Abstract

We develop a theory of endogenous uncertainty in which the ability of investors to learn about firm-level fundamentals is impaired during financial crises. At the same time, higher uncertainty reinforces financial distress. Through this two-way feedback loop, a temporary financial shock can cause a persistent reduction in risky lending, output, and employment that coincides with increased uncertainty, default rates, risk premia and disagreement among forecasters. We embed our mechanism into a standard real business cycle model and show how it manifests as an endogenous and highly internally persistent process for aggregate productivity.

Keywords: Endogenous uncertainty, financial crises, internal persistence.

JEL Classification: D83, E32, E44, G01.
1 Introduction

Financial crises often entail deep and long-lasting recessions (Reinhart and Rogoff, 2009; Hall, 2014; Ball, 2014). A common view gives a central role to uncertainty, both as an amplifier of financial distress and a source of slow recovery.\(^1\) This paper explores this idea, developing a theory that formalizes the interaction between financial constraints and uncertainty.

Our theory provides a narrative of how a temporary shock emanates from the financial sector, is reinforced by endogenously rising uncertainty, and ultimately develops into a long-lasting crisis of the real economy. The theory is consistent with a number of stylized facts from previous financial crises, such as the one in 2008/09: (i) persistently depressed employment and output; (ii) large credit spreads; (iii) a rise in default rates; (iv) an increased cross-sectional dispersion of firm sales; (v) the contemporaneous increase in measured uncertainty; and (vi) high levels of disagreement among forecasters.\(^2\)

At the core of our theory is a two-way interaction between firms’ access to external funds and information. Firms require bank loans to operate their businesses at potential. The ability of a firm to obtain funding hinges on how banks perceive the quality of its potential. The more pessimistic or uncertain banks are about firm potential, the less likely the firm obtains funding. Vice versa, when a firm is unable to operate at potential, less information about its quality is being generated, increasing uncertainty. Jointly, these two forces imply that an exogenous, but temporary, reduction in funding can morph into a persistent spiral of increased uncertainty about a firm’s potential, heightened credit risk premia, and the firm operating below its potential.

We embed this mechanism in a neoclassical general equilibrium model with a representative household and heterogeneous firms, which are funded by a competitive banking sector. We show that the amplification and internal persistence inherent in the mechanism carries over to aggregate financial shocks that hit banks’ capacity to lend (henceforth, credit shocks). Calibrating our model to U.S. data, we simulate an aggregate credit shock with a half-life

---

\(^1\)For example, Olivier Blanchard (2009) speculated at the height of the recent financial crisis that “the crisis would largely go away” if it were not for uncertainty, whereas Bloom et al. (2018) document how uncertainty was repeatedly recognized by the Federal Open Market Committee as a driver of both, the recession that followed the dot-com bubble in 2001, and the recent Great Recession. An increasing number of empirical studies further substantiates these ideas, pointing to the Great Recession being likely “an acute manifestation of the toxic interaction between uncertainty and financial shocks” (Caldara et al., 2016; see also Stein and Stone, 2013, Stock and Watson, 2012, and Gilchrist, Sim and Zakrajšek, 2016).

\(^2\)The rise in credit spreads during the 2008/09 financial crisis has been documented in, e.g., Gilchrist and Zakrajšek (2012); the rise in default rates has been documented in, e.g., Gourio (2014); the rise in sales dispersion has been documented by Bloom et al. (2018); unusually high levels of uncertainty have been documented using a variety of different approaches, including Jurado, Ludvigson and Ng (2015), Born, Breuer and Elstner (2018), and the studies cited in Footnote 1; and the increase in disagreement has been documented by Senga (2018).
of 4 quarters. We find that the persistence of the output response in our model is much greater than that, with a half-life of 22 quarters. The discrepancy is caused entirely by the interaction between endogenous uncertainty and financial frictions: when shutting down the former, the half life of the output response falls to 4 quarters, mirroring the half-life of the exogenous financial shock.

For illustrative purposes, our baseline model is stylized and does not feature capital. Nevertheless, as we demonstrate in two extensions, it is straightforward to incorporate our mechanism into richer environments. First, we explore a variant of our model, in which a fraction of firms does not rely on external funds to finance their projects. While the presence of such firms scales down the overall impact of credit shocks, we find that it changes little about their propagation through endogenous uncertainty and does not shorten the implied persistence.

Second, we extend our baseline model to include investment and capital. Interestingly, we show that our model—with its two-way interaction between lending and information and the cross-sectional heterogeneity in beliefs about firm potential—is formally equivalent to a standard real business cycle (RBC) model with an endogenous process for TFP, or an “efficiency wedge”, in the spirit of Chari, Kehoe and McGrattan (2007). The efficiency wedge arising from our mechanism is different from the ones in existing models based on financial frictions such as Buera and Moll (2015) in its internal persistence after a financial shock.

While the aggregate response to a credit shock is fully captured by an endogenous efficiency wedge, our model also has implications at the firm level. In particular, rising uncertainty helps explain a variety of financial market characteristics associated with financial crises: increased credit spreads, a rise in default rates, an increased cross-sectional dispersion of firm sales, and high levels of disagreement among forecasters about firm-level profitability.

Finally, we provide direct evidence for the link between credit access and information predicted by the model. Using a combination of firm-level survey data and accounting data to proxy for firm-level uncertainty and firms’ financial health, we compute the correlation between these measures. Supporting the main mechanism of our model, we find a significantly positive correlation between financial constraints and uncertainty.

**Related literature.** Our paper is related to a large and growing literature that introduces dispersed information into macroeconomics (e.g., Lorenzoni, 2009; Angeletos and La’O, 2010, 2013; Amador and Weill, 2010, 2012; Mackowiak and Wiederholt, 2015; Hassan and Mertens, 2014, 2017; Acharya, 2013; Hellwig and Venkateswaran, 2014; Chahrour and Gaballo, 2021). La’O (2010) shares with us the combination of dispersed information with financial frictions, but considers a static model with a constant level of uncertainty. David, Hopenhayn and
Venkateswaran (2016) also analyze information frictions as a source for factor misallocation, but focus on long-run consequences rather than fluctuations driven by financial shocks.

Our paper also contributes to a recent literature that explores the role of endogenous fluctuations in uncertainty for business cycles, including van Nieuwerburgh and Veldkamp (2006), Ordoñez (2013), and Fajgelbaum, Schaal and Tasscherau-Dumouchel (2017). In these papers, the level of aggregate investment determines the amount of information and hence aggregate uncertainty. An important distinction relative to these papers is this paper’s focus on uncertainty regarding firm-specific fundamentals rather than economic aggregates (see Senga 2018 for a similar approach). On the one hand, this allows us to explain the aforementioned stylized facts regarding the cross-sectional distribution of firm sales and investor beliefs. On the other hand, this also helps overcome a challenge in the endogenous uncertainty literature; namely, that it is often hard to generate meaningful endogenous fluctuations in uncertainty. In our model, by contrast, learning slows down when a firm is unable to obtain funding to operate at its potential, not when aggregate economic activity comes to a stand-still. Thus, even small aggregate perturbations can get severely amplified.

A second important difference to the existing endogenous uncertainty literature is that this paper links financial crises and uncertainty through a novel mechanism, explaining why high levels of uncertainty are particularly prevalent during financial crises.

In our model, the emergence of uncertainty due to financial distress interacts with the propagation of uncertainty through the financial sector. In support of such a financial transmission channel, Gilchrist, Sim and Zakrajšek (2016) present evidence that uncertainty strongly affects investment via increasing credit spreads, but has virtually no impact on investment when controlling for credit spreads. The financial transmission of uncertainty relates our model to a recent literature around Christiano, Motto and Rostagno (2014), Arellano, Bai and Keohoe (2019), and Gilchrist, Sim and Zakrajšek (2016), which stresses the importance of uncertainty or risk shocks in the financial sector, but treats these shocks as exogenous.

Ilut and Saijo (2021) propose a related mechanism based on ambiguity aversion. Studies of endogenous uncertainty in financial market settings include Veldkamp (2005), Yuan (2005), Albagli (2011), and Sockin (2017). However, none of these papers considers spillovers from financial distress on the real economy that are at the core of this paper. Benhabib, Liu and Wang (2019) and Gaballo and Marimon (2021) study settings, in which an informational interdependence between financial markets and learning gives rise to self-fulfilling fluctuations. Finally, Gorton and Ordoñez (2014) and Asriyan, Laeven and Martin (2021) study the reverse scenario in which a depletion of information makes the economy prone to fall into credit crises.

We illustrate the discrepancy between idiosyncratic and aggregate uncertainty in Section 6, where we explore a variant of our model, in which agents learn from aggregate output about aggregate productivity, as in van Nieuwerburgh and Veldkamp (2006).

We show in Appendix B that our mechanism does not cause meaningful increases in uncertainty after exogenous shocks to aggregate productivity. This is consistent with the evidence in Figure 11.

Two other related strands of the literature study the propagation of exogenous uncertainty through real
The predictions of our model are also broadly consistent with a recent empirical literature that measures the effects of tightening financial constraints. Giroud and Mueller (2017) show that establishments of firms that are more likely to be financially constrained were heavily affected by falling collateral values (house prices). In fact, they show that the entire correlation of employment loss and house prices is explained by these arguably financially constrained firms. Similar in spirit, Chodorow-Reich (2013) and Huber (2018) document that firms borrowing from less healthy lenders experience relatively steeper declines in employment during the financial crisis, consistent with the interpretation that these firms faced tighter financial constraints. Our model clarifies how an intense but relatively short-lived financial crisis can still translate into persistent financial constraints for firms, making it much harder for them to weather such periods and retain their employment and capital.

Outline The plan for the rest of the paper is as follows. The next section introduces the model economy and characterizes equilibrium. Section 3 explores our main mechanism focusing on the partial equilibrium dynamics of a single firm. Section 4 analyzes the general equilibrium response to an aggregate financial shock. Section 5 introduces heterogeneity in firms’ reliance on external funds, and explores the model with capital. Section 6 studies a variant of our model, in which uncertainty is about aggregate productivity. Section 7 tests the mechanism at the core of the paper using micro data. Section 8 concludes and offers a few policy insights.

2 Baseline Model

We study our mechanism in a neoclassical economy with a representative household, a competitive final goods sector, and a continuum of monopolistically competitive intermediate-goods firms, which are partially funded by a competitive banking sector. Time is discrete with an infinite horizon and is indexed by \( t \). To illustrate the mechanism, our baseline model is stylized. We study two extensions in Section 5.

Options as in Bloom (2009), Bloom et al. (2018), and Bachmann and Bayer (2013), and through risk premia as in the time-varying (disaster) risk literature (e.g., Gabaix, 2012; Gourio, 2012). Related to the latter, Kozlowski, Veldkamp and Venkateswaran (2020) explore a model where agents learn about tail-risks and where belief revisions after short-lived financial shocks can have long-lasting effects. Similar, Nimark (2014) presents a mechanism that increases uncertainty after rare events, if news selectively focus on outliers.


2.1 Environment

Firms. A competitive final-good sector combines intermediate goods, \( \{Y_{i,t}\}_{i \in [0,1]} \), to produce final output, \( Y_t \), using the technology

\[
Y_t = \left( \int_0^1 Y_{i,t}^{\frac{\xi - 1}{\xi}} \, di \right)^{\frac{\xi}{\xi - 1}},
\]

where \( \xi > 1 \) is the elasticity of substitution between input varieties. Profit maximization yields the demand for input \( i \) with price \( p_{i,t} \),

\[
Y_{i,t} = Y_t p_{i,t}^{-\xi},
\]

where the aggregate price index \( P_t = \left( \int_0^1 p_{i,t}^{1-\xi} \, di \right)^{1/(1-\xi)} \) has been normalized to 1.

