Journal of International Money and Finance 122 (2022) 102529

Contents lists available at ScienceDirect

Journal of International Money and Finance

journal homepage: www.elsevier.com/locate/jimf

Financial frictions in macroeconomics ☆

Lawrence Christiano

Department of Economics, Northwestern University, United States

ARTICLE INFO

Article history:
Available online 25 October 2021

Keywords:
Risk shock
Business cycles
Rollover crisis
Banking
Interest rate spread

ABSTRACT

I review two examples that show how the nature of the financial system can play a central role in shaping the behavior of the aggregate economy. In the first example, variations over time in the cross-sectional dispersion of a productivity shock, which would have no aggregate effect in a frictionless model, produce effects that look like business cycles because of the nature of financial (and nominal) frictions. The second example suggests how a shock originating outside the financial system, which ordinarily might not be expected to have a large aggregate effect, can lead to a systemic banking collapse. The relevance of the examples to the US economy is discussed.

1. Introduction

Research on Macroeconomics and Finance has exploded since the Great Financial Crisis in 2007–2008, and a complete overview is of course impossible. With such a literature, many perspectives are possible and so it's important for me to stress at the outset that the perspective offered here is my own. Among other things, this means that when I refer to empirical analysis I will be relying on data from my own country, the United States (US). To organize the literature, I find Fig. 1 useful. That figure represents the balance sheet of the whole financial system, including parts that are not directly under the protective umbrella of the Federal Reserve (Fed). This includes, for example, what some have come to call the 'shadow banking system': a part of the banking system which does not directly enjoy the 'lender of last resort' and other backstop facilities provided by the Fed.

The liability side of the banking system summarizes the borrowing that the US financial system does from the public. Frictions on that side of the balance sheet emphasize the banking system itself as a source of financial disturbance. This includes models of systemic banking crises such as the bank run crisis model of Diamond and Dybvig (1983) or the rollover crisis model of the type emphasized in Gertler and Kiyotaki (2015).[1] These types of crises are the major focus of 'macro prudential policy'. Fortunately, these crises are relatively rare events in countries like the US. But when they happen, the Great Depression and the Great Recession are testimony to the massive harm they can do to the economy.

The left side of banks' balance sheet pertains to the relationship between banks and the non-financial entities to whom they lend. The financial frictions that arise here typically originate in the non-financial sector. Examples of this include

☆ Keynote address, 5th HenU/INFER Workshop on Applied Macroeconomics, March 2019. Remarks are based on the material in Christiano et al. (August 2018), as well as on-going work with Husnu Dalgic, Xiaoming Li and Yuta Takahashi.

E-mail address: l-christiano@northwestern.edu

[1] Papers like Gertler and Karadi (2011),Gertler et al. (2016) and Gertler et al. (2020) also belong in this category.

https://doi.org/10.1016/j.jimonfin.2021.102529
0261-5606/

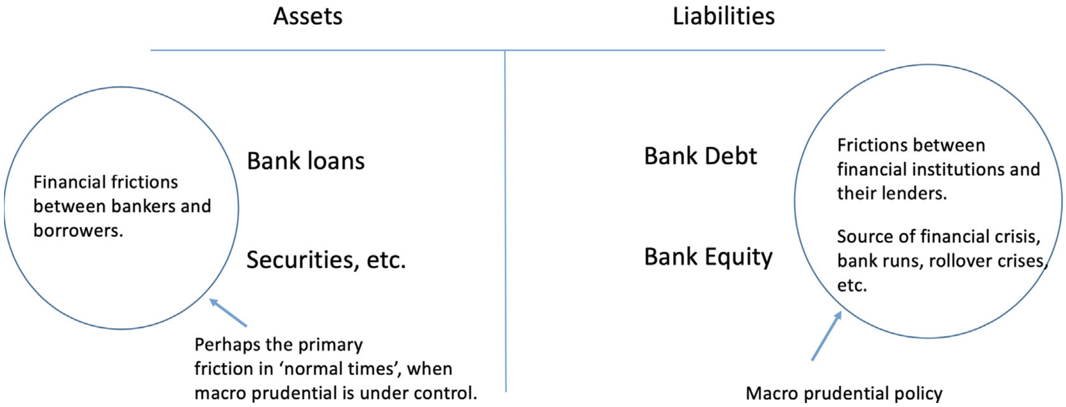

Fig. 1. Balance Sheet, Financial System.

Bernanke et al. (1999), Christiano et al. (2014, 2003) and many others, including Liu et al., 2013; Di Tella and Hall (2020).[2] I will argue that the disturbances on the left side of the balance sheet are the sort of shocks that occur in 'normal times' and provide a natural theory of business cycle fluctuations.

I will now briefly discuss the normal-times shock on the left of the balance sheet and the crisis-times shocks on the right hand side. I begin with the normal-times shocks.

2. Normal times

Consider Fig. 2, which graphs US data on GDP and on an indicator of the interest rate spread. The interest rate spread is measured by the difference between what BAA-rated private firms pay to borrow and what the Federal Government pays to borrow. One interpretation of this spread is that it reflects the relatively greater riskiness of non-financial firm debt compared to US government debt. Fig. 2 suggests that riskiness in the private sector increases in recessions and falls in expansions. This association could reflect that GDP falls in response to an increase in private sector risk. Or it could reflect that something else drives GDP down, which then makes firms more risky. Here, we explore the former idea. In particular, we explore the idea that, because of the way changes in risk propagate through the financial system, changes in private sector risk are an important source of business cycle fluctuations.

2.1. Risk shocks

The analysis is done by an extensive econometric study which makes use of a particular model. For the model, we adopt a modification of the one in the seminal paper by Bernanke et al. (1999). In the model there are agents (entrepreneurs) that buy and use capital and those purchases drive investment activity as well as aggregate output. Each entrepreneur finances their capital purchases in part by borrowing from a financial intermediary. The financial friction stems from the assumption that no one can really know how productive a capital project will be until it is actually up and running. That is, the outcome of any project, no matter how carefully planned in advance, contains an element of surprise *ex post*. This observation is captured in the model by the assumption that after an entrepreneur buys capital and puts it to work, an idiosyncratic productivity shock, ω, affects the capital's productivity. Although the realized value of ω is observed by the entrepreneur, it can only be observed by the financial intermediary by paying a monitoring cost. This assumption captures a natural asymmetry between entrepreneurs and their creditors that arises because of the close involvement of the entrepreneur with the complexities and details of a project. Creditors cannot see these things without expending time and resources. An equity-type contract between the entrepreneur and the bank is not feasible because the entrepreneur has an incentive to understate profits. So, following the seminal contribution of Townsend (1979), credit is assumed to be extended to entrepreneurs in the form of a *standard debt contract*.

The standard debt contract specifies a loan amount and the interest to be paid. At the time the loan is made, neither the entrepreneur nor the bank knows what value of ω the entrepreneur will experience. The contract internalizes that a known fraction of entrepreneurs will experience a value of ω so low that they cannot fully repay the interest and principal on their loan. The contract specifies that, in cases like this, the entrepreneur is monitored by the lender, who then takes whatever

[2] There are many other examples, indeed too many to provide an exhaustive list here. Not all important financial friction papers fit neatly into my division of the literature into models that locate the source of frictions in non-financial firms or in financial firms. Buera and Moll (2015) is an example of a model that does not fit into my taxonomy. That paper studies the macroeconomic consequences of a credit tightening. But it is not clear whether the tightening originates with rules imposed by regulators, a change in the risk appetite of borrowers or something about lenders. That the paper does not take a stand on this question is actually a strength because it allows one to think about macroeconomic effects of 'deleveraging' without getting into the details of why the deleveraging occurs.

