The Tail that Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation*

Julian Kozlowski† Laura Veldkamp‡ Venky Venkateswaran§
NYU NYU Stern NYU Stern

June 9, 2015

Abstract

The "great recession,"was a deep downturn with long-lasting effects on credit markets, labor markets and output. In the wake of this recession, many economists explored new sources of business cycle fluctuations, such as news, sentiment, disaster risk or uncertainty shocks. But these theories have difficulty explaining why post-recession output would remain persistently low. We propose a business cycle model where firms do not know the true distribution of economic shocks. Each period, they observe a new shock and re-estimate its distribution, using standard econometric tools. Once observed, a shock remains forever in an agent's data set. That observation alters beliefs about the probability of future shocks, which permanently shifts credit spreads, hiring and output. Highly leveraged firms, which are sensitive to tail risk, amplify this effect. Because data on tail risk is scarce, one extreme observation can "wag the tail"of the distribution. One negative shock that permanently increases estimated tail risk can stifle investment, hiring and output for years to come.

JEL Classifications: E32

Keywords: Stagnation, tail risk, belief-driven business cycles

---

*Veldkamp and Venkateswaran gratefully acknowledge financial support from the NYU Stern Center for Global Economy and Business.

†kozjuli@nyu.edu.

‡lveldkam@stern.nyu.edu.

§vvenkate@stern.nyu.edu.

1 Introduction

The recent “great recession” was a deep recession with long-lasting effects on the real economy. After observing that event, many economists began to look beyond standard theories of business cycles. New theories explore shocks to news, sentiment, disaster risk or uncertainty shocks. While these theories have provided new insights about the depth of the recession, most do not explain the length. They lack internal propagation mechanisms that would deliver persistent effects[1]. If a model’s predictions are only as persistent as the shocks are assumed to be, then the theory does not teach us anything about why this downturn had longer-lasting effects than others. So, if beliefs matter for business cycles, why did output remain below trend, long after the bad news had passed?

We propose a quantitative, belief-driven business cycle model where transitory shocks have large, persistent effects on real output. In our model, firms do not know the true distribution of shocks to capital quality. Each period, they observe a new shock realization, add it to their data set, use a standard kernel-density estimator to re-estimate the shock distribution, and then choose investment, labor and debt. The reason shocks have persistent effects is that once a shock is observed, it is in the agents’ data set forever after. The direct effect of the shock may pass quickly. But the observation of that event permanently alters the estimated probability of future events. Some types of shocks have small permanent effects and others have large permanent effects on beliefs. When the permanent effect on beliefs is large relative to the transitory direct effect of the shock, the business cycle effect is more persistent.

When capital investments are financed with defaultable debt, the effect of belief revisions is amplified. Debt is an asset whose payoffs are insensitive to output in most of the state space, but are very sensitive in the case of a left-tail outcome that triggers default. Thus, firms’ use of debt financing makes the economy sensitive to changes in estimated tail risk. The fact that we have so few tail observations makes estimates of tail risk particularly sensitive to new data. Therefore, the cost of issuing debt depends on perceived tail risk, which in turn, is sensitive to new observed shocks. Since real investment depends on the cost of credit, the combination of debt financing and re-estimating tail risks amplifies small shocks and allows them to generate large fluctuations in investment and output.

Our theory of time-varying disaster risk builds on existing models that trace out the macro consequences of an exogenous increase in disaster risk, e.g. Gourio (2012) and is similar in many respects to the existing theories of shocks to beliefs that drive business cycles[2]. Our

---

[1] See Backus et al. (2015) for a formal analysis of propagation in business cycle models with belief shocks.

[2] Papers on news driven business cycles include papers on news shocks, such as, Beaudry and Portier (2004), Lorenzoni (2009), Veldkamp and Wolfers (2007), papers on uncertainty shocks, such as Jaimovich and Rebelo (2006), Bloom et al. (2014), Nimark (2014) and papers on higher-order belief shocks, such as Angeletos and

approach based on re-estimating tails risks offers two advantages. First, our belief shocks are not exogenous. Without any discipline on the possible time-series of beliefs, many macroeconomic outcomes are rationalizable. Our agents' beliefs are the outcome of a standard kernel-density estimation using actual macroeconomic data. The second advantage is that our model offers an explanation for why some shocks are more persistent than others - events that trigger larger revisions in our perception of tail risk will have more persistent effects. This can help us understand why many recessions have rapid recoveries and yet, some do not.

The model features a continuum of firms that produce output with capital and labor. Each period, the firms' capital is hit by aggregate shocks to the capital quality as well as idiosyncratic shocks. Because the complete distribution of idiosyncratic shocks is observed every period, it is easy to learn and is therefore assumed to be common knowledge to our agents. But, only one realization of the aggregate shock is observed each period. As a result, agents re-estimate the distribution for aggregate shocks, adding one observation each time. Given an estimated distribution of future capital quality shocks, each firm chooses its capital investment. That capital investment can be financed with debt, which yields a a tax advantage to the firm, but also subjects it to bankruptcy costs if it defaults. We calibrate model parameters to match average leverage, including operating leverage, investment and default rates. The cost of issuing debt (the credit spread or risk premium) depends on the probability of default, which in turn, depends on the probability of adverse aggregate shocks. Thus, when the probability of a left tail event rises, the credit spread rises, debt issuance and real investment fall, and output declines. We explore the magnitude and persistence of this effect both with and without risk aversion. This mechanism is not novel - see, for example, Gourio (2013). The contribution of the paper lies in the way in which it gets tail risks to fluctuate, by tying these fluctuations to observable data, thereby generating endogenous persistence of such fluctuations.

We estimate the subjective distribution of capital quality shocks at each point in time using historical data on replacement and market value of the non-financial capital stock from the Flow of Funds report. Following the large negative shock to capital in 2008-'09, agents revise their estimates of the probability of similar shocks. This increase in tail risk triggers a decline in investment, hiring and output. Between 2009 and 2014, the model predicts a cumulative drop in capital of 15-20% and in output of about 12%. These effects are economically large, but perhaps more importantly, persistent. In fact, the model predicts that the long-run fall in output from the change in beliefs is about 13%.

These predicted effects of the financial crises, are similar in magnitude to the real effects observed in the data. Hall (2014) estimates that the U.S. capital stock and U.S. real GDP are each 13% lower than they would be if the economy had continued to grow at its pre-crisis rate

---

La'O (2013) or Huo and Takayama (2015).

of trend growth. Our model, which is calibrated to capital quality, default and leverage data, but not to GDP data, matches the size of the output drop. Hall also argues that the depressed rate of business capital formation was the single largest contributor to the persistent depressed output (often referred to as secular stagnation) in the post-crisis period. We use our model output to do the same decomposition as Hall does. We show that the symptoms of secular stagnation, as seen in changes in investment and labor between 2008 and 2015, resemble those in the data.

Our findings reveal that belief-driven theories of business cycles are capable of explaining both deep and long-lasting recessions. The missing ingredient is relaxing the assumption that agents know the distribution of aggregate economic shocks. Allowing agents to estimate this shock distribution in real time can add persistence to any model where beliefs play an important role. In a model of debt-financed firms, learning about the probability distribution of capital quality shocks delivers a quantitatively plausible theory of secular stagnation. It suggests that the recovery from the great recession has been slow because we learned that financial crises are still possible in the U.S. and this new knowledge permanently changed our assessment of macroeconomic risk.

Comparison to the literature A small number of uncertainty-based theories of business cycles also deliver persistent effects from transitory shocks. In Straub and Ulbricht (2013) and Van Nieuwerburgh and Veldkamp (2006), a negative shock to output raises uncertainty, which feeds back to lower output, which in turn creates more uncertainty. To get even more persistence, Fajgelbaum et al. (2014) combine this mechanism with an irreversible investment cost, a combination which can generate multiple steady-state investment levels.

These uncertainty-based explanations leave two questions unanswered. First, why were credit markets and investment the hardest hit and the most persistently impaired after the crisis? Second, why did the depressed level of economic activity continue long after the VIX and other measures of uncertainty had recovered? Our theory is based on tail risk. Like uncertainty, tail risk is a moment of the perceived distribution of outcomes. But the value of debt is particularly sensitive to this moment. While a rise in tail risk will often increase conditional variance (uncertainty), the two do not always move in lock step. Figure 1 plots the values of the VIX and the SKEW indices from 1990-2014. These are an implied volatility index and a 2-standard-deviation tail risk index, both constructed using option prices by the Chicago Board of Options. The indices show that while post-crisis uncertainty measures (VIX) recovered quickly, the tail risk (SKEW) index did not. It started rising as the VIX was peaking and continued to rise throughout the post-crisis period. Thus, financial data suggests that the effects of uncertainty-based theories should have dissipated by now. Even if some initial shocks had persistent effects, if those effects work through uncertainty, when the uncertainty has passed, the effects should have as well. The same is true of our tail risk effects. The difference is that the data reveal that tail risk has lingered, making it a better candidate for explaining continued depressed output.