Each input, \( i \in [0,1] \), is produced by a monopolistically competitive firm that operates a linear production technology,

\[
Y_{i,t} = A_{i,t}^{\frac{1}{\xi - 1}} L_{i,t},
\]

subject to a fixed operating cost \( \phi > 0 \) (in units of the final good). Here, the exponent on \( A_{i,t} \) is chosen so that \( A_{i,t} \) corresponds to firm \( i \)'s revenue productivity, which simplifies the exposition below. In any given period \( t \), \( A_{i,t} \) takes one of two values, \( \{A_{i,t}, \bar{A}\} \). We refer to \( A_{i,t} \) as the productivity of the risky technology (or risky productivity), and to \( \bar{A} \) as the productivity of the baseline technology (or baseline productivity). We interpret the risky technology as capturing a firm's potential.

While the baseline technology has a constant productivity \( \bar{A} > 0 \), the log productivity of the risky technology \( \log A_{i,t} \) evolves according to an AR(1) process,

\[
\log A_{i,t} = \rho \log A_{i,t-1} + (1 - \rho) \log \bar{A} + \epsilon_{i,t},
\]

with persistence \( \rho \in (0, 1) \), a long-run mean \( \log \bar{A} \), and i.i.d. (across firms and time) Gaussian innovations \( \epsilon_{i,t} \) with zero mean and variance \( \sigma^2 \).

Conditional on period-\( t \) productivities, firms choose \( p_{i,t} \) to maximize operating profits,

\[
\Pi_{i,t} \equiv p_{i,t} Y_{i,t} - w_t L_{i,t}
\]

subject to (1) and (2).

The determination of which firm produces using which of the two productivity levels is subject to two interacting frictions: a financial friction and an informational friction. We explain them next, beginning with the financial friction.
Financial friction. Each period has two sub-periods, a morning and an afternoon. In the morning, firms choose their technology subject to an information set $I_t$, which is introduced below. Importantly, $I_t$ does not contain the realization of $A_{i,t}$, which is why the “risky technology” is indeed risky. Firms are then required to finance the fixed cost $\phi$ by borrowing from banks. In the afternoon, all firms produce, all goods are sold, wages are being paid, and any loans are repaid.

We assume that a liquidity constraint prevents firms from using their revenues to pay for the fixed cost $\phi$. Instead, firms borrow from a competitive banking sector in the morning, at an interest rate $r_{i,t}$, and repay their loans in the afternoon. When a firm is unable to do so due to its profits falling short of the repayment,

$$\Pi_{i,t} < (1 + r_{i,t})\phi,$$

it defaults on its loan. We assume that in case of default, banks need to pay a cost verifying the firms’ default a la Townsend (1979), amounting to the firm’s profits $\Pi_{i,t}$. For simplicity, we assume that these costs are not resource costs and instead transfer from banks to households. If a firm defaults, it gets a bankruptcy flag that precludes it from obtaining risky loans, and thus precludes it from operating the risky technology. At the beginning of each period, bankruptcy flags are removed with an exogenous recovery probability $\eta \in (0, 1]$.

The interest rate $r_{i,t}$ compensates banks for the default risk. It is determined as the solution to the zero profit condition

$$1 + r_{i,t} = 1 + \lambda_t.$$

The left-hand side of (6) corresponds to the expected return on lending one unit: $(1 + r_{i,t})$ is the return if the loan is repaid, and $1 - \mathbb{P}_t(\Pi_{i,t} < (1 + r_{i,t})\phi)$ is the probability of repayment. The right-hand side is the cost of funds for banks and is subject to an exogenous “credit shock” $\lambda_t > -1$. We assume $\lambda_t = 0$ in steady state, corresponding to a zero cost of funds for within-period lending. When $\lambda_t$ rises above 0, this indicates that banks face an increased shadow cost of lending and are therefore required to raise lending rates. As such, our credit shock is similar to shocks to intermediary net worth, as in Gertler and Karadi (2011) and Gertler and Kiyotaki (2011).

A direct implication of the zero profit condition (6) is that when the productivity $A_{i,t}$ is known in advance, which is the case for firms operating the baseline technology, the interest

\footnote{Costly state verification can also be used to show that debt contracts are optimal in this setting.} 

\footnote{In case there are multiple solutions, $r_{i,t}$ is given by the smallest one. In case there are no solutions, a firm is unable to operate the risky technology and is forced to choose the baseline technology.}
rate is simply equal to the banks’ cost of funds; that is, \( r_{i,t} = \lambda_t \).\(^9\) For firms operating the risky technology, there is a positive risk premium, or credit spread, \( 1 + r_{i,t}^p \equiv (1 + r_{i,t})/(1 + \lambda_t) > 1 \). When \( \lambda_t \neq 0 \), the banking sector transfers its surplus \( T_t^{\text{banks}} = \lambda_t \phi \) to the representative household.

**Representative household.** The representative household maximizes expected utility over consumption \( C_t \) and labor \( L_t \),

\[
E_t \sum_{\tau=t}^{\infty} \beta^{\tau-t} \left( \frac{1}{1 - \sigma^{-1}} C_t^{1-\sigma^{-1}} - \frac{1}{1 + \zeta^{-1}} L_t^{1+\zeta^{-1}} \right),
\]

where \( \beta \in (0, 1) \) is the discount factor, \( \nu > 0 \) is a scale parameter, \( \zeta > 0 \) is the Frisch elasticity of labor supply, and \( \sigma > 0 \) is the intertemporal elasticity of substitution. The flow budget constraint is given by

\[
C_t + B_{t+1} = w_t L_t + (1 + r_{t-1})B_t + T_t.
\]

Here, \( B_{t+1} \) is the household’s end-of-period \( t \) holding of a risk-less asset in zero net supply, \( w_t \) is the real wage, \( r_t \) is the real interest rate, and \( T_t \) represents various lump-sum payments. Specifically, each period, the representative household receives firms’ profits \( \Pi_{i,t} \) (of both defaulting and non-defaulting firms) and transfers from banks \( T_t^{\text{banks}} \) if there is a financial shock \( \lambda_t \neq 0 \). Together, \( T_t \) can be written as

\[
T_t = \int_0^1 \Pi_{i,t} di - \phi.
\]

**Information friction.** We consider a simple information structure where all learning is public and there is no aggregate uncertainty; i.e., agents have complete information about the history of \( \lambda_t \) and the cross-sectional distribution over \( A_{i,t} \). The only source of uncertainty is a lack of information about the productivities of the risky technology of each individual firm. Specifically, each period, agents observe \( A_{i,t} \) for all firms adopting the risky technology. By contrast, the risky productivities \( A_{i,t} \) of firms using the baseline technology are only observed with an exogenous probability \( \theta \in [0, 1) \), independent across firms and time, and remain otherwise unknown. Let \( B_t \) denote the union of firms that adopt the risky technology in period \( t \) and for which \( A_{i,t} \) is exogenously revealed. Then the information available to agents in the morning of date \( t \) is

\[
\mathcal{I}_t = \lambda_t \cup \{A_{i,t-1}\}_{i \in B_{t-1}} \cup \mathcal{I}_{t-1}.
\]

\(^9\)Here we tacitly assume that operating the baseline technology is always profitable, which we impose more formally below.
These assumptions imply that the common belief entertained about each firm’s risky productivity is log-normal at all times, allowing us to track the public beliefs in terms of each firms’ expected log-productivity and the corresponding uncertainty,

$$\mu_{i,t} \equiv \mathbb{E}_t[\log A_{i,t}|I_t] \quad \Sigma_{i,t} \equiv \text{Var}_t[\log A_{i,t}|I_t].$$

**Timing and market clearing.** The timing of events within each period can be summarized as follows:

- **Morning:** Bankruptcy flags are removed with probability $\eta$; firms choose their technology; firms approach banks for funding and pay the fixed cost $\phi$.

- **Afternoon:** Risky productivities $A_{i,t}$ are revealed for all firms operating the risky technology, and with probability $\theta$ for all other firms; firms hire labor, produce, set prices, and repay loans; if firms are unable to repay, they default and get a bankruptcy flag; dividends and transfers are paid; the household consumes.

In equilibrium, the representative household chooses $C_t, L_t$ and $B_t$ to maximize utility (7), firms choose $p_{i,t}$ to maximize profits, banks lend if their zero profit condition can be satisfied at the competitive risk premium, and markets clear: labor markets satisfy $\int_0^1 L_{i,t} \, di = L_t$, goods markets satisfy $Y_t = C_t + \phi$, and asset markets satisfy $B_t = 0$ at all times $t$.

Below, we work with a parametrization of the model in which firms using the baseline technology always make positive profits; and in which firms that can get a bank loan for the risky technology always find it optimal to do so. The former assumption ensures that firms prefer operating the baseline technology to exiting; the latter assumption ensures that the financial friction has an impact on firms’ technology choice.

**Discussion.** Two ingredients are at the core of our model. First, firms rely, at least in part, on external finance, and access to external finance hinges on the perceived quality and risk of their potential. We model this by assuming that there is a fixed operating cost of $\phi$ that needs to be financed through loans. In this environment, more pessimistic and/or uncertain beliefs by financial markets naturally reduce access to loans, because they translate into greater default risk, raising risk premia. In our baseline model all firms have ex-ante the same reliance on external finance. Ex-post, the more productive ones have no issues

---

10We can state the former assumption formally as $A^\xi (\xi - 1)^{\xi - 1} \cdot (1 + \lambda_t) > (1 + \lambda_t) \phi$. The latter assumption is more complex as it involves future trajectories of firms. We verify that it holds numerically in our calibration.

11As we have explored in an earlier working paper version of this paper, a similar logic applies if firms are funded through equity (Straub and Ulbricht, 2018).

8
securing funding at minimal risk premia. In Section 5.1, we demonstrate that our mechanism is robust to also allowing for ex-ante heterogeneity in reliance on external funding. We do so by letting some firms to fund $\phi$ frictionlessly (e.g., due to equity, retained earnings or available safe collateral).

Second, a lack of funding leads to a lack of information about firms’ potential productivity. In our model, firms that do not operate the risky technology generate less information about its productivity $A_{i,t}$. In reality, the risky technology captures a firm’s potential, which is ex-ante unclear. The longer a firm remains underfunded, unable to reach and test its potential, the less clear it becomes how profitable it actually is. Observe that $A_{i,t}$ need not correspond to productivity in reality. It could equally well capture firm-specific demand shifters; the two are isomorphic from a modeling perspective.

### 2.2 Equilibrium Characterization

We next characterize the equilibrium in the economy, starting with the price setting choice of intermediate-goods firms.

**Price setting.** Observe that, conditional on productivity $A_{i,t}$, intermediate-goods firms solve a conventional monopolistic competition problem, charging a price with a constant markup over marginal cost,

$$p_{i,t} = \frac{\xi}{\xi - 1} \frac{w_t}{A_{i,t}^{\xi-1} \phi}.$$  

At the optimum, operating profits are then given by

$$\Pi_{i,t} = A_{i,t} \xi - \xi (\xi - 1)^{\xi-1} \frac{Y_t}{w_t^{\xi-1}}.$$  

(8)

**Technology choice.** Given our assumptions, firms always prefer to operate the risky technology as long as banks are willing to fund it, and otherwise operate the baseline technology. To see when banks are willing to fund the risky technology, we rewrite the default condition (5) using (8),

$$\log A_{i,t} < \log \left( (1 + r_{i,t}) \phi \right) - \log \left( Y_t / w_t^{\xi-1} \right) + \log \left( \xi^{\xi} (\xi - 1)^{1-\xi} \right).$$  

(9)

That is, firms with the risky technology default on their loan when $A_{i,t}$ falls below a certain threshold, which is more likely when the outstanding debt $(1 + r_{i,t})\phi$ is greater, the real wage $w_t$ is greater, and aggregate demand $Y_t$ is weaker. Using (9) together with the standard Normal cdf $\Phi(\cdot)$, we can express the probability of repayment, $1 - \mathbb{P}_t(\Pi_{i,t} < (1 + r_{i,t})\phi)$, in
terms of the belief at the time of the funding choice, $\mu_{i,t}$ and $\Sigma_{i,t}$. Substituting into (6), this gives us a zero-profit condition of

$$
\Phi\left(\frac{\mu_{i,t} - \log\left((1 + r_{i,t}) \phi\right) + \log\left(Y_t/w_i^{\xi-1}\right) - \log\left(\xi^\xi (\xi - 1)^{1-\xi}\right)}{\sqrt{\Sigma_{i,t}}}\right) = \frac{1 + \lambda_t}{1 + r_{i,t}}. \tag{10}
$$

If the zero profit condition holds for some $1 + r_{i,t} \geq 0$ (and a firm has no bankruptcy flag), the firm receives funding for the risky technology; else it does not. Reformulating (10), we obtain the following result.

