_j^0 denotes the $N-1$ points analog of C_j where we collapse points x_N and x_{N-1} by reducing their distance δ to zero. Similar results emerge for D_j, r_2 and D_j^0. By the induction hypothesis we must have $D_j^0 \geq C_j^0$. Note that r_1 and r_2 are independent of the actual levels of p_N and p_{N-1}. From (21) we can infer that $p_N(1-p_N) \leq q_N(1-q_N)$, thus if C_j was bigger than D_j, we would certainly need $r_1 > r_2$.

Now, let us study how these expressions change if we shift mass from x_{N-1} to x_N. Since p_{N-1} does not enter C_j directly,[42] a mass increase towards x_N by one percent yields

$$p_N \frac{\partial C_j}{\partial p_N} = C_j - C_j^0 - \delta^2 p_N^2 \geq 0 \tag{22}$$

$$q_N \frac{\partial D_j}{\partial q_N} = D_j - D_j^0 - \delta^2 q_N^2. \tag{23}$$

Here, the positivity of the upper derivative relies on the fact that $p_N \leq 1 - p_1 \leq 1/2$ and therefore $\delta^2 p_N(1-2p_N) \geq 0$. In virtue of the positivity, we increase p_N to make C_j as large as possible without violating MLRP, i.e. we set

$$p_N = q_N \chi \equiv q_N \frac{1 - \sum_{i<N-1} p_i}{1 - \sum_{i<N-1} q_i}.$$

The ratio χ here guarantees that $p_N/q_N = p_{N-1}/q_{N-1}$. In other words, we used the positivity of (22) to reduce the two degrees of freedom of the two mass shifts to one. Precisely by percentage shifts up or down, we can now control how much weight *both* distributions lay on x_N, thereby keeping our equated MLRP condition $p_N/q_N = p_{N-1}/q_{N-1}$ intact. The overall effect of these remaining percentage shifts on $D_j - C_j$ is a comparison of (22) and (23).

First, assume $p_N \partial C_j/\partial p_N \leq q_N \partial D_j/\partial q_N$. Then, adding the two inequalities $C_j^0 \leq D_j^0$ and $\delta^2 p_N^2 \leq \delta^2 q_N^2$ yields $C_j \leq D_j$ and we are done. Now, suppose $p_N \partial C_j/\partial p_N > q_N \partial D_j/\partial q_N$. This means, we increase $C_j - D_j$ by moving more weight from x_{N-1} to x_N until there is nothing left at x_{N-1}, i.e. $p_{N-1} = q_{N-1} = 0$. This is equivalent to omitting x_{N-1}, so, again, we know by the induction hypothesis that $D_j \geq C_j$. In sum, we have shown that there is no way in which C_j could exceed D_j, for any $j \geq 2$ and any $N \geq 2$. □

A.3 Proof of Proposition 2

It is straightforward to show that H is increasing concave. Hence, given Properties 1 and 2, we can apply Theorem 1. In case (i) of Property 2, the claim then follows directly from the theorem. In case (ii), the theorem yields that $\text{Var}\{\tilde{\chi}_t^m | \chi_t^s\}$ is increasing in χ_t^s, implying that $\text{Var}\{\tilde{\theta}_t | \chi_t^s\} = \text{Var}\{\tilde{\theta}_t - \log(\bar{q}) | \chi_t^s\} = \text{Var}\{\tilde{\chi}_t^m | \chi_t^s\}$ is also increasing in χ_t^s.

[42] Notice that C_j^0 only the depends on the sum $p_{N-1} + p_N$ and is invariant under mass shifts from one to the other.

A.4 Proof of Lemma 1

First, note that $\tilde{\chi}_t^m = \tilde{\theta}_t - \log(q_t)$, so that given a flat prior[43] over $\tilde{\theta}_t$, the posterior belief $\tilde{\theta}_t|\chi_t^s$ is normally distributed around $\alpha_1(\chi_t^s) + \log(\bar{q})$ with variance $\alpha_2(\chi_t^s)^2$ by Property 2 (ii). This is exactly the same posterior belief a Bayesian updater would hold after observing a Gaussian signal $\tilde{\theta}_t + \tilde{\upsilon}_t$ with realization $\alpha_1(\chi_t^s) + \log(\bar{q})$ and variance $\tau_\upsilon^{-1} = \alpha_2(\chi_t^s)^2 \equiv \alpha_{21}(\alpha_1(\chi_t^s)) = \alpha_{21}(\theta_t + \upsilon_t - \log(\bar{q}))$. Note that τ_υ is increasing and has the desired limiting properties by Lemma 3. Thus, an observer with a flat prior updates his information given χ_t^s *as if* the signal he receives is Gaussian with a constant variance that happens to be τ_υ^{-1}. The crucial step is now to show that given this informational equivalence holds for a flat prior distribution of $\tilde{\theta}_t$, it continues to hold for *any* normal prior over $\tilde{\theta}_t$.

Suppose an observer holds a normal prior over $\tilde{\theta}_t$ as given by a pdf $p(\theta) = \phi_{\theta_0, \tau_0^{-1}}(\theta)$ and receives some signal s with pdf $q(s|\theta)$ such that he would have updated to an s-dependent normal posterior $p_0(\theta|s) = \phi_{\mu(s), \tau(s)^{-1}}(\theta)$ had he held a flat prior $p_0(\theta) = 1$ over $\tilde{\theta}_t$. This means,

$$p_0(\theta|s) = \frac{q(s|\theta)}{\int q(s|z)\,\mathrm{d}z}.$$

Therefore, the updated posterior pdf $p(\theta|s)$ given a normal prior can be written as

$$p(\theta|s) = \frac{q(s|\theta)p(\theta)}{\int q(s|z)p(z)\,\mathrm{d}z} = \frac{p_0(\theta|s)p(\theta)}{\int p_0(z|s)p(z)\,\mathrm{d}z},$$

which is just a normal pdf with mean $(\tau_0 + \tau(s))^{-1}(\tau_0\theta_0 + \tau(s)\mu(s))$ and variance $(\tau_0 + \tau(s))^{-1}$. This is exactly the posterior distribution a Bayesian updater infers from observing the realization $\mu(s)$ of a Gaussian signal $\tilde{\theta}_t + \tilde{\upsilon}_t$ where $\tau_\upsilon = \tau(s)$.

A.5 Proof of Lemma 2

To verify the fixed point note that given the law of motion (8), $\tilde{q}_t$ is informationally equivalent to a signal

$$\tilde{\theta}_t + \delta_\eta \tilde{\eta}_t + \delta_\upsilon \tilde{\upsilon}_t,$$

where $\delta_\eta = \frac{\tau_\xi + \tau_\eta}{\tau_\xi + \tau_\eta + \tau_\upsilon}$ and $\delta_\upsilon = \frac{\tau_\upsilon}{\tau_\xi + \tau_\eta + \tau_\upsilon}$. Straightforward application of Bayes rule yields

$$b_{it} = \bar{\pi}_t^{-1} \times \begin{bmatrix} \tau_\xi & \tau_\eta & \tau_\upsilon & \hat{\pi}_{t-1} \end{bmatrix} \times \begin{bmatrix} \theta_t + \xi_{it} \\ \theta_t + \eta_t \\ \theta_t + \upsilon_t \\ \rho b_{t-1} \end{bmatrix} \tag{24}$$

when solving the inference problem including the private signal $\tilde{s}_{it}$, and yields b_t as stated in the lemma when considering only the publicly observable history of prices. Aggregating over i,

[43] We use the flat prior merely for simplicity. The argument goes through for any normal prior with a variance larger than some (constant) upper bound.

substituting into the equilibrium price as pinned down by the marginal trader

$$q_t = \mathbb{E}\{\tilde{A}_t | b_{it}, \bar{\pi}_t\} = \exp\{\gamma_0 + \gamma_1(\sigma_t \sqrt{\tau_\xi}\,(\tilde{\eta}_t - \mu) + \bar{b}_t + \gamma_1/(2\bar{\pi}_t))\},$$

noting that (24) implies a cross-sectional variation $\sigma_t^2 = \text{Var}\{b_{it}\} = \tau_\xi/\bar{\pi}_t^2$, and using $\mu = \gamma_1/(2\tau_\xi)$ verifies that the mapping (8) is indeed a fixed point.