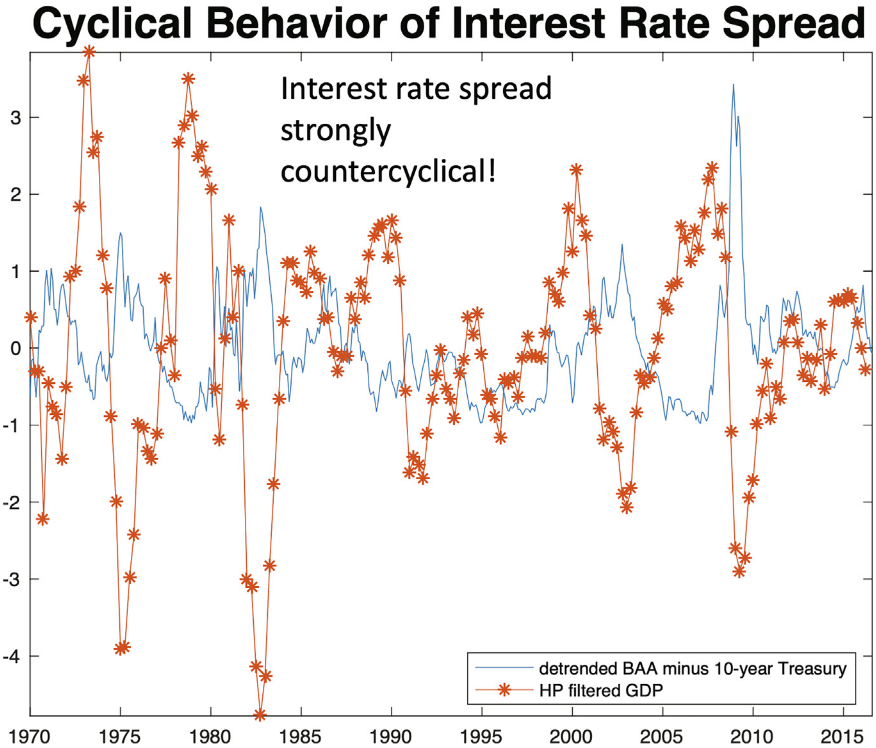

Fig. 2. Source of HP filtered data and detrended BAA minus 10-year Treasury return data: Federal Reserve Bank of St. Louis data base, FRED.

assets the entrepreneur has. This arrangement is called a standard debt contract because it resembles the kind of debt contracts observed in practice.[3] The standard deviation of $\ln \omega$, σ, is constant in Bernanke et al. (1999), but fluctuates stochastically in Christiano et al. (2003, 2014).[4] I follow Christiano et al. (2014) in calling σ a *risk shock*. Because σ is the realization of a stochastic process, I will refer to it as σ_t.

Christiano et al. (2014) show (using standard Bayesian estimation methods) that exogenous time series variations in the risk shock, σ_t, can account for roughly 60 percent of business cycle fluctuations in US output.[5] At the same time, variations in σ_t can account for the counter-cyclicality of bankruptcy and interest rate spreads; the pro-cyclicality of the stock market; the comovement of consumption, investment and credit over the business cycle; and many other features of US business cycle fluctuations.

Of course, for a new shock to be important in a data set that has also been analyzed in many other papers requires that some shocks estimated in previous models play a smaller role. In Christiano et al. (2014), the financial frictions and risk shock are incorporated into a variant of the New Keynesian model in Christiano et al. (2005). Smets and Wouters (2007) found that this model required substantial shocks to markups and to labor supply to fit standard macroeconomic data. Chari et al. (2009) criticized the model on the grounds that labor supply and markup shocks are not plausible as significant sources of business cycle fluctuations. Christiano et al. (2014) show that when financial frictions and the risk shock are introduced, the risk shock takes over the role of the labor supply and markup shocks. In this sense, by introducing financial

[3] The fact that the lender takes *all* the assets of entrepreneurs who cannot pay the interest on their debt reflects that entrepreneurs in the model behave as though they are risk neutral. They do this even though they are risk averse because they have access to insurance against idiosyncratic shocks via a large family assumption in Christiano et al. (2014). If entrepreneurs did not have access to insurance markets then presumably the equilibrium contracts would bear an even stronger resemblance to actual loan contracts, which do not transfer all the assets of a bankrupt borrower to the lender.

[4] The Christiano et al. (2014) model with a time-varying σ is also used in the analysis of Simon Gilchrist et al. (2012). In addition, Di Tella and Hall (2020) pursue the time-varying variance idea, but by a modeling approach substantially different from the one in Christiano et al. (2003, 2014). Finally, a variant on the Christiano et al. (2003, 2014) model is used in Del Negro and Schorfheide (2013). For an earlier paper that proposed variations in σ as a driving force of the business cycle, see Williamson (1987).

[5] The data used for estimation in Christiano et al. (2014) include the standard 8 variables used in business cycle analysis - GDP, inflation, the real wage, a risk-free interest rate, employment, consumption, investment and the relative price of investment goods. In addition, Christiano et al. (2014) use four financial variables - credit to non-financial firms, an indicator of the slope of the term structure of interest rates, an indicator of equity (the Wilshire 5000), and the interest rate spread data in Figure spread between BAA-rated corporate bonds and ten-year US government bonds.

frictions and risk shocks into a standard New Keynesian model we end up with a more plausible narrative of what drives business cycle fluctuations.[6]

2.2. Risk shocks and good in-sample fit

The mechanism by which a jump in σ_t propagates through a model like the one in Christiano et al. (2014) is simple and intuitive, and can be described with reference to Fig. 3. On the horizontal axis we have the possible realizations of the idiosyncratic shock, ω, to an entrepreneur's capital. This shock is assumed to obey a lognormal distribution with mean normalized to unity, so that $\omega \geqslant 0$ and $E\omega = 1$. Also, the standard deviation of $\ln \omega$ is σ_t, and $\ln \sigma_t$ is itself assumed to be the realization of a first-order autoregressive process.

The solid line in Fig. 3 is the distribution of ω for a baseline value of σ_t and the dot-dashed line is a perturbed distribution in which σ_t is 20 percent higher. The financial system, concerned about the greater fraction of entrepreneurs with bad luck (i.e., low ω) when σ_t is higher, charges entrepreneurs higher interest rates on their loans. In this way the lucky entrepreneurs with high ω, who are able to pay the interest on their standard debt contracts, pay enough to cover the greater losses to the banking system associated with entrepreneurs with bad luck. With the higher interest rate, all entrepreneurs borrow less from the financial system and the fraction of non-performing loans rises.[7]

The fall in borrowing results in a decline in entrepreneurial purchases of capital, which produces a fall in investment and a fall in the price of capital. The fall in the value of capital imposes capital losses on entrepreneurs, so that their net worth falls. I interpret this net worth as the value of the stock market. The underlying economy has sticky wages and prices so that monetary policy influences how the risk shock propagates through the rest of the economy. In the model, policy is governed by a standard Taylor rule and this is characterized, among other things, by interest rate smoothing. With interest rates and inflation relatively unchanged after a risk shock, and with income falling due to the drop in investment, households cut back on consumption. In this way, a rise in risk leads to an increase in interest rate spreads and non-performing loans, as well as declines in credit to non-financial businesses, investment, consumption and the stock market. Because a jump in risk in effect represents a negative shock to the demand for goods, the model implies that inflation falls. In short, a jump in the risk shock generates responses that resemble what we observe in actual recessions.