**Proposition 1.** Define the lending threshold as

$$
\nu_t \equiv \log\left((1 + \lambda_t) \phi\right) - \log\left(Y_t/w_i^{\xi-1}\right) + \log\left(\xi^\xi (\xi - 1)^{1-\xi}\right).
$$

Firm $i$ obtains funding for the risky technology if and only if (i) it has no bankruptcy flag, and (ii) the beliefs $(\mu_{i,t}, \Sigma_{i,t})$ satisfy

$$
\mu_{i,t} - \mathcal{V}(\Sigma_{i,t}) \geq \nu_t \tag{11}
$$

where $\mathcal{V}(\Sigma)$ is defined as

$$
\mathcal{V}(\Sigma) \equiv \min_{x \in [0,1]} \left\{ \Phi^{-1}(x) \sqrt{\Sigma} - \log x \right\}.
$$

Banks are willing to fund all projects for which $\mu_{i,t} - \mathcal{V}(\Sigma_{i,t})$ exceeds a time-varying threshold $\nu_t$, which we henceforth refer to as lending threshold. For realistic calibrations, we have $\mathcal{V}(0) = 0$ and $\mathcal{V'} > 0$, capturing that default becomes more likely as uncertainty increases, which in turn increases risk premia and reduces the willingness of banks to lend.\textsuperscript{12}

**Belief dynamics.** The distribution of beliefs $(\mu_{i,t}, \Sigma_{i,t})$ about productivities $A_{i,t}$ across firms $i$ are crucial state variables in our economy. From (3) we can derive the law of motion

---

\textsuperscript{12}In the degenerate case where default is more likely than repayment, $\mathcal{V}$ may decrease in $\Sigma$. However, this requires levels of uncertainty that are way beyond the unconditional variance of log revenue productivity documented in the data. In line with our calibration, we focus on the realistic case where $\mathcal{V}$ is increasing throughout its support, $[\sigma^2_2, \sigma^2_2/(1 - \rho^2)]$. For details on properties of $\mathcal{V}(\Sigma)$, see Appendix A.2.
of beliefs as

\[
\begin{align*}
\mu_{i,t+1} &= \begin{cases} 
\rho \log A_{i,t} + (1 - \rho) \log \bar{A} & \text{if } i \in B_t \\
\rho \mu_{i,t} + (1 - \rho) \log A & \text{if } i \notin B_t
\end{cases} \\
\Sigma_{i,t+1} &= \begin{cases} 
\sigma_{\epsilon}^2 & \text{if } i \in B_t \\
\rho^2 \Sigma_{i,t} + \sigma_{\epsilon}^2 & \text{if } i \notin B_t.
\end{cases}
\end{align*}
\]

(12)

As long as a firm adopts the risky technology, learning about \( A_{i,t} \) is perfect; uncertainty only reflects current innovations to \( A_{i,t} \). By contrast, when a firm switches to the baseline technology, uncertainty about the productivity of its risky technology accumulates and beliefs converge towards the unconditional prior.

**General equilibrium and steady state.** Each firm \( i \) has an idiosyncratic state that is given by \( S_{i,t} \equiv (A_{i,t}, \mu_{i,t}, \Sigma_{i,t}, \delta_{i,t}) \) where \( \delta_{i,t} \in \{0, 1\} \) is firm \( i \)'s bankruptcy flag. In any given period, firm \( i \)'s output and labor demand, \( Y_{i,t} \) and \( L_{i,t} \), are functions of its state \( S_{i,t} \) as well as of the aggregates \( (\lambda_t, w_t, Y_t) \),

\[
Y_{i,t} = A_{i,t}^{\xi/(\xi-1)} \left( \frac{\xi}{\xi-1} \right)^{-\xi} \frac{Y_t}{w_t^\xi} \quad \text{and} \quad L_{i,t} = A_{i,t} \left( \frac{\xi}{\xi-1} \right)^{-\xi} \frac{Y_t}{w_t^\xi}.
\]

where \( A_{i,t} \) is firm \( i \)'s technology, determined by Proposition 1. Aggregating across firms, we find that

\[
w_t = (1 - \xi^{-1}) A_t \quad \text{and} \quad Y_t = A_t L_t
\]

(14)

where

\[
A_t \equiv \left( \int_0^1 A_{i,t} \, di \right)^{\frac{1}{1-\xi}}
\]

(15)

corresponds to the efficiency wedge in the economy, in the spirit of Chari, Kehoe and McGrattan (2007), and \( 1 - \xi^{-1} \) stems from the monopoly distortion induced by monopolistic competition. Together with the first order condition for household labor supply, \( w_t = v L_t^{1/\xi} (Y_t - \phi)^{1/\sigma} \), we find

\[
(1 - \xi^{-1}) A_t = v L_t^{1/\xi} (A_t L_t - \phi)^{1/\sigma}
\]

(16)

Conditional on \( A_t \), this equation admits a unique positive solution for \( L_t \). The solution always satisfies \( A_t L_t > \phi \). Thus, output \( Y_t \) is uniquely determined given \( A_t \).

However, in general equilibrium, the composition of the technology used by firms, and thus \( A_t \), is endogenous to \( Y_t \) and \( w_t \). Specifically, \( A_t \) is a decreasing function of the lending
threshold \( \nu_t \), which in turn is a function of \( Y_t \) and \( w_t \) (Proposition 1), capturing how in general equilibrium tighter credit constraints translate into suboptimal project choices. In Appendix A.3, we show that taking into account this feedback from \( Y_t \) and \( w_t \) to \( A_t \), (16) admits a unique fixed point if

\[ \xi - 1 > \frac{\zeta^{-1} + 1}{\zeta^{-1} + \sigma^{-1}}, \]

which is easily satisfied in our calibration below. Conditional on (17), and starting with a given initial distribution \( \{S_{i,0}\} \), the general equilibrium of our economy is unique, pinning down unique paths of aggregates, such as \( Y_t, L_t, w_t \), as well as a unique path of the distribution of idiosyncratic states \( \{S_{i,t}\} \).

We next analyze the model from the perspective of a single firm, holding the wage \( w_t \) and aggregate output \( Y_t \) constant. Then, in Section 4, we study how the general equilibrium of the model responds to an aggregate credit shock.

### 3 Endogenous Uncertainty and Lending

We are now ready to study the interaction between credit and learning that is at the core of our mechanism. In this section, we illustrate this interaction, focusing on the partial equilibrium dynamics of a single firm. From Proposition 1, a firm \( i \) is denied funding for the risky technology (“risky funding”) if

\[ \mu_{i,t} - V(\Sigma_{i,t}) < \nu_t, \]

where the lending threshold \( \nu_t \) captures the combined impact of \( w_t, Y_t \) and the credit shock \( \lambda_t \) on firms’ access to credit. From (12) and (13), uncertainty about firms without risky funding at \( t \) increases at \( t + 1 \); and expectations will remain anchored around \( \mu_{i,t} \), slowly converging back to the unconditional mean, irrespective of the actual realization of \( A_{i,t} \). Through this decoupling of belief dynamics from fundamentals, credit constraints get reinforced over time and may outlast any shock to \( \nu_t \).

#### 3.1 Interaction Between Credit and Learning

To illustrate, suppose productivities are at their mean, \( A_{i,s} = \bar{A} \); there is constant lending threshold \( \nu_s = \nu \) for all \( s \geq 0 \); firm \( i \) has no bankruptcy flag at \( t = 0 \); and the exogenous informative signal does not materialize. In that case, the joint dynamics of beliefs \( (\mu_{i,t}, \Sigma_{i,t}) \) are deterministic, and can be captured in a simple, but useful phase diagram. We show two examples for such phase diagrams in Figure 1, for two different values of \( \nu \).
Figure 1: Phase diagram for firm-level beliefs in the absence of shocks

Note. Thin gray lines depict $(\mu - V(\Sigma) = \nu)$-contours; Z-shaped blue lines are the constant-$\Sigma_{i,t}$ locus; vertical red lines are the constant-$\mu_{i,t}$ locus. Arrowheads represent distinct points in time along the plotted trajectories. Left: Case with a unique steady state ($\nu < \bar{\nu}$). Right: Case with multiple steady states ($\nu < \nu < \bar{\nu}$).

The thin gray line depicts the contour where (18) holds with equality, dividing the state space into a region where the firm has access to risky bank loans (southeastern) and one where it has not (northwestern). The red line corresponds to the constant expectations locus, where $\mu_{i,t} = \mu_{i,t-1} = \log \bar{A}$. The blue lines correspond to the constant uncertainty locus, where $\Sigma_{i,t} = \Sigma_{i,t-1}$.

The uncertainty locus consists of two separate pieces. First, when uncertainty is at its lowest, $\Sigma_{i,t} = \sigma^2_\varepsilon$, and the expected risky productivity $\mu_{i,t}$ is sufficiently large for funding, $\mu_{i,t} > V(\sigma^2_\varepsilon) + \nu$, uncertainty remains constant at $\sigma^2_\varepsilon$. This is the flat blue line at the bottom of Figure 1. Second, and symmetrically, when uncertainty is at its highest, $\Sigma_{i,t} = \sigma^2_\varepsilon/(1 - \rho^2)$, and $\mu_{i,t}$ is sufficiently low, $\mu_{i,t} < V(\sigma^2_\varepsilon/(1 - \rho^2)) + \nu$, there is no funding and uncertainty remains at its highest level. For intermediate levels of uncertainty, the firm is marginally funded when $\mu_{i,t} - V(\Sigma_{i,t}) = \nu$. Starting from this curve, slightly greater uncertainty or slightly more pessimistic expectations raise uncertainty, while points on or below the curve reduce uncertainty.

The “Z” shaped pattern visible in Figure 1 captures the self-reinforcing nature of endogenous uncertainty in our model. When uncertainty is high today, a firm is less likely to receive funding for the risky technology, which further increases uncertainty going forward. When $\nu$ is neither too low nor too high, this effect can be sufficiently strong to generate two steady states in the phase diagram. As our next proposition shows, and Figure 1 illustrates, this happens when the Z-shape intersects with the vertical red line.

**Proposition 2.** There exist two thresholds $\nu < \bar{\nu}$, such that for all $\nu \leq \nu \leq \bar{\nu}$, there are two
(non-stochastic) steady states at the firm-level, and otherwise there is a unique (non-stochastic) steady state. The thresholds are given by $\nu = \log \bar{A} - \mathcal{V}(\sigma^2_e)$ and $\tau = \log \bar{A} - \mathcal{V}(\sigma^2_e/(1 - \rho^2))$.

For intermediate levels of the lending threshold $\nu$, a denial in funding for the risky technology is infinitely persistent in the absence of shocks. Accordingly, a one-time disruption in a firm’s access to risky funding can cut it off indefinitely from future risky funding. This is certainly a stylized result, but it neatly illustrates the forces that are active in the model. Of course, even when the steady state is unique, disruptions in credit access may be persistent (though not indefinite), as illustrated by the gray trajectories in the left panel of Figure 1. Along these trajectories, each arrowhead represents a distinct point of time. This means that the distance between two consecutive arrowheads is an inverse measure of the speed at which beliefs are evolving.