Uniqueness follows from following the same steps above, but leaving δ_η and δ_υ unspecified. Solving the resulting system of equations yields two solutions. The first one being the one stated in the lemma and the second one being $\delta_\eta = \delta_\upsilon = 0$. Note that the second solution implies that rational beliefs and, hence, market prices are invariant to the realization of noisy asset demand $\tilde{\eta}_t$. Therefore, $\delta_\eta = \delta_\upsilon = 0$ clearly violates market clearing for almost all realizations of $\tilde{\eta}_t$, implying uniqueness of the first solution.

A.6 Proof of Proposition 3

Please note that the notation in this proof is currently inconsistent with the notation in the main body of the paper and needs to be adjusted. We are sorry for any confusion arising from that.

Existence

To show the existence of a solution to the fixed point problem (9), we first derive the inverse of g_q. We used g_q to describe the functional form of the log-linear equilibrium asset price q_t, defined in equation (8). It can be rewritten in terms of τ_υ, $r \equiv \log(q_t)|_{\tau_\upsilon \to 0}$, and $s \equiv \gamma_1(\theta_t + \upsilon_t) + \gamma_0$,

$$\log(q_t) = g_q(\tau_\upsilon, \Omega_t) = (1 - \frac{\tau_\upsilon}{\bar{\pi}_t})r + \frac{\tau_\upsilon}{\bar{\pi}_t}s.$$

Dropping the second argument Ω_t for simplicity from now on, the inverse reads

$$g_q^{-1}(\log(q_t)) = (\bar{\pi}_t|_{\tau_\upsilon \to 0})\,\frac{\log(q_t) - r}{s - \log(q_t)}. \tag{25}$$

Evidently, g_q^{-1} is only defined for $\log(q_t)$ between r and s whenever $r \neq s$. Then, its range is found to be $[0, \infty)$, i.e. $g_\tau(\log(q_t)) - g_q^{-1}(\log(q_t))$ attains $g_\tau(r) > 0$ for $\log(q_t) = r$ and converges to $-\infty$ for $\log(q_t) \to s$. By the intermediate value theorem, there exists a value for $\log(q_t)$ such that $g_\tau(\log(q_t)) - g_q^{-1}(\log(q_t)) = 0$. For $r = s$, g_q^{-1} is not well-defined, so we cannot study the problem in terms of (9). Instead, we consider the usual form of the fixed point equation, $g_q(g_\tau(\log(q_t))) = \log(q_t)$. Obviously, since $g_q = r$, $\log(q_t) = r$ is the unique fixed point if $r = s$.

Uniqueness

Let $\delta = s - r$. For $\delta < 0$, g_q^{-1} is strictly decreasing and $g_\tau - g_q^{-1}$ is strictly increasing. Thus, a unique fixed point exists if $\delta < 0$. The case $\delta = 0$ is discussed above. Given these considerations, we see that there must be a non-empty set $\Xi \subset \mathbb{R}_+ \times \mathbb{R}^4$ characterizing all parameter constellations Ω_t that lead to unique equilibria. Bear in mind that r and s are just combinations of different components of Ω_t that we use to describe the influence of Ω_t on $g_q(\log(q_t))$. Since the set $\{\Omega_t \mid \delta \leq 0\}$ entirely lies in Ξ, we now derive bounds for the case $\delta > 0$.

Bound 1 In virtue of (25), the fixed-point equation (9) is equivalent to

$$(s - \mathring{s} - z)(1 + \frac{\tau_\upsilon(-z)}{\bar{\pi}_t|_{\tau_\upsilon \to 0}}) - s = -r, \tag{26}$$

where $\mathring{s} \equiv (s - \gamma_0)/\gamma_1$ and $z \equiv \log(q_t) - \mathring{s} \in (r - \mathring{s}, s - \mathring{s})$. This has a unique solution for any value of r if the left hand side is strictly decreasing for all $z \in (r, s)$, i.e.

$$s - \mathring{s} < f(s - \mathring{s}) \equiv \inf_{z \in (-\infty, s - \mathring{s})} \frac{\bar{\pi}_t|_{\tau_\upsilon \to 0} + \tau_\upsilon(-z)}{-\tau'_\upsilon(-z)} + z. \tag{27}$$

It is evident that the function f is increasing. Apart from that, we show in Lemma 7 below that there exists a unique bound s^* such that below s^*, $f(s) > s$ while $f(s^*) = s^*$. Note that this also implies that $f(s) > -\infty$ for any s. Using s^* the condition for uniqueness independent of r becomes $s - \mathring{s} < s^*$, or, split according to the signs of $\gamma_1 - 1$,

$$\begin{aligned} \gamma_1(\theta_t + \upsilon_t) + \gamma_0 = s < M, &\quad \text{for } \gamma_1 > 1 \\ \gamma_1(\theta_t + \upsilon_t) + \gamma_0 = s > M, &\quad \text{for } \gamma_1 < 1 \\ M_0 > 0, &\quad \text{for } \gamma_1 = 1, \end{aligned}$$

where $M = (\gamma_1 s^* - \gamma_0)/ \mid \gamma_1 - 1 \mid$ and $M_0 = \gamma_1 s^* - \gamma_0$. Interestingly, the uniqueness condition for $\gamma_1 = 1$ is independent of s, i.e. whenever $M_0 > 0$, *any* equilibrium is unique. On the other hand, for $M_0 < 0$, we always find *some* values for r such that the equilibrium is not unique, irrespective of s. Linearly redefining M establishes the desired bounds on $\theta_t + \upsilon_t$.

Bound 2 Similar to (26), the fixed-point equation (9) can be rewritten to

$$\delta \frac{\bar{\pi}_t|_{\tau_\upsilon \to 0}}{\tau_\upsilon(-z) + \bar{\pi}_t|_{\tau_\upsilon \to 0}} + z = s - \mathring{s}, \tag{28}$$

with $\delta, z, s, \mathring{s}$ as above. Now suppose $\gamma_1 \neq 1$. To find an upper bound δ^* for $s - r$ that establishes uniqueness independent of s, we set δ^* to the largest level need to have that the left hand side of (28) is increasing, i.e.

$$\delta^* \equiv \inf_{z \in \mathbb{R}} \frac{(\tau_\upsilon(-z) + \bar{\pi}_t|_{\tau_\upsilon \to 0})^2}{-\tau'_\upsilon(-z)\, \bar{\pi}_t|_{\tau_\upsilon \to 0}} > 0.$$

It now holds that, whenever $\delta = \gamma_1(\theta_t + \upsilon_t) + \gamma_0 - \log(q_t)|_{\tau_\upsilon \to 0} < \delta^*$, the left hand side of (28) is strictly increasing and hence has a unique solution. The bound δ^* is by construction the largest with this property. The inequality stated in the proposition can be derived by a linear transformation.

Let us regard the special case $\gamma_1 = 1$. Here, we can do better than δ^*. Since $s - \mathring{s}$ is constant at γ_0/γ_1, to ensure uniqueness δ must only be small enough for (28) to have a unique solution if the right hand side equals γ_0/γ_1, i.e.

$$\delta < \delta^{**} \equiv \sup\{\delta \mid (28) \text{ has a unique solution for } s - \mathring{s} = \gamma_0/\gamma_1\}.$$

Clearly, $\delta^{**} \geq \delta^* > 0$.

Lemma 7. *The following two statements hold:*

1. *As $x \to \infty$, $x\,\tau_v'(x) \to 0$.*
2. *Let f be as defined in* (27)*. Then,* $\lim_{s\to-\infty} f(s) - s > 0$ *while* $\lim_{s\to\infty} f(s) - s < 0$.