We can see the quantitative responses in the estimated model to a jump in σ_t in Fig. 4. Panel I shows the dynamic response in the risk shock to an innovation. The risk shock jumps 10 percent in the initial period and then slowly returns to its steady state value. The jump in risk leads to a jump in the interest rate spread (panel A), and an immediate decline in credit (panel B). Interestingly, credit continues to decline over four years even though the gap between σ_t and its steady state declines monotonically over time. Consumption drops on impact, and that decline accelerates over time, showing little sign of returning steady state even after 4 years.[8] This is so even though by this time about 60 percent of the gap between σ and its steady state value has been closed. Evidently, the model exhibits a substantial amount of internal propagation to the risk shock.

Since consumption and investment have a U-shaped response to an innovation in σ_t, GDP does too. The effect on GDP, in percent terms, is smaller than investment but larger than consumption. This corresponds roughly to the relative volatility observed in the data on these variables. Note that the impact on the stock market (entrepreneurial net worth) is largest in the period of the shock, even though all the other real variables take time to exhibit their maximal response. Thus, the stock market has an important high frequency component, something that we also see in the data. That inflation drops is not surprising because a jump in risk is a negative demand shock. The fact that these impulse responses 'look' so much like an actual recession helps explain why the Bayesian estimation procedure assigns so much responsibility to the risk shock.

Another way to see why the econometric analysis reported in Christiano et al. (2014) assigns such a major role to the risk shock is to ask: 'what would the past macroeconomic data have looked like if the risk shock had been the *only* shock?' We get the answer by feeding only the estimated risk shock to the model and comparing the simulated data with the actual data. Through the eyes of the model, the actual data are the economy's response to the many shocks (twelve) included in the estimation. If the datadriven only by the estimated value of the risk shock 'looks' like the actual data, this suggests that the other shocks play a relatively small role. Fig. 5 displays the growth rate of GDP and credit, as well as the value of the stock market, the interest rate spread ('premium') and the slope of the term structure. According to the model, the variable that accounts for virtually 100% of the observed spread is the risk shock. This is because, although all shocks theoretically move the interest rate spread, only the risk shock does so in a quantitatively substantial way. In effect, the econometric procedure 'identifies' the risk shock in the interest rate spread and de-emphasizes other shocks because the risk shock makes the model-simulated data move roughly as it moves in the actual data (compare the dotted and solid lines in the figure).

[6] The risk shock also greatly reduces the role of technology shocks in business cycles. Other shocks included in the analysis include temporary and permanent neutral shocks to technology as well as shocks to the marginal efficiency of investment. The estimated model assigns a greatly reduced role to those shocks in favor of the risk shock. Christiano et al. (2014) show that their results depend on including news about risk shocks in the model. A conjecture about why this is so is suggested by the results in Bachmann and Moscarini (2012), who show that Granger causality goes from measures of cross-sectional uncertainty to GDP. News shocks as in Christiano et al. (2014) can explain this statistical phenomenon and may be part of the reason why news shocks in σ are so important in the Bayesian estimation of the Christiano et al. (2014) model.

[7] That a rise in risk leads to an equilibrium fall in the interest and a cut-back in borrowing is more subtle than suggested in the text and is, for example, in part due to the assumption of a log normal distribution for ω.

[8] Of course, eventually consumption does return to its steady state value.

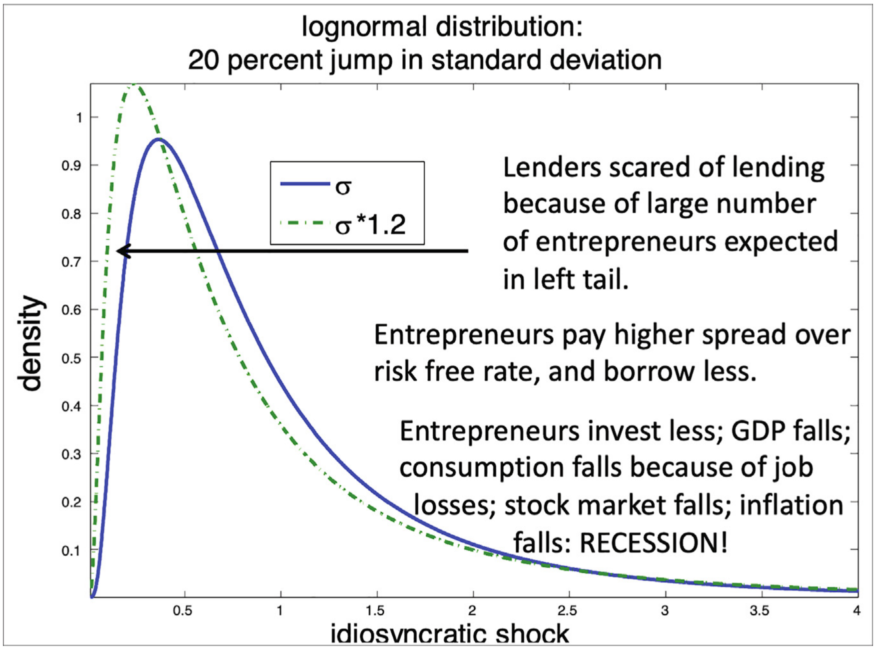

Fig. 3. Economic Impact of Risk Shock.

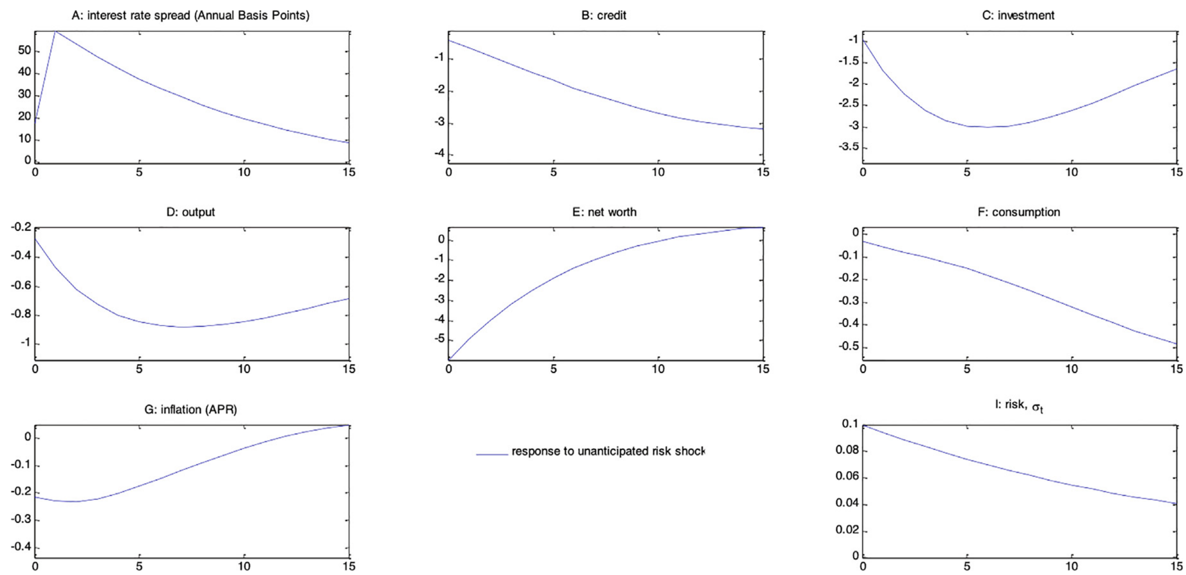

Fig. 4. Dynamic Responses to Unanticipated Risk Shock. Source: author's computations.