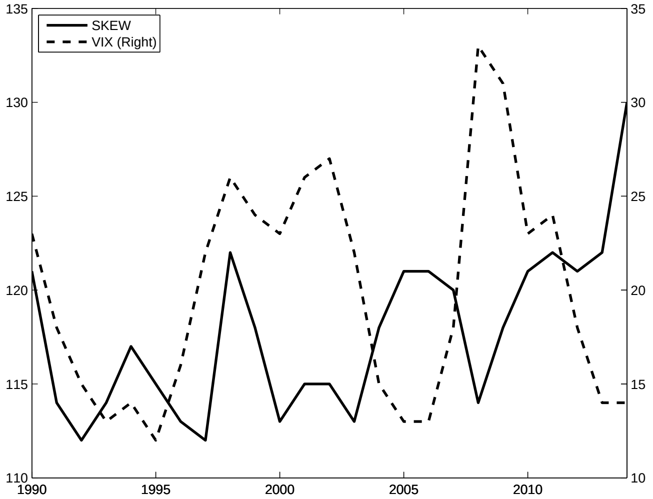

Figure 1: Post-2007, tail risk remained persistently high. Volatility did not.

This belief formation process in our model is similar to the parameter learning models by Johannes et al. (2015), Cogley and Sargent (2005) and Orlik and Veldkamp (2014), but none of these papers has a production economy or considers persistent shocks to output. This paper adds a story about how parameter learning can generate large and persistent real effects in a standard macroeconomic setting with production and investment.

Finally, the paper is related to many popular theories of the great recession, such as Gertler et al. (2010), Gertler and Karadi (2011), Brunnermeier and Sannikov (2014) and Gourio (2012). Moriera and Savov (2015) is similar to our model in that agents learn and it changes their demand for shadow banking (debt) assets. But their agents learn about a hidden two-state Markov process, which has persistence built in. While this literature has taught us an enormous amount about the mechanisms that triggered declines in lending and output in the financial crisis, it also hard-wired in persistent shocks. Our model aims to complement these theories by describing a simple mechanism that is compatible with many existing frameworks, is easy to implement, and delivers persistence. Rather than substituting for these existing narratives about the mechanics of financial crisis, our belief-formation mechanism adds another layer to the story, by explaining why some shocks deliver more persistent responses than others.

2 Model

We explore the quantitative implications of parameter learning for persistence in a standard business cycle framework widely used in recent work on the financial crisis and the Great Recession. We begin by laying out a general model along the lines of Gertler and Karadi (2011) and Gourio (2012). Firms finance investment and payroll expenses using a combination of debt and equity financing and are subject to aggregate and idiosyncratic shocks. Our main innovation is to introduce real-time model estimation - specifically, agents in the model use the observed shocks and standard econometric tools (kernel density estimators) to estimate the shape of the distribution of aggregate shocks.

Preferences and technology: An infinite horizon, discrete time economy has a representative household, with preferences over consumption and labor supply, following

$$U_t = \left[(1-\beta) \left(C_t - \frac{\zeta L_t^{1+\gamma}}{1+\gamma} \right)^{1-\psi} + \beta E_t \left(U_{t+1}^{1-\eta} \right)^{\frac{1-\psi}{1-\eta}} \right]^{\frac{1}{1-\psi}} \tag{1}$$

where ψ is the inverse of the intertemporal elasticity of substitution, η indexes risk-aversion and γ is inversely related to the elasticity of labor supply.

The economy is also populated by a unit measure of firms, indexed by i and owned by the representative household. Firms produce output with capital and labor, according to a standard Cobb-Douglas production function $Ak_{it}^{\alpha} l_{it}^{1-\alpha}$, where A is total factor productivity (TFP), which is the same for all firms and constant over time. Firms are subject to an aggregate shock to capital quality ϕ_t. A firm that enters the period with capital $\hat{k}_{it}$ and is hit by a shock ϕ_t has effective capital $k_{it} = \phi_t \hat{k}_{it}$.

Our focus on capital quality shocks as the source of aggregate fluctuations is in the tradition of a number of recent papers on financial frictions, crises and the Great Recession - for example, Gertler et al. (2010), Gertler and Karadi (2011), Brunnermeier and Sannikov (2014) and Gourio (2012). These shocks work to *permanently* scale up or down the effective capital stock. However, such shocks by themselves are not enough to generate long-lived output responses. An adverse quality shock creates incentives to invest rapidly and return to a steady state level of capital. To deter this investment boom, Gertler and Karadi (2011), Gourio (2012) and others add persistence to the shock process - a bad shock not only wipes out a fraction of today's capital, but also makes it more likely that any investments today will be hit by bad shocks tomorrow. This persistence then spills over to aggregate outcomes, allowing them to generate long-lived output responses.

Importantly, we assume that the shock ϕ_t is i.i.d. The independence assumption ensures our persistence arises endogenously, from changes to beliefs, and distinguishes our results from those with exogenous persistence (autocorrelated ϕ_t). The distribution of these shocks G is unknown to agents. Learning about G is the key novel feature of our setup and is the focus of the paper. The following subsection will describe how agents use the observed history of ϕ_t to update their beliefs about G.

Firms are also subject to an idiosyncratic shock v_{it} . These shocks scale up and down the total resources available to each firm (before paying debt, equity or labor):

$$\Pi_{it} = v_{it}\left[Ak_{it}^{\alpha}l_{it}^{1-\alpha} + (1-\delta)k_{it}\right] \tag{2}$$

where δ is the rate of capital depreciation. The shocks v_{it} are i.i.d across time and firms and are drawn from a known distribution[3], F. The mean of the idiosyncratic shock is normalized to be one: $\int v_{it} \;\; di = 1$

Information Sets: The key innovation in the model is the assumption that agents must estimate the aggregate shock distribution G. Their common information set is $\mathcal{I}_t$, which includes all aggregate and idiosyncratic variables observed up to and including time-t. At each point in time, they use the empirical distribution of ϕ_t up to that point to construct an estimate $\hat{G}_t$ of the true distribution G. Formally, at every date t, agents construct the following kernel density estimator of the pdf g [4]:

$$\hat{g}_t\left(\phi\right) = \frac{1}{n_t\kappa}\sum_{s=0}^{n_t-1}\Omega\left(\frac{\phi-\phi_{t-s}}{\kappa}\right)$$

where $\Omega\left(\cdot\right)$ is the standard normal density function, κ is the bandwidth parameter and n_t is the number of available observations of at date t. As new data arrives, agents update their estimates, generating a sequence of beliefs $\left\{\hat{G}_t\right\}$.

To keep our problem tractable, we follow most previous work on learning (Cogley and Sargent (2005), Piazzesi et al. (2015), Johannes et al. (2015)) in using anticipated utility (Kreps, 1998). Under this notion, agents are myopic with respect to changes in their future beliefs but otherwise make optimal decisions. In other words, at each date t, agents act as if the true distribution is $\hat{G}_t$. While the alternative of acknowledging and adjusting actions for the possibility of future changes in beliefs may be appealing, it adds considerable opacity and computational complexity and typically generates similar dynamics.

---

[3] This is a natural assumption - with a continuum of firms and a stationary shock process, firms can learn the complete distribution of any idiosyncratic shocks after one period.

[4] In our numerical implementation, we fit a smooth density function to the empirical distribution. We also studied a handful of flexible parametric specifications, which yielded similar results.

Labor, credit markets and default: Firms hire labor in advance, i.e. before observing the realizations of the aggregate and idiosyncratic shocks. Importantly, wages are non-contingent - in other words, workers are promised a non-contingent payment and face default risk.

Firms have access to a competitive non-contingent debt market, where lenders offer bond price (or equivalently, interest rate) schedules as a function of all relevant aggregate and idiosyncratic states.