The three trajectories differ in the persistence of beliefs and the amount of uncertainty induced along the path. Along the rightmost trajectory, the firm is initially unconstrained and beliefs immediately adjust to the unique steady state. By contrast, along the two trajectories starting to the left of the gray contour line, the firm is initially denied risky funding so that learning breaks down. Accordingly, expectations only slowly converge to the unconditional prior, whereas uncertainty accumulates to higher and higher levels over time, because information about past levels of $A_{i,t}$ becomes less and less useful for predicting current productivity. This, in turn, reinforces tight credit constraints. Hence, even though the steady state is unique, a firm can find itself lacking full access to credit for a significant period of time, unable to invest in their risky technology.

More generally, the duration without access to risky funding is governed by a “race” between the inherited mean-reversion in $\mu_{i,t}$ and rising uncertainty. Consider a marginally constrained firm with $\mu_{i,t}$ just below $\mathcal{V}(\sigma^2_e) + \nu$. Stepping forward in time by one period, it will be constrained at $t+1$ if and only if

$$\rho \mathcal{V}(\sigma^2_e) - \mathcal{V}((1 + \rho^2)\sigma^2_e) < (1 - \rho)(\nu - \log \bar{A}).$$

Hence, the marginally constrained firm will stay constrained longer than the shock if either aggregate credit conditions are sufficiently bad ($\nu$ is sufficiently large) or if $A_{i,t}$ is sufficiently persistent ($\rho$ is sufficiently large).

### 3.2 Temporary Disruption in Credit

The phase diagram in Figure 1 studied permanent shocks to $\nu_t$. We next relax this assumption and study the average effect of a one-time shock to $\nu_t$, first in a phase diagram, then as a
Figure 2: Dynamic response to a credit shock at \( t = 1 \) and a subsequent recovery at \( t = 2 \)

*Note.* Arrowheads depict distinct points in time along the plotted trajectory.

simulation in our model that includes idiosyncratic shocks such as the exogenous revelation of information with probability \( \theta \).

To do so, we fix an initial productivity \( A \) and lending threshold \( \nu \), such that the firm is just active absent a shock, \( \rho \log A = \nu + \mathcal{V}(\sigma^2) \). Now suppose that at \( t = 0 \), \( i \)'s access to risky funding is curtailed by an exogenous increase in the lending threshold to \( \nu_0 > \nu \), which mechanically reverts back to \( \nu \) at \( t = 1 \). For concreteness, we interpret the shift in \( \nu_0 \) as stemming from a credit shock \( \lambda_0 \), but we note that in general equilibrium \( \nu_t \) also reflects shifts in aggregate demand \( Y_t \) and wages \( w_t \).

Figure 2 illustrates the dynamics using the phase diagram developed above. In the diagram, the shock to the banking sector results in a rightward-shift of the \((\mu - \mathcal{V}(\Sigma) = \nu)\)-contour (the dashed gray line). For sufficiently large \( \nu_0 \), the firm is denied funding for the risky technology, setting in motion a feedback loop between uncertainty and continued inability to fund the risky technology. Once uncertainty has passed the original \((\mu - \mathcal{V}(\Sigma) = \nu)\)-contour line (the solid gray line), even a reversal of \( \nu_t \) to \( \nu \) does not end the feedback loop, generating internal persistence of the shock.\(^{13}\)

\(^{13}\)Here we initialized the firm close enough to the constraint so that uncertainty surpasses the original \((\mu - \mathcal{V}(\Sigma) = \nu)\)-contour line after one period. In general, an exogenous disruption in credit lasting for \( T - 1 \)
Figure 3: Impact of temporary credit shock on firm dynamics

Note. Blue solid line: Effect of one time disruption in credit, $\nu_0 > \nu$, in period $t = 0$ on the average evolution of a firm close to the funding threshold, $\rho \log A = \nu + V(\sigma^2)$. Red dashed line: Same evolution, but fixing uncertainty exogenously at $\Sigma = \sigma^2$. 

Figure 3 repeats the experiment in our model with all firm-level shocks active, showing how the average evolution across different sample paths is affected by a one-period long disruption in credit. To isolate the contribution of the endogenous-uncertainty channel, we contrast the model’s response (solid blue lines) with a counterfactual response, in which the firm suffers the same exogenous credit shock but uncertainty is fixed at its lower bound, $\Sigma = \sigma^2$ (dashed red lines).\(^{14}\) We call this the exogenous uncertainty model as contrast to our endogenous uncertainty model. The exogenous uncertainty model will serve as a useful benchmark for the remainder of this paper.

In both cases, output initially drops due to the switch in technologies for the duration of the credit shock. The difference between our model and the exogenous-uncertainty counterfactual emerges at $t = 1$. Whereas output recovers in the counterfactual once access to credit is restored, the firm continues to be denied funding in the presence of endogenously increased uncertainty. The disruption in credit continues until either $\mu_{i,t}$ crosses the $(\mu - V(\Sigma) = \nu)$-contour in Figure 2 or the potential productivity $A_{i,t}$ is exogenously revealed (with probability $\theta$). In both cases, uncertainty drops to $\sigma^2_\epsilon$ and the firm switches back to the risky technology.

The dynamics shown in Figure 3 are reminiscent of the evidence in Huber (2018), who shows that a quasi-exogenous temporary credit shock can have a long-lasting effect on firm performance. In particular, Huber (2018) shows that the gap in employment between firms that were exposed to the shock and firms that were not remains elevated for two years after the shock.\(^{14}\)

periods cause internal persistence beyond the exogenous shock if $\rho^T \log A < \nu_0 + V\left(\frac{1-\rho^T}{1-\rho^2} \sigma^2_\epsilon\right)$.\(^{14}\)

\(^{14}\)The parametrization of the model is detailed in Section 4.
3.3 Informational Externalities

We conclude this section with a brief discussion of efficiency. Our specification of credit constraints implies two sources of inefficiency. First, credit access is \textit{statically inefficient} due to the presence of default costs, which give rise to the usual static wedge between supply and demand for credit.\footnote{In general, as long as banks cannot recover the full share of operating profits in the event of default, the maximal repayment that banks can generate (under any interest rate) is strictly smaller than $E_t[\Pi_{i,t}]$ and, hence, the statically marginally profitable project with $E_t[\Pi_{i,t}] \to \phi$ will never be funded by banks.} Second, the combination of endogenous learning and external funding introduces a novel \textit{dynamic inefficiency} that arises because atomistic banks do not internalize the option value of learning about a firm’s risky technology. In our setup, this is because firms and banks cannot write contracts that are contingent on productivity realizations in future periods. This leads banks to lend too little.

The two inefficiencies suggest welfare gains from subsidizing bank lending. Interestingly, by mitigating the dynamic inefficiency, subsidized bank lending generates new information about firms’ risky productivities $A_{i,t}$, helping the market identify which firms are creditworthy. Hence, public lending may in fact crowd \textit{in} future private lending, raising the social returns.

To illustrate the two inefficiencies, we numerically solve for the steady state distribution of firms in our model, and compute the private and social gains from lending to firms with various initial expectations $\mu_{i,t}$ at the lowest uncertainty level $\Sigma_{i,t} = \sigma^2_\epsilon$ at some time $t$. Specifically, for each $\mu_{i,t}$, we compute: (i) the private benefit from lending, as the largest loan size a bank would be willing to lend, given by

$$\phi^\text{private}_{i,t} = \phi \exp \{ \mu_{i,t} - V(\sigma^2_\epsilon) - \nu \} ,$$

(ii) the static social benefit from lending, defined as

$$\Delta \Pi^\text{static}_{i,t} = \left( E_{\{A_{i,t}|\mu_{i,t}, i \in B_t\}} - \tilde{A} \right) \xi^{-\xi} (\xi - 1)^{\xi - 1} \frac{Y}{w^{\xi - 1}} ,$$

and (iii) the dynamic social benefit from lending

$$\Delta \Pi^\text{dynamic}_{i,t} = \Delta \Pi^\text{static}_{i,t} + \sum_{s=1}^\infty (1 + r)^{-s} \left( E_{\{A_{i,t+s}|\mu_{i,t}, i \in B_t\}} - E_{\{A_{i,t+s}|\mu_{i,t}, i \notin B_t\}} \right) \xi^{-\xi} (\xi - 1)^{\xi - 1} \frac{Y}{w^{\xi - 1}} .$$

After time $t$, firms follow their equilibrium behavior. We note that the static social benefit from lending corresponds to the increase in date-$t$ utility (in units of date-$t$ consumption) from lending to firm $i$ at date $t$. The dynamic social benefit from lending corresponds to the
increase in welfare across all future periods (in units of date-\(t\) consumption) from lending to firm \(i\) at date \(t\).

Figure 4 plots \(\phi_{i,t}^{\text{private}}, \Delta \Pi_{i,t}^{\text{static}}, \text{ and } \Delta \Pi_{i,t}^{\text{dynamic}}\) against \(\mu_{i,t}\). As expected, the maximum private loan size exceeds the fixed cost \(\phi\) precisely to the right of the lending threshold (vertical dotted line), where \(\mu_{i,t} > \nu + \mathcal{V}(\sigma_{t}^{2})\). The static social benefit from lending \(\Delta \Pi_{i,t}^{\text{static}}\) lies consistently above the private willingness to fund, illustrating the standard static inefficiency induced by default risk. The dynamic social benefit \(\Delta \Pi_{i,t}^{\text{dynamic}}\) is closely aligned with the static benefit for productivity expectations \(\mu_{i,t}\) far away from the threshold. For those, the actual technology used in later periods \(A_{i,t+a}\) is close to independent of whether the firm was funded, \(i \in \mathcal{B}\), or not, \(i \notin \mathcal{B}\), in period \(t\).

For values of \(\mu_{i,t}\) right around the lending threshold, however, there is a large dynamic gain. Information revelation in period \(t\) on a firm’s risky productivity \(A_{i,t}\) allows banks to accurately assess on which side of the funding threshold a firm’s future expectation \(\mu_{i,t+1}\) lies. This leads to more accurate lending by banks in the future and therefore increases the present discounted value of public lending.\(^{16}\)

\(^{16}\)The region in which the dynamic gains lie below the static one indicates firms for which lending today may reveal that their productivity lies below the lending threshold in the future, whereas absent public lending access to credit markets would be restored soon, reducing the dynamic gains from lending.
4 Aggregate Credit Crunches

Having studied the impact of a credit shock on a single firm, we next study the general equilibrium response of the economy to an aggregate credit shock.

4.1 Parametrization

We interpret one period as a quarter, and set the discount factor $\beta$ to 0.99. The intertemporal elasticity of substitution $\sigma$ and the Frisch elasticity of labor supply $\zeta$ are set to 2, the elasticity of substitution between consumption goods is set to 5, and the scaling parameters $v$ and $\bar{A}$ are set to normalize steady state employment and output to 1. The revenue productivity parameters are set to $\rho = 0.9440$ and $\sigma_c = 0.0726$, consistent with the revenue productivity process estimated by Foster, Haltiwanger and Syverson (2008) (converted to a quarterly frequency). Our choice for the baseline productivity $\bar{A}$ is based on Chodorow-Reich (2013, online appendix). Using bank lending relations to instrument for the credit access of firms, Chodorow-Reich finds that an exogenous preclusion from credit results in an employment decline of 53%, but qualifies that this number is likely overstating the true effect due to the first-stage coefficients being biased towards zero. Keeping in mind this qualification, we set $\bar{A}$ so that the marginal employment effect of having access to risky credit equals 30%.

It remains to choose values for $\phi$, $\eta$ and $\theta$. We set $\phi$ to pin down the share of firms without access to risky funding at the steady state. Lacking more precise data, we inform our choice of $\phi$ using two proxies for the fraction of firms that lack sufficient funding. First, among Compustat firms in the years 1976–1999, the sales-weighted average fraction of firms that does not pay dividends is 8%. Second, among small businesses, the 2020 Fed Small Business Credit Survey documents that 21% of all firms were (partially) denied credit and an additional 9% of firms were discouraged from seeking credit, because they believed they would be turned down. Based on this evidence, we set $\phi$ so that 25% of firms do not have access to risky funding at the steady state, implying a sales-weighed average of 8%. Appendix B explores alternatives. Next, we set the rate of bankruptcy removals $\eta$ to 0.35, targeting an average bankruptcy length of 2.86 quarters, consistent with the average duration of Chapter 11 bankruptcy negotiations documented by Teloni (2015). Finally, we set the exogenous information revelation rate $\theta$ to implement an average duration of a denial in risky funding of 2 years, consistent with the impact on firm growth documented in Huber (2018).