2.3. Out-of-sample evidence for the risk shock

The fact that the risk shock helps the model fit the data well is not sufficient to conclude that the risk shock is important, because there is a possibility that inference is distorted by overfitting. To guard against this possibility, I investigate whether there is more direct evidence on the risk shock and I also examine the model's predictions for variables not used in the estimation.

We do not have direct observations on the risk shock. But data on quarterly equity returns in a panel of non-financial firms from the Center for Research on Securities Prices (CRSP) are suggestive.[9] Fig. 6 displays (solid line) the time series on

[9] CRSP is affiliated with the University of Chicago, Booth School of Business.

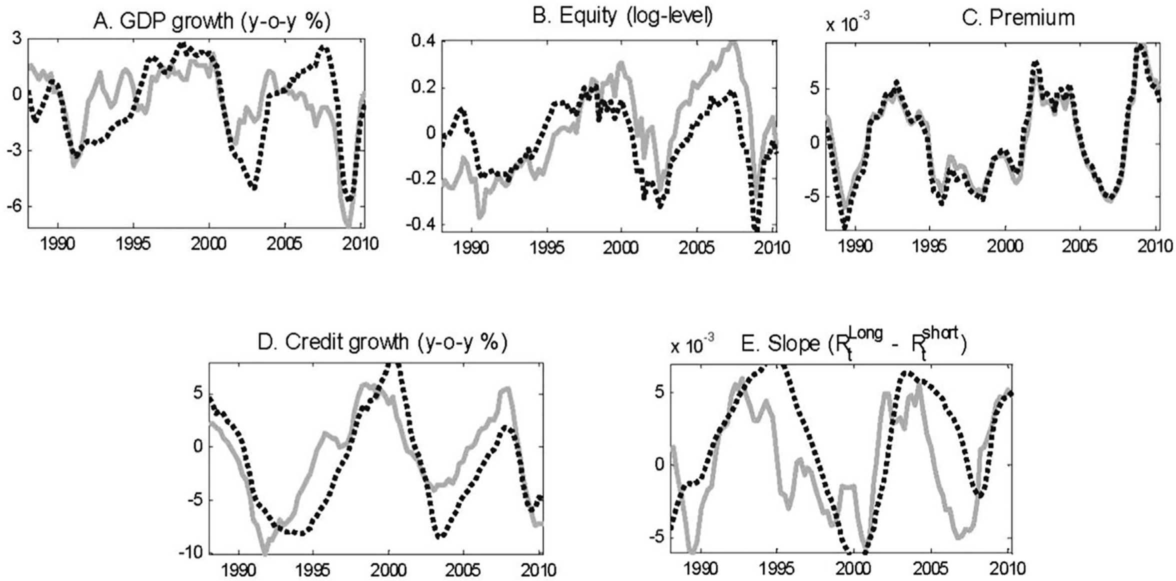

Fig. 5. Role of the Risk Shock in Macro and Financial Variables. Note: author's calculations. Solid line corresponds to actual year over year growth rate of indicated variable. Dotted line indicates what the 'actual' data would have been had only the estimated risk shock occurred. For further details see Christiano et al. (2014).

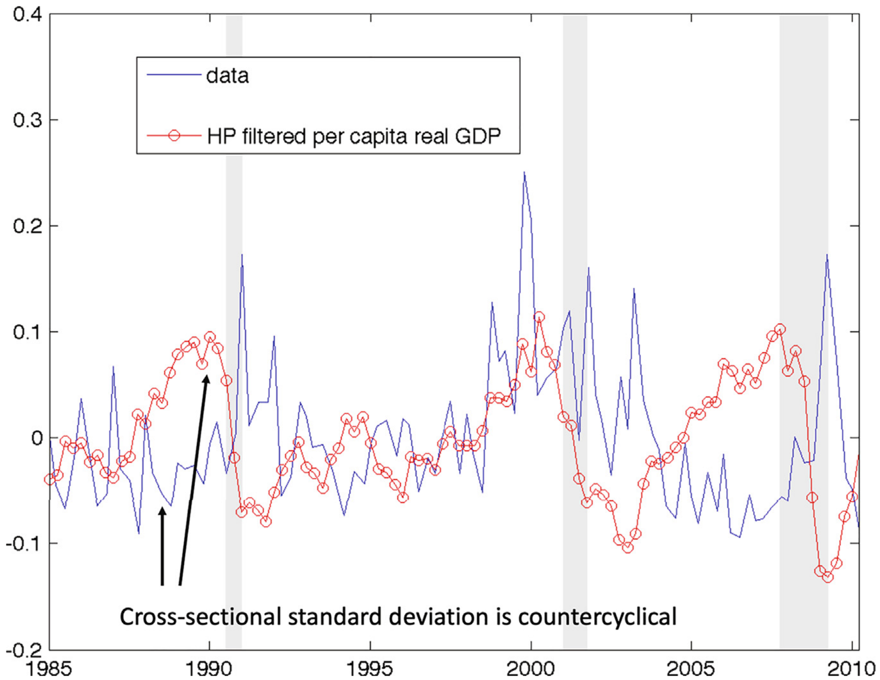

Fig. 6. Cross-sectional Standard Deviation, Quarterly Rate of Return on Non-financial Firm Equity, CRSP Data. Data sources: Solid line, the CRSP data were kindly converted to quarterly frequency and given to us by Thiago Teixeira Ferreira. Dot-circle line, US real GDP data, after taking logs and Hodrick-Prescott filtering, with smoothing parameter, 1600. The GDP data were obtained from the St. Louis Feds' website.

Λ_t. Here, Λ_t denotes the cross-sectional standard deviation of quarter t equity returns for non-financial firms in the CRSP dataset.[10] This cross-sectional standard deviation is an indirect indicator of σ_t.[11] The circle-dash line denotes log, HP detrended GDP. Note that in the periods before 1995 and after 2000, there is a negative relationship between the two series: in those periods, dispersion in equity returns tends to be high when GDP is low relative to trend. Thus, consistent with the spirit of the model, periods of low economic activity appear to be associated a greater dispersion in firm productivity, here measured by the return on equity.[12]

To get a better sense of the compatibility of Λ_t with the model, Christiano et al. (2014) reports the model's best guess for Λ_t based on the 12 variables used in estimation.[13] Fig. 7 displays the actual data (solid line) against the model's prediction (solid line with dots). Note that these series exhibit reasonably similar behavior, particularly if we take into account that the model-based data are computed using the posterior mode of the model parameters, while the probability interval around those parameters is non-trivial (though not reported here).

We also performed this type of out-of-sample forecasting exercise by comparing data on non-performing loans with the prediction of those data produced by the model. See Christiano et al. (2014) for a discussion of these results.

