Booms and Banking Crises

Frédéric Boissay

Bank for International Settlements

Fabrice Collard

University of Bern

Frank Smets

European Central Bank and Katholieke Universiteit Leuven

Banking crises are rare events that break out in the midst of credit-intensive booms and bring about deep and long-lasting recessions. This paper presents a textbook dynamic stochastic general equilibrium model to explain these phenomena. The model features a nontrivial banking sector, where bank heterogeneity gives rise to an interbank market. Moral hazard and asymmetric information in this market may lead to sudden market freezes, banking crises, credit crunches, and severe "financial" recessions. Those recessions follow credit booms and are not necessarily triggered by large exogenous adverse shocks.

We are thankful to two anonymous referees and the editor for their very insightful comments. We benefited from discussions with F. Alvarez, F. Canova, L. Christiano, S. Claessens, R. Cooper, V. Curdìa, M. Darracq-Paries, H. Dellas, W. Den Hann, B. Diba, F. Dufourt, H. Degryse, M. Ellison, A. Ellul, X. Freixas, D. Garcia-Macia, M. Gertler, S. Gilchrist, J. Gomes, P.-O. Gourinchas, J. Henry, I. Jaccard, P. Jacquinot, P. Karadi, N. Kiyotaki, R. Kollmann, L. Laeven, G. Lombardo, K. Nikolov, G. Nuño, J. C. Rochet, G. Saint-Paul, H.-S. Shin, J. Stein, G. Ströbl, J. Suarez, M. Trabandt, H. Uhlig, P. Weil, as well as participants at various workshops and seminars. Frank Smets acknowledges financial support from the EU FP7-Project MACFINROBODS, grant 612796. Disclaimer: The views expressed in this paper are our own and should not be interpreted as reflecting the views of the Bank for International Settlements, the European Central Bank, or the Eurosystem. Data are provided as supplementary material online.

Electronically published March 4, 2016
[*Journal of Political Economy*, 2016, vol. 124, no. 2]

I. Introduction

Recent empirical research on banking crises has highlighted the existence of common patterns across diverse episodes. Banking crises are rare events. They come along recessions that are deeper and last longer than other recessions. And, more importantly for the purpose of this paper, banking crises follow credit-intensive booms; they are "credit booms gone wrong."[1] In the recent macroeconomic literature, banking crises mostly result from the propagation and the amplification of random adverse financial shocks. Rare, large enough, adverse financial shocks can account for the first two properties (see, e.g., Gertler and Kiyotaki 2010). However, by implying that banking crises may break out at any time in the business cycle, they do not seem in line with the fact that the occurrence of a banking crisis is closely linked to credit conditions (Gorton 2010, 2012). The third stylized fact therefore calls for an alternative approach.

In this paper, banking crises result from the procyclicality of bank balance sheets that emanates from interbank market funding. During expansions, bank market funding and credit supply increase, pushing down the rates of return on corporate and interbank loans. The lower rates aggravate agency problems in the interbank market, which lead to a reduction in market funding and contractions. The larger the credit boom relative to the possibilities for productive use of loans, the larger the fall in interest rates, and the higher the probability of a bank run in—and therefore of a disastrous freeze of—the interbank market.[2] As in Shin (2009) and Hahm, Shin, and Shin (2011), the behavior of banks (credit in our case) during good times sows the seeds of the next banking crisis.

In our model, banks are heterogeneous in terms of—non–publicly observed—intermediation efficiency. They finance their activities with funds obtained from depositors/shareholders or raised in the interbank market. There exists the usual agency problem in this market as borrowing banks can always divert some of the funds toward low return assets that cannot be recovered by the lending banks. The incentives for diver-

[1] See Borio and Drehmann (2009) and Schularick and Taylor (2012); the notion that banking crises are endogenous and follow prosperous times is also present in Minsky (1977).

