4

The Tail That Keeps the Riskless Rate Low

Julian Kozlowski, *Federal Reserve Board of St. Louis*
Laura Veldkamp, *Columbia University and NBER*
Venky Venkateswaran, *New York University and NBER*

I. Introduction

Interest rates on safe assets fell sharply during the 2008 financial crisis. This is not particularly surprising; there are many reasons, from an increased demand for safe assets to monetary policy responses, why riskless rates fall during periods of financial turmoil. However, even after the financial markets calmed down, this state of affairs persisted. In fact, by 2017, several years after the crisis, government bond yields still showed no sign of rebounding. In figure 1, we show the change in longer-term government yields in a number of countries since the financial crisis. Looking at longer-term rates allows us to abstract from transitory monetary policy and interpret the graph as evidence of a persistent decline in the level of riskless interest rates.

The decline in interest rates following the financial crisis took place in the context of a general downward trend in real rates since the early 1980s. Obviously, this longer-run trend cannot be attributed to the financial crisis. Instead, it may have come from a gradual change in expectations following the high inflation in the 1970s or a surge in savings from emerging markets seeking safe assets to stabilize their exchange rates. This longer-run trend taking place in the background is hugely important but distinct from our question. We seek to explain the fact that interest rates fell (relative to this long-run trend) during the financial crisis and failed to rebound.

We explore a simple explanation for this phenomenon: before 2008, no one believed that a major recession sparked by a financial crisis with market freezes, failure of major banks, and so forth could happen in the United States. The events in 2008 and 2009 taught us that this is more likely than we thought. Today, the question of whether the financial crisis might repeat itself arises frequently. Although we are no longer on the precipice, the knowledge we gained from observing 2008–9 stays with us and re-

978-0-226-64572-8/2018/2018-0401$10.00

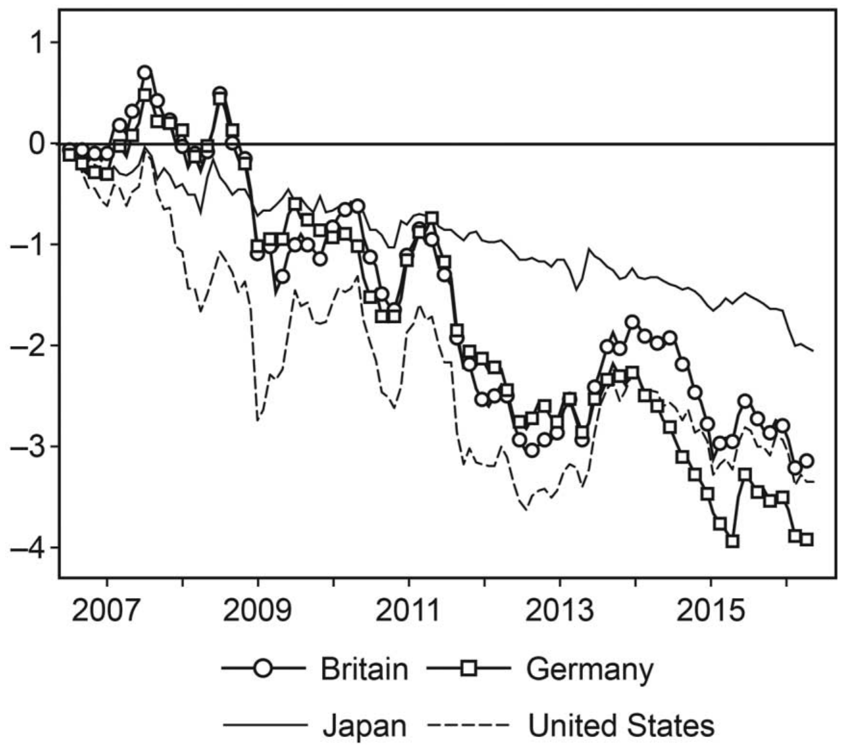

Fig. 1. Low interest rates are persistent. Change in percentage points of 10-year government bond yield since July 3, 2006 (Irwin 2016). A similar pattern emerges, even if we control for inflation.

shapes our beliefs about the probability of large adverse shocks. This persistent increase in perceived "tail risk" makes safe liquid assets more valuable, keeping their rates of return depressed for many years. The contribution of this paper is to measure how much tail risk rose, explain why it remains elevated, and quantitatively explore its consequences for riskless interest rates.

At its core is a simple theory about how agents form beliefs about the probability of rare tail events. Our agents do not know the distribution of shocks hitting the economy and use macro data and standard econometric tools to estimate the distribution in a flexible, nonparametric way. Transitory shocks have persistent effects on beliefs, because once observed, the shocks remain forever in the agents' data set. Then, we embed our belief formation mechanism into a standard production economy with liquidity constraints. When we feed a historical time series of macro data for the postwar United States into our model and let our agents reestimate the distribution from which the data are drawn each period, our belief revision mechanism goes a long way in explaining the persistent postcrisis decline in government bond yields since 2008–9.

The link between heightened tail risk and rates of return in the model comes from two intuitive mechanisms. First, the increase in consumption risk makes safe assets more valuable, lowering the required return on riskless government bonds. The second stems from the fact that government bonds also provide liquidity services that are particularly valuable in very bad states. Intuitively, in those states, the liquidity available from other sources falls. The main contribution of this paper is to combine these standard forces with the aforementioned theory of beliefs in a simple, tractable, and empirically disciplined framework and show that rare events like the 2008–9 recession generate large, persistent drops in riskless interest rates.

Apart from being quantitatively successful, our explanation is also consistent with other evidence of heightened tail risk. For example, in their value at risk (VAR) analysis, Del Negro et al. (2017) find that most of the decline in riskless rates is attributable to changes in the value of safety and liquidity. From 2007 to 2017, they estimate a 52-basis point change in the convenience yield of US Treasury securities (which is about 80% of the estimated drop in the natural riskless real rate over the same time period). A second piece of evidence comes from the SKEW index, an option-implied measure of tail risk in equity markets. Figure 2 shows a clear rise since the financial crisis, with no subsequent decline.[1] In our quantitative analysis, we show that the implied changes in tail probabilities are roughly in line with the predictions of our calibrated model. Finally, popular narratives about stagnation emphasize a change in "attitudes" or "confidence," which we capture with belief changes, and the reductions in debt financing that result: "Years after US investment bank Lehman Brothers collapsed, triggering a global financial crisis and shattering confidence worldwide . . . 'The attitude toward risk is permanently reset.' A flight to safety on such a global scale is unprecedented since the end of World War II" (Condon 2013).

In many macro models, including belief-driven ones, deviations of aggregate variables from trends inherit the exogenously specified persistence of the driving shocks (see, e.g., Maćkowiak and Wiederholt 2010; Angeletos and La'O 2013).[2] Therefore, these theories cannot explain why interest rates remain persistently low. In our setting, when agents repeatedly reestimate the distribution of shocks, persistence is endogenous and state dependent. Extreme events like the recent crisis are rare and thus lead

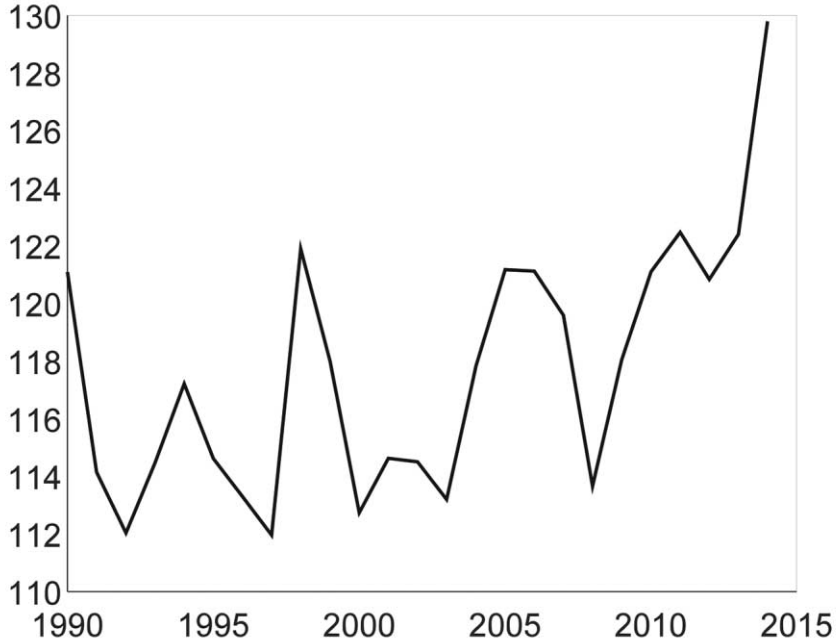

Fig. 2. The SKEW index. A measure of the market price of tail risk on the S&P 500, constructed using option prices (CBOE 2019).

to significant belief changes (and, through them, aggregate variables such as riskless rates) that outlast the events themselves. More "normal" events (e.g., milder downturns), in contrast, show up relatively more frequently in the agents' data set, and therefore additional realizations have relatively small effects on beliefs. In other words, although all changes to beliefs are, in a sense, long-lived, rarer events induce larger, more persistent belief changes and interest rate responses. Rare event beliefs are more persistent because rare event data are scarce. It takes many observations without a rare event to convince observers that the event is much more rare than they thought.