Proof. We first show part 1. Recall that $\tau_v(x) = 1/\sigma_v(x)^2$ is defined by the following implicit equation,

$$\operatorname{Var}\{H(\sigma_v(x)\tilde{X} + x)\} = \sigma_c^2, \tag{29}$$

where $\tilde{X}$ is standard normally distributed and we denote $\operatorname{Var}\{\tilde{\chi}^d \mid \chi^s\}$ by σ_c^2. Note that H is similar to a standard kink function $\mathring{H} = \min\{\cdot, 0\}$ in that they are almost equal outside the area around the kink where H is smooth while $\mathring{H}$ is not. Now, for x large enough, the probability mass of the distribution of $\sigma_v(x)\tilde{X} + x$ that is assigned to the area around the kink becomes arbitrarily small. Therefore, for large x, the solution σ_v of (29) behaves exactly like the solution $\mathring{\sigma}_v$ to the implicit equation

$$\operatorname{Var}\{\mathring{H}(\mathring{\sigma}_v(x)\tilde{X} + x)\} = \sigma_c^2. \tag{30}$$

In contrast to (29), we can analytically compute the variance on the left hand side of (30),

$$\operatorname{Var}\{\mathring{H}(\mathring{\sigma}_v(x)\tilde{X} + x)\} = \mathring{\sigma}_v(x)^2 F(x/\mathring{\sigma}_v(x)),$$

where $F(\beta) = 1 - \Phi(\beta) - \phi(\beta)^2 + \beta^2(\Phi(\beta) - \Phi(\beta)^2) - \beta\phi(\beta)(2\Phi(\beta) - 1)$. Substituting this in (30), the implicit definition for $\tau_v(x)$ reads,

$$F\left(\sqrt{\tau_v(x)\,x^2}\right) = \sigma_c^2 \tau_v(x).$$

Differentiating this with respect to x yields,

$$\frac{F'}{2}(\tau_v(x))^{-1/2}(x\,\tau_v'(x) + 2\tau_v(x)) = \sigma_c^2 \tau_v'(x). \tag{31}$$

Regard the signs on both sides of (31): $F'(\beta) = -2(\phi(\beta) - \beta(1 - \Phi(\beta)))\Phi(\beta) < 0$ due to the standard bound on the tails of Φ, $1 - \Phi(\beta) < |\beta|^{-1}\phi(\beta)$, and $\tau_v'(x) < 0$ as we saw in Proposition 2. This immediately implies that $x\,\tau_\upsilon'(x) + 2\tau_\upsilon(x)$ must be positive, or in other words,

$$0 > x\,\tau_\upsilon'(x) > -2\tau_\upsilon(x),$$

which establishes part 1 of the lemma for we know $\tau_\upsilon(x)$ converges to 0 as x tends to infinity (see Lemma 1).

In (27), we defined the function f such that

$$f(-s) = \inf_{z>-s} \frac{\bar{\pi}_t|_{\tau_\upsilon\to 0} + \tau_\upsilon(z)}{-\tau_\upsilon'(z)} - z.$$

The first claimed property $\lim_{s\to-\infty} f(s) - s = \lim_{s\to\infty} f(-s) + s > 0$ means that there exists an $\bar{s} > 0$ such that for any $s > \bar{s}, z \geq s$,

$$s > \frac{\bar{\pi}_t|_{\tau_\upsilon\to 0} + \tau_\upsilon(z)}{\tau_\upsilon'(z)} + z \iff -(z-s)\tau_\upsilon'(z) < \bar{\pi}_t|_{\tau_\upsilon\to 0} + \tau_\upsilon(-z).$$

But the left hand side of the second equation is smaller than $-z\,\tau_\upsilon'(z)$ —which we know tends to 0 from part 1—and therefore has to be smaller than $\bar{\pi}_t|_{\tau_\upsilon\to 0}$ for large values of z.

The second property $\lim_{s\to\infty} f(s) - s < 0$ is equivalent to finding a $\bar{s} > 0$ such that for any $s > \bar{s}$ there exists a $z \leq s$ for which

$$\frac{\bar{\pi}_t|_{\tau_\upsilon\to 0} + \tau_\upsilon(-z)}{-\tau_\upsilon'(-z)} + z < s \iff \bar{\pi}_t|_{\tau_\upsilon\to 0} + \tau_\upsilon(-z) - z\,\tau_\upsilon'(-z) < -s\,\tau_\upsilon'(-z).$$

Trivially, for $z = 0$ the right hand side of the second equation diverges to ∞ for large values of s while the rest remains constant. Thus we can just pick a $\bar{s} > 0$ that is large enough. This establishes the second property of part 2. □

A.7 Proof of Proposition 4

For the most part, the proof follows from the discussion in the main body of the paper. It remains to be shown that $\text{sign}\{\tau_\upsilon^1 - \mathring{\tau}_\upsilon\} = \text{sign}\{\tau_\upsilon^* - \mathring{\tau}_\upsilon\}$. Given the definition of $\mathring{\tau}_\upsilon$, τ_υ^1, and τ_υ^*, and given that g_τ is increasing (see Lemma 1), proving the claim is equivalent to showing that

$$\log(\mathring{q}_t) > -\gamma_0 \iff \log(q_t^*) > -\gamma_0. \tag{32}$$

We distinguish two cases. First, consider the case where g_q is decreasing. Then we have that

$$\log(\mathring{q}_t) = g_q(g_\tau(-\gamma_0)) > -\gamma_0 \iff g_\tau(-\gamma_0) < g_q^{-1}(-\gamma_0).$$

But given that $g_q' < 0$ and $g_\tau' > 0$, it follows that g_q^{-1} and g_τ intersect to the right of $-\gamma_0$ if and only if $g_q^{-1}(-\gamma_0) > g_\tau(-\gamma_0)$. Hence, (32) holds at the unique fixed point q_t^*.

Now consider the case where g_q is increasing. Then

$$\log(\mathring{q}_t) = g_q(g_\tau(-\gamma_0)) > -\gamma_0 \iff g_\tau(-\gamma_0) > g_q^{-1}(-\gamma_0).$$

From Lemma 2, it follows that $\lim_{x\to\gamma_0+\gamma_1(\theta_t+\upsilon_t)} g_q^{-1}(x) = \infty$, while by Lemma 1 we have that $g_\tau(x)$ is finite for all x. Hence, whenever $g_\tau(-\gamma_0) > g_q^{-1}(-\gamma_0)$, there necessarily exists an intersection between g_q^{-1} and g_τ to the right of $-\gamma_0$. Further, since by Lemma 2 $\lim_{x\to-\infty} g_\tau(x) = 0$, while $\lim_{x\to-\infty} g_q^{-1}(x) < 0$ (since $g_q(0)$ is finite by Lemma 2), we also have that there necessarily exists an intersection between g_q^{-1} and g_τ to the left of $-\gamma_0$ whenever $g_\tau(-\gamma_0) > g_q^{-1}(-\gamma_0)$. Hence, there always exists at least one fixed point q_t^* such that (32) holds. In the case considered in the main body of the text where there always exists a unique equilibrium, this concludes the proof. Moreover, when there are multiple equilibria, in a given state Ω_t, our analysis continues to apply to each equilibrium that satisfies (32) (for further details, see Appendix B.4).