In order to characterize these schedules, we need to analyze the firm's default decision. A firm enters the period with an obligation, b_{it+1} to bondholders and a promise of $w_{it+1}l_{it+1}$ to its workers. The shocks are then realized and the firm (i.e. its shareholders) decide whether to repay their obligations or default. A firm that defaults makes no payments to equity holders. Formally, default is optimal for shareholders if

$$\Pi_{it+1} - b_{it+1} - w_{it+1}l_{it+1} + \Gamma_{t+1} < 0$$

where Γ_{t+1} is the present value of continued operations (we characterize this object later in this section - specifically, we will show that, since idiosyncratic shocks are iid, this is the same for all firms and, in equilibrium, equal to 0). Thus, the default decision is a function of the resources available to the firm (Π_{it+1}) and the *total* obligations of the firm to both bondholders and workers ($b_{it+1} \ + w_{it+1}l_{it+1} \ \equiv B_{it+1}$). The former is a function of the capital and labor choices, as well as the realizations of shocks. Let $r_{it}\left(\Pi_{it+1}, B_{it+1}, S_t\right) \in \{0,1\}$ denote the default decision of the firm, where we make explicit the dependence on the aggregate state S_t. Importantly, S_t includes the estimated distribution function $\hat{G}_t$.

In the event of default, the workers and bondholders take over the firm. The productive resources of a defaulting firm are sold to an identical new firm at a discounted price, equal to a fraction $\theta < 1$ of the value of the defaulting firm. The proceeds are distributed *pro-rata* among the creditors (both bondholders and unpaid workers). Note that the claims of both bondholders and workers have equal seniority[5].

Let $q\left(\hat{k}_{it+1},\ l_{it+1}, w_{it+1},\ b_{it+1}, S_t\right)$ denote the bond price schedule faced by a firm in period t. In other words, the firm receives $q\left(\cdot\right)$ in exchange for a promise to pay one unit of output at date $t+1$. Note that the bond price is determined before the next period's capital quality shocks are known. So, the price depends on the amount of capital invested $\hat{k}_{it+1}$, but it cannot be made contingent on the effective capital that will be available for production k_{it+1} or the profit shock v_{it+1}. The dependence on the other variables follows from the discussion on the

---

[5] Note also that this means that default does not destroy resources - the penalty is purely private.

default decision,

$$q\left(\hat{k}_{it+1},\ l_{it+1}, w_{it+1},\ b_{it+1}, S_t\right) = \mathbb{E}M_{t+1}\left[r\left(\Pi_{it+1}, B_{it+1}\right) + \left(1 - r\left(\Pi_{it+1}, B_{it+1}\right)\right)\frac{\theta\tilde{V}\left(\Pi_{it+1}, S_{t+1}\right)}{B_{it+1}}\right]. \tag{3}$$

where $\tilde{V}\left(\Pi_{it+1}, S_{t+1}\right)$ is the value of the assets of the firm (to be characterized later) and M_{t+1} is the stochastic discount factor of the representative household, which, given our Epstein-Zin specification takes the form

$$M_{t+1} = \left(\frac{dU_t}{dC_t}\right)^{-1}\frac{dU_t}{dC_{t+1}} = \beta\left[E_t\left(U_{t+1}^{1-\eta}\right)\right]^{\frac{\eta-\psi}{1-\eta}} U_{t+1}^{\psi-\eta}\left(\frac{u\left(C_{t+1}, L_{t+1}\right)}{u\left(C_t, L_t\right)}\right)^{-\psi} \tag{4}$$

We assume that debt is associated with a tax advantage, which creates incentives for firms to borrow. A firm which issues debt at price q_{it} and promises to repay b_{it+1} in the following period, receives a total date-t payment of $\chi q_{it} b_{it+1}$, where $\chi > 1$. This subsidy to debt issuance along with the cost of default introduces a trade-off in the firm's capital structure decision, breaking the Modigliani-Miller theorem[6].

For a firm that does not default, the dividend payout is its total available resources times output shock, minus its payments to debt and labor, minus the cost of building next period's capital stock (the undepreciated current capital stock is included in Π_{it}), plus the revenue earned from issuing new debt, including its tax subsidy:

$$d_{it} = \Pi_{it} - B_{it} - \hat{k}_{it+1} + \chi q_{it} b_{it+1} \tag{5}$$

Importantly, we do not restrict dividends to be positive, with negative dividends interpreted as (costless) raising of equity. Thus, firms are not financially constrained, ruling out another potential source of persistence.

Workers, who are also members of the representative family, evaluate their wage claims using the stochastic discount factor, M_{t+1}. This implies that the present value of a promise of wage w_{it+1} is given by

$$w_{it+1}\mathbb{E}M_{t+1}\left[r\left(\Pi_{it+1}, B_{it+1}\right) + \left(1 - r\left(\Pi_{it+1}, B_{it+1}\right)\right)\frac{\theta\tilde{V}\left(\Pi_{it+1}, S_{t+1}\right)}{B_{it+1}}\right] = w_{it+1} q_{it}$$

where the expectation is taken over aggregate and idiosyncratic shocks. From the household's

---

[6] The subsidy is assumed to be paid by a government that finances it through a lumpsum tax on the representative household.

problem, we can derive the following optimality condition for labor supply:

$$\begin{aligned} w_{it+1} q_{it} \frac{dU_t}{dC_t} &= \frac{dU_t}{dL_{t+1}} \\ w_{it+1} q &= \left(\frac{dU_t}{dC_t}\right)^{-1} \frac{dU_t}{dL_{t+1}} \\ &\equiv \mathcal{W}_t \end{aligned} \tag{6}$$

In other words, the expected value of wages, weighted by the economy-wide stochastic discount factor M_{t+1} is the same for all firms and is equal to the marginal rate of substitution of the representative household. The wage promise,w_{it+1}, must offer workers compensation for default risk. Since the risk is identical for bonds and wage payments, this risk adjustment involves simply multiplying the promised wage by the equilibrium bond price. In other words, the workers are essentially paid through bonds.

Timing, value functions and equilibrium: The timing of events in each period t is as follows:

1. Firms enter the period with a capital stock $\hat{k}_{it}$, labor l_{it}, outstanding debt b_{it}, and a wage obligation $w_{it} l_{it}$.
2. The aggregate capital quality shock ϕ_t and the firm-specific profit shock v_{it} are realized. Production takes place.
3. The firm decides whether to default or repay ($r_{it} \in \{0, 1\}$) its bond and labor claims.
4. The firm makes capital $\hat{k}_{it+1}$, debt b_{it+1} choices for the following period, along with wage/employment contracts w_{it+1} and l_{it+1}. Workers commit to next-period labor supply l_{it+1}. Note that all these choices are made concurrently.

Value functions: In recursive form, the expected value of a firm is

$$V\left(\Pi_{it}, B_{it}, S_t\right) = \max \left[0, \max_{d_{it}, \hat{k}_{it+1}, b_{it+1}, w_{it+1}, l_{it+1}} d_{it} + \mathbb{E}_t M_{t+1} V\left(\Pi_{it+1}, B_{it+1}, S_{t+1}\right)\right] \tag{7}$$

where d_{it} is the dividend payment given by (5) and the realized effective capital stock is $k_{t+1} = \phi_t \hat{k}_{t+1}$. The first max operator in (7) captures the firm's option to default if the value of the firm is negative. The expectation is taken over the idiosyncratic and aggregate shocks, taking the estimated aggregate shock distribution S_t as given. Substituting for d_{it} from (5), the second expression inside the square brackets can be written as

$$
\begin{aligned}
& \max_{\hat{k}_{it+1},\ b_{it+1}, w_{it+1}, l_{it+1}} \quad \Pi_{it} - B_{it} - \hat{k}_{it+1} + \chi q_{it} b_{it+1} + \mathbb{E} M_{t+1} V\left(\Pi_{it+1}, B_{it+1}, S_{t+1}\right) \\
= \; & \Pi_{it} - B_{it} + \max_{\hat{k}_{t+1},\ b_{t+1}, w_{t+1}, l_{t+1}} \; -\hat{k}_{it+1} + \chi q_{it} b_{it+1} + \mathbb{E}_t M_{t+1} V\left(\Pi_{it+1}, B_{it+1}, S_{t+1}\right) \\
\equiv \; & \Pi_{it} - B_{it} + \Gamma_t
\end{aligned}
$$

Finally, the value of the assets of a defaulting firm $\tilde{V}\left(\Pi_{it}, S_t\right)$ is simply the value of a firm with no external obligations, i.e. $V\left(\Pi_{it}, 0, S_t\right) = \tilde{V}\left(\Pi_{it}, S_t\right)$.