This mechanism for generating persistent responses to transitory shocks is simple to execute and can be easily combined with a variety of sophisticated, quantitative macro models to introduce persistent effects of rarely observed shocks. Although our focus in this paper is on interest rates, it could be applied to other phenomena, including labor force participation rates, corporate debt issuance and cash hoarding, house prices, export decisions, and trade credit. The crucial ingredients of the model are some nonlinearity (typically, a constraint that binds in bad states), some actions that compromise efficiency in the current state but that hedge the risk of this binding state, and then a large, negative shock. If those ingredients are present, then adding agents who learn like econometricians is likely to induce sizeable, persistent responses.

Because the novel part of the paper is using this belief formation mechanism to explore interest rates, Section II starts by examining the belief formation mechanism in a simple context. We construct a time series of "quality" shocks to US nonresidential capital and use it to show how our nonparametric estimation works. Agents estimate the underlying distribution of the shocks by fitting a kernel density function to the data in their information set. When they see extreme negative realizations from the financial crisis, it raises their estimation of large negative outcomes. More important, this effect is persistent. The theoretical underpinning of the persistence is the martingale property of beliefs. Intuitively, once observed, an event stays in the agents' data set and informs their probability assessment, even after the event itself has passed. Decades later, the probability distribution still reflects a level of tail risk that is higher than it was precrisis. Knowing that a crisis is possible influences risk assessment for many years to come.

We embed this mechanism into a standard production economy with a liquidity friction. Every period, in addition to their usual production, firms have access to an additional investment opportunity. However, to exploit this opportunity, they need liquidity in the form of pledgeable collateral. Both capital and government bonds act as collateral, but only a fraction of the former's value can be pledged. An adverse shock lowers the value of pledgeable capital and, therefore, liquidity.

Section V presents our quantitative results. We perform two sets of exercises. The first involves long-run predictions under the assumption that crises continue to occur. Specifically, we simulate long-run outcomes (i.e., stochastic steady states) drawing shocks from the updated beliefs. Our calibrated model predicts that the increase in tail risk is associated with a 1.45% drop in interest rates on government bonds in the long run. Most of this drop can be attributed to the liquidity mechanism. The modest degree of risk aversion in our calibration implies that the increase in consumption risk by itself induces only a very small change in interest rates. We also show that the implications of the model for changes in equity market variables (e.g., equity premium, tail risk implied by options) line up reasonably well with the data.

Next, we generate time paths for the economy under the assumption that the financial crisis we saw in 2008–9 was a one-off event and will never recur. Then, the economy eventually returns to its precrisis stochastic steady state, but we show that this occurs at a very slow rate. Even after several years, interest rates on safe assets remain depressed. Intuitively, learning about rare events is, in a sense, "local": probabilities in the tail respond sharply to extreme realizations but slowly to realizations from elsewhere in the support. As a result, it takes a very long period without extreme events to convince agents that such events can be safely ignored. Finally, to demonstrate that belief revisions are key to the model's ability to generate sustained drops in interest rates, we also generate counterfactual time paths with the belief mechanism turned off. In other words, we endow agents with knowledge of the true distribution from the very beginning. We find that the initial impact of the shock on interest rates is similar, but they start to rebound almost immediately and return to precrisis levels at a much faster rate. In other words, without changes to beliefs, the financial crisis would induce a fairly transitory fall in interest rates.

Comparison to the Literature

Our paper speaks to a large body of work that focuses on the macroeconomic consequences of beliefs.[3] Most of these papers focus on uncertainty (or second-moment changes) and, perhaps more important, rely on exogenous assumptions about the persistence of shocks for propagation. Essentially, beliefs about time-varying states are only persistent to the extent that the underlying states are assumed to be persistent.[4] Our mechanism, on the other hand, generates persistence endogenously and helps explain why some recessions have long-lasting effects while others do not. A second advantage of our contribution is that by tying beliefs to observable data, we are able to impose considerable empirical discipline on the role of belief revisions, a key challenge for this whole literature.

The nonparametric belief formation process specified in this paper is similar to other adaptive learning approaches. Kozlowski, Veldkamp, and Venkateswaran (2017) use a similar belief formation mechanism to explain persistence in real output fluctuations. That paper, however, abstracts from liquidity, an important amplification mechanism, and therefore cannot match the large decline in the riskless rate. In constant gain learning (Sargent, 1999), agents combine last period's forecast with a constant times the contemporaneous forecast error. Such a process gives recent observations more weight, similar to the behavior of agents in Malmendier and Nagel (2011) following the Great Depression. The reason why we use a nonparametric belief formation process is that we want to model time-varying changes in perceived tail risk, which requires a richer specification of the distribution of state variables.

Our belief formation process also has similarities to the parameter learning models by Johannes, Lochstoer, and Mou (2016) and Orlik and Veldkamp (2015) and is advocated by Hansen (2007). Similarly, in least-squares learning (Marcet and Sargent 1989), agents have bounded rationality and use past data to estimate the parameters of the law of motion for the state variables. However, these papers do not have meaningful changes in tail risk and do not analyze the potential for persistent effects on interest rates. Pintus and Suda (2015) embed parameter learning into a production economy but feed in persistent leverage shocks and explore the potential for amplification when agents hold erroneous initial beliefs about persistence. Sundaresan (2018) generates persistence by deterring information acquisition. Weitzman (2007) shows that the parameter uncertainty about the variance of a thin-tailed distribution can help resolve many of the asset pricing puzzles confronted by the rational expectations paradigm.

Finally, our paper contributes to a growing literature on low interest rates. Recent contributions include Bernanke et al. (2011), Barro et al. (2014), Bigio (2015), Caballero, Farhi, and Gourinchas (2016), Carvalho, Ferrero, and Nechio (2016), Del Negro et al. (2017), and Hall (2017). To this body of work, we add a novel mechanism, one that predicts persistent drops in riskless interest rates in response to rare transitory shocks, and demonstrate its quantitative and empirical relevance.

II. How Belief Updating Creates Persistence

The main contribution of this paper is to explain why tail risk fluctuates and to show how an extreme event like the Great Recession can induce a persistent drop in riskless rates. Before laying out the whole model, we begin by explaining the novel part of the paper: how agents form beliefs and the effect of tail events on beliefs. This will highlight the broader insight that unusual events induce larger and more persistent belief changes. Later, we layer the economic model on top to show how this mechanism affects interest rates.

The story that this model formalizes is that before the financial crisis hit, most people in the United States thought that such crises only happened elsewhere (e.g., in emerging markets) and that bank runs were a topic for historians. Observing the events of 2007–9 changed those views. Many journalists, academics, and policy makers now routinely ask whether the financial architecture is stable. But formalizing this story requires a departure from the standard rational expectations paradigm, where the distributions of all random events are assumed to be known. Then, observing an unusual event should not change one's probability assessment of that event in the future. Instead, we need a machinery that allows agents to not know the true distribution, so that upon seeing something they thought should not happen, they can revise their beliefs. There are many ways to depart from full knowledge of distributions. One that is realistic, quantifiable, and tractable is treating agents like classical econometricians. The agents in our model have a finite data set—the history of all realized shocks—and they estimate the distribution from which those shocks are drawn, using tools from a first-year econometrics class.

Learning models are not new to the macro literature. A common approach is to assume a normal distribution and to estimate its mean and variance. However, the normal distribution has thin tails, making it less useful to think about changes in the risk of extreme negative realizations. We could choose an alternative distribution with more flexibility in higher moments. However, this will raise obvious concerns about the sensitivity of results to the specific functional form used. To minimize such concerns, we take a nonparametric approach and let the data inform the shape of the distribution.

Instead, our agents take all the data they have observed and use a kernel density procedure to estimate the probability distribution from which these data were drawn. One of the most common approaches in nonparametric estimation, a kernel density essentially takes a histogram of all observed data and draws a smooth line over that histogram. There are a variety of ways to smooth the line. The most common is called the "normal kernel." It does not result in normal (Gaussian) distributions. We also studied a handful of other kernels and (sufficiently flexible) parametric specifications, which yielded similar results.[5] The kernel density approach allows for flexibility in the shape of the distribution while strictly tying the learning process to data that we, as economists, can observe. We do not need to guess or calibrate the precision of some signal. Instead, we take a macro data series, apply this econometric procedure to it, and read off the agents' beliefs.