A.8 Proof of Proposition 5

Let $(\theta_t + \upsilon_t, \theta_t + \eta_t) = a\mathcal{S}_t + \mathcal{O}_t \equiv a \cdot (x, y) + (0, z)$, with $z \equiv -\breve{b}_{t-1}/(\tau_\xi + \tau_\eta)$. Then, differentiating $\Delta \equiv \log(q_t^*) - \log(\mathring{q}_t) = g_q(\tau_\upsilon^*, \Omega_t) - g_q(\mathring{\tau}_\upsilon, \Omega_t)$ with respect to a yields

$$\frac{\mathrm{d}\Delta}{\mathrm{d}a} = \frac{\partial\Delta}{\partial a} + \frac{\partial g_q(\tau_\upsilon^*, \Omega_t)}{\partial\tau_\upsilon^*}\frac{\mathrm{d}\tau_\upsilon^*}{\mathrm{d}a}. \tag{33}$$

Consider the first term first. From Lemma 2, we have that

$$g_q(\tau_\upsilon^*, \Omega_t) = a\gamma_1(y(\tau_\xi + \tau_\eta) + x\tau_\upsilon^*)/\bar{\pi}^*,$$

and analogous for $\mathring{\tau}_\upsilon$. Substituting in Δ, differentiating, and rearranging, we get

$$\frac{\partial\Delta}{\partial a} = \frac{\gamma_1}{\bar{\pi}^*\mathring{\bar{\pi}}(\tau_\xi + \tau_\eta)} \times \frac{1}{\mathring{\tau}_\upsilon - \tau_\upsilon^*} \times [y + z - \delta_1(x)].$$

From the definitions of x, y, and z, $y + z - \delta_1(x)$ is positive if and only if $\theta_t + \eta_t > \delta_1(\theta_t + \upsilon_t)$. Moreover, $\mathring{\tau}_\upsilon - \tau_\upsilon^*$ is positive if and only if $\theta_t + \eta_t < \delta_2(\theta_t + \upsilon_t)$. Thus $\partial\Delta/\partial a < 0$ if and only if sign$\{\theta_t$+η_t-$\delta_1(\theta_t$+$\upsilon_t)\}$=sign$\{\theta_t$+η_t-$\delta_2(\theta_t$+$\upsilon_t)\}$.

Consider now the second term of (33). Substituting $\log(q_t^*) = g_q(\tau_\upsilon^*, \Omega_t)$ into (9) and implicit differentiating yields

$$\frac{\mathrm{d}\tau_\upsilon^*}{\mathrm{d}a} = \frac{x - \frac{\partial g_q}{\partial a}}{\frac{\partial g_q}{\partial\tau_\upsilon} - \frac{\partial g_\tau^{-1}}{\partial\tau_\upsilon}}. \tag{34}$$

As illustrated in Figure 2, $\frac{\partial g_q}{\partial\tau_\upsilon} < \frac{\partial g_\tau^{-1}}{\partial\tau_\upsilon}$ at the fixed point, so that the denominator is necessarily negative. Again, substituting for $\partial g_q(\tau_\upsilon^*, \Omega_t)/\partial a$, the numerator simplifies to

$$x - \frac{\partial g_q}{\partial a} = -\frac{\gamma_1(\tau_\xi + \tau_\eta)}{\bar{\pi}^*} \times [y + z - \delta_2^*(x)],$$

where

$$\delta_2^*(x) = (\tau_\xi + \tau_\eta)^{-1}(-\breve{b}_{t-1} + \gamma_1^{-1}(\breve{\pi}_{t-1} + (1-\gamma_1)\tau_\upsilon^*)x)$$

is defined such that $\theta_t + \eta_t < \delta_2^*(\theta_t + \upsilon_t)$ if and only if $\mathring{\tau}_\upsilon - \tau_\upsilon^* > 0$. Hence, since it also holds that $\mathring{\tau}_\upsilon - \tau_\upsilon^* > 0$ if and only if $\mathring{\tau}_\upsilon - \tau_\upsilon^1 > 0$, we have that from the definitions of x, y, and z, $-[y + z - \delta_2^*(x)]$—and, hence, the numerator of (34)—is positive if and only if $\theta_t + \eta_t < \delta_2(\theta_t + \upsilon_t)$. Taken together, we thus have that $\mathrm{d}\tau_\upsilon^*/\mathrm{d}a$ is negative if and only if $\theta_t + \eta_t < \delta_2(\theta_t + \upsilon_t)$. Moreover, in Section 5.1, we show that $\partial g_q/\partial \tau_\upsilon > 0$ if and only if $\theta_t + \eta_t < \delta_1(\theta_t + \upsilon_t)$. Hence, like the first term, the second term of (33) is negative if and only if sign$\{\theta_t$+η_t-$\delta_1(\theta_t$+$\upsilon_t)\}$=sign$\{\theta_t$+η_t-$\delta_2(\theta_t$+$\upsilon_t)\}$, completing the proof.

A.9 Proof of Proposition 6

We consider a generic information structure given by

$$s_t = H'\theta_t + \psi_t, \qquad \psi_t \sim N(0, \Psi_t),$$

where H' is a $m \times 1$-vector and Ψ_t is a positive-semidefinite, symmetric $m \times m$ matrix. Note that s_t is informationally equivalent to $\bar{s}_t = B_t H'\theta_t + \bar{\psi}_t$ with $\bar{\psi}_t \sim N(0, B_t\Psi_t B_t')$ for all invertible B_t which match the number of rows in H'. In particular, we can choose B_t, such that $B_t H' = (1, \ldots, 1)'$ and $\bar{\Psi}_t \equiv B_t \Psi_t B_t'$ is diagonal.[44]

Accordingly, suppose without loss of generality that

$$s_t = (1, \ldots, 1)'\theta_t + \psi_t, \qquad \psi_t \sim N(0, \mathrm{diag}(\tau_t)^{-1}),$$

where τ_t is a $m \times 1$-vector of strictly positive signal precisions. Then posterior beliefs at time t are given by

$$b_t = \frac{1}{\pi_t}\sum_{i=0}^{m} \tau_{t,i} s_{t,i} \qquad \text{and} \qquad \pi_t = \sum_{i=0}^{m} \tau_{t,i},$$

where the subscript i denotes the i-th element of vectors τ_t and s_t with the convention that $\tau_{t,0} = \hat{\pi}_{t-1} \equiv (\rho^2\tau_\epsilon + \pi_{t-1})^{-1}\tau_\epsilon\pi_{t-1}$ and $s_{t,0} = \rho b_{t-1}$. Accordingly, $\lambda_t = \rho\hat{\pi}_{t-1}/\pi_t$, matching exactly the definition in the special case discussed in the main body of the text.

Without loss of generality, consider a generic change of $\tau_{t,1}$ by a differential $\mathrm{d}\tau_{t,1}$. Then for

[44] To see this, let L_t be the lower Cholesky factor of Ψ_t. Then for all diagonal A_t, we have that $A_t L_t^{-1}\Psi_t L_t'^{-1}A_t'$ is diagonal. Hence, setting $A_t = \mathrm{diag}(L_t^{-1}H')^{-1}$ and defining $B_t \equiv A_t L_t^{-1}$ yields the desired result.

$r, s > 0$,

$$\frac{\mathrm{d}\Lambda_{t-r,t+s}}{\mathrm{d}\tau_{t,1}} = \Lambda_{t-r,t+s} \times \left\{ \frac{1}{\lambda_t}\frac{\partial \lambda_t}{\partial \pi_t} + \sum_{q=t+1}^{t+s} \left(\frac{1}{\lambda_q}\frac{\partial \lambda_q}{\partial \pi_{q-1}} \prod_{p=t+1}^{q-1} \frac{\partial \pi_p}{\partial \pi_{p-1}} \right) \right\},$$

or, after computing the individual terms and dividing by $\Lambda_{t-r,t+s} > 0$,

$$-\frac{1}{\pi_t} + \sum_{q=t+1}^{t+s} \left(\frac{\hat{\pi}_{q-1}}{\pi_{q-1}^2 \pi_q} \left(\rho^2 \sum_{i=1}^{m} \tau_{q,i} \right) \prod_{p=t+1}^{q-1} \left(\frac{\rho \hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 \right). \tag{35}$$

To show that (35) is negative, we proceed in two steps.