[2] Our representation of banking crises as market-based bank runs is consistent with what happened during the 2007–8 financial crisis (see Uhlig 2010). Shin (2010, chap. 8), e.g., depicts the demise of Northern Rock—a UK bank—in 2007 as primarily originating from the sudden freeze of the short-term funding market, what he refers to as a "modern bank run." A traditional, deposit-based, run on the bank took place as well, but it did so 1 month later, accounted for only 10 percent of the bank's fall in total funding, and rapidly stopped because, following the news of the run, the UK authorities pledged 100 percent deposit guarantees.

sion are stronger for less productive banks and depend on the level of interest rates in the economy. The lower the return on corporate loans, the greater the incentive to engage in fund diversion and hence the greater counterparty risk in the interbank market. The typical run of events leading to a banking crisis is as follows. A sequence of favorable, nonpermanent, supply shocks hits the economy. The resulting increase in productivity leads to an expansion of credit. The more efficient banks expand their corporate loan operations by drawing funds from the less efficient banks and the size of the banking sector as a whole increases. The economy booms. But as supply shocks run their course, the probability of imminent reversion to average productivity increases. This slows down corporate demand for loans while at the same time inducing households to accumulate savings in order to smooth consumption. As this saving glut develops, interest rates go down. The rate of return on interbank loans declines, making the less efficient banks more prone to borrow and divert those funds. The identity of these banks being unknown to lenders, counterparty risk in the interbank market increases, interbank loans decline, and market finance recedes. We show that there is a threshold value of interest rates below which the interbank market freezes, corporate credit collapses, and the economy tanks. This threshold can be alternatively expressed in terms of the level of banking assets relative to the level of productivity (output) in the economy, which we call the absorption capacity of the banking sector. Excessive credit creation shifts the economy beyond its absorption capacity, triggering a banking crisis.

Interbank market freezes in our model are reminiscent of Mankiw's (1986) credit market collapses. In Mankiw's analysis, changes in the interest rate also alter the composition and the quality of the pool of borrowers. One important difference though is that, in our case, banks can be on either side of the interbank loan market and choose which side they stand on. So, following an increase in the interest rate, low-quality banks switch to the supply side and the average quality of the remaining borrowers improves, whereas in a typical adverse selection model like Mankiw's, quality would deteriorate. In this sense, our model features a "positive" selection of borrowers, which explains why low—not high—interest rates are conducive to market freezes. These mechanisms are also present in Boissay (2011), but his partial equilibrium approach still misses the link between credit conditions and banking crises. By embedding the positive borrower selection mechanism into a general equilibrium model, where low interest rates may be induced by a—supply-driven—boom in credit, we are able to account for why credit booms may lead to banking crises.

The present paper is related to the literature that introduces financial frictions in macroeconomic models. For instance, Bernanke, Gertler, and Gilchrist (1999), Gertler and Karadi (2011), Jermann and Quadrini (2012), and Christiano, Motto, and Rostagno (2014) show how adverse wealth or financial shocks can be amplified by financial market frictions and generate deep and long-lasting recessions. We depart from these approaches in that we do not analyze a linearized version of the model but instead characterize the full equilibrium dynamics, inclusive of important and critical nonlinearities such as the freezing of interbank markets. This is an important difference because, near the steady state, our model features the traditional financial accelerator that the previous models possess. But, away from it, our model may exhibit banking crises and severe recessions. Moreover, large departures from the steady state are the endogenous outcome of a boom-bust endogenous cycle rather than a big shock.

In this respect, our approach shares similarities with the work of Brunnermeier and Sannikov (2012) and He and Krishnamurthy (2012), which also features powerful nonlinear amplification mechanisms due to financial frictions, as well as with the literature on sudden stops in emerging market economies. For example, as in Mendoza and Smith (2005) and Mendoza (2010), we find that when agents are highly leveraged, then even standard (mild) negative productivity shocks may trigger crises (in their case, capital flow reversals) and severe recessions. One important difference with sudden stops models, though, is that they typically consider small open economies, in which the interest rate is exogenous, whereas in our case the endogenous dynamics of interest rates play a central role.