Next, we describe the Gaussian kernel. Consider the shock ϕ_t whose true density g is unknown to agents in the economy. The agents do know that the shock ϕ_t is independent and identically distributed (i.i.d.). The information set at time t, denoted $\mathcal{I}_t$, includes the history of all shocks ϕ_t observed up to and including t. They use this available data to construct an estimate $\hat{g}_t$ of the true density g. Formally, at every date t, agents construct the following normal kernel density estimator of the probability density function g:

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

where $\Omega(\cdot)$ is the standard normal density function, κ_t is the smoothing or bandwidth parameter, and n_t is the number of available observations at date t. As new data arrive, agents add the new observations to their data set and update their estimates, generating a sequence of beliefs $\{\hat{g}_t\}$.

Finally, back to the main point: Why does this estimated distribution change in such a persistent way in response to a tail event? We will explain this graphically and then mathematically. Figure 3 shows three panels. Figure 3*a* is the histogram of a data series. In this case, the data series happens to have some measures of capital quality, which we will describe in detail later. For right now, this is an arbitrary sequence of data generated from an unknown distribution. The smooth line over the histogram is the estimated normal kernel. Figure 3*b* shows what happens when two data points that are negative outliers are observed. The locations of the two new observations are highlighted (black rectangles) in the histogram. Notice that the new kernel estimator, and thus agents' beliefs, now places greater probability weight on the possibility of future negative outcomes. If in the next period the state returns to normal, those two rectangular data points are still in the histogram and still create the bump on the left. Although the tail event has passed, tail risk remains elevated. Figure 3*c* adds 30 years of additional observations, drawn to look just like the preceding years, except without any crisis events. The kernel on the right still shows a left bump. Smaller than it was before, but still present, elevated tail risk still persists 30 years after the tail event was observed.

The persistence in figure 3 has its origins in the so-called martingale property of beliefs—that is, conditional on time t information ($\mathcal{I}_t$), the estimated distribution is a martingale. Thus, on average, the agent expects her future beliefs to be the same as her current beliefs. This property holds exactly if the bandwidth parameter κ_t is set to zero.[6] In our empirical implementation, in line with the literature on nonparametric assumption, we use the optimal bandwidth (see Hansen 2015). This leads to smoother density but also means that the martingale property does not hold exactly. Numerically, the deviations are minuscule for our application. In other words, the kernel density estimator is, for all practical purposes, a martingale $\mathbb{E}_t[\hat{g}_{t+j}(\phi)|\mathcal{I}_t] \approx \hat{g}_t(\phi)$.

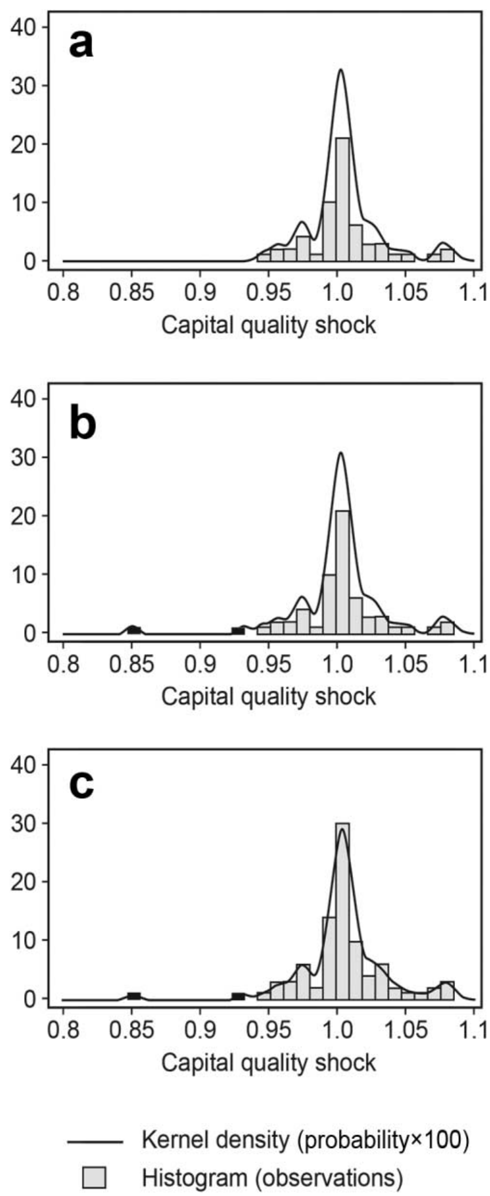

Fig. 3. The persistence of estimated probabilities. (*a*) Period 1950–2007. (*b*) Period 1950–2009. (*c*) Period 1950–2039. Data in the histograms are capital quality shocks, measured as described in Sec. IV. Kernel densities are constructed with the normal kernel in (3) and the optimal bandwidth.

Now, in the simulations underlying figure 3*c*, we drew future shock sequences from the precrisis distribution (i.e., $\hat{g}_{2007}$ instead of the revised belief $\hat{g}_{2009}$). This implies that beliefs will revert, namely, the bump in the left tail will eventually disappear. However, the rate at which this occurs is very slow. This has to do with the fact that under our nonparametric approach, outlier observations play a crucial role in learning about the frequency of tail events. Ordinary events are just not very informative about those tail probabilities. And since data on tail events are scarce, observing one makes the resulting belief revisions large and extremely persistent (even if they are ultimately transitory). It is worth pointing out that this slow convergence need not necessarily obtain with a parametric specification of the learning process. For example, suppose there is uncertainty about the standard deviation of a thin-tailed distribution, as in Weitzman (2007). Because all realizations are informative about standard deviations, the effect of observing a tail event is more muted (i.e., there is a lot more relevant data) and relatively less persistent (convergence to the true distribution occurs at a faster rate).

III. Model

Preferences and Technology

The economy is populated by a representative firm, which produces output with capital and labor, according to a standard Cobb-Douglas production function:

$$Y_t = AK_t^{\alpha}N_t^{1-\alpha}, \tag{1}$$

where A is total factor productivity, which is the same for all firms and constant over time. The firm is subjected to an aggregate shock to capital quality ϕ_t: formally, it enters the period with capital $\hat{K}_t$ and is hit by a shock ϕ_t, leaving it with "effective" capital K_t:

$$K_t = \phi_t\hat{K}_t. \tag{2}$$

These capital quality shocks are i.i.d. over time and are the only aggregate disturbances in our economy. The i.i.d. assumption is made to avoid an additional exogenous source of persistence.[7] They are drawn from a distribution $g(\cdot)$: this is the object agents are learning about.

As we see from equation (2), these shocks scale the effective capital stock up or down. This is not to be interpreted literally—it is hard to visualize shocks that regularly wipe out fractions of the capital or create it out of thin air. Instead, these shocks are a simple, if imperfect, way to model the extreme and unusual effects of the 2008–9 recession on the economic value and returns to nonresidential capital. It allows us to capture the idea that a hotel built in 2007 in Las Vegas may still be standing after the Great Recession but may deliver much less economic value. The use of such shocks in macroeconomics and finance goes back at least to Merton (1973), but they have become more popular recently (precisely to generate large fluctuations in the returns to capital), for example, in Gourio (2012) and in a number of recent papers on financial frictions, crises, and the Great Recession (e.g., Gertler and Kiyotaki 2010; Gertler and Karadi 2011; Brunnermeier and Sannikov 2014).

Finally, the firm is owned by a representative household, the preferences of which over consumption C_t and labor supply N_t are given by a flow utility function $U(C_t, N_t)$, along with a constant discount rate β.

Liquidity

We now introduce liquidity considerations, which will act as an amplification mechanism for tail risk changes. We model them in a stylized but tractable specification in the spirit of Lagos and Wright (2005): firms have access to a productive opportunity but require liquidity in the form of pledgeable collateral in order to exploit it. As in Venkateswaran and Wright (2014), both capital and riskless government bonds can be pledged, albeit to different degrees. Bonds are fully pledgeable, but only a fraction of the effective capital can be used as collateral. An increase in tail risk now has an additional effect—it reduces the liquidity value of capital, increasing the demand for an alternative source of liquidity, namely, riskless government bonds, amplifying the interest rate response.

Formally, at the beginning of each period, firms can invest in a project that costs X_t and yields a payoff $H(X_t)$ (both denominated in the single consumption/investment good). The function H is assumed to be strictly increasing and concave, which implies that the net surplus from the project, namely, $H(X) - X$, has a unique maximum at X^*. In the absence of other constraints, therefore, every firm presented with this opportunity will invest X^*. However, the firm faces a liquidity constraint:

$$X_t \leq B_t + \eta K_t$$

In other words, the investment in the project cannot exceed the sum of pledgeable collateral, which comprises a fraction η of its effective capital K_t and the value of its liquid assets (riskless government bonds) B_t.[8]

Therefore,

$$X_t = \min(X^*, B_t + \eta K_t).$$

After this stage, production takes place according to equation (1).