Step 1 We claim that (35) is maximized by setting $\tau_{o,1} \to \infty$ for all $o > t$. We prove this claim by proceeding recursively. For $o = s$, the term is obviously increasing in $\tau_{s,1}$ since

$$\frac{\partial}{\partial \tau_{s,1}} \left\{ \frac{\sum_{i=1}^{m} \tau_{s,i}}{\pi_s} \right\} = \frac{\partial}{\partial \tau_{s,1}} \left\{ \frac{\sum_{i=1}^{m} \tau_{s,i}}{\sum_{i=0}^{m} \tau_{s,i}} \right\} > 0.$$

Hence, suppose that $\tau_{o,1} \to \infty$ for all $o > n$. Then, $\hat{\pi}_o \to \tau_\epsilon$ and $\pi_o \to \infty$. Thus, differentiating (35) with respect to $\tau_{n,1}$, $n > t$, simplifies to

$$\begin{aligned} &\frac{\partial}{\partial \tau_{n,1}} \left\{ \sum_{q=n}^{n+1} \left(\frac{\hat{\pi}_{q-1}}{\pi_{q-1}^2 \pi_q} \left(\rho^2 \sum_{i=1}^{m} \tau_{q,i} \right) \prod_{p=t+1}^{q-1} \left(\frac{\rho \hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 \right) \right\} \\ &\qquad = \frac{\partial}{\partial \pi_n} \left\{ \frac{\pi_n - \hat{\pi}_{n-1}}{\hat{\pi}_{n-1}\pi_n} + \frac{\rho^2 \hat{\pi}_n}{\pi_n^2} \right\} \times \prod_{p=t+1}^{n} \left(\frac{\rho \hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 \\ &\qquad = \left(\frac{\pi_n - \rho^2 \hat{\pi}_n}{\pi_n^2} \right)^2 \times \prod_{p=t+1}^{n} \left(\frac{\rho \hat{\pi}_{p-1}}{\pi_{p-1}} \right)^2 > 0, \end{aligned}$$

verifying the claim.

Step 2 By step 1, it is sufficient to show that (35) is negative if $\tau_{o,1} \to \infty$ for all $o > t$. Accordingly, (35) simplifies to

$$-\frac{\pi_t - \rho^2 \hat{\pi}_t}{\pi_t^2} = -\frac{\hat{\pi}_t}{\tau_\epsilon \pi_t} < 0,$$

completing the proof.

B Supplementary material

B.1 Pooled asset market

First, we prove that in any equilibrium the market clearing price must be the same in both occurrences of the asset market. Suppose the contrary holds and the market clearing price in the first asset market is larger than the price in the second market.[45] Then, given that no informational gains are possible between the two market instances, all gatherers find it optimal to sell their asset (tree) in the first market. Yet, total asset demand is always smaller than $1 - m$ in the first market, a contradiction. Suppose now the price in the first market is smaller than the price in the second market. Consequently, all agents find it optimal not to sell assets in the first market. Again, this is incompatible with market clearing since total asset demand is always strictly positive. Thus, in any equilibrium, the two occurrences of the asset market share the same market price.

To show the equivalence of the two separated markets to one pooled market we need to show that any equilibrium in the separated markets is also an equilibrium in the pooled market and vice versa. Consider an equilibrium in the separate markets. From above we know that there is a single market clearing price equating supply and demand in both markets. This price must also be an equilibrium price in the pooled market for it obviously equates total supply and total demand. Vice versa, suppose a price is an equilibrium price in the pooled market. We construct an equilibrium candidate for the separate markets by letting all gatherers who trade in the pooled equilibrium trade in the first market and all farmers who trade in the pooled equilibrium trade in the second market. Indeed, this is an equilibrium with the pooled price since the fraction of trading agents must be the same across farmers and gatherers and nobody has an incentive to change his marketplace.

The last argument also implies that there is an equilibrium where the relevant collateral constraint for all farmers is given by $n_{it} \leq q_t/w_t$. Since selling assets on the first market and ending up constrained is dominated for all farmers by waiting and selling assets on the second market, we conclude that every separate asset market equilibrium can be represented by a pooled asset market that operates parallel to the labor market in which $n_{it} \leq q_t/w_t$ is exogenously imposed on all farmers.

B.2 Endogenous learning with alternative financial frictions

For demonstrating how endogenous learning interacts with other financial frictions, we consider four toy models. To be consistent with the structure of the model, all frictions are sited at the firm-level. However, one could also shift constraints to a separate financial sector, which then constrains the real sector depending on the state of the economy. In the following, our strategy is to set up simple versions of these alternative frictions and solve these model fragments up to a point where Theorem 1 is applicable.

Cash-in-advance constraints Consider a continuum of entrepreneurs with an investment opportunity that for an initial investment of k_i pays

$$F(\tilde{A}_i, k_i) = \tilde{A}_i \log(k_i) + \tilde{A}_i \gamma,$$

[45]Note that the price in the second market can be forecasted by the time the first market operates because (i) no exogenous information realizes between the two markets, and (ii) all information aggregated by the second market price is already aggregated by the first.

where $\log(\tilde{A}_i) \sim \mathcal{N}(\theta, 1/\tau_\xi)$. Let p denote the price per unit of investment and assume that investments have to be paid in advance using cash. For the purpose of raising cash, entrepreneurs may sell claims on the investment return on a financial market. For simplicity, assume that each entrepreneur is exogenously endowed with γ units and that claims can only be written on the return $\tilde{A}_i\gamma$ of these units (e.g., because k_i is unobservable or noncontractable). The timing is as follows:

1. Entrepreneurs choose to sell claims on any fraction $x_i \in [0,1]$ of γ on the financial market.

2. The financial market operates and yields an equilibrium price q per claim. (Each share entitles its owner to claim the return of γ units).

3. Entrepreneurs learn the realization of $\tilde{A}_i$ and decide how much to invest, subject to the cash-in-advance constraint $k_i \leq (q/p)\, x_i$.

For any reasonable specification of the financial market, it should be clear that whenever agents on the financial market and entrepreneurs start out with a common prior, then entrepreneurs optimally set $x_i = 1$. Then given p, each entrepreneur optimally sets $k_i = \min\{A_i, q\}/p$, so that aggregated demand k^d resembles (6). Without specifying the details of the capital supply side, assume that there exists a noisy pricing function $\tilde{p} = f(\tilde{k}^d, \tilde{\psi})$ that clears the market and which is increasing in both arguments ($\tilde{\psi}$ being some random variable). Given these assumption, there trivially exists a transform of $\tilde{\psi}$ which gives rise to a concave signal structure that, given the appropriate assumptions on the random variable $\tilde{\psi}$, is isomorphic to the one resulting from our baseline model.

Skin-in-the-game constraints A common generalization of the above cash-in-advance approach is to allow entrepreneurs to give out claims on profits, but assume that in order to provide the right incentives, entrepreneurs must have some "skin-in-the-game" that exogenously restricts the maximal number of shares that can be issued to $x_i \leq \bar{x} < 1$. Keeping the timing identical to our cash-in-advance setup, the difference is now that claims on the financial market are defined on expected entrepreneurs profits:

$$\mathbb{E}\{\Pi(\tilde{A}_i, x_i, \tilde{p}, q)|\mathcal{I}_j\} \equiv \mathbb{E}\{F(\tilde{A}_i, k_i^*(\tilde{A}_i, \tilde{p})) - k_i^*(\tilde{A}_i, \tilde{p})\tilde{p} + x_i q|\mathcal{I}_j\}.$$

Assuming risk-neutrality on the financial market (and some bounds on traders' asset demands that, as in our baseline setup, ensure the existence of a market clearing price), the equilibrium price will be given by the marginal trader m's expectation

$$\begin{aligned} q &= \mathbb{E}\{\Pi(\tilde{A}_i, x_i, \tilde{p}, q)|\mathcal{I}_m\} \\ &= (1-x_i)^{-1}\mathbb{E}\{F(\tilde{A}_i, k_i^*(\tilde{A}_i, \tilde{p})) - k_i^*(\tilde{A}_i, \tilde{p})\tilde{p}|\mathcal{I}_m\}. \end{aligned}$$

Suppose again that prior information ensures that entrepreneurs optimally set $x_i = \bar{x}$. For any $\bar{x} < 1$, we thus have that q amounts to a finite number that for well-behaved F, $\tilde{A}_i$ and $\tilde{p}$,[46] is increasing in $\mathbb{E}(\tilde{A}_i|\mathcal{I}_m)$. Based on the previous cash-in-advance setting, our results can therefore also be extended to such more general skin-in-the-game settings.