For a given $\hat{G}_t$, a recursive equilibrium is a set of (i) functions for aggregate consumption and labor that maximize (1) subject to a budget constraint, (ii) firm value functions and associated policy functions that solve (7) , taking the bond price and wage functions (3) , and (6) and the stochastic discount factor (4) as given. (iii) aggregate consumption and labor are consistent with individual choices.

3 Solving the Model

We now characterize the equilibrium and explore how tail events and the subsequent changes in beliefs affect the persistence and level of macro and financial outcomes. We present only the key equations here and relegate the detailed derivations to Appendix A.

We first note that, for any given aggregate capital shock ϕ_t, we can represent the optimal repayment policy as a threshold rule in the idiosyncratic output shock v_{it},

$$
r\left(\Pi_{it}, B_{it}, S_t\right) = \begin{cases} 0 & \text{if } \; v_{it} \; < \; \underline{v}\left(S_t\right) \\ 1 & \text{if } \; v_{it} \; \geq \; \underline{v}\left(S_t\right) \end{cases}
$$

Working from the first-order condition for the firm's capital choice, we find that the optimal $\hat{k}_{t+1}$ choice solves[7]

$$
1 + \chi \mathcal{W}_t \frac{l_{t+1}}{\hat{k}_{t+1}} = \mathbb{E}\left[M_{t+1} R^k_{t+1} J^k(\underline{v})\right] \tag{8}
$$

$$
\begin{aligned}
\text{where} \qquad R^k_{t+1} &= A \phi^{\alpha}_{t+1} \left(\frac{\hat{k}_{t+1}}{l_{t+1}}\right)^{\alpha - 1} + (1-\delta)\, \phi_{t+1} \\
J^k(\underline{v}) &= 1 + h\left(\underline{v}\right)(\theta\chi - 1) + \underline{v}\left(1 - F\left(\underline{v}\right)\right)(\chi - 1) \\
h\left(\underline{v}\right) &\equiv \int_{-\infty}^{\underline{v}} v f(v) dv
\end{aligned}
$$

---

[7] Since all firms are identical, they make symmetric choices and accordingly, we suppress the i subscript.

The term R^k_{t+1} is related to the return on capital, augmented with the capital quality shock, ϕ_{t+1}. The term $J^k(\underline{v})$ reflects the net effect of distortions induced by the tax advantage and default penalties associated with debt. In the absence of any these distortions (e.g. if $\chi = 1$), $J^k(\underline{v}) = 1$, reducing (8) to a standard Euler equation. In general, however, the wedge $J^k(\underline{v})$ distorts the equilibrium choice of capital away from the choices of a planner.

The optimality condition for labor looks quite similar. Just like with capital, firms equate the marginal cost of an additional unit of labor, namely $\mathcal{W}_t$, with the expected marginal product of labor, adjusted for the effect of additional promised wages on the cost of default:

$$\chi \mathcal{W}_t = \mathbb{E}\left[M_{t+1}\left(1-\alpha\right)A\phi^{\alpha}_{t+1}\left(\frac{\hat{k}_{t+1}}{l_{t+1}}\right)^{\alpha} J^l(\underline{v})\right] \tag{9}$$

$$\text{where} \qquad J^l(\underline{v}) = 1 + h\left(\underline{v}\right)\left(\theta\chi - 1\right) - \underline{v}^2 f\left(\underline{v}\right)\chi\left(\theta - 1\right)$$

Finally, the firm's optimality condition with respect to leverage

$$\left(1-\theta\right)\mathbb{E}_t\left[M_{t+1}\underline{v}f\left(\underline{v}\right)\right] = \left(\frac{\chi - 1}{\chi}\right)\mathbb{E}_t\left[M_{t+1}\left(1 - F\left(\underline{v}\right)\right)\right] \tag{10}$$

The left hand side is the marginal cost of increasing leverage - it raises the expected losses from the default penalty (a fraction $(1-\theta)$ of the firm's value). The right hand side is the marginal benefit - the tax advantage times the value of debt issued.

The three optimality conditions, $(8) - (10)$, along with those from the household side - in particular, the labor supply condition (6) - characterize the equilibrium of this economy and can be solved numerically.

In our numerical analysis, we solve two variants of the model

1. **Version 1: Epstein-Zin preferences, fixed leverage** The first maintains the EZ specification for preferences, but, for tractability, exogenously fixes leverage (defined as the ratio of total obligations to capital). This is equivalent to replacing (10) with

$$\frac{B_{it+1}}{\hat{k}_{it+1}} = lev^{\text{Target}}$$

2. **Version 2: Quasi-linear preferences, endogenous leverage** The second variant preserves the endogenous choice of leverage i.e. uses (10) but sets the utility parameters $\psi = \eta = 0$, reducing the utility function (1) to a quasi-linear specification:

$$\left(1-\beta\right)\mathbb{E}_t\left[\sum_{s=0}^{\infty} C_{t+s} - \frac{\zeta}{1+\gamma}L^{1+\gamma}_{t+s}\right]$$

This eliminates consumption smoothing incentives and therefore, any time variation in the discount factor, i.e. $M_{t+1} = \beta$. This isolates the effect coming from changes in expected returns to investing.

4 Measurement and Calibration

One of the key strengths of our belief-driven theory is that, by assuming that agents form beliefs as an econometrician would, we allow the data to discipline beliefs. In this section, we parameterize the model to match key features of the US economy. We then subject the model economy to the realized time series of capital quality shocks from US post-war data and evaluate the predictions for aggregates that we did not calibrate to, such as investment, output and consumption.

4.1 Measuring capital quality shocks

The next step is to construct a time series of $\{\phi_t\}$. We use annual data on non-financial assets of non-financial corporations in the US economy. The Flow of Funds reports published by the Federal Reserve contain two such series - one evaluated at historical cost and the other at replacement cost or market value. We interpret the latter as corresponding to effective capital. Letting P_t^k denote the nominal price of capital in t, we can then map these two series into model objects as follows:

$$\begin{aligned} NFA_t^{HC} &= \text{Historical cost of non-financial assets in } t = P_{t-1}^k X_{t-1} + (1-\delta)\, NFA_{t-1}^{HC} \\ NFA_t^{RC} &= \text{Replacement cost of non-financial assets in } t = P_t^k K_t \end{aligned}$$

where X_{t-1} is investment in period $t-1$. Now,

$$\begin{aligned} \frac{NFA_t^{RC}}{NFA_{t-1}^{RC}} - 1 &\approx \ln\ NFA_t^{RC} - \ln\ NFA_{t-1}^{RC} \\ &= \ln\ P_t^k - \ln\ P_{t-1}^k + \ln\ K_t - \ln\ K_{t-1} \end{aligned}$$

Thus, the change in the NFA_t^{RC} series reflects both changes in effective capital (K_t) as well as price changes. We then use the change in price index for non-residential investment from the National Income and Product Accounts (denoted $\ln\ PINDX_t^{Inv}$) to control for the latter[8]:

$$\ln\ P_t^k - \ln\ P_{t-1}^k = \ln\ PINDX_t^{Inv} - \ln\ PINDX_{t-1}^{Inv}$$

---

[8] Our results are robust to alternative measures of price changes.

Substituting and rearranging, we get

$$\frac{K_t}{K_{t-1}} \approx 1 + \ln\ K_t - \ln\ K_{t-1} = 1 + \ln\ \frac{NFA_t^{RC}}{NFA_{t-1}^{RC}} - \ln\ \frac{PINDX_t^{Inv}}{PINDX_{t-1}^{Inv}} \tag{11}$$

Finally, from the law of motion for $K_t = \phi_t \left(X_{t-1} + K_{t-1} \left(1 - \delta\right)\right),$

$$\begin{aligned} \frac{K_t}{K_{t-1}} &= \phi_t \left(\frac{X_{t-1}}{K_{t-1}} + 1 - \delta \right) = \phi_t \left(\frac{P_{t-1}^k X_{t-1}}{P_{t-1}^k K_{t-1}} + 1 - \delta \right) \\ &= \phi_t \left(\frac{NFA_t^{HC} - \left(1 - \delta\right) NFA_{t-1}^{HC}}{NFA_{t-1}^{RC}} + 1 - \delta \right) \end{aligned}$$

Combining with (11), we get

$$\phi_t = \frac{\frac{K_t}{K_{t-1}}}{\frac{NFA_t^{HC} - (1-\delta) NFA_{t-1}^{HC}}{NFA_{t-1}^{RC}} + 1 - \delta} = \frac{1 + \ln\ \frac{NFA_t^{RC}}{NFA_{t-1}^{RC}} - \ln\ \frac{PINDX_t^{Inv}}{PINDX_{t-1}^{Inv}}}{\frac{NFA_t^{HC} - (1-\delta) NFA_{t-1}^{HC}}{NFA_{t-1}^{RC}} + 1 - \delta} \tag{12}$$

Using (12) as a measurement equation, we construct an annual time series for capital quality shocks for the US economy over the last few decades. The left panel of Figure 2 plots the resulting series. For most of the sample period, the shock realizations are in a relatively tight range around 1, but at the onset of the recent Great Recession, we saw two large adverse realizations: 0.93 in 2008 and 0.84 in 2009. We will use the model to simulate the responses of the economy over time to negative shocks of this magnitude.