2.4. What can we do with the risk shock model?

Both in-sample and out-of-sample considerations support the idea that changes over time in risk can, because of the presence of financial frictions, explain a large portion of US business cycle fluctuations. What can we do with such a model? Ultimately, models (at least the ones that match the data well) are needed to answer important policy questions. A small list of questions that the type of model described here can address is:

- How should monetary policy response to an increase in the interest rate spread? We saw large increases in such spreads in 2008. But at that time the standard macroeconomic toolbox did not include models that could help provide guidance on this question. Models of the kind described here can do so.
- Extensions of the model described here have been used to study the relationship between monetary policy and credit growth. They have even been used to show that standard formulations of monetary policy that put too much focus on current inflation could inadvertently cause the central bank to become complicit in an inefficient stock market boom (see Christiano et al., 2010).
- Open economy versions of the kind of model considered here can be used to think about the desirability of smoothing exchange rates as well as government policies that can achieve this.[14]

3. Financial frictions on the liability side of the balance sheet

In response to the 2007–2008 financial crisis and the ensuing Great Recession, much work has been done to better understand how vulnerabilities in the financial system can transform a shock that might have had only small effects into a disaster. We discuss one such line of research here.

3.1. Proximate cause of financial crisis: housing boom and bust

A consensus has emerged that the proximate cause of the crisis was the housing boom that led to a sharp run-up in housing prices. In real time, it was generally understood at the Fed that the price rise might be 'excessive' and that at some point prices might falter and perhaps even undergo a correction. Two considerations underly the Fed's decision not to take a pro-active stance against the housing price rise. The first was a conviction that no one can really be sure what the 'right' price of housing is. Indeed, the idea that prices were not demonstrably out of line with fundamentals receives support from the fact that inflation-adjusted prices are now higher than they were at their 2006 peak (see Fig. 8). Second, there was a belief that policymakers could contain the macroeconomic fallout if the housing price boom came to the end.

In retrospect, the latter conviction was clearly wrong. When prices began to falter in late 2006, a sequence of events unfolded that led not only to a disastrous collapse in housing prices, but also to a systemic financial collapse (see Bernanke, 2010; Gorton, 2010). Policymakers were seemingly unable to stop the disaster. This fact was a surprise to virtually

[10] To be specific, let i denote a firm and t denote a date. Let $r_{i,t}$ denote the rate of return on equity from quarter $t-1$ to quarter t. Then, the t^{th} observation in Fig. 6 is $\Lambda_t = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}(r_{i,t}-r_t)^2}$, where N_t denotes the number of firms in the cross section in period t in the (unbalanced) CRSP data set; and $r_t = \sum_{i=1}^{N_t} r_{i,t}/N_t$. The mean of Λ_t has been removed in Fig. 6. The equity returns were aggregated from CRSP's monthly frequency to quarterly in Ferreira (2016) and I am grateful to Thiago Ferreira for providing the data. The equity return data are similar to those studied by Bloom (2009), with the difference that I consider quarterly observations for non-financial firms, while Bloom (2009) examines monthly observations on all firms.

[11] For a formal discussion of the link between the variance of equity returns and the risk shock, see Ferreira (2013) and also Christiano et al. (2014, ftn 45).

[12] The results in Fig. 6 are similar to what is reported in Bloom (2009), who first noticed the association between greater dispersion in a variety of variables, not just the rate of return on equity, and business cycle recessions.

[13] The calculations were done using the Kalman smoother.

[14] Early examples of such models include Gertler et al. (2007) and Copaciu et al. (2015). The kinds of questions about exchange rates I have in mind are increasingly being addressed. For an example, see Fanelli and Straub, 2021.

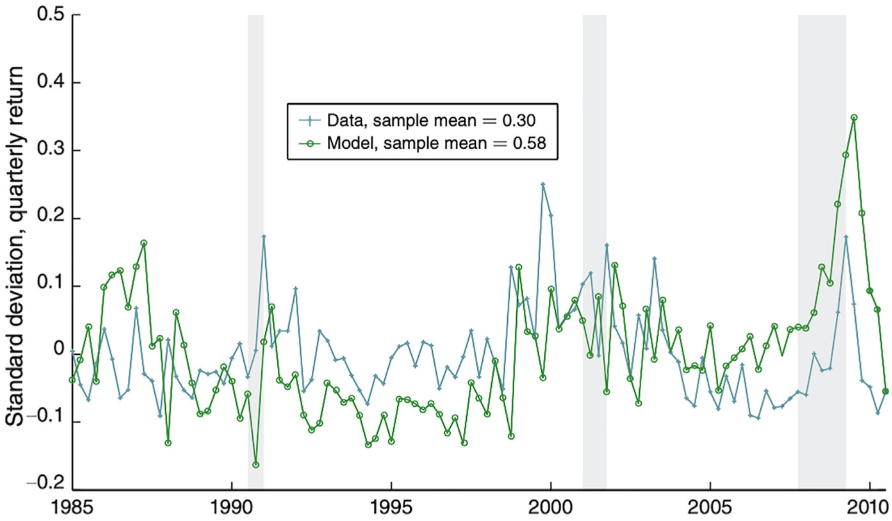

Fig. 7. Time Series, Cross Section Standard Deviations of Quarterly Rates of Return Model and Data. Note: replicated from Christiano et al. (2014, fig. 7).

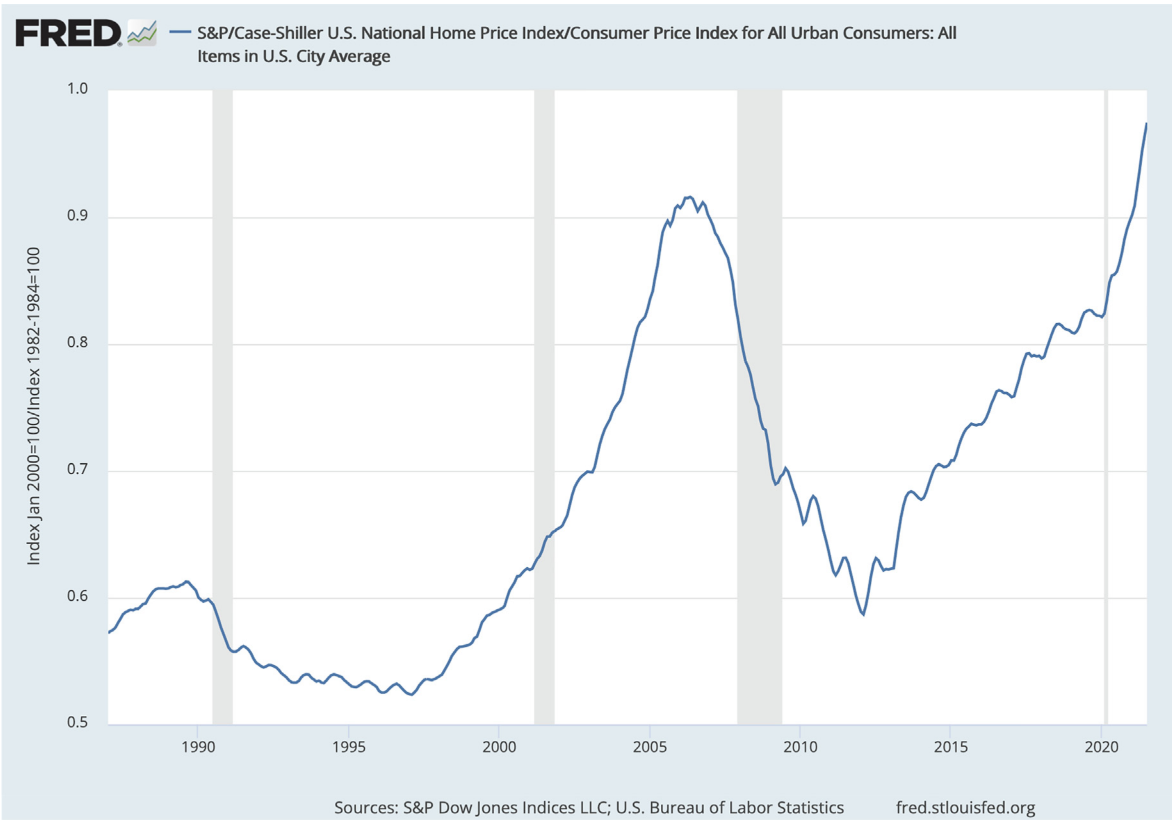

Fig. 8. Housing Prices, Adjusted for Inflation.

everyone, including policymakers, market participants and academics. Observers feared that the US might actually be headed towards a truly major disaster, another Great Depression. Fortunately, the authorities reacted vigorously enough to head off this really bad scenario.