Costly state verification We argue on an intuitive level. As firms' internal funding decreases, standard auditing models imply that the markup over the risk-free rate increases, implying that firms invest less and are less responsive to the price of output (see, e.g., Carlstrom and Fuerst, 1997). It is straightforward to extent such frameworks to the case where the return to investments also depends on an unobserved (to the financial market) state. To fix ideas, consider the case where $F(k_i, A_i) = A_i k_i$ with $\log(A_i) \sim \mathcal{N}(\theta, \sigma^2)$. With such log-normal specification, the production function transforms to the "standard" setting where $F(k_i, \theta, \omega_i) = p\omega_i k_i$ with $p = \exp(\theta + \sigma^2/2)$ and $\omega_i \sim \mathcal{N}(-\sigma^2/2, \sigma^2)$. That is, one can absorb θ into the output price. Assuming that firms are matched to lenders with zero bargaining power, and further assuming that lenders learn θ upon matching and that lenders finance themselves through an exogenous financial market, this is equivalent to the setting at the core of Carlstrom and Fuerst (1997). In particular, higher effective interest rates imply that firms' investment choices respond less to θ as firms have less internal funds n_i for the purpose of financing k_i; i.e., as $(k_i - n_i)$ increases. Aggregating over k_i gives rise to a concave relation between k and θ. Accordingly, based on this paper's analysis the ability of any outside observer (like the financial market), who observes a noisy signal of (aggregate) k, is impeded during financial crises.

Adverse selection We argue on an intuitive level. Suppose there are two types of firms that differ in their probability of defaulting. Then for standard adverse selection setups, good firms are crowded out of the market in crisis times. But if good firms are more likely to succeed, then they will also be more respondent to any change in fundamentals that affect profits in the non-default state. This reduces overall responsiveness to the fundamental among market-financed firms during financial crises, so that based on this paper's analysis the information aggregation becomes less efficient.

B.3 Distributional assumptions

To clarify the conditions under which Theorem 1 is applicable to our model, we state them in terms of properties of the posterior distribution $\tilde{\chi}_t^d|\chi_t^s$. Here we illustrate how these properties of $\tilde{\chi}_t^d|\chi_t^s$ can be mapped into assumptions on the distribution of $\tilde{\Psi}_t$.

From $\tilde{\chi}_t^s|\chi_t^s$ to $\tilde{\Psi}_t$ Consider an arbitrary posterior density function $P(\chi_t^d|\chi_t^s)$ given a flat prior over $\tilde{\chi}_t^d$. Trivially, we can choose $P(\chi_t^d|\chi_t^s)$ to satisfy Properties 1 and 2 (an example is the quasi-Gaussian

[46]Here, we implicitly assume that as long as the "skin-in-the-game" constraint is fulfilled, entrepreneurs choose $k_i^*(\theta_i) = \arg\max\{F(\theta, k_i) - k_i p\}$. Then, we in particular require that F is increasing in θ_i and that $F - kp$ is concavely increasing in k and has an interior solution.

posterior we use throughout most of the paper). By Bayes' law, reverse engineering gives the corresponding conditions on $P(\chi_t^s|\chi_t^d)$:

$$P(\chi_t^d|\chi_t^s) = \frac{P(\chi_t^s|\chi_t^d)}{\int P(\chi_t^s|\hat{\chi}^d)\,\mathrm{d}\hat{\chi}^d}.$$

Rearranging yields

$$P(\chi_t^s|\chi_t^d) = g(\chi_t^s)P(\chi_t^d|\chi_t^s), \tag{36}$$

where g is indeterminate (i.e., arbitrary).

That is, any $P(\chi_t^s|\chi_t^d)$ that is consistent with (36) implements the chosen posterior $P(\chi_t^d|\chi_t^s)$. In particular note that because of the indeterminacy of g, there are infinite many *improper* conditional distributions $P(\chi_t^s|\chi_t^d)$ that are consistent with our assumptions on $P(\chi_t^d|\chi_t^s)$. Lets for now not worry whether there exists any g which guarantees the existence of a *proper* signal structure (but see below). Then in order to transform $P(\chi_t^s|\chi_t^d)$ into a distribution of $\tilde{\psi}_t$, recall that

$$\tilde{\chi}_t^s = \tilde{\chi}_t^d + \tilde{\psi}_t.$$

Then, from (36),

$$P(\psi_t|\chi_t^d) = P(\chi_t^s - \chi_t^d|\chi_t^d),$$

yielding the following CDF for $\tilde{\psi}_t|\chi_t^d$

$$\begin{aligned} P(\psi_t \leq z|\chi_t^d) &= \int_{-\infty}^{z+\chi_t^d} P(\chi^s|\chi_t^d)\,\mathrm{d}\chi^s \\ &= \int_{-\infty}^{z+\chi_t^d} g(\chi^s)P(\chi_t^d|\chi^s))\,\mathrm{d}\chi^s. \end{aligned} \tag{37}$$

Condition (37) defines the distributional assumptions on $\tilde{\psi}_t$ that lead to the posterior distribution $P(\chi_t^d|\chi_t^s)$. It can be seen that any consistent distribution of $\tilde{\psi}_t$ necessarily varies with the state of the world χ_t^d. But since $\tilde{\psi}_t$ is explicitly allowed to be dependent on $\tilde{\theta}_t$ and $\tilde{q}_t$ (and thus on $\tilde{\chi}_t^d = H(\tilde{\theta}_t - \log(\tilde{q}_t)))$, the above distribution is well in line with our model specifications.[47] Moreover, by setting $z \to \infty$, we see that the distribution of $\tilde{\psi}_t$ is also proper whenever $P(\chi_t^s|\chi_t^d)$ is proper.

Existence of a proper signal structure For a class of conditional distributions to be proper, we need that there exists some function g, such that:

$$\int P(\chi^s|\chi_t^d)\,\mathrm{d}\chi^s = 1,$$

[47]Because $\tilde{w}_t$ and $\tilde{n}_t^d$ can both be written as functions of $\tilde{\theta}_t$ and $\tilde{q}_t$ and $\tilde{\Psi}_t$, we can also transform $\tilde{\psi}_t$ back to $\tilde{\Psi}_t = (\exp(\tilde{\psi}_t) - 1)\tilde{n}_t^s$.

or

$$\int g(\chi^s) P(\chi_t^d|\chi^s)\, \mathrm{d}\chi^s = 1 \tag{38}$$

for all $\chi_t^d \in \text{supp}(\chi_t^d)$. This problem turns out to be quite challenging (it is equivalent to solving a Fredholm integral equation of kind 1). For the special case where Property 2b holds with $\tilde{X} \sim \mathcal{N}(0,1)$ (the baseline setup in the paper), we can verify the existence numerically. More specifically, we use an algorithm that ensures that with a probability arbitrary close to 1, the economy realizes in a state such that $P(\chi_t^s|\chi_t^d)$ integrates arbitrary close to 1.[48] Based on our algorithm, we conjecture that—even if there does not exist an exact solution—there always exists a signal structure which is in this sense "almost" proper.