Timing and Value Functions

The timing of events in each period t is as follows: (i) the firm enters the period with capital stock $\hat{K}_t$ and liquid assets B_t, (ii) the aggregate capital quality shock ϕ_t is realized, (iii) the firm chooses X_t subject to the liquidity constraint, (iv) the firm chooses labor and production takes place, and (v) the firm chooses capital and liquid asset positions for $t + 1$.

Denoting the aggregate state S_t (described in detail later in this section), the economy-wide wage rate W_t, the price of the riskless bond P_t, and the stochastic discount factor M_{t+1}, we can write the problem of the firm in recursive form as follows:

$$\begin{aligned} V(K_t, B_t, S_t) = \max_{X_t, N_t, B_{t+1}, \hat{K}_{t+1}} & H(X_t) - X_t + F(K_t, N_t) - W_t N_t + K_t(1-\delta) + B_t \\ & - P_t B_{t+1} - \hat{K}_{t+1} + \beta \mathrm{E}_t M_{t+1} V(K_{t+1}, B_{t+1}, S_{t+1}) \\ \text{s.t.} \quad X_t & \leq B_t + \eta K_t, \\ K_{t+1} & = \phi_{t+1} \hat{K}_{t+1}. \end{aligned} \tag{3}$$

The stochastic discount factor M_{t+1} and the wage W_t are determined by the marginal utility of the representative household:

$$W_t = -\frac{U_2(C_t, N_t)}{U_1(C_t, N_t)}, \tag{4}$$

$$M_{t+1} = \frac{U_1(C_{t+1}, N_{t+1})}{U_1(C_t, N_t)}. \tag{5}$$

The aggregate state S_t consists of $(\Pi_t, \mathcal{I}_t)$, where $\Pi_t \equiv H(X_t) - X_t + AK_t^{\alpha} L_t^{1-\alpha} + (1-\delta)K_t$ is the aggregate resources available and $\mathcal{I}_t$ is the economy-wide information set. Standard market clearing conditions yield:

$$C_t = \Pi_t - \hat{K}_{t+1}, \tag{6}$$

$$B_t = \bar{B}. \tag{7}$$

where $\bar{B}$ is the exogenous supply of the riskless government bond. The interest expenses on these bonds is financed through lump-sum taxes.

Information, Beliefs, and Equilibrium

The set $\mathcal{I}_t$ includes the history of all shocks ϕ_t observed up to and including time t. For now, we specify a general function, denoted Ψ, which maps $\mathcal{I}_t$ onto an appropriate probability space. The expectation operator $\mathbb{E}_t$ is defined with respect to this space. In the following section, we make this more concrete using the kernel density estimation procedure to map the information set into beliefs.

For a given belief function Ψ, a recursive equilibrium is a set of functions for (i) aggregate consumption and labor supply that maximize household utility subject to a budget constraint; (ii) a bond price that clears the market for bonds; (iii) firm values and policies that solve equation (3), taking as given the stochastic discount factor and wages according to equations (4)–(5) and the bond price; and (iv) aggregate consumption and labor that are consistent with individual choices and thus the bond market clears.

Characterization and Solution

The equilibrium of the economic model is a solution to a set of nonlinear equations, namely, the optimality conditions of the firm and the household, along with resource constraints. The optimality conditions of the firm (3) are:

$$1 = \beta\mathbb{E}_t\{M_{t+1}\phi_{t+1}[F_1(K_{t+1}, N_{t+1}) + 1 - \delta + \eta\mu_{t+1}]\}, \tag{8}$$

$$P_t = \beta\mathbb{E}_t\{M_{t+1}(1 + \mu_{t+1})\}, \tag{9}$$

$$\mu_t = H'(X_t) - 1, \tag{10}$$

$$W_t = F_2(K_t, N_t), \tag{11}$$

where μ_t is the Lagrange multiplier on the liquidity constraint. The first two equations are the Euler equations for capital and liquid assets, respectively. The value of liquidity services is reflected on the right-hand side (in the term involving μ_t). The third equation characterizes μ_t. In states of the world where liquidity is sufficiently abundant, $X_t = X^*$ and $\mu_t = 0$. Otherwise, $\mu_t > 0$. The expectation of μ_{t+1} (weighted by the SDF M_{t+1}) raises the price of the liquid bond P_t, or, equivalently, lowers the risk-free rate. An increase in tail risk—namely, the likelihood of large adverse realizations of ϕ_{t+1}—makes the constraint more likely to bind and thus raises the liquidity premium on the riskless bond.

Belief Formation

Next, we choose a particular estimation procedure for how agents form beliefs. Specifically, we employ the kernel density estimation procedure, which we described in Section II.

Consider the shock ϕ_t for which true density g is unknown to agents in the economy. The agents do know that the shock ϕ_t is i.i.d. The information set at time t, denoted $\mathcal{I}_t$, includes the history of all shocks ϕ_t observed up to and including t. They use these available data to construct an estimate $\hat{g}_t$ of the true density g. Formally, at every date t, agents construct the following normal kernel density estimator of the probability density function g:

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

where $\Omega(\cdot)$ is the standard normal density function, κ_t is the smoothing or bandwidth parameter, and n_t is the number of available observations at date t. As new data arrive, agents add the new observations to their data set and update their estimates, generating a sequence of beliefs $\{\hat{g}_t\}$.

IV. Measurement and Calibration

In this section, we describe how we use macro data to construct a time series for ϕ_t and pin down beliefs. A key strength of our belief-driven theory is that by assuming that agents form beliefs as an econometrician would, we can use observable data to discipline those beliefs. We also parameterize the model to match key features of the US economy and describe key aspects of our computational approach.

Measuring Capital Quality Shocks

To construct a time series of ϕ_t, we follow the approach in Kozlowski et al. (2017). They used data on nonfinancial assets in the US economy, reported in the "Flow of Funds" tables, both at historical cost, which we will denote NFA_t^{HC}, and at market value, NFA_t^{MV}. The latter series corresponds to the nominal value of effective capital, K_t in the model. Letting X_{t-1} denote investment in period $t-1$ and P_t^k denote the nominal price of capital goods in t, the two time series can be mapped onto their model counterparts as follows:

$$\begin{aligned} P_t^k K_t &= \mathrm{NFA}_t^{MV} \\ P_{t-1}^k \hat{K}_t &= (1-\delta)\mathrm{NFA}_{t-1}^{MV} + P_{t-1}^k X_{t-1} \\ &= (1-\delta)\mathrm{NFA}_{t-1}^{MV} + \mathrm{NFA}_t^{HC} - (1-\delta)\mathrm{NFA}_{t-1}^{HC}. \end{aligned}$$

To adjust for changes in nominal prices, we use the price index for nonresidential investment from the National Income and Product Accounts (denoted PINDX_t).[9] This allows us to recover the quality shock ϕ_t:

$$\begin{aligned} \phi_t &= \frac{K_t}{\hat{K}_t} = \left(\frac{P_t^k K_t}{P_{t-1}^k \hat{K}_t}\right)\left(\frac{P_{t-1}^k}{P_t^k}\right) \\ &= \left(\frac{\mathrm{NFA}_t^{MV}}{(1-\delta)\mathrm{NFA}_{t-1}^{MV} + \mathrm{NFA}_t^{HC} - (1-\delta)\mathrm{NFA}_{t-1}^{HC}}\right)\left(\frac{\mathrm{PINDX}_{t-1}^k}{\mathrm{PINDX}_t^k}\right), \end{aligned} \tag{12}$$

where the second line replaces P_{t-1}^k/P_t^k with $\mathrm{PINDX}_{t-1}^k/\mathrm{PINDX}_t^k$.

Using the measurement equation (12) (and a value of $\delta = 0.03$), we construct an annual time series for capital quality shocks for the US economy since 1950, plotted in figure 4*a*. For most of the sample period, the shock realizations were 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. To put these numbers in context, the mean and standard deviation of the series from 1950–2007 were 1 and 0.03, respectively.

We then apply our kernel density estimation procedure to this time series to construct a sequence of beliefs. In other words, for each t, we construct $\{\hat{g}_t\}$ using the available time series until that point. The resulting estimates for two dates, 2007 and 2009, are shown in figure 4*b*. They show that the Great Recession induced a significant increase in the perceived likelihood of extreme negative shocks. The estimated density for 2007 implies almost zero mass below 0.90, whereas the one for 2009 attaches a nontrivial (approximately 2.5%) probability to this region of the state space.