3.2. The rise of the shadow banking system

The question on everyone's mind was: "what were the vulnerabilities in the financial system that converted a shock that 'should' have been containable into a disaster?" Although many factors were at play, there emerged a consensus that, unbeknownst to most policymakers and academics, a large portion of the banking system had become vulnerable to a type of bank run. In effect, the financial system was an 'accident waiting to happen'.

The bank run that occurred was obviously not the same as the type of run that occurred in the US Great Depression. We did not see the large crowds of forlorn people assembling in front of banks hoping to get their money, as in old photographs of bank runs from the 1930s. Another image of a bank run is the one depicted in the 1930s movie that is popular even today, 'Gone With the Wind'. These were obviously not the type of bank run that occurred in the US in 2007–2008. The runs in the Great Financial crisis occurred inside hermetically sealed office buildings in financial centers, and involved highly-paid professionals with eyes glued to computer terminals.[15]

Gradually, it became apparent that the financial vulnerability was due to a shadow banking system that had been growing since the early 1990s. The shadow banking system is the part of the banking system that is not under the protective umbrella of the Fed. Fig. 9 shows the evolution of the size (measured by the dollar value of its non-equity liabilities) of the shadow banking system (see 'Shadow liabilities'). In 1990 that system was smaller than the deposit-taking part of the banking system which has deposit insurance and which enjoys the protection of the Fed in exchange for being regulated (see 'Banking liabilities'). At that time, and for the next 17 years, those deposit-taking banks constituted the entire 'banking system' as far as many observers were concerned. But note in Fig. 9 how rapidly the shadow banking system grew in the 2000s. By 2007 that sector was substantially larger than the regular deposit-taking banking system.

The Fed was not entirely unaware of the shadow banking system, since its policies had in some ways explicitly encouraged its growth as part of a policy that in effect relaxed regulatory restrictions on the financial system. Much of this happened under Alan Greenspan, who was chairman of the Fed from August 1987 to January 2006. Greenspan was famous for his view that one can rely on bankers' sophistication and devotion to shareholders to take the appropriate precautions against risk. The heavy hand of regulation would only stand in the way of robust economic growth, according to Greenspan.

The conviction that the banking system could, in effect, regulate itself gave rise, for example, to the phenomenon whereby banks originated mortgages, and then resold them to financial entities outside the Federal Reserve universe (the shadow banks). The shadow banks in turn found cheap funding to finance the transaction because loan-originating banks (called 'sponsoring banks') offered implicit guarantees: if the mortgages went sour, the sponsoring bank would buy back the mortgages. Guarantees by sponsoring banks had some credibility because of their access to the financial protections of the Fed. At the same time, because the guarantees were only informal, reasonable people might have had doubts about them if market conditions changed.

Why did these off-balance sheet transactions occur? The implicit nature of the guarantees by sponsoring banks allowed those banks to exclude the assets explicitly from their balance sheets for the purpose of assessing capital requirements. If sponsoring bank guarantees had been explicit, then the banks would have been required to retain the mortgages on their balance sheet and 'pay' a tax on them in the form of heavier capital requirements. In this case, there would have been no particular advantage for sponsoring banks to re-sell the mortgages to shadow banks. But, because of the implicit nature of sponsoring-bank guarantees, whether they had to keep those assets on their balance sheets when they 'sold' them fell into a kind of 'grey zone' in the regulations. This gave discretion to regulators, and given the prevailing view from Fed leadership that regulatory burdens on the financial system should be minimized, regulators tended to allow banks to remove 'sold' mortgage assets from their balance sheets. This arrangement allowed sponsoring banks and shadow banks to in effect enjoy rents from sponsoring banks' access to Fed protections.[16]

Since the regulators at the Fed knowingly participated in the growth of the shadow banking system, they understood that it existed. Still, it was not recognized until too late just how big the shadow banking system had become. In some respects, it is not surprising that the size was not obvious to everyone. For example, a large part of the system was actually outside the US. Many European financial firms effectively became part of the US financial system by issuing short term dollar obligations in the US, and using the proceeds to purchase mortgage-backed securities in the US. Their strong preference for securities specifically backed by mortgages is part of the reason that so much finance poured into housing.[17]

The large size of the shadow banking system is what made the financial system an 'accident waiting to happen', and it was the failure to understand that size that accounts for the failure to understand the system's vulnerability. The accident that happened was the housing price correction that started in the summer of 2006. The idea is that it was the vulnerability of the banking system that converted the price correction into a rout. Note from Fig. 8 that real house prices adjusted for inflation are now roughly 6 percent higher than they were in 2006. Perhaps without the growth of the shadow banking system, housing prices would not have risen so much in the first place and, if there had been a correction, we might not have seen the massive drop that actually occurred (see Fig. 8).

Famously, in Congressional testimony in October 2008, Greenspan acknowledged that there had been a flaw in his thinking about the ability of financial markets to self-regulate, saying (see Beattie and Politi, 2008):

[15] See the popular Hollywood movies, 'Wall Street', or 'Margin Call'.

[16] There remains a question of why the shadow banks, which fell under the regulatory authority of the Securities and Exchange Commission (SEC), weren't regulated more carefully. One hypothesis is that, with regulatory authority split between different parts of the financial system, no one regulator had a broad enough perspective to see the systemic vulnerability of the entire system.

[17] See Justiniano et al. (2014) for a review of the literature on the large gross financial flows, and very small net flows, between Europe and the US, as well as the preference of European institutions for mortgage backed securities.

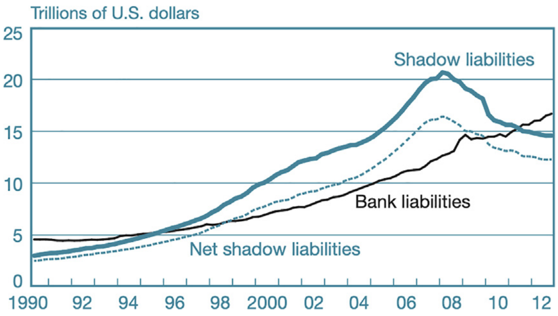

Fig. 9. Shadow Bank Liabilities versus Traditional Bank Liabilities. Notes: figure taken from Pozsar et al. (2013, page 7).

"I made a mistake in presuming that the self-interest of organizations, specifically banks and others, was such that they were best capable of protecting their own shareholders."

Note in Fig. 9, how sharply the shadow banking system contracted in the wake of the financial crisis. Ordinary deposit-taking commercial banks generally continued on trend and even expanded a little, providing a (small) buffer against the collapse of shadow banks.