In summary, we conclude that there always exists an improper signal structure in line with Properties 1 and 2. The question whether there also exists a proper signal structure is analytically unclear. However, even if there does not exist a proper signal structure, then for the quasi-Gaussian case covered in most of the paper, our numerical algorithm suggests that there always exists an "almost proper" one. This would then suggest that there also exists an almost identical proper signal structure, which does not satisfy our assumptions, but for which the results in our paper nevertheless describe an arbitrary accurate solution. Or, alternatively, that there exists an almost identical proper signal structure, which does not satisfy our assumptions, but for which the results in our paper describe the exact solution given that agents make arbitrary small errors by erroneously holding Gaussian beliefs.

B.4 Sunspot regimes

By Proposition 3, a necessary condition for sunspot regimes is that $\theta_t + \upsilon_t$ is sufficiently large (small) if $\gamma_1 > 1$ ($\gamma_1 < 1$) in absolute terms and also relative to $\theta_t + \eta_t$ (specifically, sunspots require that $\theta_t + \eta_t < \delta_1(\theta_t + \upsilon_t)$). If both of these conditions hold, then the economy could potentially (but not necessarily) be in a sunspot state ($\Omega_t \in \bar{\Xi}$). Figure 9 plots the set $\bar{\Xi}$ of sunspot regimes for the parameter set underlying Figure 3. In the figure, region E defines the set $\bar{\Xi}$. There are three equilibria in the interior of this set and two equilibria at the boundary. It can be shown that at any boundary of $\bar{\Xi}$, one of these equilibria is the continuation of the unique equilibrium outside Ξ. Moreover, by the proof of Proposition 4, there always exists at least one equilibrium within $\bar{\Xi}$ for which our analysis in Section 5.1 applies; i.e., Propositions 4 and 5 continue to hold for this equilibrium. For instance, in Figure 9, there always exists on equilibrium in the intersection of regions A and E in which $\mathring{q}_t > q_t^*$. Similarly, there always exists one equilibrium in the intersection of regions B and E in

[48]For any $\epsilon, \delta > 0$, we first define a set $A \subset \text{supp}(\tilde{\chi}_t^d)$, such that $Pr(\chi_t^d \in A) > 1 - \epsilon$. Given this set, we then ensure that

$$\left| \int g(\chi^s) P(\chi_t^d|\chi^s)\, \mathrm{d}\chi^s - 1 \right| < \delta$$

for all $\chi_t^d \in A$. Given that this condition does not need to hold for $\chi_t^d \in \text{supp}(\chi_t^d) \setminus A$, we have infinite many degrees of freedom in the tails of $P(\chi_t^s|\chi_t^d)$, which allow us to design g such that (38) holds for all $\chi_t^d \in A$ with arbitrary precision.

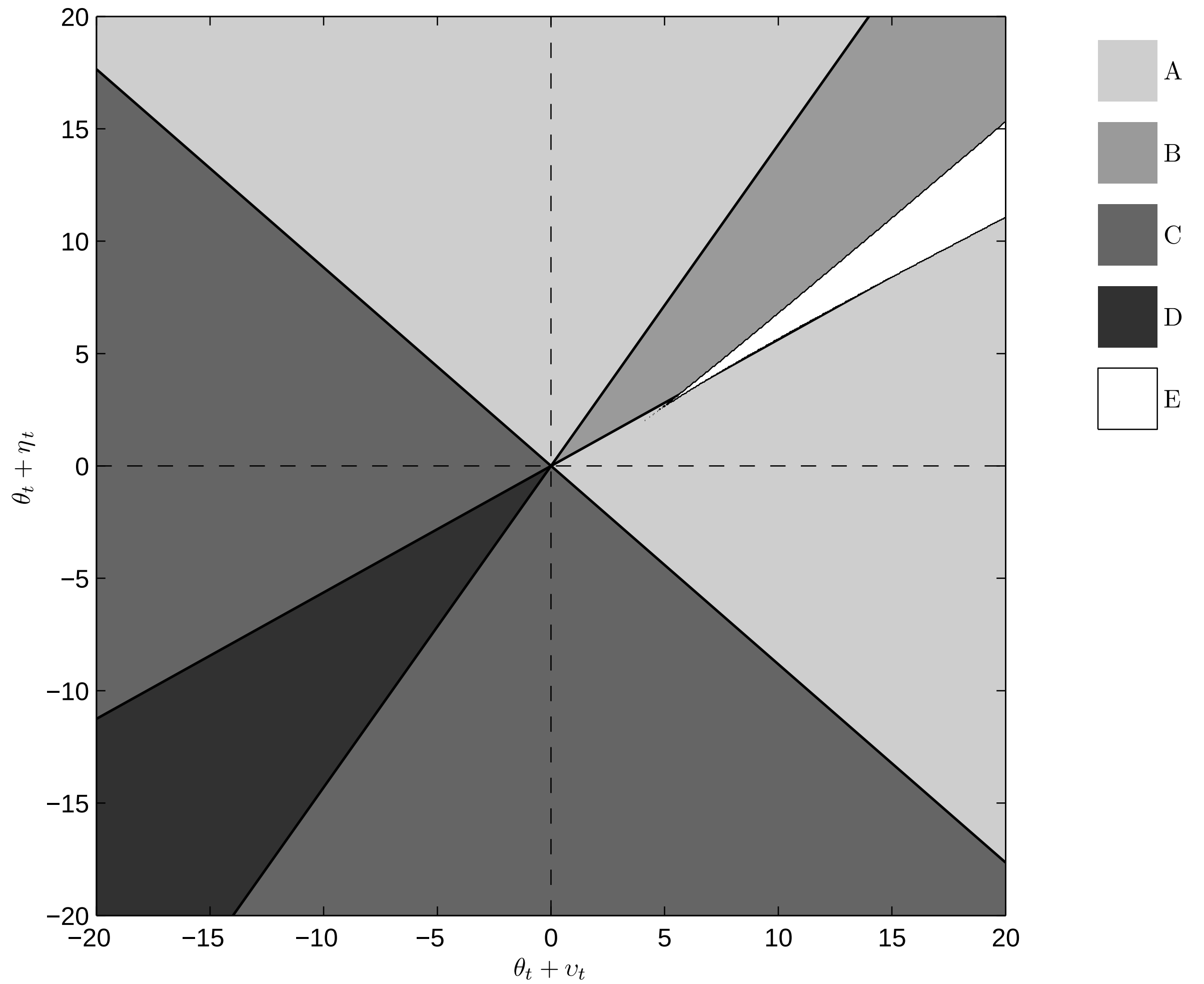

Figure 9: Sunspot regimes. *Note:* Shocks in regions A and B have an overall positive impact on $\mathring{q}_t$, shocks in regions C and D have an overall negative impact. Shocks in regions B and C are endogenously amplified, shocks in regions A and D are endogenously de-amplified. Region E defines the set $\bar{\Xi}$ of sunspot regimes.

which $q_t^* > \mathring{q}_t$.[49] For these equilibria, all results in this paper apply without any adjustment. Among the other equilibria, our results also continue to hold, but require an adjustment of the conditions that define the respective cases. For instance, for some equilibria in the intersection of regions A and E it holds that $q_t^* > \mathring{q}_t$. Accordingly, these equilibria are described by our characterization of region B rather than the one for region A. Accounting for that, all further results continue to hold.

References

Adam, Klaus. 2007. "Optimal monetary policy with imperfect common knowledge." *Journal of Monetary Economics*, 54(2): 267–301.

Albagli, Elias. 2011. "Amplification of Uncertainty in Illiquid Markets." *mimeo.*

[49]Since $\mathring{q}_t$ is uniquely pinned down by Ω_t, $\mathring{q}_t > -\gamma_0$ holds for all equilibria in E.

Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski. 2011. "A Theory of Asset Pricing Based on Heterogeneous Information." *mimeo.*

Amador, Manuel, and Pierre-Olivier Weill. 2008. "Learning from prices: Public communication and welfare." *mimeo.*

Amato, Jeffery D., and Hyun Song Shin. 2006. "Imperfect common knowledge and the information value of prices." *Economic Theory*, 27(1): 213–241.