Calibration

We begin by specifying the functional form of preferences and technology. The period utility function of the household is $U(C, N) = \{C - [N^{1+\gamma}/(1+\gamma)]\}/1-\sigma$. The risk aversion parameter σ is set to 0.5. The payoff

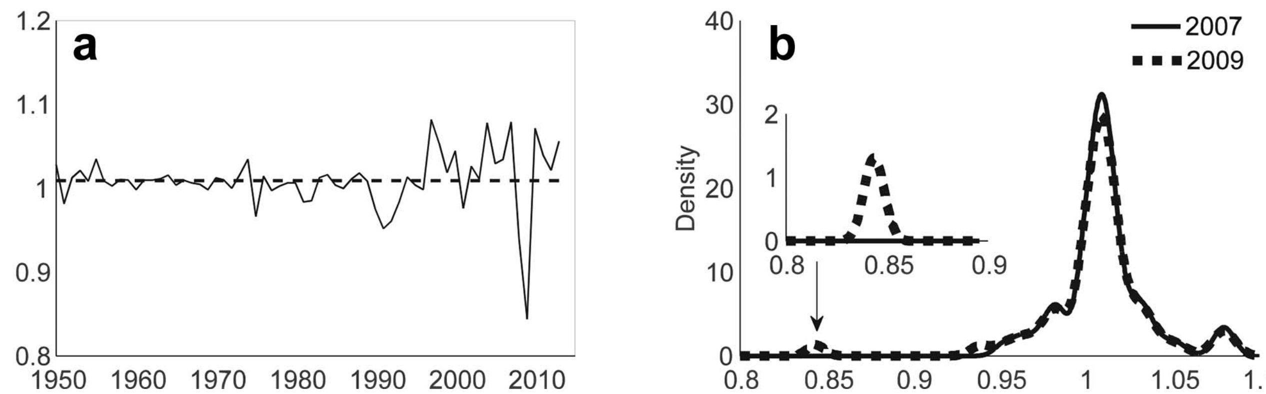

Fig. 4. Capital quality shocks and beliefs. (*a*) Capital quality shocks. (*b*) Beliefs. Panel (*a*) shows the capital quality shocks measured from US data using (12). Panel (*b*) shows the estimated kernel densities in 2007 (solid) and 2009 (dashed). The change in the left tail shows the effect of the Great Recession.

from the project is $H(X) = 2\zeta\sqrt{X} - \xi$. The labor supply parameter γ is set to 0.5, corresponding to a Frisch elasticity of 2 in line with Midrigan and Philippon (2011). The labor disutility parameter π is normalized to 1. The parameter ξ acts like a fixed cost and separates the liquidity premium (which is a function only of $H'[X]$) from the level of the net surplus, a flexibility that proves helpful in the calibration process.

A period is interpreted as 1 year. Accordingly, we choose the discount factor $\beta = 0.95$ and depreciation $\delta = 0.03$. The share of capital in the production is set to 0.40, while the total factor productivity parameter A is normalized to 1.

Next, we turn to the liquidity-related parameters. The parameter governing the pledgeability of capital η is set to match the ratio of short-term obligations of US nonfinancial corporations to the capital stock in the Flow of Funds. Short-term obligations comprise commercial paper (BGFRS 2017, table B.103, row 27), bank loans (row 31), and trade payables (row 34). Capital stock is the market value of nonfinancial assets (row 2). This ratio stood at 0.16 in 2007.[10]

There are three other parameters to be determined: the supply of liquid assets $\bar{B}$ and the technology parameters ζ and ξ. These are chosen to jointly target the following moments: (i) the ratio of liquid asset holdings of US nonfinancial corporations, which stood at 0.082 in 2007;[11] (ii) an interest rate of 2% on government bonds (which corresponds to the precrisis average for real interest rates in the United States); and (iii) a capital-output ratio of 3.5. In the model, the analogous objects are averages in the stochastic steady state under the precrisis belief distribution. Though this calibration is done jointly, a heuristic argument can be made for identification—the first moment is informative about $\bar{B}$, the second about ζ, and the third helps us pin down ξ. Table 1 summarizes the resulting parameter choices.

V. Results

Our main goal in this section is to quantify the size and persistence of the response of risk-free rates to a large but transitory shock ϕ_t in an economy where agents are learning about the distribution. We begin by computing the stochastic steady state associated with $\hat{g}_{2007}$, the distribution estimated using precrisis data.[12] Then, starting from this steady state, we subject the model economy to the two adverse realizations observed in 2008 and 2009, namely, 0.93 and 0.84. As we saw in the previous section, this leads to a revised estimate for the distribution $\hat{g}_{2009}$, which shows an increase in perceived tail risk.

Table 1
Parameters

Parameter	Value	Description
Preferences:		
β	.95	Discount factor
γ	.50	1/Frisch elasticity
π	1	Labor disutility
σ	.5	Risk aversion
Technology:		
α	.40	Capital share
δ	.03	Depreciation rate
Liquidity:		
η	.16	Pledgeability of capital
$\bar{B}$	4.93	Supply of liquid assets
ζ	3.93	Investment technology (affects liquidity)
ξ	9.00	Investment fixed cost

We perform two exercises to demonstrate the quantitative bit of our belief revision mechanism. First, we compare the stochastic steady states implied by the two distributions $\hat{g}_{2007}$ and $\hat{g}_{2009}$ for both aggregate macroeconomic quantities (e.g., output, capital, and labor) and asset prices. This corresponds to the long-term behavior of the US economy under the assumption that crises continue to occur with the same likelihood as the updated beliefs (formally, if future shocks are drawn from the postcrisis distribution $\hat{g}_{2009}$). Second, we simulate time paths for the economy under the assumption that there are no future crises, namely, with future shocks drawn from the precrisis distribution $\hat{g}_{2007}$. In other words, we assume that the 2008–9 recession was a one-off adverse realization. As a result, beliefs will eventually revert to their precrisis levels. However, the effects of the tail events in 2008–9 on beliefs (and, therefore, aggregate outcomes) turn out to be quite persistent and remain significant over a relatively long horizon.

Long-Run Analysis

The results from the first exercise, where we compare long-run averages under $\hat{g}_{2007}$ and $\hat{g}_{2009}$, are reported in table 2. As the table shows, the rise in tail risk causes the economy to invest and produce less, leading to lower output and capital. This occurs because investing now has a lower mean return but is also significantly riskier. The change in beliefs leads to a sharp drop in the risk-free rate—in the new steady state, government bond yields are almost 1.3% lower. Two forces contribute to this

Table 2
Steady State Interest Rates and Macro Aggregates, Pre- and Postcrisis

	$\hat{g}_{2007}$	$\hat{g}_{2009}$	Change
ln $F(K, N)$	2.39	2.36	−.03
ln X	2.68	2.65	−.03
ln K	4.10	4.06	−.04
Riskless rate (R^f)	2.31	.86	−1.45
Return on capital (R^v)	5.30	5.29	−.01
Premium ($R^v - R^f$)	2.99	4.43	1.44

Note: R^f is the interest rate on government bonds, whereas R^v is the average expected returns on unlevered claims to the firm.

drop. First, future consumption is riskier, which has the usual effect of lowering the required return on risk-free claims. Second, the liquidity premium rises. This is partly because there is less liquidity in the economy (due to the lower levels of capital in the new steady state), but also due to the increase in liquidity risk. A tail event also implies states with very low levels of liquidity, which translates to a higher premium on liquid assets.

How do these predictions compare with the post-2008 data? The first row of table 3 compares the drop in interest rates predicted by the model to different measures of changes in risk-free rates since the Great Recession. The second row reports the change from 2007 to 2017 in short-term real rate. This is defined as the difference between 1-year nominal Treasury yield (taken from the H15 release [BGFRS 2019, table H15]) and 1-year expected inflation from the Federal Reserve Bank of Cleveland's inflation forecasting model. The next three rows contain estimates of changes in

Table 3
Interest Rates, Model and Data

	Change, %
Model:	
Riskless rate, R^f	−1.45
Data:	
1-year real rate	−2.48
5-year real rate, 5 years forward	−1.57
5-year real rate, 5 years forward (HP trend)	−1.78
Natural real rate*	−.66
Liquidity premium*	.52

Note: The changes in the data panel are differences between average levels in 2017 and 2007.
*From Del Negro et al. (2017).

longer-term real rates. The third row shows the change in the 5-year real rates 5 years forward. To estimate this, we use the nominal 5-year rate 5 years forward (computed from the constant maturity nominal Treasury yield curve) and the corresponding expected inflation (i.e., the expected 5-year inflation rate 5 years forward, which can be computed using the 5- and 10-year expected inflation series from the Federal Reserve Bank of Cleveland). The fourth row reports the change in the Hodrick–Prescott-trend component of the 5-year real rate 5 years forward (computed using annual data from 1982–2017 with a smoothing parameter of 6). The fifth row shows the change in the estimate of the long-run natural rate from Del Negro et al. (2017), who use a flexible VAR specification to extract the permanent component of the real interest rate from data on nominal bond returns, inflation, and their long-run survey expectations (from the Survey of Professional Forecasters).[13] Taken together, the data show that belief revisions can go long a way in explaining the drop in interest rates since the financial crisis.