Table 1 summarizes the calibrated parameters.

---

17 We only use data until 1999, excluding the burst of the dot-com bubble and the 2008/09 financial crisis, in line with our interpretation of the steady state as normal times.

18 The targeted duration is documented for 2005–2014, after the Bankruptcy Abuse Prevention And Consumer Protection Act went into effect, and includes pre-negotiated deals.
Table 1: Calibrated parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>β</th>
<th>σ</th>
<th>ζ</th>
<th>ξ</th>
<th>υ</th>
<th>̅A</th>
<th>̃A</th>
<th>ρ</th>
<th>σ_x</th>
<th>φ</th>
<th>η</th>
<th>θ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.99</td>
<td>2.00</td>
<td>2.00</td>
<td>5.00</td>
<td>0.860</td>
<td>1.005</td>
<td>0.697</td>
<td>0.944</td>
<td>0.073</td>
<td>0.135</td>
<td>0.35</td>
<td>0.117</td>
</tr>
</tbody>
</table>

Figure 5: General equilibrium response to an AR(1) credit shock

*Note.* Panels (a)–(e) are denominated in percentage deviations from the steady state; panel (f) shows the increase in risk premia relative to the steady state denominated in percentage points. Solid blue lines are model responses; dashed red lines are exogenous uncertainty counterfactuals where uncertainty is fixed at $\Sigma_{i,t} = \sigma^2$ for all $i$ and $t$.

4.2 Simulation of an Aggregate Credit Shock

We are now ready to explore the economy’s response to an aggregate credit shock. The economy is initialized at its stochastic steady state where $\lambda_{t-s} = 0$ for $s > 0$ and the cross-section of firms is at its ergodic distribution. We simulate the economy’s response to a credit shock $\lambda_t$ to first order in aggregate variables (remaining fully nonlinear in idiosyncratic variables), as in Reiter (2009). We assume $\lambda_t$ follows an AR(1) process with a half-life of four quarters (implying an autocorrelation of 0.84). All responses are normalized to induce a 1% drop in output at impact.\(^{19}\)

Figure 5 displays the model’s responses to the perturbation at $t = 0$, depicted by solid

---

\(^{19}\)Since our equilibrium is backward-looking (for now), we do not yet need to take a stand on whether the shock is anticipated or not. In Section 5.2, after introducing capital, we take the stand that the shock is unanticipated.
blue lines. To isolate the role of the endogenous amplification through rising uncertainty from the exogenous impact of tightened credit constraints, we again contrast the model’s responses with the exogenous uncertainty benchmark for which uncertainty is fixed at $\Sigma_{i,t} = \sigma^2_\epsilon$ for all $i$ and $t$ (dashed red lines). The responses in panels (a)–(e) are in percentage deviations from their respective steady state values; the response in panel (f) shows the increase in risk premia relative to the steady state denominated in percentage points.

**Amplification and persistence.** Because uncertainty is predetermined, the impact response at date 0 is fully explained by tightened credit constraints. Starting at date 1, however, the adverse effects of rising uncertainty start to both amplify and prolong the crisis relative to the exogenous uncertainty case. At $t = 4$, output in the exogenous uncertainty counterfactual has recovered exactly half of its impact losses. By comparison, at $t = 4$, output in the endogenous uncertainty economy is more than 2 percent below steady, surpassing the impact effect. In terms of half-lives, recovery takes more than four times as long in the endogenous uncertainty economy (22 quarters) compared to the exogenous uncertainty counterfactual (4 quarters).

The main reason for the persistent increase in uncertainty is the feedback between uncertainty and lending we laid out in Section 3. As the credit shock hits the economy, some firms lose access to funding for the risky technology. This can lead to several periods without new information, and hence greater uncertainty. Figure 6 illustrates this at $t = 5$, plotting the cross-sectional distribution of uncertainty. Compared to the steady state distribution, there are increased masses of firms in the states corresponding to 1–5 periods without new information. Those states are associated with significantly greater uncertainty (right y-axis), raising the overall level of uncertainty. Since it takes a while until the distribution of firms across states has converged back to its steady state distribution (see Section 3) there is a significant amount of persistence in the impulse responses in Figure 5.

**Risk premia, default rates, and dispersion.** Rising uncertainty also helps explain a few financial market characteristics typically associated with financial crises. First, rising uncertainty increases average risk premia, defined as before as the gap between the rate charged to firms and the internal discount rate by banks (which is equal to the lending rate on safe loans), $1 + r^p_{i,t} = (1 + r_{i,t})/(1 + \lambda_t)$. Second, default rates increase as uncertainty rises over the course of the response. Third, consistent with the evidence in Bloom et al. (2018), increased uncertainty at the firm-level translates into an increased cross-sectional dispersion of firm sales, plotted in panel (d) of Figure 5. This is caused by an increase in firms that are unable to finance the risky technology, and therefore have to resort to using the baseline
Figure 6: The distribution of uncertainty 5 periods into the impulse response and at the steady state

Note. Solid black line: level of uncertainty corresponding to $s$ periods without new information (right axis).

technology. The increase is significantly stronger and more persistent in the endogenous uncertainty model.\textsuperscript{20}

\textbf{Disagreement.} With a few extra ingredients, our model also has predictions for the beliefs of market observers. Intuitively, as the publicly available information about firms that are denied risky lending diminishes, market observers will rely more on other sources to form their beliefs. As long as those other sources are partially dispersed across observers, (average) disagreement among market observers increases when an increasing number of firms becomes constrained.

To formalize this prediction in a simple extension, consider a set of outside observers (or forecasters) $j \in [0, 1]$. In addition to $\mathcal{I}_t$, these forecasters each observe a private signal $\omega_{ij,t} = \log A_{i,t} + \psi_{ij,t}$, where $\psi_{ij,t}$ is normally distributed with zero mean and variance $\sigma^2_{\psi}$, i.i.d. across $i, j, \text{and } t$. For simplicity, we assume that at each date $t$, the previous generation of forecasters is replaced by a new one. The belief of forecaster $j$ about firm $i$'s productivity

\textsuperscript{20}Gourio (2014) also provides a theory which predicts countercyclical default and a countercyclical dispersion of firm sales.
at date $t$ is given by

$$\tilde{\mu}_{ij,t} = \frac{\Sigma_{i,t}^{-1} \mu_{i,t} + \sigma^{-2}_\psi \omega_{ij,t}}{\Sigma_{i,t}^{-1} + \sigma^{-2}_\psi}.$$  \hfill (19)

In Section 7, we use these forecasters’ beliefs to compare the model’s predictions with micro data from a survey of professional forecasters. To keep this extension tractable, we assume that forecasters do not interact with the rest of the economy.\textsuperscript{21} From (19), the degree of “disagreement” among market observers is given by

$$sd_j[\tilde{\mu}_{ij,t}] = \frac{\sigma^{-1}_\psi}{\Sigma_{i,t}^{-1} + \sigma^{-2}_\psi},$$  \hfill (20)

Thus, according to our model, there should be a tight empirical link between disagreement (as measure of uncertainty) and the degree to which a firm is financially constrained. We explore this link in Section 7.

## 5 Extensions

We next present two extensions of our baseline model from Section 2.

### 5.1 Introducing Equity-Financed Firms

So far, all firms equally relied on bank credit in order to fund their projects, exposing their ability to operate to the beliefs of the financial market. We now explore the case in which some firms are equity-financed and do not need bank credit to fund the fixed cost $\phi$.\textsuperscript{22} This allows them to always operate their preferred technology. To make it even starker, we assume away any information frictions for those firms as well. That is, equity-financed firms are able to observe $A_{i,t}$ at the end of each period, irrespective of the technology that was actually used in production.\textsuperscript{23} To explore robustness of our mechanism to this extension of our model, we assume two thirds of all firms are equity-financed and thus never face any financial constraints.

Figure 7 shows the aggregate responses to a credit shock, normalized again to induce a one-percent reduction in output. For comparison, we include the responses from the baseline

\textsuperscript{21}This assumption spares us from dealing with infinite regress (Townsend, 1983). Nevertheless it is worth pointing out that our predictions about disagreement are not limited to professional forecasters, but would more generally carry over to any set of agents with dispersed signals (or priors).

\textsuperscript{22}Through other mechanisms, equity financing may also subject firms to the beliefs of the financial market, giving rise to a similar mechanism as the one in this paper. We explored this in a previous working paper version (Straub and Ulbricht, 2018).

\textsuperscript{23}This assumption is not necessary for the results in this section.
Figure 7: General equilibrium response in the model with both bank-financed and equity-financed firms.

The responses are mostly similar to the ones in the baseline model. The similarity comes from the fact that, conditional on observing a financial crisis of a set magnitude, driven by a credit shock, the two models behave very similarly. The only difference is that the credit shock in the background needs to be stronger in the model with equity-financed firms.

5.2 Introducing Capital

Our second extension introduces capital to the baseline model in Section 2 and compares it to a standard real business cycle (RBC) model. To do so, we modify the production function of firm \( i \) to a Cobb-Douglas aggregate of capital and labor:

\[ Y_{i,t} = A_{i,t}^{\frac{1}{1-\alpha}} K_{i,t}^{\alpha} L_{i,t}^{1-\alpha} \]

where capital \( K_{i,t} \) is rented at the competitive rate \( 1 + r^K_t > 0 \) from households. The representative household is now allowed to not only save in bonds \( B_t \) (which are still in zero net supply) but also in capital \( K_t \). The date-\( t \) budget constraint now reads:

\[ C_t + B_{t+1} + K_{t+1} = w_t L_t + (1 + r_{t-1}) B_t + (1 + r^K_t - \delta) K_t + T_t. \]

As usual, capital \( K_t \) is determined one period in advance. Market clearing,

\[ K_t = \int_0^1 K_{i,t} \, di, \]

\(^{24}\) The labor response is significantly smaller in the model with equity-financed firms, as those firms can benefit from lower wages and invest in their risky technologies at higher rates. This improves their productivity, implying reduced employment levels for a given impact response of output.
determines the rental rate $1 + r_t^K$ in equilibrium. All other model elements remain unchanged. Households still maximize utility (7), firms maximize profits (4) and require a bank loan to finance the fixed cost $\phi$. The financing condition (11) still applies, only that the lending threshold $\nu_t$ is now given by

$$
\nu_t = \log \left( (1 + \lambda_t) \phi \right) - \log \left( \frac{Y_t}{(1 + r_t^K)^{\alpha (1 - \xi) w_t^{(1 - \alpha)(\xi - 1)}}} \right) + \log \left( \xi^\xi (\xi - 1)^{1 - \xi} \right).
$$

We next show that the model with capital is equivalent to an RBC model, with an endogenous process for TFP corresponding to the efficiency wedge introduced in (15).

**Proposition 3.** Conditional on the efficiency wedge $\{A_t\}$, defined in (15), the equilibrium behavior of $\{C_t, K_t, L_t\}$ (and therefore also of other aggregates, such as $Y_t, w_t, r_t^K$) is described by a standard RBC model,

$$
C_t^{-1/\sigma} = \mathbb{E}_t \left[ \beta \left( (1 - \xi^{-1}) \alpha A_t K_{t+1}^{\alpha - 1} L_{t+1}^{1 - \alpha} + 1 - \delta \right) C_{t+1}^{-1/\sigma} \right],
$$

$$
u L_t^{1/\zeta} = (1 - \xi^{-1})(1 - \alpha) C_t^{-1/\sigma} A_t K_t^{\alpha} L_t^{-\alpha},
$$

$$
C_t + K_{t+1} = A_t K_t^{\alpha} L_t^{1 - \alpha} + (1 - \delta) K_t - \phi,
$$

with monopoly distortion $1 - \xi^{-1}$, and fixed cost of production $\phi$.