Angeletos, George-Marios, and Alessandro Pavan. 2004. "Transparency of information and coordination in economies with investment complementarities." *The American Economic Review*, 94(2): 91–98.

Angeletos, George-Marios, and Alessandro Pavan. 2007. "Efficient use of information and social value of information." *Econometrica*, 75(4): 1103–1142.

Angeletos, George-Marios, and Alessandro Pavan. 2009. "Policy with dispersed information." *Journal of the European Economic Association*, 7(1): 11–60.

Angeletos, George-Marios, and Jennifer La'O. 2012*a*. "Optimal Monetary Policy with Informational Frictions." *mimeo.*

Angeletos, George-Marios, and Jennifer La'O. 2012*b*. "Sentiments." *mimeo.*

Angeletos, George-Marios, Guido Lorenzoni, and Alessandro Pavan. 2010. "Beauty Contests and Irrational Exuberance: A Neoclassical Approach." *mimeo.*

Bachmann, Ruediger, and Christian Bayer. 2009. "Firm-Specific Productivity Risk over the Business Cycle: Facts and Aggregate Implications." *mimeo.*

Bachmann, Ruediger, and Giuseppe Moscarini. 2011. "Business Cycles and Endogenous Uncertainty." *mimeo.*

Bernanke, Ben, and Mark Gertler. 1989. "Agency Costs , Net Worth , and Business Fluctuations." *The American Economic Review*, 79(1): 14–31.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist. 1996. "The Financial Accelerator and the Flight to Quality." *The Review of Economics and Statistics*, 78(1): 1–15.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist. 1999. "The financial accelerator in a quantitative business cycle framework." In *Handbook of Macroeconomics*. Vol. 1, , ed. John B. Taylor and Michael Woodford, 1341–1393. Amsterdam, New York and Oxford:North-Holland.

Bigio, Saki. 2011. "Endogenous Liquidity and the Business Cycle." *mimeo.*

Bigio, Saki. 2012. "Financial Risk Capacity." *mimeo.*

Bloom, Nicholas. 2009. "The Impact of Uncertainty Shocks." *Econometrica*, 77(3): 623–685.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry. 2012. "Really Uncertain Business Cycles." *mimeo.*

Boissay, Frédéric, Fabrice Collard, and Frank Smets. 2012. "Booms and Systemic Banking Crises." *mimeo.*

Brunnermeier, Markus K., and Yuliy Sannikov. 2012. "A macroeconomic model with a financial sector." *mimeo.*

Carlstrom, Charles T., and Timothy S. Fuerst. 1997. "Agency costs, net worth, and business fluctuations: A computable general equilibrium analysis." *The American Economic Review*, 87(5): 893–910.

Christiano, Lawrence, Roberto Motto, and Massimo Rostagno. 2009. "Financial Factors in Economic Fluctuations." *mimeo.*

Coibion, Olivier, and Yuriy Gorodnichenko. 2012. "Information Rigidity and the Expectations Formation Process : A Simple Framework and New Facts." *mimeo.*

Gertler, Mark, and Nobuhiro Kiyotaki. 2011. "Financial intermediation and credit policy in business cycle analysis." In *Handbook of Monetary Economics, Vol. 3A.* , ed. Benjamin M. Friedman and Michael Woodford. Amsterdam, New York and Oxford:North-Holland.

Gertler, Mark, and Peter Karadi. 2011. "A model of unconventional monetary policy." *Journal of Monetary Economics*, 58(1): 17–34.

Gilchrist, Simon, Jae W. Sim, and Egon Zakrajšek. 2010. "Uncertainty, Financial Frictions, and Investment Dynamics." *mimeo.*

Goldstein, Itay, Emre Ozdenoren, and Kathy Yuan. 2011. "Trading Frenzies and Their Impact on Real Investment." *mimeo.*

Hamilton, James D. 1994. *Time Series Analysis.* Princeton University Press.

Hart, Oliver, and John Moore. 1994. "A Theory of Debt Based on the Inalienability of Human Capital." *The Quarterly Journal of Economics*, 109(4): 841–879.

Hart, Oliver, and John Moore. 1998. "Default and Renegotiation: A Dynamic Model of Debt." *The Quarterly Journal of Economics*, 113(1): 1–41.

Hassan, Tarek A., and Thomas M. Mertens. 2011. "The social cost of near-rational investment." *mimeo.*

Hellwig, Christian, and Laura L. Veldkamp. 2009. "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies*, 76(1): 223–251.

Hellwig, Martin F. 1980. "On the aggregation of information in competitive markets." *Journal of Economic Theory*, 22(3): 477–498.

Iacoviello, Matteo. 2005. "House prices, borrowing constraints, and monetary policy in the business cycle." *The American Economic Review*, 95(3): 739–764.

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

Kiyotaki, Nobuhiro, and John Moore. 2008. "Liquidity, Business Cycles, and Monetary Policy." *mimeo.*

Krishnamurthy, Arvind. 2003. "Collateral constraints and the amplification mechanism." *Journal of Economic Theory*, 111(2): 277–292.

Kurlat, Pablo. 2010. "Lemons, Market Shutdowns and Learning." *mimeo.*

La'O, Jennifer. 2010. "Collateral Constraints and Noisy Fluctuations." *mimeo.*

Lorenzoni, Guido. 2009. "A Theory of Demand Shocks." *American Economic Review*, 99(5): 2050–2084.

Lorenzoni, Guido. 2010. "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information." *Review of Economic Studies*, 77(1): 305–338.

Lucas, Robert E. 1978. "Asset prices in an exchange economy." *Econometrica*, 46(6): 1429–1445.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99(3): 769–803.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2010. "Business cycle dynamics under rational inattention." *mimeo.*

Mertens, Thomas M. 2011. "Volatile Stock Markets: Equilibrium Computation and Policy Analysis." *mimeo.*

Milgrom, Paul R. 1981. "Good news and bad news: Representation theorems and applications." *The Bell Journal of Economics*, 12(2): 380–391.

Morris, Stephen, and Hyun Song Shin. 2002. "Social value of public information." *The American Economic Review*, 92(5): 1521–1534.

Morris, Stephen, and Hyun Song Shin. 2006. "Inertia of forward-looking expectations." *The American economic review*, 96(2): 152–157.

Ordoñez, Guillermo L. 2010. "Larger crises, slower recoveries: the asymmetric effects of financial frictions." *mimeo.*

Sim, Jae W. 2008. "Uncertainty driven business cycle." *mimeo.*

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

Sims, Christopher A. 2006. "Rational inattention: Beyond the linear-quadratic case." *The American economic review*, 96(2): 158–163.

Townsend, Robert M. 1979. "Optimal contracts and competitive markets with costly state verification." *Journal of Economic Theory*, 21(2): 265–293.

Townsend, Robert M. 1983. "Forecasting the forecasts of others." *The Journal of Political Economy*, 91(4): 546–588.

Van Nieuwerburgh, Stijn, and Laura L. Veldkamp. 2006. "Learning asymmetries in real business cycles." *Journal of Monetary Economics*, 53(4): 753–772.

Veldkamp, Laura L. 2005. "Slow boom, sudden crash." *Journal of Economic Theory*, 124(2): 230–257.

Woodford, Michael. 2003. "Imperfect common knowledge and the effects of monetary policy." In *Knowledge, Information and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps.* , ed. Philippe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford, 25–58. Princeton:Princeton University Press.

Woodford, Michael. 2009. "Information-constrained state-dependent pricing." *Journal of Monetary Economics*, 56: 100–124.

Yuan, Kathy. 2005. "Asymmetric Price Movements and Borrowing Constraints: A Rational Expectations Equilibrium Model of Crises, Contagion, and Confusion." *The Journal of Finance*, 60(1): 379–412.