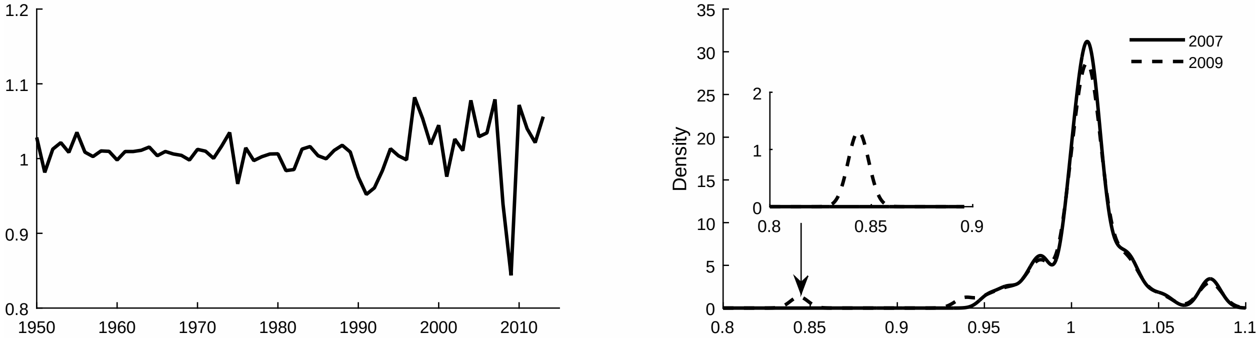

Figure 2: **Distribution of capital quality shocks.** The left tail shows the effect of the Great Recession.

We then apply a standard kernel density estimation procedure to this time series to construct a sequence of beliefs $\{\hat{g}_t\}$. The resulting estimates for 2007 and 2009 are shown in the right panel of Figure 2. They show a significant increase in the perceived tail risk post–2007. The density function for 2007 implies almost zero mass below 0.95. After the great recession, agents revise their estimates and now attach a non-trivial probability to significantly worse realizations.

4.2 Calibration

We first choose parameter values/targets that are common for the two versions of our model. A period is interpreted as a year and the discount factor β is set to 0.9. The share of capital in the production, α, is set to 0.40. The recovery rate upon default, θ, is set to 0.70, following Gourio (2013). The distribution for the idiosyncratic shocks, v_{it} is assumed to be lognormal, i.e. $\ln v_{it} \sim N\left(-\frac{\hat{\sigma}^2}{2}, \hat{\sigma}^2\right)$ with $\hat{\sigma}^2$ chosen to target a default rate[9] of 0.02. The labor supply parameter, γ, is set to 0.5, in line with Midrigan and Philippon (2011), corresponding to a Frisch elasticity of 2. The labor disutility parameter ζ and the TFP term in production are normalized to 1.

Next, we turn to the version-specific parameters.

Version 1 (EZ preferences, fixed leverage): For this version, we use preference parameters from the asset pricing literature[10] and set $\psi = 0.5$ (or equivalently, an intertemporal elasticity of substitution of 2) and $\eta = 10$. Our target for steady state leverage is 0.70, obtained by adding the wage bill (approximately 0.2 of the steady state capital stock) to a target for financial leverage (the ratio of external debt to capital of 0.5- from Gourio (2013)). Since leverage is exogenous, the tax advantage χ is a free parameter. We set it to a baseline value of 1.06 and verified numerically that our results are not particularly sensitive to this choice.

Version 2 (Quasi-linear preferences, endogenous leverage): In this version, $\psi = \eta = 0$, but leverage is chosen optimally. We now choose the tax advantage parameter χ to generate a leverage of 0.7. This leads to a value of 1.06 for χ.

Table 1 summarizes all parameter choices.

5 Quantitative Results

Our main goal in the quantitative exercise is to explore the size and persistence of the macroeconomic response to large, transitory negative shocks. We want to understand how belief updating, debt and risk aversion each contribute to this response. With this goal in mind, we perform the following experiment using historical data on ϕ_t realizations from 1950-2009, measured using the strategy outlined in Section 4. We begin by estimating $\hat{G}_{2007}$ using the data through 2007. Then, starting from the steady state associated with this estimate[11], we subject the model economy to a sequence of two adverse realizations - 0.93 and 0.84, which correspond to the shocks that we observed in 2008 and 2009. This leads to a revised estimate

---

[9] This is in line with the target in Khan et al. (2014), though higher than the one in Gourio (2013). We verified that our quantitative results are not sensitive to this target.

[10] See discussion in Gourio (2013).

[11] The steady state is obtained by simulating the model for 1000 periods using the $\hat{G}$ and the associated policy functions, discarding the first 500 observations and time-averaging across the remaining periods.

Parameter	Value	Description
β	0.91	Discount factor
η	10	Risk aversion
ψ	0.50	1/Intertemporal elasticity of substitution
γ	0.50	1/Frisch elasticity
ζ	1	Labor disutility
α	0.40	Capital share
δ	0.03	Depreciation rate
A	1	TFP
χ	1.06	Tax advantage of debt
θ	0.70	Recovery rate
$\hat{\sigma}$ (version 1)	0.33	Idiosyncratic volatility
$\hat{\sigma}$ (version 2)	0.26	Idiosyncratic volatility
lev^{Target}	0.70	Leverage ratio

Table 1: **Parameters**

for the distribution, $\hat{G}_{2009}$. We solve the model using this revised estimate and use those policy functions to generate a time path for aggregate variables, under the assumption that the shock realizations from 2010 onward are equal to their average value.

We then compare these paths to an otherwise identical economy without learning, i.e. one where agents are assumed to know the final distribution $\hat{G}_{2009}$ from the very beginning and so, do not revise their beliefs. In other words, we generate the same impulse response functions starting from the steady state associated with $\hat{G}_{2009}$ and holding beliefs constant throughout. This corresponds to a standard rational expectations approach, where agents are assumed to know the true distribution. The econometrician estimates this distribution using all the available data.

We perform these exercises for both versions of our model. This allows us to analyze the interaction of belief revisions with risk aversion (which creates consumption-smoothing incentives) and endogenous leverage choice. We also compare our baseline predictions (from the version with risk averse preferences) to predictions from a model with the same belief updating process, but without debt.

5.1 Results: Risk aversion, fixed leverage

The impulse response functions for the first version are shown in Figure 3. The first panel shows the time path for ϕ_t (as deviations from its average value of 1). The solid and dashed lines in the remaining panels show the response of aggregate variables with and without learning respectively. In the absence of belief revisions, a negative capital quality shock prompts the firm to increase investment to replenish the lost effective capital. While the curvature in the utility function moderates the speed of this transition to an extent, the overall pattern of a steady recovery back to the original steady state is clear. Under learning, however, the new information leads firms to permanently adjust their estimates of the expected return to investment and labor. As a result, the shocks induce a prolonged stagnation, with the economy trending towards a new, lower steady state level. Output in the new steady state is about 13% lower than the original one. The corresponding drops in capital and labor are about 19% and 9%. Thus, even though the ϕ_t shocks were transitory, they were so large that the resulting change in beliefs permanently reduces economic activity[12].