For macroeconomic quantities, the predicted drops in table 2 generally underpredict the deviations from precrisis trends observed in the data. For example, at the end of 2017, output was about 14% below the 1952–2007 trend. This suggests a need for additional amplification mechanisms. In our related work in Kozlowski et al. (2017), we explore two such mechanisms—Epstein-Zin utility (which allows us to separate risk aversion and intertemporal elasticity of substitution) and defaultable debt (higher tail risk makes default debt less attractive, curtailing borrowing and investment)—and show that they help bring the model's predictions much closer to the data. Here, given our focus on interest rates, we abstract from these modifications. This allows us to highlight, in a more transparent fashion, the interaction of tail risk with liquidity considerations.

Role of Liquidity

To understand the role played by liquidity, we repeat the analysis above, setting the pledgeability of capital to 0. This implies that shocks to capital do not directly affect the available liquidity in the economy (because bonds are the only liquid asset in the economy). The remaining parameters are calibrated using the same strategy as before. The results are shown in table 4. The table shows that without liquidity effects, the increase in tail risk has a very small effect on the riskless rate. The interest rate on government bonds in the new steady state is only 2 basis

Table 4
Interest Rates and Macro Aggregates in the Long Run, without Liquidity Effects

	$\hat{g}_{2007}$	$\hat{g}_{2009}$	Change
ln $F(K, N)$	2.27	2.19	−.09
ln X	1.29	1.29	.00
ln K	3.93	3.80	−.13
Riskless rate (R^f)	2.31	2.29	−.02
Risky return (R^v)	5.28	5.27	−.01
Risk premium ($R^v - R^f$)	2.97	2.98	.01

Note: R^f is the interest rate on government bonds, whereas R^v is the average expected returns on unlevered claims to the firm.

points lower. In other words, almost all of the drop in our baseline analysis comes from the interaction of tail risk and liquidity.[14] This finding is consistent with that of Del Negro et al. (2017), who find that most of the change in the natural real rate comes from a rise in the convenience yield associated with US government bonds. Their VAR estimate is reported in the last row of table 3 (labeled "liquidity premium")—the change in convenience yield since 2007 constitutes almost 80% of the drop in real rates.[15]

Comparing the implications for macroeconomic aggregates in tables 2 and 4 shows that liquidity dampens the effect of increased tail risk on capital and output (the predicted drops in table 4 are smaller). Intuitively, when capital also provides liquidity, an increase in tail risk induces a precautionary response—firms hold more capital to buffer against the drop in liquidity due to an adverse shock. As a result, steady-state capital (and, therefore, output) does not fall by as much as it would have in the absence of liquidity considerations.

Evidence from Equity Markets

Our model stays relatively close to the standard neoclassical paradigm and inherits many of its limitations when it comes to matching asset pricing facts, particularly asset price volatility.[16] With that caveat in mind, we confront the model's predictions for equity markets with the data in table 5.

To do this, we interpret equity as a levered claim on the value of the firm in the model. The main role of leverage is to amplify the volatility of equity returns. We use a leverage (defined as the ratio of debt to total assets) of 0.8. This is higher than most estimates in the literature—for example, Kozlowski et al. (2017) use 0.7, an estimate that combines operating and financial leverage. We discuss the reasons behind the higher leverage assumption later.

Table 5
Implications for Equity Markets

	Changes	
	Model	Data
Return on equity, $\mathbb{E}(R^e)$ (%)	−.065	−.184
ln equity/capital	.010	.225
$\mathbb{E}(R^e - \bar{R}^e)^3$	−.002	−.002
$\Pr(R^e - \bar{R}^e \leq -0.30)$	.022	.015

Note: The model changes represent the difference between the average value under $\hat{g}_{2009}$ and that under $\hat{g}_{2007}$. The change in the data is the difference between the average value from 2013 through 2017 and the precrisis average (from 2005 to 2007).

The implications of the model for various equity market variables are shown in table 5. The increase in tail risk leads to a slight fall in the expected return on equity claims in the new steady state. Because rates on riskless assets drop significantly, this implies a big rise in the equity premium. In data, expected returns on the S&P 500, computed following the methodology of Cochrane (2011) and Hall (2015), also show a small drop relative to precrisis levels.[17] The small drop in expected returns also means that the model-implied value of equity claims (per unit capital) is actually higher in the new steady state. In the data, we observe a much larger run-up in equity prices over the past few years. We are not claiming that the model can rationalize such a large increase, but it is worth noting that increased tail risk does not necessarily imply a precipitous fall in valuations.

Evidence of returns and valuations is at best a very indirect measure of tail risk. We therefore turn to options prices, arguably a better source of evidence of changes in tail risk. The model, even with the relatively high leverage adjustment, does not generate sufficient variability in equity returns. The model-implied value for the Chicago Board Options Exchange (CBOE) Volatility Index (VIX) under the pre-2008 beliefs is 8.37 (the average from 1990–2007 in the data was 19). Furthermore, in the data, the VIX spiked in the immediate aftermath of the crisis, averaging 32 during 2008–9, but then fell sharply to historically low levels in 2017. The model, on the other hand, predicts a more modest but persistent increase (from 8.37 to 11.35). The SKEW index reported by the CBOE (2019) is a transformation of the standardized third moment:

$$\mathrm{SKEW}_t = 100 - 10\frac{\mathbb{E}(R^e - \bar{R}^e)^3}{(\mathrm{VIX}_t/100)^3}.$$

Because this is a function of the VIX, the model's difficulty in matching the time variation in the VIX also spills over to the SKEW index. For example, the SKEW spiked in part due to the sharp drop in VIX. Fixing these issues, that is, matching the levels and time variation in volatility measures, would require adding more shocks—almost certainly with heteroskedastic processes—and mechanisms to address the well-known excess volatility puzzles, an exercise beyond the scope of the current paper. Instead, we use the two reported indexes to construct two indicators of tail risk—namely, the nonstandardized third moment of the risk-neutral distribution (the numerator in the second term of the SKEW equation above) and the (risk-neutral) probability of an extreme negative return realization (defined as 30% below the mean).[18] As table 5 shows, the model predicts significant increases in both objects. These predictions line up reasonably well with the changes in their empirical counterparts relative to their precrisis levels. In other words, while the model cannot exactly match the time path of asset market variables, the evidence from asset markets appears to be broadly consistent with the idea that tail risk rose sharply since 2008.

What If There Are No More Crises?

Next, we compute time paths for riskless interest rates, starting from the average long-run values under $\hat{g}_{2007}$. These paths are generated using two different assumptions about future shocks. The first corresponds to the stochastic steady-state analysis from earlier and draws shocks from the updated belief distribution $\hat{g}_{2009}$. The second assumes that crises do not recur—that is, the shock sequences drawn from $\hat{g}_{2007}$. For each sequence of shocks, we compute beliefs, equilibrium prices, and quantities at each date. Finally, we average over all these paths and plot the mean change in interest rates (relative to the starting level) in figure 5*a* and 5*b* ("learning" line). It shows that under both assumptions, they remain depressed for a prolonged period. In the no-crisis version (fig. 5*b*), although the economy eventually returns to its precrisis stochastic steady state, learning about tail probabilities is sufficiently slow that interest rates are almost 1% lower 20 years after the crisis. This occurs because learning about tail events is "local" under our nonparametric approach: beliefs about the likelihood change a lot when such events are observed but are less responsive to realizations elsewhere in the support of the distribution. In contrast, if we imposed a parametric assumption (e.g., a normal distribution), then all realizations contain information about parameters (mean and variance) and so beliefs (and, therefore, interest rates) would converge back to their precrisis levels relatively quickly.