3.3. Why a large shadow banking system creates vulnerability to a rollover crisis

Much progress has been made on the task of integrating the type of banking vulnerability described above into models.[18] Ultimately, the ideas are similar to the logic in Diamond and Dybvig (1983), though there is a difference in the details. As a result, I will refer to the banking crisis as a rollover crisis rather than a bank run.[19]

To begin, consider the bank balance sheet in Fig. 10. This bank is solvent. It has taken 100 in deposits and combined that with 20 units of its own resources (net worth) and used it to acquire 120 units of assets. In principle, if the depositors wanted their money back, the bank could give it to them by selling 100 units of assets.

We add three assumptions into the above example to create the possibility of a crisis. The first assumption is that there is a maturity mismatch on the balance sheet: the non-equity liabilities ('deposits') are short-term and the assets are long-term. Most real-world banks satisfy this assumption. In the case of the shadow banks discussed above, imagine the liabilities are short-term mortgage-backed commercial paper, with a duration perhaps of 6 months. The assets are mortgages with, say, 30-year duration. In the above example, the bank would pay off its short-term liabilities of 100 units every 6 months using money raised by issuing new short-term liabilities (this is called 'rolling over' its liabilities). The interest on the short-term liabilities can be paid using the interest earned on the assets. If the bank for some reason could not roll over its liabilities, this is no problem. It could just sell the assets. Especially by the mid 2000s, the market for mortgages was gigantic and one bank stepping into sell would quickly find a buyer. Still, we do not have enough assumptions to give us anything like a crisis.

We now turn to our second assumption. Suppose that in case *all* banks find themselves having to sell their assets at the same time, then the value of the assets would fall. Why assets would trade at fire-sale prices in such a case has been discussed in a variety of places (see, for example, Shleifer and Vishny, 2011). Although the model in Gertler and Kiyotaki (2015) is fairly abstract, the spirit of their mechanism can be conveyed intuitively as follows. We do not have to imagine the unrealistic scenario that literally *all* creditors to *all* banks in the entire financial system choose not to roll over. Suppose instead that there are different sub-sectors of the financial system which specialize in specific kinds of credit, e.g., car loans, mortgages, business loans for equipment purchases. One can even imagine finer sub-sectors (e.g., mortgages for large estates versus for modest condos). Because each sector is somewhat different from the other, over time they evolve different norms and conventions for writing their loans. In this way, idiosyncrasies in the debt contracts of one sector appear opaque to traders in other sectors, but are well understood by the people in the sub-sector, who are the ones that ordinarily trade the assets.

However, if *all* participants in a sub-sector are attempting to sell the assets of that sector, say because creditors refuse to roll over the sector's short-term liabilities, then there is a problem. Now there are important asymmetries in market trades. All the insiders in the market experiencing the financial crisis are sellers and all the potential buyers are in other markets and they find the assets to be opaque. These kinds of asymmetries between buyers and sellers have the consequence that the assets can only be sold for a low price, a *fire-sale price*. Thus, a bank with a balance sheet like in Fig. 10 could look perfectly sound, but then suddenly become bankrupt in case of a rollover crisis. For an eloquent description of the start of the rollover crisis in the summer of 2007, see the section "Triggers of the Crisis" in Bernanke (2010).

This is where size is crucial. If the rollover crisis is in a tiny sector, the amount of assets that must be sold to other sectors is not very large. Presumably, prices do not have to fall an enormous amount and in any case the small amount of assets

[18] For one strand of that research, see Gertler and Karadi (2011), Gertler and Kiyotaki (2015), Gertler et al. (2016) and Gertler et al. (2020).

[19] In Diamond and Dybvig (1983) the bank run occurs because of a 'sequential service constraint': people are given their deposits on a first-come-first-serve basis. This gives depositors an incentive to not end up at the back of the line when they hear that others are on their way to the bank to get their deposits. In rollover crisis models there is no sequential service constraint and everyone acts simultaneously. Still, the fundamental logic of both ideas is very similar.

Assets	Liabilities
120	Deposits: 100
	Banker net worth 20

Fig. 10. Solvent Bank.

involved does not matter. But, if a large sector (say, the market for mortgage-backed securities or the entire shadow banking system) experiences a rollover crisis, then prices will fall a lot and the spillovers onto the rest of the economy will be substantial.

For a rollover crisis to actually occur, we need a third assumption. The assumption is required for a crisis to be an *equilibrium*. For a rollover crisis to be a (Nash) equilibrium requires that each creditor believes it is in his or her own private interest to refuse to roll over, if he or she thinks all the other creditors have chosen not to roll over. That a rollover crisis might *not* be a Nash equilibrium is not hard to see. An individual who is contemplating that all other creditors refuse to roll over understands there will be fire sales of assets, and bankruptcy. Thus, the individual understands that their own credit will not be fully repaid. But there is nothing the individual can do about that. That is in effect a sunk cost ('water under the bridge') for the individual creditor. The important thing is that the creditor knows that very few investment projects will be financed in case all other creditors refuse to rollover. This means that the return on the dollar he or she deposits in the bank could be used to fund a very high return project. Under this line of reasoning, a rollover crisis is not a Nash equilibrium.

So, more is required for a roll over crisis to be a Nash equilibrium. In their model of a rollover crisis, Gertler and Kiyotaki (2015) adopt an assumption in Gertler and Karadi (2011) which has the (intuitively appealing) implication that no rational person would extend credit to a bankrupt financial institution. With this third and final assumption, we now have the possibility for a rollover crisis to be an equilibrium.

To understand this possibility, consider Fig. 11 which refers to the same bank as in Fig. 10. The critical difference is that the value of its assets is lower, indeed even lower than its non-equity liabilities because of a fire-sale drop in asset values. The net worth of this bank is zero and non-equity claims can only get 90 cents on the dollar.

3.4. Narrative of the financial crisis

We are now in a position to describe a highly stylized narrative of the financial crisis. Suppose the left panel of Fig. 12 characterizes the situation before the price correction began in the summer of 2006. Numbers not in parentheses are the balance sheet in case all creditors rollover their liabilities to the bank. Each bank has 120 in assets, easily enough to cover the 100 in non-equity liabilities. The numbers in parentheses indicate what would happen in the event that all creditors refuse to rollover. In this case, the fire sale value of the assets is 105, still enough to cover the non-equity liabilities. Because the bank is not bankrupt in the event of a failure of creditors to roll over, then it is not in the interest of an individual creditor to refuse to roll over, given the third assumption mentioned above. So, according to that model, the probability of a rollover crisis is zero.

Now consider the right panel in Fig. 12. This is the balance sheet after the initial housing price corrections in the summer of 2006. Note that the value of the assets falls in the no rollover equilibrium. So, there is an equilibrium in which creditors roll over their liabilities because the banks have positive net worth in that case. Consider the alternative scenario in which everyone refuses to roll over. Now, unlike previously, the fire sale dip in the value of assets wipes out the bank's net worth, and it is therefore a Nash equilibrium for everyone to not roll over. Thus, after the price correction there are two equilibria, one in which there is no crisis and one in which there is a crisis. According to this model, the economy is now in a region of

Fire sale value of assets:

Assets	Liabilities
90	Deposits: 100
	Banker net worth -10

Fig. 11. Bank in a Rollover Crisis with Assets at Fire-sale Values.

Pre-housing market correction

Assets	Liabilities
120 (105)	Deposits: 100
	Banker net worth 20 (5)

Post-housing market correction

Assets	Liabilities
110 (95)	Deposits: 100
	Banker net worth 10 (-5)

Fig. 12. Model of Pre- Crisis Period and Crisis.

vulnerability so that, with some bad luck, the banking system collapses. This narrative says that bad luck struck a vulnerable banking system after the price correction in summer 2006. It came as a surprise, in part because the vulnerability of the banking system was not generally known.