Our paper is also related to the literature that investigates the role of pecuniary externalities in generating excessive financial fragility, such as Bianchi and Mendoza (2010) or Bianchi (2011). In these papers, financial frictions affect nonfinancial firms and entail excess credit demand (i.e., “overborrowing”). In our case, pecuniary externalities arise because the household does not internalize the effects of her saving decisions on the interest rate, which leads to an excess supply of credit.

Finally, our model departs from the existing literature in that banking crises are not due to occasionally binding collateral constraints (see, among others, Bianchi and Mendoza [2010], Brunnermeier and Sannikov [2012], and Perri and Quadrini [2011]), but to occasional financial market runs.[3]

[3] Perri and Quadrini (2011) recently proposed a model that combines collateral constraints and self-fulfilling expectations, where large recessions may also follow credit booms even in the absence of fundamental shocks. In their model, whether or not the constraint binds emerges endogenously as multiple self-fulfilling equilibria. Gertler and Kiyotaki (2013) also develop a model in which bank runs induce regime switches. One important difference is that in their model, bank runs are deposit based and unanticipated, while

We calibrate the model on post–World War II US data and the financial cycles in 14 OECD countries (1870–2008) and assess its quantitative properties. Most of the time, bank assets remain below the threshold for banking crises and the model behaves like a standard financial accelerator model. But once in a while—on average, every 40 years—a banking crisis bursts. Potentially, our model can generate banking crises following an exogenous, large, and negative productivity shock around the steady state; following an endogenous overaccumulation of assets by the household and excess credit supply by banks; or following a combination of both exogenous and endogenous factors. Simulations of the model indicate that the typical crisis is triggered by a moderate negative productivity shock toward the end of an unusual boom in credit. Thus, as in the data, banking crises in our model are closely linked to credit conditions. They also bring about more severe and longer-lasting recessions.

The paper proceeds as follows. Section II documents stylized facts about financial recessions in 14 OECD countries for the period 1870–2008. Section III describes our theoretical framework, the microfoundations of interbank market freezes, and the dynamic implications of such events. In Section IV we solve a calibrated version of the model and describe the typical dynamics to banking crises. Section V discusses the ability of the model to account for the stylized facts and the sensitivity of the results to alternative parameterizations. Section VI evaluates the welfare costs associated with the existence of financial frictions. Section VII presents conclusions.

II. Facts on Financial Recessions

The recent empirical literature has put forward a set of facts that seem to indicate that financial recessions—recessions that are concomitant with banking crises—are special events that significantly differ from other types of recessions. Reinhart and Rogoff (2009, 2013) report that banking crises have historically been followed by profound declines in output and employment, with output falling by about 9 percent from peak to trough and the duration of the downturn averaging roughly 2 years. This is much more than during ordinary recessions, possibly because financial recessions come along with a credit crunch that amplifies the downturn, as Claessens, Kose, and Terrones (2008, 2011) show. Mendoza and Terrones (2012) and Schularick and Taylor (2012), among others, also

in ours they are market based and agents are fully rational: they perfectly know and take into account the probability that runs will occur in the future. Angeloni and Faia (2013) also propose a model with bank runs, but their runs are idiosyncratic and frequent, while ours are systemic and rare.

provide evidence that banking crises follow credit booms. For instance, Schularick and Taylor show that high credit growth is a robust predictor of future banking crises in the case of advanced economies. The predictive power of credit may not necessarily reflect a causal relationship, as credit is strongly correlated with other financial phenomena that have also been found to precede banking crises such as asset price bubbles or large capital inflows (Kaminski and Reinhart 1999; Reinhart and Rogoff 2009, 2013). However, these other factors seem to be more relevant for emerging market economies than for advanced countries (Reinhart and Reinhart 2008) where credit remains a strong predictor even once past current account deficits, asset prices, or GDP growth are controlled for (Jordà, Schularick, and Taylor 2011, 2013; Schularick and Taylor 2012). The different evolution of GDP and credit gaps around financial and nonfinancial recessions, as shown in figure 1, illustrate these facts, which we document in more detail below using the data compiled by Jordà et al. and Schularick and Taylor.