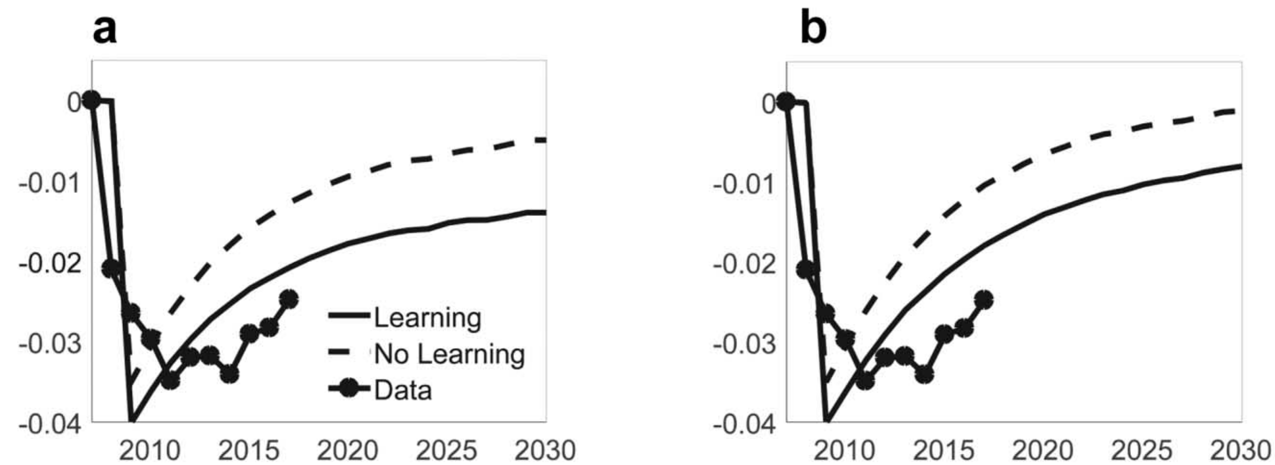

Fig. 5. Risk-free rate. (*a*) With crisis. (*b*) No more crisis. Panel (*a*) shows the risk-free rate when the data generating process is $\hat{g}_{2009}$. Panel (*b*) shows an identical model in which future shocks are drawn from $\hat{g}_{2007}$. The "learning" line in both panels shows the solution when agents update their beliefs, and the "no learning" line shows the model with no learning. The circles show changes in 1-year real rates from US data for the period 2008–17.

Turning Off Belief Updating

To demonstrate the central role of learning, we also plot average simulated outcomes from an otherwise identical economy in which agents know the final distribution $\hat{g}_{2009}$ with certainty from the very beginning ("no learning" line in fig. 5). These agents do not revise their beliefs. This corresponds to a standard rational expectations econometrics approach, in which agents are assumed to know the true distribution of shocks hitting the economy and econometricians estimate this distribution using all available data. The post-2009 paths are simulated as follows: Each economy is assumed to be at its stochastic steady state in 2007 and is subjected to the same sequence of shocks—two large negative ones in 2008 and 2009. After 2009, the sequence of shocks is drawn from the estimated 2009 distribution.

In the absence of belief revisions, the negative shock causes the real rate to surge and then recover. The interest rate rises because as the economy recovers to the previous steady state, there is a lower demand for debt.[19] This shows that learning is what generates long-lived reductions in economic activity.

VI. 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. 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. In this paper, we present one such application: the effect of large, unusual events on riskless interest rates.

The central mechanism is that observing tail events like the Great Recession leads agents to assign a higher likelihood to such events going forward. Importantly, this change in beliefs is relatively persistent, even if crises never recur. As a result, assets that are safe and liquid, such as government bonds, become more valuable.

When we quantify this mechanism and use capital price and quantity data to directly estimate beliefs, the model predicts large, persistent drops in interest rates, similar to the observed decline in government yields in the years following the Great Recession. These results suggest that perhaps persistently low interest rates took hold because after seeing how fragile our financial sector is, market participants will never think about tail risk in the same way again.

Appendix

Role of Risk Aversion

In table A1, we show how higher risk aversion translates to a larger sensitivity of interest to tail risk, even in the absence of liquidity effects.

Computing Option-Implied Tail Probabilities

To compute tail probabilities, we follow Backus, Foresi, and Wu (2008) and use a Gram-Charlier expansion of the distribution function. The CBOE also follows this method in its white paper on the SKEW index to compute implied probabilities. This yields an approximate density function for the standardized random variable, $\omega = (x - \mu)/\sigma$:

$$f(\omega) = \varphi(\omega)\left[1 - \gamma\frac{(3\omega - \omega^3)}{6}\right] \quad \text{where} \quad \gamma = E\left[\frac{x-\mu}{\sigma}\right]^3,$$

where $\varphi(\omega)$ is the density function of a standard normal random variable and γ is the skewness. (The Gram-Charlier expansion also includes a term for the excess kurtosis, but it is omitted from the expansion because, as shown by Bakshi, Kapadia, and Madan [2003], it is empirically not significant.)

Table A1
Interest Rates in the Long Run, without Liquidity Effects

Risk Aversion	$\hat{g}_{2007}$	$\hat{g}_{2009}$	Change
σ = 2	2.31	2.23	−.08
σ = 10	2.31	1.67	−.64

Endnotes

Authors' email addresses: Kozlowski (kozjuli@gmail.com), Veldkamp (lveldkam@stern.nyu.edu), Venkateswaran (vvenkate@stern.nyu.edu). We thank Jonathan Parker, Marty Eichenbaum, and Mark Gertler for helpful comments and suggestions, and François Gourio and Robert E. Hall for their insightful discussions. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c14073.ack.

1. This index is constructed from market prices of out-of-the-money options on the S&P 500 (CBOE 2019). It is designed to mimic movements in the skewness of risk-adjusted skewness; a higher level indicates a more negatively skewed distribution. Note that this is different from the Chicago Board Options Exchange Volatility Index (VIX), which measures implied volatility (i.e., the second moment). The VIX rose dramatically in the immediate aftermath of the crisis but came down quite sharply afterward.

2. Backus, Ferriere, and Zin (2015) analyze propagation in business cycle models.

3. These include papers on news shocks (e.g., Beaudry and Portier 2004; Lorenzoni 2009; Veldkamp and Wolfers 2007), uncertainty shocks (e.g., Jaimovich and Rebelo 2006; Bloom et al. 2014; Nimark 2014; Berger, Dew-Becker, and Giglio 2017), and higher-order belief shocks (e.g., Angeletos and La'O 2013; Huo and Takayama 2015).

4. For example, in Moriera and Savov (2015), learning about a hidden two-state Markov process with exogenously known persistence changes the demand for shadow banking (debt) assets.

5. Other kernels we explored included other nonparametric kernels (e.g., Epinechnikov), kernels designed to better capture tail risk (e.g., Champernowne), and semiparametric kernels with Pareto tails and the g-and-h family, which covers several transformations of the normal distribution. Each alternative yielded similar economic predictions because new data increased the tail probabilities of each distribution in a similar way. For a detailed discussion of nonparametric estimation, see Hansen (2015).

6. As $\kappa_t \to 0$, the CDF of the kernel converges to $\hat{G}_t^0(\phi) = (1/n_t)\Sigma_{s=0}^{n_t-1} 1\{\phi_{t-s} \leq \phi\}$. Then, for any ϕ, $j \geq 1$, $\mathbb{E}_t[\hat{G}_{t+j}^0(\phi)|\mathcal{I}_t] = \mathbb{E}_t\{[1/(n_t+j)]\Sigma_{s=0}^{n_t+j-1} 1\{\phi_{t+j-s} \leq \phi\}|\mathcal{I}_t\}$ and $\mathbb{E}_t[\hat{G}_{t+j}^0(\phi)|\mathcal{I}_t] = [n_t/(n_t+j)]\hat{G}_t^0(\phi) + [j/(n_t+j)]\mathbb{E}_t[1\{\phi_{t+1} \leq \phi\}|\mathcal{I}_t]$. Thus, future beliefs are, in expectation, a weighted average of two terms: the current belief and the distribution from which the new draws of the data ϕ_t are made. When shocks are also drawn from the current belief distribution, the two terms are exactly equal, implying $\mathbb{E}_t[\hat{G}_{t+j}^0(\phi)|\mathcal{I}_t] = \hat{G}_t^0(\phi)$.

7. The i.i.d. assumption also has empirical support. In the next section (Liquidity), we use macro data to construct a time series for ϕt. We estimate an autocorrelation of .15, statistically insignificant.

8. It is straightforward to allow for some unsecured debt, and this has a negligible effect on our results.

9. Our results are robust to alternative measures of nominal price changes, e.g., computed from the price index for gross domestic product or personal consumption expenditure.

10. Calibrating to the average values during 1950–2007 yields almost identical results.

11. Liquid assets are defined as total financial assets (BGFRS 2017, table B.103, row 7) less long-term financial assets (rows 21–24).

12. The steady state is obtained by simulating the model for 1,000 periods using $\hat{g}_{2007}$ and the associated policy functions, discarding the first 500 observations and time-averaging across the remaining periods.

13. We thank Del Negro et al. (2017) for sharing their estimates with us.

14. This is in part due to the low level of risk aversion in our parameterization. In the appendix, we repeat this analysis with higher risk aversion (specifically, $\sigma = 2$ and $\sigma = 10$). Then, tail risk has a somewhat larger effect on interest rates, even in the absence of liquidity.