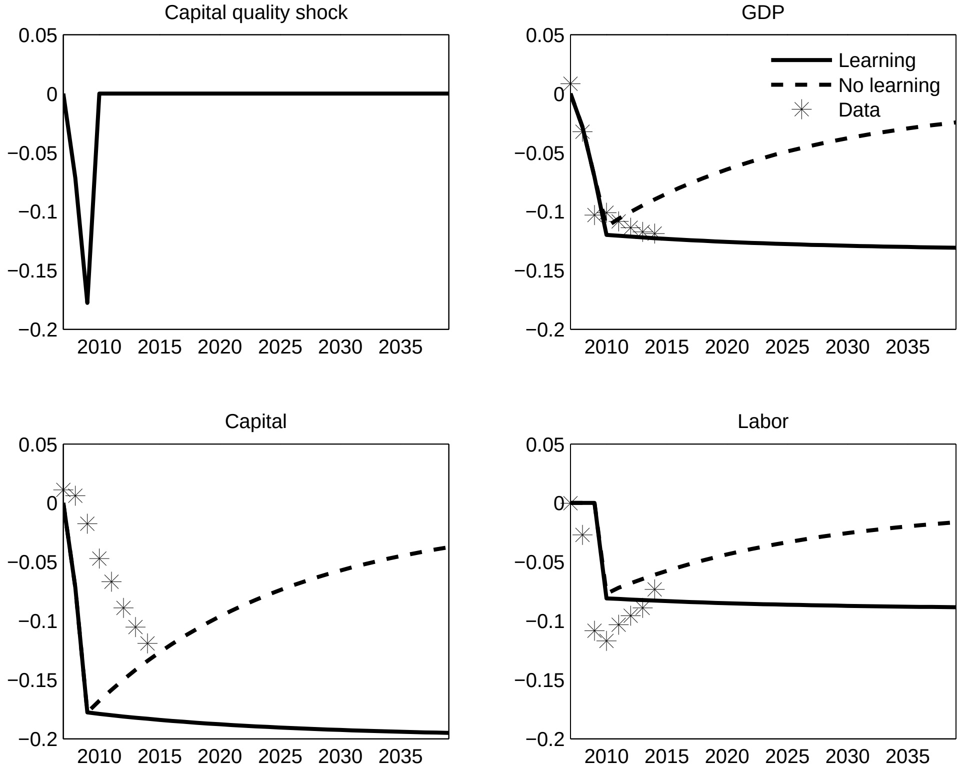

Figure 3: **Re-estimating shock distribution creates persistent responses in output, investment and labor.** Impulse responses with Epstein-Zin preferences and fixed leverage. Solid line (learning) is the model with belief updating. Dashed line (no learning) is an identical model where agents believe that shocks are drawn from the distribution estimated on the full sample of data and never revise those beliefs. Zero is the steady state level in each economy.

Figure 3 also plots the actual data on the deviations of output, capital stock and labor for the US economy from pre-crisis trends[13]. As the graph shows, the model's predictions for the drop in output line up remarkably well with the data. The predicted path for capital

---

[12] In levels, the no learning and learning economies converge at the end. But they start at different steady states and we are normalizing each series' steady state to zero here.

[13] We use data on output, capital and labor input from Fernald (2014). Each series is adjusted for growth in working age population and then detrended using a log-linear trend estimated using data from 1950-2007.

and labor, however, show some divergence with the data, though they are both in the right ballpark. The former exhibits a sharper drop than the observed drop in capital input - which is not particularly surprising, given that our model abstracted from adjustment costs and other frictions that could induce a more sluggish response of capital. Similarly, the model's predictions for labor underpredict the actual change in employment. In the data, employment dropped sharply in 2008-'09, almost contemporaneously with the negative shocks and then recovered slowly. In the model, however, the drop occurs later, but that is largely due to the assumption that labor is chosen in advance. Bringing the model closer to the data along these dimensions is no doubt important and will require a richer model with additional features and frictions, but the considerable improvement in performance that learning brings in a very standard business cycle setting is apparent from the figure.

Isolating the Role of Debt: A No-Debt Benchmark Next, we focus on the role of debt by comparing our results to an identical economy where all investment is financed through equity but beliefs are updated over time. Formally, we set the tax advantage parameter χ to 1 and the leverage target to 0. This implies that $J^k(\underline{v}) = J^l(\underline{v}) = 1$, i.e. the debt-related distortions in capital and labor choice disappear.

Figure 4 plots the time path for aggregate variables for this variant of our model, along with our baseline version from Figure 3. The graph shows that the effects of belief revisions are similar in both cases - the economy converges to a new steady state with lower output, capital and employment. The drop in aggregates is smaller in the model without debt - for example, a 10% reduction in output (compared to 13% in the baseline version with debt). In this sense, the presence of debt amplifies the effects of changes in tail risk.

Thus, in the absence of debt, the effect of the great recession shock cannot be more persistent than the shock itself. In Figure 4, the time series of output and capital look just like the capital quality shock itself. The reason for this lack of additional persistence is that most of the revisions in beliefs occur in the tails. Without debt, real economic variables are not very sensitive to tail probabilities. So, even though there are permanent changes in tail probabilities, they have small effects on the real economy, which are swamped by the transitory direct effect of the capital quality shock. This exercise reveals why increases in perceived tail risk alone cannot generate large, persistent contractions of the sort that we saw in the US economy over the last 6 years- for that, we need both belief revisions and debt financing.

5.2 Results: Quasi-linear utility, endogenous leverage

Next, we repeat the experiment from the previous section in the second version of the model, with quasi-linear utility and endogenous leverage. This eliminates the desire for consumption

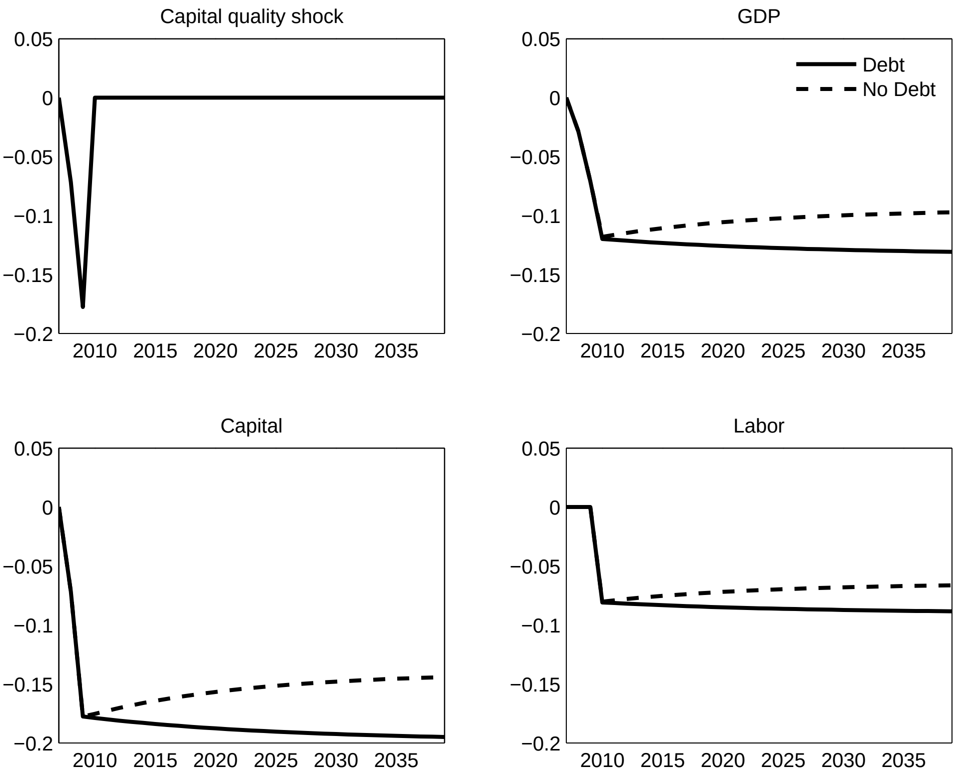

Figure 4: **Debt amplifies the effects of belief revisions on output, investment and labor.** Impulse responses with Epstein-Zin preferences and fixed leverage. Solid line (debt) is the model in section 2. Dashed line (no debt) is an identical model with $\chi = 0$, where firms choose zero debt. Zero is the steady state level in each economy.

smoothing, which was responsible for the relatively smooth and slow transitions in the previous version. While such smoothing effects may well be an important part of reality, suppressing it allows us to gain a clearer understanding of how much persistence belief updating, by itself, can produce. This is a simpler framework, computationally, which allows leverage to be chosen endogenously.

The impulse-response exercise is the same as before, starting the economy in its steady state, shocking it with two large, negative capital quality shocks and then returning the shock process to its average level. Figure 5 presents the time path for aggregate variables in this version, both with and without learning. As we would expect, the absence of curvature in consumption means that the economy transitions immediately to the new steady state. However, belief revisions still have substantial, permanent effects on the level of economic activity. For example, they lead to a drop in steady state output of about 7%. Recall that, under risk aversion, learning led to a bigger drop in steady state output, about 13%. The difference stems from the fact, with risk aversion, the return to capital (and labor) includes a risk premium. Changes in beliefs, and in particular tail risk, lead to an increase in this risk premium, further dampening firms' incentives to invest (and hire). These results indicate that this effect is quite strong and accounts for a significant portion of the long-run drop in output.