Someone familiar with the Diamond and Dybvig (1983) model will see the similarities and differences between the Gertler and Kiyotaki (2015)-style here and the Diamond and Dybvig (1983) analysis. In Diamond and Dybvig (1983) there is a run and a no-run equilibrium and luck determines which one occurs. However, the details about how and why the 'run' equilibrium exists are different.

4. Conclusion

I have sought to review some of the literature on macro-finance that has developed in recent years. One class of models locates the origin of financial turbulence in non-financial firms. The policy implication of this model is that policy should subsidize non-financial businesses in a recession through some kind of fiscal policy or by an expansionary monetary policy. The latter could be accomplished by cutting the interest rate when interest rate spreads rise.[20] Models of financial crises suggest two things. First, it is vital that the regulatory authorities have an integrated perspective of the overall financial system. To the extent that macroeconomists looked at financial data at all, they tended to divide the financial sector into banks and bond and equity markets. The banking sector, however, was limited to the deposit-taking commercial banks regulated by the Fed. The experience of the financial crisis indicates that it is important to also integrate the shadow banking system. Second, models must be constructed to identify the economic inefficiencies in banking and the macroprudential policies required to correct those inefficiencies. Work on both these projects is well under way.

References

Bachmann, Rudiger, Moscarini, Giuseppe, 2012. Business Cycles and Endogenous Uncertainty. Manuscript.
Beattie, Alan, Politi, James, 2008. "'I made a mistake', admits Greenspan". Financial Times, October 23 2008. https://www.ft.com/content/aee9e3a2-a11f-11dd-82fd-000077b07658.
Bernanke, Ben S., 2010. Statement Before the Financial Crisis Inquiry Commission, Washington, D.C. https://www.federalreserve.gov/newsevents/testimony/bernanke20100902a.pdf.
Bernanke, Ben S., Gertler, Mark, Gilchrist, Simon, 1999. The Financial Accelerator in a Quantitative Business Cycle Framework. Handbook Macroecon. 1, 1341–1393.
Bloom, Nicholas, 2009. The Impact of Uncertainty Shocks. Econometrica 77 (3), 623–685.
Buera, Francisco J., Moll, Benjamin, 2015. Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity. Am. Econ. J. Macroecon. 7 (3), 1–42.
Chari, V.V., Kehoe, Patrick J., McGrattan, Ellen R., 2009. New Keynesian Models: Not Yet Useful for Policy Analysis. Am. Econ. J. Macroecon. 1(1), 242–266.
Christiano, Lawrence J., Motto, Roberto, Rostagno, Massimo, 2003. The Great Depression and the Friedman-Schwartz Hypothesis. J. Money Credit Banking 35 (6), 1119–1197.
Christiano, Lawrence J., Eichenbaum, Martin S., Evans, Charles L., 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. J. Polit. Econ. 113 (1), 1–45.
Christiano, Lawrence J., Ilut, Cosmin L., Motto, Roberto, Rostagno, Massimo, 2010. Monetary Policy and Stock Market Booms. In: Proceedings of the Conference in Jackson Hole, Macroeconomic Challenges: The Decade Ahead.
Christiano, Lawrence J., Motto, Roberto, Rostagno, Massimo, 2014. Risk Shocks. Am. Econ. Rev. 104 (1), 27–65.
Christiano, Lawrence J., Eichenbaum, Martin S., Trabandt, Mathias, August 2018. On DSGE Models. J. Econ. Perspect. 32 (3), 113–140.
Copaciu, Mihai, Nalban, Valeriu, Bulete, Cristian, 2012. R.E.M. 2.0, An estimated DSGE model for Romania. Dynare Working Paper 48. CEPREMAP November 2015.
Diamond, Douglas W., Dybvig, Philip H., 1983. Bank runs, deposit insurance, and liquidity. J. Polit. Econ. 91 (3), 401–419.
Di Tella, Sebastian, Hall, Robert E., 2020. Risk Premium Shocks Can Create Inefficient Recessions. NBER Working Paper No. 26721.
Ferreira, Thiago R.T., 2013. Financial Volatility and Economic Activity. Manuscript.
Ferreira, Thiago R.T., 2016. Financial Volatility and its Economic Effects. Working Paper. Federal Reserve Board, January 2016.
Gertler, Mark, Karadi, Peter, 2011. A Model of Unconventional Monetary Policy. J. Monetary Econ. 58 (1), 17–34.
Gertler, Mark, Kiyotaki, Nobuhiro, 2015. Banking, Liquidity, and Bank Runs in an Infinite Horizon Economy. Am. Econ. Rev. 105 (7), 2011–2043.
Fanelli, Sebastián, Straub, Ludwig, 2021. A Theory of Foreign Exchange Interventions. Rev. Econ. Stud. 88 (6), 2857–2885.
Gertler, Mark, Gilchrist, Simon, Natalucci, Fabio M., 2007. External Constraints on Monetary Policy and the Financial Accelerator. J. Money Credit Bank. 39, 295–330.
Gertler, Mark, Kiyotaki, Nobuhiro, Prestipino, Andrea, 2016. Wholesale Banking and Bank Runs in Macroeconomic Modeling of Financial Crises. Handbook Macroecon. 2, 1345–1425.
Gertler, Mark, Kiyotaki, Nobuhiro, Prestipino, Andrea, 2020. A Macroeconomic Model with Financial Panics. Rev. Econ. Stud. 87 (1), 240–288.
Gorton, Gary, 2010. B, Slapped by the Invisible Hand: The Panic of 2007. Oxford University Press.
Justiniano, Alejandro, Primiceri, Giorgio E., Tambalotti, Andrea, 2014. The Effects of the Saving and Banking Glut on the U.S. Economy. J. Int. Econ. 92, S52–S67.
Liu, Zheng, Wang, Pengfei, Zha, Tao, 2013. Land-Price Dynamics and Macroeconomic Fluctuations. Econometrica 81 (3), 1147–1184.
Negro, Marco Del, Schorfheide, Frank, 2013. DSGE Model-Based Forecasting. Handbook Econ. Forecast.
Pozsar, Zoltan, Adrian, Tobias, Ashcraft, Adam, Boesky, Hayley, 2013. Shadow Banking. FRBNY Econ. Policy Rev.
Shleifer, Andrei, Vishny, Robert, 2011. Fire Sales in Finance and Macroeconomics. J. Econ. Perspect. 25 (1), 29–48.
Simon Gilchrist, Zakrajsek, Egon, Acharya, Viral, Adrian, Tobias, Chen, Hui, Krishnamurthy, Arvind, Li, Haitao, Swanson, Eric, Wei, Min, 2012. Credit Spreads and Business Cycle Fluctuations1. Am. Econ. Rev. 102(4), 1692–1720.
Smets, Frank, Wouters, Rafael, 2007. Shocks and Firctions in US Business Cycle: A Bayesian DSGE Approach. Am. Econ. Rev. 97, 586–606.
Townsend, Robert M., 1979. Optimal Contracts and Competitive Markets with Costly State Verification. J. Econ. Theory 21 (2), 265–293.
Williamson, Stephen D., 1987. Financial Intermediation, Business Failures, and Real Business Cycles. J. Polit. Econ. 95 (6), 1196–1216.

[20] See also Di Tella and Hall (2020) for a discussion of the policy implications of a risk shock.