15. They add the spread between Baa corporate bonds and Treasuries to their VAR to identify the convenience yield component.

16. However, the model actually implies a sizable equity premium even in the pre-2008 steady state. This stems almost entirely from liquidity considerations, which drive down the required return on government bonds relative to all illiquid assets (e.g., equity). This is essentially the mechanism in Lagos (2010), who shows that a model with liquidity considerations can help rationalize many asset pricing anomalies, including the equity premium puzzle.

17. The 1-year-ahead forecast of returns is obtained using a regression where the left-hand variable is the 1-year real return on the S&P and the right-hand variables are a constant, the log of the ratio of the S&P at the beginning of the period to its dividends averaged over the prior year, and the log of the ratio of real consumption to disposable income in the month prior to the beginning of the period.

18. Details of the computation are in the appendix.

19. Since the no-learning economy is endowed with the same end-of-sample beliefs as the learning model, they both ultimately converge to the same level. But they start at different steady states (normalized to 0 for each series).

References

Angeletos, G.-M., and J. La'O. 2013. "Sentiments." *Econometrica* 81 (2): 739–79.

Backus, D., A. Ferriere, and S. Zin. 2015. "Risk and Ambiguity in Models of the Business Cycle." *Journal of Monetary Economics* 69:42–63.

Backus, D. K., S. Foresi, and L. Wu. 2008. "Accounting for Biases in Black-Scholes." SSRN Working Paper Series. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=585623.

Bakshi, G., N. Kapadia, and D. Madan. 2003. "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options." *Review of Financial Studies* 16 (1): 101–43.

Barro, R. J., J. Fernández-Villaverde, O. Levintal, and A. Mollerus. 2014. "Safe Assets." Working Paper no. 20652, NBER, Cambridge, MA.

Beaudry, P., and F. Portier. 2004. "An Exploration into Pigou's Theory of Cycles." *Journal of Monetary Economics* 51:1183–216.

Berger, D., I. Dew-Becker, and S. Giglio. 2017. "Uncertainty Shocks as Second-Moment News Shocks." Working Paper no. 23796, NBER, Cambridge, MA.

Bernanke, B. S., C. C. Bertaut, L. Demarco, and S. B. Kamin. 2011. "International Capital Flows and the Return to Safe Assets in the United States, 2003–2007." International Finance Discussion Papers no. 1014, Board of Governors of the Federal Reserve System, Washington, DC.

BGFRS (Board of Governors of the Federal Reserve System). 2017. "Financial Accounts of the United States—Z.1: B.103 Balance Sheet of Nonfinancial Corporate Business." https://www.federalreserve.gov/releases/z1/current/html/b103.htm.

BGFRS (Board of Governors of the Federal Reserve System). 2019. "Selected Interest Rates (Daily)—H.15: H.15 Selected Interest Rates." https://www.federalreserve.gov/releases/h15/.

Bigio, S. 2015. "Endogenous Liquidity and the Business Cycle." *American Economic Review* 105 (6): 1883–927.

Bloom, N., M. Floetotto, N. Jaimovich, I. Sapora-Eksten, and S. Terry. 2014. "Really Uncertain Business Cycles." Working Paper no. 13385, NBER, Cambridge, MA.

Brunnermeier, M., and Y. Sannikov. 2014. "A Macroeconomic Model with a Financial Sector." *American Economic Review* 104 (2): 379–421.

Caballero, R. J., E. Farhi, and P.-O. Gourinchas. 2016. "Safe Asset Scarcity and Aggregate Demand." *American Economic Review* 106 (5): 513–18.

Carvalho, C., A. Ferrero, and F. Nechio. 2016. "Demographics and Real Interest Rates: Inspecting the Mechanism." *European Economic Review* 88: 208–26.

CBOE (Chicago Board Options Exchange). 2019. "Cboe SKEW Index (SKEW)." http://www.cboe.com/products/vix-index-volatility/volatility-indicators/skew.

Cochrane, J. H. 2011. "Presidential Address: Discount Rates." *Journal of Finance* 66 (4): 1047–108.

Condon, Bernard. 2013. "AP Impact: Families Hoard Cash 5 yrs after Crisis." *Associated Press*, October 6. https://apnews.com/10697d7ba73645eabe3e39adff0c190c.

Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti. 2017. "Safety, Liquidity, and the Natural Rate of Interest." *Brookings Papers on Economic Activity* 2017 (1): 235–316.

Gertler, M., and P. Karadi. 2011. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics* 58:17–34.

Gertler, M., and N. Kiyotaki. 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." In *Handbook of Monetary Economics*, vol. 3, ed. Benjamin Friedman and Michael Woodford, 547–99. Amsterdam: Elsevier.
Gourio, F. 2012. "Disaster Risk and Business Cycles." *American Economic Review* 102 (6): 2734–66.
Hall, R. E. 2015. "High Discounts and High Unemployment." Working paper, Hoover Institution, Stanford University.
———. 2017. "Low Interest Rates: Causes and Consequences." *International Journal of Central Banking* 13 (3): 103–17.
Hansen, B. E. 2015. "Econometrics." Working paper, University of Wisconsin. https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics2015.pdf.
Hansen, L. 2007. "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk." *American Economic Review* 97 (2): 1–30.
Huo, Z., and N. Takayama. 2015. "Higher Order Beliefs, Confidence, and Business Cycles." Working paper, University of Minnesota.
Irwin, Neil. 2016. "'Brexit' Is Locking In the Forces That Already Haunt the Global Economy." *New York Times*, June 27. https://www.nytimes.com/2016/06/28/upshot/brexit-is-strengthening-the-forces-that-already-haunt-the-global-economy.html.
Jaimovich, N., and S. Rebelo. 2006. "Can News about the Future Drive the Business Cycle?" *American Economic Review* 99 (4): 1097–18.
Johannes, M., L. Lochstoer, and Y. Mou. 2016. "Learning about Consumption Dynamics." *Journal of Finance* 71 (2): 551–600.
Kozlowski, J., L. Veldkamp, and V. Venkateswaran. 2017. "The Tail That Wags the Economy: Belief-Driven Business Cycles and Persistent Stagnation." Working paper, New York University.
Lagos, R. 2010. "Asset Prices and Liquidity in an Exchange Economy." *Journal of Monetary Economics* 57 (8): 913–30.
Lagos, R., and R. Wright. 2005. "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy* 113 (3): 463–84.
Lorenzoni, G. 2009. "A Theory of Demand Shocks." *American Economic Review* 99 (5): 2050–84.
Maćkowiak, B., and M. Wiederholt. 2010. "Business Cycle Dynamics under Rational Inattention." *Review of Economic Studies* 82 (4): 1502–32.
Malmendier, U., and S. Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" *Quarterly Journal of Economics* 126 (1): 373–416.
Marcet, A., and T. J. Sargent. 1989. "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models." *Journal of Economic Theory* 48 (2): 337–68.
Merton, R. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica* 41 (5): 867–87.
Midrigan, V., and T. Philippon. 2011. "Household Leverage and the Recession." Working Paper no. 16965, NBER, Cambridge, MA.
Moriera, A., and A. Savov. 2015. "The Macroeconomics of Shadow Banking." Working paper, New York University.
Nimark, K. 2014. "Man-Bites-Dog Business Cycles." *American Economic Review* 104 (8): 2320–67.
Orlik, A., and L. Veldkamp. 2015. "Understanding Uncertainty Shocks and the Role of the Black Swan." Working paper, New York University. https://www0.gsb.columbia.edu/faculty/lveldkamp/papers/uncertaintyOV.pdf.
Pintus, P., and J. Suda. 2015. "Learning Financial Shocks and the Great Recession." Working paper, Aix-Marseille School of Economics.

Sargent, T. J. 1999. *The Conquest of American Inflation*. Princeton, NJ: Princeton University Press.
Sundaresan, S. 2018. “Optimal Preparedness and Uncertainty Persistence.” Working paper, Imperial College, London. https://docs.wixstatic.com/ugd/0da7cb_361e147bfbe24f27a92b7bf0b363ccd3.pdf.
Veldkamp, L., and J. Wolfers. 2007. “Aggregate Shocks or Aggregate Information? Costly Information and Business Cycle Comovement.” *Journal of Monetary Economics* 54:37–55.
Venkateswaran, V., and R. Wright. 2014. “Pledgability and Liquidity: A New Monetarist Model of Financial and Macroeconomic Activity.” *NBER Macroeconomics Annual* 28 (1): 227–70.
Weitzman, M. L. 2007. “Subjective Expectations and Asset-Return Puzzles.” *American Economic Review* 97 (4): 1102–30.