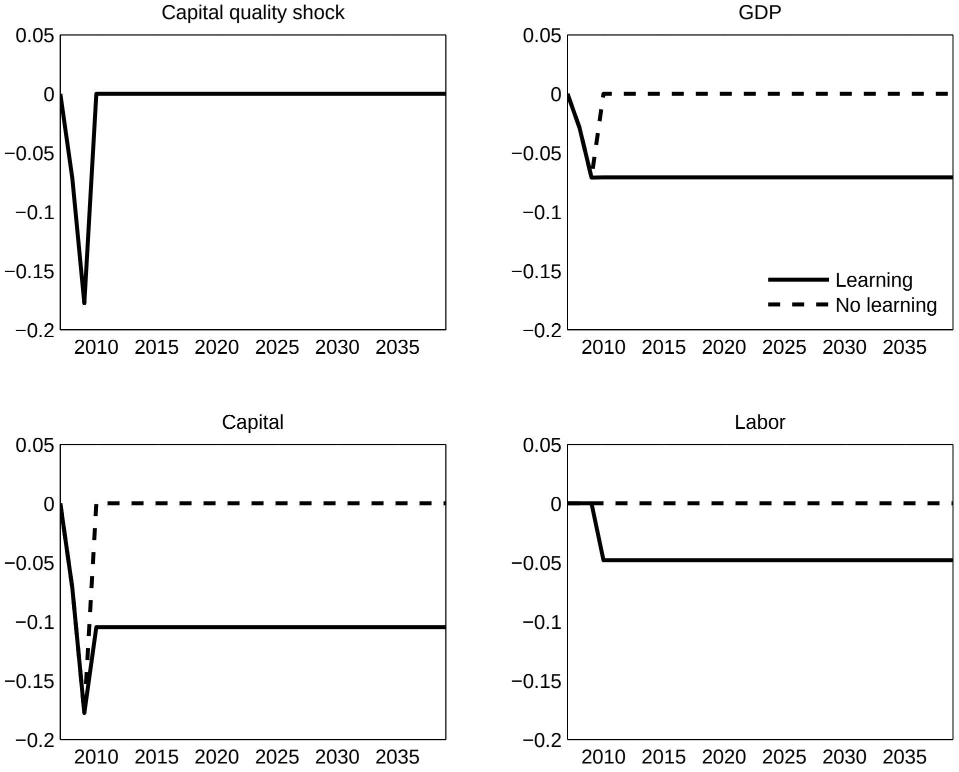

Figure 5: **Impulse responses with quasi-linear preferences, endogenous leverage: Effect of learning.** Without curvature, transitions are starker, but the effect of learning is still substantial.

6 Conclusion

No one knows the true distribution of shocks to the economy. Economists typically assume that agents in their models do know this distribution as a way to discipline beliefs. But assuming that agents do the same kind of real-time estimation that an econometrician would do is equally disciplined and more plausible. For many applications, assuming full knowledge has little effect on outcomes and offers tractability. But for outcomes that are sensitive to tail probabilities, the difference between knowing these probabilities and estimating them with real-time data can be large. The estimation error can be volatile and can introduce new, persistent dynamics into a model with otherwise transitory shocks. The essence of the persistence mechanism is this: Once observed, a shock (a piece of data) stays in one's data set forever and therefore permanently affects belief formation.

When firms finance investments with debt, they make investment and output sensitive to tail risk. Debt is an asset whose payoffs are flat throughout most of the state space, but very sensitive to the state for left-tail, default events. Therefore, the cost of debt depends precisely on the probabilities of a tail event, which are hardest to estimate and whose estimates fluctuate greatly. When debt (leverage) is low, the economy is not very sensitive to tail risk, and economic shocks will be more transitory. The combination of high debt levels and a shock that is a negative outlier makes tail risk surge, investment fall and depresses output in a persistent way.

When we quantify this mechanism and use capital price and quantity data to directly estimate beliefs, our model's predictions resemble observed macro outcomes in the wake of the great recession. These results suggests that perhaps persistent stagnation arose because, after seeing how fragile our financial sector is, market participants will never think about tail risk in the same way again.

References

Angeletos, G.-M. and J. La'O (2013). Sentiments. *Econometrica 81*(2), 739–779.

Backus, D., A. Ferriere, and S. Zin (2015). Risk and ambiguity in models of the business cycle. *Journal of Monetary Economics forthcoming.*

Beaudry, P. and F. Portier (2004). An exploration into pigou's theory of cycles. *Journal of Monetary Economics 51*, 1183–1216.

Bloom, N., M. Floetotto, N. Jaimovich, I. Sapora-Eksten, and S. Terry (2014). Really uncertain business cycles. NBER working paper 13385.

Brunnermeier, M. and Y. Sannikov (2014). A macroeconomic model with a financial sector. *American Economic Review 104*(2).

Cogley, T. and T. Sargent (2005). The conquest of us inflation: Learning and robustness to model uncertainty. *Review of Economic Dynamics 8*, 528–563.

Fajgelbaum, P., E. Schaal, and M. Taschereau-Dumouchel (2014). Uncertainty traps. NBER Working Paper No. 19973.

Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. *Federal reserve bank of San Francisco working paper 19.*

Gertler, M. and P. Karadi (2011). A model of unconventional montary policy. *Journal of Monetary Economics 58*, 17–34.

Gertler, M., N. Kiyotaki, et al. (2010). Financial intermediation and credit policy in business cycle analysis. *Handbook of monetary economics 3*(3), 547–599.

Gourio, F. (2012). Disaster risk and business cycles. *American Economic Review 102(6)*, 2734–66.

Gourio, F. (2013). Credit risk and disaster risk. *American Economic Journal: Macroeconomics 5(3)*, 1–34.

Hall, R. (2014). Quantifying the lasting harm to the u.s. economy from the financial crisis. *NBER Macroeconomics Annual 29*, forthcoming.

Huo, Z. and N. Takayama (2015). Higher order beliefs, confidence, and business cycles. University of Minnesota working paper.

Jaimovich, N. and S. Rebelo (2006). Can news about the future drive the business cycle? *American Economic Review 99(4)*, 1097–1118.

Johannes, M., L. Lochstoer, and Y. Mou (2015). Learning about consumption dynamics. *Journal of Finance forthcoming.*

Khan, A., T. Senga, and J. K. Thomas (2014). Default risk and aggregate fluctuations in an economy with production heterogeneity. Technical report, Ohio State University working paper.

Kreps, D. (1998). *Anticipated Utility and Dynamic Choice*. Cambridge University Press.

Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review 99*(5), 2050–84.

Midrigan, V. and T. Philippon (2011). Household leverage and the recession.

Moriera, A. and A. Savov (2015). The macroeconomics of shadow banking. New York University working paper.

Navarro, G. (2014). Financial crises and endogenous volatility. *manuscript, New York University*.

Nimark, K. (2014). Man-bites-dog business cycles. *American Economic Review 104(8)*, 2320–67.

Orlik, A. and L. Veldkamp (2014). Understanding uncertainty shocks and the role of the black swan. Working paper.

Piazzesi, M., J. Salomao, and M. Schneider (2015). Trend and cycle in bond risk premia. Stanford University working paper.

Straub, L. and R. Ulbricht (2013). Credit crunches, information failures, and the persistence of pessimism. Toulouse School of Economics working paper.

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

Veldkamp, L. and J. Wolfers (2007). Aggregate shocks or aggregate information? costly information and business cycle comovement. *Journal of Monetary Economics 54*, 37–55.