Proposition 3 is intuitive. The path of the lending threshold $\{\nu_t\}$ shapes the process of the distribution of productivities $\{A_{i,t}\}$. Given $\{A_{i,t}\}$, the process of efficiency wedges $\{A_t\}$ can be computed from (15). The lending threshold $\nu_t$ itself is determined by the interacting financial and information frictions. The credit shock $\lambda_t$ acts by shifting $\nu_t$. However, as before, $\nu_t$ is also a function of output $Y_t$. This is especially relevant here, as it generates a two-way feedback between $\nu_t$ and the capital stock $K_t$. A tighter lending threshold $\nu_t$ reduces the efficiency wedge $A_t$ going forward, as more firms have trouble obtaining bank funding. Reduced efficiency, in turn, leads to reduced investment in capital, which further tightens the lending threshold (raising $\nu_t$).

Figure 8 illustrates these dynamics for the same credit shock as in Figure 5 and compares the responses to the ones in the baseline model without capital. While the responses in the economy with exogenous uncertainty are still modest, the endogenous uncertainty economy generates more amplification and persistence than before. The reason why is the strong reduction in investment in Panel (b), which further reduces the efficiency wedge in Panel (d).

---

25 We choose $\alpha = 0.30$ and a quarterly depreciation rate of $\delta = 0.02$. 
A crucial assumption of our model is that agents are uncertain about firm-level productivity, implying that the feedback between credit and uncertainty operates entirely at the firm level. In this section, we explore the alternative case in which learning is about aggregate productivity instead. While the model has qualitatively similar predictions, we find that the scope for amplification due to aggregate uncertainty is quantitatively very small.

To explore learning about aggregate productivity, suppose that both the baseline and risky productivity have a common time-varying mean. Specifically, suppose that \( \bar{A}_t = Z_t^{\xi-1} \) and \( \tilde{A}_t = (\bar{A}/\bar{A}) \times Z_t^{\xi-1} \). Here \( Z_t \) is the common aggregate component. We use \( Z_t^{\xi-1} \) so as to simplify the expressions below. The model in Section 2 corresponds to the case where \( Z_t = \bar{A}^{1/(\xi-1)} = \text{const.} \) We assume that \( Z_t \) follows an AR(1) process given by

\[
Z_t = \rho_Z Z_{t-1} + (1 - \rho) \bar{Z} + u_t,
\]

where \( \rho_Z \in [0, 1] \) and the innovations \( u_t \) are i.i.d., normal, with zero mean and variance \( \sigma_u^2 \).\(^{26}\)

\(^{26}\)Here we assume that \( \bar{Z} \) is sufficiently large so \( Z_t > 0 \) with near certainty. Notice that for large \( \bar{Z} \), the implied process for \( Z_t \) is virtually identical to the case where \( Z_t \) follows an AR(1) in logs. Assuming that \( Z_t \) follows an AR(1) in levels allows us to provide an exact characterization of the relevant uncertainty dynamics in closed form.

Figure 8: General equilibrium response in the model with capital
To isolate the contribution of aggregate uncertainty, we assume that all firm-level productivities are perfectly observable at the end of the period, as in our exogenous uncertainty counterfactual. Let
\[
\Delta a_{i,t} \equiv \begin{cases} 
\log A_{i,t} - \log \bar{A}_t & \text{if risky technology is adopted} \\
0 & \text{else}
\end{cases}
\]
denote the firm-level component to productivity. The information set at the beginning of period \(t\) is then given by
\[
\mathcal{I}_t = \{\lambda_t, s_t\} \cup \{\Delta a_{i,t-1}\}_{i \in [0,1]} \cup \mathcal{I}_{t-1},
\]
where \(s_t\) is a noisy signal of last period’s aggregate output,
\[
s_t = Y_{t-1} + \omega_t,
\]
with \(\omega_t\) i.i.d., normal, with zero mean and variance \(\sigma^2_{\omega}\). To maximize the potential for aggregate uncertainty fluctuations, we assume that agents do not infer any information about \(Z_t\) from the cross-sectional distributions of prices, outputs, etc.

To characterize the uncertainty dynamics in this economy, define
\[
X_t = \frac{Y_t}{Z_t} = \left(\int_0^1 e^{\Delta a_{i,t}/\xi} L_{i,t}^{\xi-1} \, di\right)^{\frac{1}{\xi}},
\]
as the aggregate input bundle in the economy, and observe that \(X_t\) is known by agents at time \(t\). Rewriting the signal (24), we have
\[
s_t = Z_{t-1}X_{t-1} + \omega_t.
\]
In this setting, aggregate financial shocks reduce the observed input \(X_t\), which similar to van Nieuwerburgh and Veldkamp (2006) then reduces the signal precision of \(s_t\), which in turn increases uncertainty.

To evaluate the quantitative potential of the induced dynamics for aggregate uncertainty, suppose that the economy is in its stochastic steady state where \(\mathbb{E}[Z_t|\mathcal{I}_t] = \bar{Z}\) and \(\Sigma^Z_t \equiv \text{Var}[Z_t|\mathcal{I}_t]\) converged to a constant. Now suppose the economy is hit by the same credit shock as before, whereas aggregate productivity and the noise shock remain at their steady state values (\(Z_{t+s} = \bar{Z}\) and \(\omega_{t+s} = 0\) for all \(s \geq 0\)). From the usual Kalman filter, it then follows that agents’ mean expectations remain unperturbed (i.e., \(\mathbb{E}[Z_{t+s}|\mathcal{I}_{t+s}] = \bar{Z}\) for all \(s\)), and
aggregate uncertainty dynamics are given by

$$\Sigma_t^Z = \frac{\rho_t^2 \Sigma_{t-1}^Z}{1 + \left(\frac{\sigma_\omega}{X_t-1}\right)^2 \Sigma_{t-1}^Z} + \sigma_t^2. \tag{26}$$

To maximize the potential impact of the aggregate uncertainty channel, we chose parameters $\rho_Z$, $\sigma_Z$, and $\sigma_\omega$ so as to maximize the percentage increase in $\Sigma_t^Z$ at the peak of the impulse response. Clearly, the response is maximized for $\rho_Z = 1$. Moreover, because any proportionate scaling of $\sigma_Z$ and $\sigma_\omega$ also scales $\Sigma_t^Z$ (and thus leaves the percentage response relative to steady state unchanged), it is sufficient to set the relative standard deviation $\sigma_\omega/\sigma_Z$. We choose this ratio to maximize the peak response of uncertainty, which gives $\sigma_\omega/\sigma_Z = 2.052$.

Figure 9 shows the response in the aggregate-uncertainty model to a credit shock, alongside the responses with endogenous firm-level uncertainty and exogenous uncertainty. The exogenous uncertainty and endogenous aggregate uncertainty models are virtually identical. This is because even for a significant crisis with an output loss of 1% on impact, the ability to learn about aggregate productivity is only marginally affected. Panel (d) shows the maximized peak increase in aggregate uncertainty, which is below 0.6% (compared to an average increase in firm-level uncertainty of above 25% in our baseline model).

To see why this is the case, consider the signal (25). By design, $X_t$ decreases by 1% on impact, decreasing the signal precision, $(X_t-1/\sigma_\omega)^2$, by only approximately 2%, severely limiting endogenous movements in aggregate uncertainty.

Would a larger crisis matter more? To explore this question, we scale up the credit shock and compute again the peak increase in uncertainty. Figure 10 shows the resulting link between the impact output loss and the peak increase in uncertainty using the above parametrization. It can be seen that for any magnitude of the shock, the peak increase in
uncertainty is proportionately smaller than the corresponding loss in output. This is markedly different in our model with endogenous firm-level uncertainty. There, no matter how small the credit shock, it always results in some firms losing risky funding at the margin, starting the adverse credit–uncertainty spiral for those firms.

7 Empirical Exploration

At the core of our model is a two-way interaction between uncertainty and financial constraints, causing both variables to comove. In this section, we explore the extent to which this comovement can be seen empirically, using a combination of firm-level survey data and accounting data.

7.1 Data

In the following we describe the data sources and define the main variables used for our empirical exploration. Technical details regarding the construction of our dataset are contained in Appendix C.1.

Proxies for uncertainty. Our proxy for uncertainty is based on data on individual forecasts about earnings per share (EPS) by financial analysts from the IBES database.27 To

27We explore alternative proxies for uncertainty based on stock returns in a previous working paper version (Straub and Ulbricht, 2018). Relatedly, Bai, Philippon and Savov (2016) explore the informativeness of
make these forecasts comparable to our model, we follow Senga (2018) and transform EPS forecasts into forecasts about returns on assets (ROA). In our dataset, median productivity as measured by ROA is 3.7 percent (−9.5 percent at the 10th percentile, 13 percent at the 90th percentile). Let \( \mu_{\text{EPS}}^{ij,t} \) denote analyst \( j \)'s expectation about firm \( i \)'s EPS at date \( t \).

Beliefs regarding returns on assets are computed as

\[
\mu_{\text{ROA}}^{ij,t} = \mu_{\text{EPS}}^{ij,t} \times \frac{\text{number of outstanding shares}_{i,t}}{\text{total assets}_{i,t-1}}.
\]

As our primary proxy for firm-level uncertainty, we look at the dispersion of forecast errors among analysts, defined by

\[
\sigma_{\text{fe}}^{i,t} \equiv \text{sd}_j [\mu_{\text{ROA}}^{ij,t} - \text{ROA}_{i,t}] .
\]

Since ROA\(_{i,t}\) is constant across all analysts \( j \), \( \sigma_{\text{fe}}^{i,t} \) can equivalently be interpreted as disagreement among forecasters.

**Proxies for financial constraint.** For the purpose of measuring financial constraints, we follow the corporate finance literature and combine various balance sheet data to proxy for firms’ access to funds. Our main measure is the “KZ-index” developed by Kaplan and Zingales (1997) and Lamont, Polk and Saá-Requejo (2001). Based on its \( k_{z_{i,t}} \)-score, we classify a firm as likely to be constrained if its current score is at or above the 95th percentile in a given calendar year. In Appendix C.2 we show similar results using dividend payouts and debt to capital ratios as proxies for financial constraints.

The resulting dataset is an unbalanced annual panel from 1976 to 2016, covering on average 1979 firms per year.

---

28 As in Senga (2018), we extract those forecasts 8 months prior to each firm’s fiscal-year end month. See Appendix C.1 for details on the timing of our variables.

29 It is worth noting that while there is a clear mapping in our model (equation (20)), it is not immediately clear that disagreement is also empirically a good measure of uncertainty. One concern stems from the fact that the precision \( \sigma_{\psi}^{-2} \) of the individual signal may change over time, too. Observe that this can go both ways, as the forecast dispersion is inverted-U shaped in the precision \( \sigma_{\psi}^{-2} \). Absent any evidence that \( \sigma_{\psi}^{-2} \) moved one way or another, we interpret dispersion as measure of disagreement.

30 The KZ-score is a weighted combination of a firm’s cash flow to total capital, its market to book ratio, its debt to capital, dividends to total capital, and cash holdings to capital (see Appendix C.1 for details). The weighting coefficients are based on an ordered logit regression relating those accounting variables to an explicit classification of firms into categories of financial constraints (Kaplan and Zingales, 1997; Lamont, Polk and Saá-Requejo, 2001). Firms with a higher \( k_{z_{i,t}} \) score are more likely to be constrained.
7.2 Financial Constraints and Uncertainty

Cross-sectional evidence. To explore whether the predicted link between financial constraints and uncertainty is present in the data, we run a simple OLS regression of forecast-error dispersion $\sigma_{fc}^{i,t}$ on the KZ-based indicator. Table 2 reports the estimated coefficients, controlling for different combinations of fixed effects. The estimated effect is roughly constant over the first three specifications where we control for a combination of year, fiscal-end year month and 4-digit sector fixed effects. In all three specifications, the forecast-error dispersion is increased by about 0.08 for firms that are classified as financially constrained. Controlling for firm-level fixed effects, the estimated difference between financially constrained and unconstrained firms is reduced to 0.031. These results lend support to the model’s predicted positive relationship between financial constraints and uncertainty.

Time series of forecast dispersion. Table 2 showed a strong significant relationship between forecast dispersion and financial constraints. What episodes in the data are responsible for this comovement? Figure 11 highlights that this is mainly driven by the recent financial crisis and previous crisis episodes. During the crises, uncertainty about constrained firms’ fundamentals increases dramatically while uncertainty about unconstrained firms’ fundamentals largely remains flat. Figure 11 also makes it obvious that forecast dispersion moves in the intuitive direction and increases significantly during the financial crisis.

8 Concluding Remarks

In this paper, we propose a theory of endogenous uncertainty and its interaction with firms’ access to funds. In the model, firms rely on external funds to finance risky projects. When the

Table 2: Financial constraints and uncertainty

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Financially constrained</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
</tr>
<tr>
<td>Observations</td>
<td>47 342</td>
</tr>
<tr>
<td>Adj. R-sq.</td>
<td>0.010</td>
</tr>
<tr>
<td>Year × month FE</td>
<td>no</td>
</tr>
<tr>
<td>Sector FE (4 digit)</td>
<td>no</td>
</tr>
<tr>
<td>Firm FE</td>
<td>no</td>
</tr>
</tbody>
</table>

Note: Standard errors clustered at the firm-level are in parenthesis.
returns to risky projects become too uncertain, firms are unable to obtain funds, resulting in a loss of information about the profitability of their projects. This further perpetuates funding problems. While present even in normal times, this feedback loop becomes especially powerful during financial crises, in which a temporary shock to bank lending entails a prolonged economic downturn.

We have so far refrained from policy analysis in this paper. There are, however, several policy insights that merit further discussion. First, recapitalizing banks (investors) is not an effective policy to restore lending in the model, once uncertainty has already increased. The critical friction that prolongs the crisis is an informational one and cannot easily be undone by transfers to banks, which would not stop the adverse feedback between uncertainty and bank lending. This, however, suggests a second policy action: direct transfers to firms as explored in Section 3.3. Even if the government has access to the same information as everyone else in the economy, providing transfers or cheap loans to inactive firms can crowd in lending in future periods. This is due to an information externality: a re-activated firm produces information that lets future private investors decide whether to resume private lending.

In this paper, we have mainly focused on the effects of financial shocks. However, the model’s internal propagation mechanism applies similarly to other types of economic shocks, such as aggregate demand shocks. For example, by reducing revenues, a fall in aggregate

\[
\begin{align*}
\text{Mean forecast error dispersion} & \quad 1990 \quad 1995 \quad 2000 \quad 2005 \quad 2010 \quad 2015 \\
\text{constrained firms} & \quad \text{unconstrained firms}
\end{align*}
\]

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure11}
\caption{Average forecast error dispersion (a proxy for uncertainty) of constrained and unconstrained firms.}
\end{figure}

\textbf{Note.} This figure shows the average forecast error dispersion among financially constrained and among financially unconstrained firms. Financially constrained firms are those whose current Kaplan and Zingales (1997) index lies in the top 5% of the distribution. Financially unconstrained firms are all other firms.
demand exacerbates the financial friction at the heart of the model, increasing the fraction of inactive firms (cf., equation 18). This raises uncertainty going forward, so that, even as demand recovers, a supply problem (uncertainty and tight financial conditions) remains. Conversely, strong demand (e.g. due to accommodative monetary policy) can help shift firms into business, and thereby reduce uncertainty and increase the allocative efficiency of the economy. We believe these are promising avenues for further research.
References


Online Appendix to
“Endogenous Uncertainty and Credit Crunches”

A Mathematical Appendix

A.1 Proof of Proposition 1

Firm $i$ at date $t$ obtains a loan operating the risky technology if there exists an interest rate $r_{i,t} \geq \lambda_t$ such that

$$
\Phi \left( \frac{\mu_{i,t} - \log ((1 + r_{i,t}) \phi) + \log \left( \frac{Y_t}{w_t \xi^{-1}} \right) - \log \left( \xi (\xi - 1)^{1-\xi} \right)}{\sqrt{\Sigma_{i,t}}} \right) = 1 + \frac{\lambda_t}{1 + r_{i,t}}
$$

(27)

Define $x \equiv \frac{1+\lambda_t}{1+r_{i,t}} \in (0, 1]$. Equation (27) is equivalent to there existing an $x \in (0, 1]$ such that

$$
\mu_{i,t} - \log ((1 + \lambda_t) \phi) + \log \left( \frac{Y_t}{w_t^{\xi^{-1}}} \right) - \log \left( \xi (\xi - 1)^{1-\xi} \right) = \Phi^{-1} (x) \sqrt{\Sigma_{i,t}} - \log x
$$

Observe that only the right hand side of this equation depends on $x$, and that it approaches infinity as $x \to 1$. Thus, the condition for a firm to be financed can be written as

$$
\mu_{i,t} - \log ((1 + \lambda_t) \phi) + \log \left( \frac{Y_t}{w_t^{\xi^{-1}}} \right) - \log \left( \xi (\xi - 1)^{1-\xi} \right) \geq \mathcal{V} (\Sigma_{i,t})
$$

with

$$
\mathcal{V} (\Sigma_{i,t}) \equiv \min_{x \in (0,1]} \left\{ \Phi^{-1} (x) \sqrt{\Sigma_{i,t}} - \log x \right\}.
$$

This proves Proposition 1.

A.2 Properties of $\mathcal{V}(\Sigma)$

We prove a few properties of $\mathcal{V}(\Sigma)$ as well, which are stated in the text.

- $\mathcal{V}(0) = \min_{x \in (0,1]} \{-\log x\} = 0$.

- There is a unique minimizer in the definition of $\mathcal{V}(\Sigma)$. To see this, note that the FOC reads

$$
\frac{1}{\phi (\Phi^{-1}(x))} \sqrt{\Sigma} = \frac{1}{x}
$$
Defining $z \equiv \Phi^{-1}(x) \in \mathbb{R}$, this can be rewritten as

$$\Phi(z)\sqrt{\Sigma} = \phi(z) \quad (28)$$

We claim that this is satisfied for a unique $z \in \mathbb{R}$. To see why, consider the derivatives of both sides

- The derivative of the left hand side (LHS) is $\phi(z)\sqrt{\Sigma}$.
- The derivative of the right hand side (RHS) is $\phi'(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} (-z)$.

As $z \to -\infty$, both sides in (28) approach zero, but the LHS does so strictly faster, as it has a flatter, less positive derivative in that limit. Thus, $\Phi(z)\sqrt{\Sigma} < \phi(z)$ for sufficiently small $z$. Next, observe that the derivative of the LHS is strictly below the one of the RHS until $z = -\sqrt{\Sigma}$. After that, it is the other way around. This implies that there can be at most a single intersection. Existence of an intersection follows from the limit $z \to \infty$, where the LHS always strictly exceeds the RHS, in combination with the intermediate value theorem. We denote the unique minimizer in the definition of $V(\Sigma)$ by $x^*(\Sigma)$, and define analogously $z^*(\Sigma) \equiv \phi^{-1}(x^*(\Sigma))$.

- We note from (28) that $z^*(\Sigma)$ is strictly decreasing in $\Sigma$, and thus $x^*(\Sigma)$ is strictly decreasing as well. It holds that $x^*(0) = 1$, $\lim_{\Sigma \to 0} z^*(\Sigma) = \infty$. Moreover, note that $z^*(2/\pi) = 0$, following straight from (28).

- $V'(\Sigma) = \Phi^{-1}(x^*(\Sigma)) \frac{1}{2\sqrt{\Sigma}} = \frac{z^*(\Sigma)}{2\sqrt{\Sigma}}$. Thus, $V'(\Sigma)$ is strictly decreasing in $\Sigma$, with $\lim_{\Sigma \to 0} V'(\Sigma) = \infty$. This shows that $V'(\Sigma) > 0$ for $\Sigma < 2/\pi$.

### A.3 Proof of Equation 17

Equation (16) can be rewritten as

$$\left(\frac{\xi - 1}{\xi}\right)^\sigma A_t^\sigma = v^\sigma \left(\frac{Y_t}{A_t}\right)^{\sigma/\zeta} (Y_t - \phi)$$
where $\mathbf{A}_t = F_t \left( \frac{Y_t}{w^{\xi-1}_t} \right)$ is an increasing function, and by (14), $w_t = (1 - \xi^{-1})\mathbf{A}_t$. Combining these equations, we find a system of two equations and two unknowns, $\mathbf{A}_t$ and $Y_t$,

$$
\left( \frac{\xi - 1}{\xi} \right) \sigma \mathbf{A}_t^\sigma = v^\sigma \left( \frac{Y_t}{\mathbf{A}_t} \right)^{\sigma/\xi} (Y_t - \phi) \tag{29}
$$

$$
\mathbf{A}_t = F_t \left( \frac{Y_t}{((1 - \xi^{-1})\mathbf{A}_t)^{\xi-1}} \right) \tag{30}
$$

First, we observe that there always exists a solution to this system of equations. The reason is that (30) implies an increasing relationship between $\mathbf{A}_t$ and $Y_t$, which remains positive and bounded for $Y_t \rightarrow 0$ (when no firm is producing using the risky technology) and $Y_t \rightarrow \infty$ (when all firms are producing using the risky technology). This means that for $Y_t \rightarrow \phi$, the right hand side in (29) is smaller than the left hand side, and for $Y_t \rightarrow \infty$, the opposite is true. Thus, by the intermediate value theorem, there always exists a $Y_t$ (and thus also an $\mathbf{A}_t$) that solve the system.

Next, we establish a condition for uniqueness. To do so, observe that, from (30) and the fact that $\mathbf{A}_t$ must be increasing in $Y_t$, we know the rate at which $Y_t$ increases with $\mathbf{A}_t$ must be bounded above by the rate that would leave $\frac{Y_t}{((1 - \xi^{-1})\mathbf{A}_t)^{\xi-1}}$ constant, that is,

$$
\frac{d \log \mathbf{A}_t}{d \log Y_t} \leq \frac{1}{\xi - 1}
$$

Substituting this bound into (29) we find that indeed there can only be a single solution $Y_t$ if

$$
\frac{1}{\xi - 1} < \frac{1}{\xi} + \frac{1}{\sigma} \frac{Y_t}{Y_t - \phi}
$$

holds for any $Y_t > \phi$. This is satisfied if

$$
\frac{1}{\xi - 1} < \frac{1}{\xi} + \frac{1}{\sigma}
$$

which is easily transformed into (17).

### A.4 Proof of Proposition 2

From (18), the contour line is upward-sloping in $\mu_{t,t}$, implying that the upper arm of the $\Sigma$-locus ($\Sigma = \sigma_t^2/(1 - \rho^2)$) is overlapping with the lower arm ($\Sigma = \sigma_t^2 < \Sigma$) for some $\mu \in [\underline{\mu}, \overline{\mu}]$. Specifically, from (18), $\underline{\mu} = \mathcal{V}(\Sigma) + \nu$ and $\overline{\mu} = \mathcal{V}(\Sigma) + \nu$. Accordingly, both arms of the
Σ-locus intersect the μ-locus ($\mu = \log \bar{A}$), if and only if
\[ \mathcal{V}(\Sigma) + \nu \leq \log \bar{A} \leq \mathcal{V}(\Sigma) + \nu \]
or
\[ \log \bar{A} - \mathcal{V}(\Sigma) \leq \nu \leq \log \bar{A} - \mathcal{V}(\Sigma). \]

Summarizing, there are two steady states whenever $\nu \in [\underline{\nu}, \overline{\nu}]$ with $\underline{\nu} = \log \bar{A} - \mathcal{V}(\sigma^2_{\epsilon})$ and $\overline{\nu} = \log \bar{A} - \mathcal{V}(\sigma^2_{\epsilon}/(1 - \rho^2))$. Otherwise, there is a unique steady state at either $\Sigma$ (for $\nu < \underline{\nu}$) or at $\bar{\Sigma}$ (for $\nu > \overline{\nu}$).