Appendix

A Optimality conditions from firm's problem

The firm's optimization problem is

$$\begin{aligned}
V\left(k_{it}, l_{it}, w_{it}, b_{it}, S_t\right) &= \max\left[0,\ \ \Pi_{it} - B_{it} + \Gamma_{it}\right] \\
\Gamma_{it} &= \max_{\hat{k}_{it+1},\ b_{it+1}, w_{it+1}, l_{it+1}} -\hat{k}_{it+1} + \chi q b_{it+1} + \mathbb{E} M_{t+1} V\left(k_{it+1}, l_{it+1}, w_{it+1}, b_{it+1}, S_{t+1}\right) \\
\Pi_{it} &= v_{it}\left(A(\phi_t k_{it})^{\alpha} l_{it}^{1-\alpha} + (1-\delta)\phi_t k_{it}\right) \\
B_{it+1} &= b_{it+1}\ + w_{it+1} l_{it+1} \\
q\left(\hat{k}_{it+1},\ l_{it+1}, w_{it+1},\ b_{it+1}, S_t\right) &= \mathbb{E} M_{t+1}\left[r\left(\Pi_{it+1}, B_{it+1}\right) + \left(1 - r\left(\Pi_{it+1}, B_{it+1}\right)\right)\frac{\theta V\left(k_{it}, l_{it}, 0, 0, S_t\right)}{B_{it+1}}\right] \\
w_{it+1} q &= \mathcal{W}_t \\
r\left(\Pi_{it}, B_{it}, S_t\right) &= \begin{cases} 0 & \text{if } \ v_{it} < \underline{v}\left(S_t\right) \\ 1 & \text{if } \ v_{it} \geq \underline{v}\left(S_t\right) \end{cases}.
\end{aligned}$$

First, note that we can write the firm's problem in term of leverage and labor capital ratio, defined as $lev_{it+1} \equiv \frac{B_{it+1}}{\hat{k}_{it+1}}$ and $\frac{l_{it+1}}{\hat{k}_{it+1}}$. Then,

$$R^k\left(\frac{l_{it+1}}{\hat{k}_{it+1}}, \phi_{t+1}\right) \equiv \frac{\Pi_{it+1}}{\hat{k}_{it+1}} = v_{it}\left(A(\phi_{t+1}{}^{\alpha}\left(\frac{l_{it+1}}{\hat{k}_{it+1}}\right)^{1-\alpha} + (1-\delta)\phi_{t+1}\right).$$

This implies that

$$\Gamma_{it} = \max_{\hat{k}_{it+1},\ lev_{it+1}, \frac{l_{it+1}}{\hat{k}_{it+1}}} \hat{k}_{it+1}\left(-1 - \chi \mathcal{W}_t \frac{l_{it+1}}{\hat{k}_{it+1}} + \chi q lev_{it+1}\ + \mathbb{E} M_{t+1} r_{t+1}\left(v_{it} R^k_{t+1} - lev_{it+1} + \frac{\Gamma_{it+1}}{\hat{k}_{it+1}}\right)\right)$$

$$q\left(\frac{l_{it+1}}{\hat{k}_{it+1}}, lev_{it+1}, S_t\right) = \mathbb{E} M_{t+1}\left[r_{t+1} + \left(1 - r_{t+1}\right)\theta \frac{v_{it} R^k_{t+1} + \frac{\Gamma_{it+1}}{\hat{k}_{it+1}}}{lev_{it+1}}\right].$$

We guess and then verify that $\Gamma_{it+1} = 0$.[14] Replacing the debt price schedule and rearranging terms yields

$$\begin{aligned}
\Gamma_{it} &= \max_{\hat{k}_{it+1},\ lev_{it+1}, \frac{l_{it+1}}{\hat{k}_{it+1}}} \hat{k}_{it+1}\left(-1 - \chi \mathcal{W}_t \frac{l_{it+1}}{\hat{k}_{it+1}} + \mathbb{E} M_{t+1} \tilde{J}^k_{t+1}\right) \\
\tilde{J}^k_{t+1} &= R^k_{t+1} + lev_{it+1}\left(\chi - 1\right) r_{t+1} + \left(\chi\theta - 1\right)\left(1 - r_{t+1}\right) v_{it} R^k_{t+1}.
\end{aligned}$$

The expectation with respect to the idiosyncratic shock implies $\mathbb{E} r_{t+1} = \left(1 - F\left(\underline{v}\right)\right)$. Also, note that the default threshold becomes $\underline{v} = \frac{lev_{it+1}}{R^k_{t+1}}$. Hence

$$\tilde{J}^k_{t+1} = R^k_{t+1}\left(1 + \underline{v}\left(\chi - 1\right)\left(1 - F\left(\underline{v}\right)\right) + \left(\chi\theta - 1\right) h\left(\underline{v}\right)\right)$$

---

[14] As the firm has constant returns to scale the problem will be linear in capital and in equilibrium $\Gamma_{it} = 0$. See Navarro (2014).

where $h(v) = \int_{-\infty}^{v} v dF(v)$. Finally, the problem is

$$\begin{aligned}
\Gamma_{it} &= \max_{\hat{k}_{it+1},\ lev_{it+1}, \frac{l_{it+1}}{\hat{k}_{it+1}}} \hat{k}_{it+1}\left(-1-\chi\mathcal{W}_t\frac{l_{it+1}}{\hat{k}_{it+1}} + \mathbb{E}M_{t+1}R^k_{t+1}J^k(\underline{v})\right) \\
J^k(\underline{v}) &= 1 + (\chi-1)\,\underline{v}\,(1-F(\underline{v})) + (\chi\theta-1)\,h(\underline{v}) \\
\underline{v} &= \frac{lev_{it+1}}{R^k_{t+1}}
\end{aligned}$$

First, note that the problem is linear in $\hat{k}_{it+1}$ therefore in equilibrium we must have that

$$1+\chi\mathcal{W}_t\frac{l_{it+1}}{\hat{k}_{it+1}} = \mathbb{E}M_{t+1}R^k_{t+1}J^k(\underline{v}),$$

which implies equation 8 in the main text and in turn it verifies the guess, $\Gamma_{it}=0$.

Next, the first order condition with respect to $\frac{l_{t+1}}{\hat{k}_{it+1}}$ is

$$\chi\mathcal{W}_t = \mathbb{E}M_{t+1}R^k\frac{\partial J^k(\underline{v})}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}} + \mathbb{E}M_{t+1}\frac{\partial R^k}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}}J^k(\underline{v}),$$

where

$$\begin{aligned}
R^k_{t+1}\frac{\partial J^k(\underline{v})}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}} &= R^k_{t+1}\frac{\partial\underline{v}}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}}\left((\chi-1)(1-F(\underline{v})) - \underline{v}\,(\chi-1)\,f(\underline{v}) + (\chi\theta-1)\frac{\partial h(\underline{v})}{\partial\underline{v}}\right) \\
\frac{\partial\underline{v}}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}} &= -\frac{lev_{it+1}}{\left(R^k\right)^2}\frac{\partial R^k}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}} = -\frac{\underline{v}^2}{lev_{it+1}}\frac{\partial R^k}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}} \\
\frac{dh(\underline{v})}{d\underline{v}} &= \underline{v}f(\underline{v}) \\
\frac{\partial R^k_{t+1}}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}} &= v_{it}A(1-\alpha)\,{\phi_{t+1}}^{\alpha}\left(\frac{l_{it+1}}{\hat{k}_{it+1}}\right)^{-\alpha}.
\end{aligned}$$

Rearranging terms yields

$$\begin{aligned}
\chi\mathcal{W}_t &= \mathbb{E}M_{t+1}\frac{\partial R^k}{\partial\frac{l_{t+1}}{\hat{k}_{it+1}}}J^l(\underline{v}) \\
J^l(\underline{v}) &= 1 + \underline{v}^2 f(\underline{v})\,\chi(1-\theta) - (1-\chi\theta)\,h(\underline{v}),
\end{aligned}$$

which is (9) in the main text.

Finally, the first order condition with respect to leverage is

$$\mathbb{E}M_{t+1}R^k_{t+1}\frac{\partial J^k_{t+1}}{\partial lev_{it+1}} = 0,$$

where

$$\begin{aligned}\frac{\partial J_{t+1}^k}{\partial lev_{it+1}} &= \frac{\partial \underline{v}}{\partial lev_{it+1}}\left(\left(\chi-1\right)\left(1-F\left(\underline{v}\right)\right)-\left(\chi-1\right)\underline{v}f\left(\underline{v}\right)+\left(\chi\theta-1\right)\underline{v}f\left(\underline{v}\right)\right)\\ &= \frac{1}{R_{t+1}^k}\left(\left(\chi-1\right)\left(1-F\left(\underline{v}\right)\right)-\chi\left(1-\theta\right)\underline{v}f\left(\underline{v}\right)\right)\end{aligned}$$

hence

$$\left(1-\theta\right)\mathbb{E}_t\left[M_{t+1}\underline{v}f\left(\underline{v}\right)\right]=\left(\frac{\chi-1}{\chi}\right)\mathbb{E}_t\left[M_{t+1}\left(1-F\left(\underline{v}\right)\right)\right],$$

which is (10) in the main text.