### A.5 Proof of Proposition 3

We derive the three equations in Proposition 3. We start with the firms’ problem. Firms maximize
\[ \Pi_{i,t} = p_{i,t}Y_{i,t} - r_tK_{i,t} - w_tL_{i,t} \]
subject to
\[ Y_{i,t} = \frac{A_{i,t}^{1/(\xi-1)}}{\alpha}K_{i,t}^{\alpha}L_{i,t}^{1-\alpha} \]
\[ Y_{i,t} = Y_t(p_{i,t} - \xi) \]

Observe that $A_{i,t}$ is the only object that is not a choice object and different across firms $i$. We conjecture that $K_{i,t}$ and $L_{i,t}$ scale with $A_{i,t}$; that $p_{i,t}$ scales with $A_{i,t}^{-1/(\xi-1)}$; and that $Y_{i,t}$ scales with $A_{i,t}^{\xi/(\xi-1)}$. To verify this conjecture, observe that: (a) profits can be rewritten as
\[ \Pi_{i,t} = A_{i,t} \cdot \left[ \frac{p_{i,t}}{A_{i,t}^{-1/(\xi-1)}} \frac{Y_{i,t}}{A_{i,t}^{\xi/(\xi-1)}} - r_tK_{i,t} - w_tL_{i,t} \right] \]
where the term in brackets only depends on scale-free objects; (b) the production function can be rewritten in scale free terms as well,
\[ \frac{Y_{i,t}}{A_{i,t}^{\xi/(\xi-1)}} = \left( \frac{K_{i,t}}{A_{i,t}} \right)^{\alpha} \left( \frac{L_{i,t}}{A_{i,t}} \right)^{1-\alpha} \]
and so can the demand for $i$’s goods,
\[ \frac{Y_{i,t}}{A_{i,t}^{\xi/(\xi-1)}} = Y_t \left( \frac{p_{i,t}}{A_{i,t}^{-1/(\xi-1)}} \right)^{-\xi}. \]

This establishes that the conjecture was correct.
From market clearing, we know that
\[ \int_0^1 K_{i,t} \, di = K_t \]
Since \( K_{i,t} \) is proportional to \( A_{i,t} \), this implies that
\[ K_{i,t} = \frac{A_{i,t}}{A_t^{\xi - 1}} K_t \]  \hfill (32)
where \( A_t \) is defined in (15). Similarly,
\[ L_{i,t} = \frac{A_{i,t}}{A_t^{\xi - 1}} L_t. \]  \hfill (33)
Using (31), (32) and (33) aggregate output is then given by
\[ Y_t = \left( \int_0^1 Y_{i,t}^{\xi - 1} \, di \right)^{\frac{\xi}{\xi - 1}} = A_t K_t^{\alpha} L_t^{1-\alpha} \]  \hfill (34)
with
\[ r_t^K = \alpha (1 - \xi^{-1}) p_{i,t} Y_{i,t} K_{i,t} = \alpha (1 - \xi^{-1}) \frac{Y_t}{K_t} = \alpha (1 - \xi^{-1}) A_t K_t^{\alpha - 1} L_t^{1-\alpha} \]  \hfill (35)
and similarly,
\[ w_t = (1 - \alpha) (1 - \xi^{-1}) A_t K_t^{\alpha} L_t^{-\alpha}. \]  \hfill (36)
The Euler equation from households is standard, and given by
\[ C_t^{-1/\sigma} = \mathbb{E}_t \left[ \beta (1 + r_{i+1}^K - \delta) C_{i+1}^{-1/\sigma} \right] \]
Substituting in (35) yields (21). The optimality condition for labor is standard and given by
\[ v L_t^{1/\zeta} = (1 - \xi^{-1}) C_t^{-1/\sigma} w_t. \]
Substituting in (36) gives (22). Finally, the resource constraint (23) follows from (34).
### Table 3: Parameters for robustness exercise

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(\beta)</th>
<th>(\sigma)</th>
<th>(\zeta)</th>
<th>(\xi)</th>
<th>(\nu)</th>
<th>(\bar{A})</th>
<th>(\bar{\bar{A}})</th>
<th>(\rho)</th>
<th>(\sigma_e)</th>
<th>(\phi)</th>
<th>(\eta)</th>
<th>(\theta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline value</td>
<td>0.99</td>
<td>2.00</td>
<td>2.00</td>
<td>5.00</td>
<td>0.860</td>
<td>1.005</td>
<td>0.703</td>
<td>0.944</td>
<td>0.073</td>
<td>0.135</td>
<td>0.35</td>
<td>0.117</td>
</tr>
<tr>
<td>20% constrained</td>
<td>0.99</td>
<td>2.00</td>
<td>2.00</td>
<td>5.00</td>
<td>0.857</td>
<td>0.996</td>
<td>0.697</td>
<td>0.944</td>
<td>0.073</td>
<td>0.129</td>
<td>0.35</td>
<td>0.110</td>
</tr>
<tr>
<td>15% constrained</td>
<td>0.99</td>
<td>2.00</td>
<td>2.00</td>
<td>5.00</td>
<td>0.854</td>
<td>0.988</td>
<td>0.692</td>
<td>0.944</td>
<td>0.073</td>
<td>0.122</td>
<td>0.082</td>
<td>0.082</td>
</tr>
</tbody>
</table>

#### Figure 12: Robustness with respect to the share of constrained firms

*Note.* All panels are denominated in percentage deviations from the steady state. Solid blue lines are the baseline model responses; dashed dark blue lines correspond to the recalibrated model with a 20% share of constrained firms; dotted light blue lines correspond to the recalibrated model with a 15% share of constrained firms.

## B Additional Results

### B.1 Robustness to the Fraction of Financially Constrained Firms

In our calibration in Section 4.1, we worked with a parametrization that targeted a steady state share of 25% of constrained firms that do not have access to funding for the risky technology. In this section, we explore robustness to shares of constrained firms of 15% and 20%. Table 3 shows the recalibrated parameters.

In Figure 12 we compare the impulse responses in our baseline model with those in the two recalibrated models. The patterns are broadly similar across models, although amplification and persistence are weaker with smaller shares of constrained firms.

### B.2 Aggregate Productivity Shocks

Our focus in this paper is on shocks to the financial sector. One may wonder, however, whether our model with its financial and information frictions also fundamentally alters the response to aggregate productivity shocks. Figure 13 shows that the endogenous and exogenous uncertainty models behave nearly identically in response to an aggregate TFP
This is because an aggregate productivity shock does not shift $\nu_t$ nearly as much as the financial shock, as the response of average uncertainty in Figure 13 shows.

C Appendix to Section 7

C.1 Data

Our dataset is a yearly panel of public US firms. Forecast data is extracted from the Institutional Brokers Estimate System (IBES). Returns are obtained from the CRSP database and are adjusted for splits and dividends. All balance sheet data is from the COMPUSTAT database. From the original sample, we exclude all financial firms (SIC codes between 6000 and 6799) and firms in the electricity sector (SIC codes between 4900 and 4999). The resulting dataset ranges from 1976 to 2016 and covers, on average, 1979 firms per year.\footnote{Due to incomplete balance sheet data and firms with less than 2 forecasters for which $\sigma_{t,e}^{\text{fce}}$ is not defined, the effective number of observations is reduced for some of our empirical tests.} All variables are winsorized at the 1 percent level.

Units of observation are defined by firm $i$ and year $t$, where $t$ refers to the year in which earnings are realized. Let $m_{i,t}$ denote the fiscal-year end month of firm $i$. All balance sheet data and realized earnings per share (EPS) for observation $(i, t)$ are extracted at $m_{i,t}$. As in Senga (2018), we match each observation $(i, t)$ with analysts’ EPS-forecasts, $\mu_{i,j,t}^{\text{EPS}}$, extracted 8 months prior to $m_{i,t}$. That is, if in 2007, firm $i$’s fiscal-year ends in March, then $\mu_{i,j,2007}^{\text{EPS}}$ would be extracted in July 2006. Similarly, returns are synced with the fiscal-year end month of each observation, computed from 12 months prior to $m_{i,t}$ to $m_{i,t}$.

The variables used for our empirical exploration are defined in the main body of the paper. The KZ-index underlying our classification of financial constraints is based on Kaplan and \footnote{Formally, this is a shock to $\bar{A}$. We use the same half-life of 4 quarters as above and normalize the impact response of output to -1%.}

Figure 13: Response to an aggregate productivity shock
Table 4: Alternative proxies for financial stress

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel a: Financial conditions measured by dividends</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of constraint</td>
<td>0.030</td>
<td>0.026</td>
<td>0.018</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Observations</td>
<td>58 737</td>
<td>58 737</td>
<td>58 735</td>
<td>57 215</td>
</tr>
<tr>
<td>Adj. R-sq.</td>
<td>0.009</td>
<td>0.022</td>
<td>0.072</td>
<td>0.700</td>
</tr>
<tr>
<td><strong>Panel b: Financial conditions measured by leverage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of constraint</td>
<td>0.016</td>
<td>0.014</td>
<td>0.015</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Observations</td>
<td>58 641</td>
<td>58 641</td>
<td>58 639</td>
<td>57 124</td>
</tr>
<tr>
<td>Adj. R-sq.</td>
<td>0.000</td>
<td>0.015</td>
<td>0.070</td>
<td>0.706</td>
</tr>
</tbody>
</table>

Year × month FE no yes yes yes
Sector FE (4 digit) no no yes no
Firm FE no no no yes

Note: Standard errors clustered at the firm-level are in parenthesis.

Zingales (1997) and Lamont, Polk and Saá-Requejo (2001). Specifically,

\[ kz_{i,t} = -1.001909 \times \frac{\text{cashflow}_{i,t}}{k_{i,t-1}} + 0.2826389 \times Q_{i,t} + 3.139193 \times \frac{\text{debt}_{i,t}}{\text{total capital}_{i,t}} - \\
- 39.3678 \times \frac{\text{dividends}_{i,t}}{k_{i,t-1}} - 1.314759 \times \frac{\text{cash}_{i,t}}{k_{i,t-1}}, \]

where cashflow\(_{i,t}\) is the sum of COMPUSTAT items “income before extraordinary items” and “depreciation and amortization”, \(Q_{i,t}\) is (“market capitalization” + “total shareholder’s equity” – “book value of common equity” – “deferred tax assets”)/“total shareholder’s equity”, debt\(_{i,t}\) is “long-term debt” + “debt in current liabilities”, total capital\(_{i,t}\) is “long-term debt” + “debt in current liabilities” + “stockholders equity“, and \(k_{i,t}\) is “total property, plants and equipment” (see the Appendix to Lamont, Polk and Saá-Requejo, 2001 for a listing of the corresponding COMPUSTAT items).

C.2 Alternative Proxies for Financial Stress

In Table 4 we show additional results using two common alternative measures for financial constraints. The first is an indicator for whether dividend payouts are zero (Panel a), the second an indicator for whether the debt to capital ratio (which is a monotone function of leverage) is in the top 5% in a given year (Panel b).

The results are qualitatively similar to the ones in Table 2. The point estimates are
somewhat similar across columns (1)–(3) but are significantly smaller when firm fixed effects are included, which, as we argued above, takes away a lot of variation in the financial constraint variables.

The magnitudes in Table 4 are somewhat smaller compared to those in Table 2. This is not surprising given that one may think of the KZ indicator as a (more or less) optimized indicator which already includes dividend payouts and leverage in its composition; and thus dividends and leverage are both relatively more noisy measures of financial constraints and therefore subject to greater attenuation bias.