The data set comprises yearly observations from 1870 to 2008 in 14 developed countries for real GDP per capita, total domestic currency loans of banks and banking institutions to nonfinancial companies and households, the dates of business cycle peaks, and the dates of banking crises. Following Laeven and Valencia (2008) and Reinhart and Rogoff (2013), a banking crisis in the data refers to an event during which the financial sector experiences bank runs, sharp increases in default rates accompanied by large capital losses that result in public intervention, bankruptcy, or the forced merger of major financial institutions. As is common in the

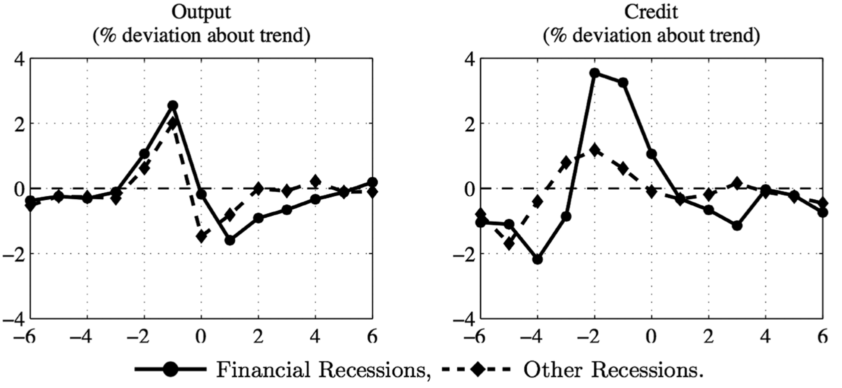


Fig. 1.—Dynamics of GDP and credit gaps around recessions. Average dynamics of the Hodrick-Prescott ($\lambda = 6.25$) cyclical component of (log) output and credit 6 years before and after the beginning of a recession (period 0) with and without a banking crisis. The data are from Jordà et al. (2011, 2013) and Schularick and Taylor (2012).

literature, a recession qualifies as a financial recession when a banking crisis breaks out between the peak year that precedes it and the trough year that ends it (see Jordà et al. 2013). Excluding wartimes, the data set includes 78 banking crises for a total of 1,736 country/year observations and includes 161 complete business cycles (i.e., periods between two business cycle peaks). We use these data and document three key facts on financial recessions.

Fact 1: Financial recessions are rare events. Among the 196 recessions identified in the sample, 41 are associated with a banking crisis, which corresponds to an average probability of a financial recession of 2.36 percent. Hence, financial recessions are rare events; by comparison, the probability of other recessions is 8.93 percent (see table 1).

Fact 2: Financial recessions are deeper and last longer than other recessions. In our sample, financial recessions last, on average, 2.32 years, about 8 months (40 percent) longer than other recessions. During these episodes, real GDP per capita falls almost twice as much, from peak to trough, as during other recessions (6.84 percent against 3.75 percent). These differences are statistically significant at the 1 percent level, and importantly, they are unlikely to be due to severe recessions causing non-

TABLE 1
Statistics on Recessions

	Financial	Other	All	Severe	Mild
Country/year observations	1,736	1,736	1,736	1,333	1,333
Events	41	155	196	54	55
Frequency (%)	2.36	8.93	11.29	4.05	4.13
Duration (years)	2.32	1.65	1.79	2.46	1.25
Magnitude ($\Delta_{p,t}y$, %)	−6.84	−3.75	−4.40	−9.28	−.89
Credit crunch:					
$\Delta_{p,t}k^{\mathrm{HP}}$ (%)	−3.80*	−2.06	−2.43	−1.48	−2.60
$\Delta_{p,p+2}k^{\mathrm{HP}}$ (%)	−3.45*	−1.22	−1.72	−1.68	−2.37
Credit boom:					
$\Delta_{p-2,p}k^{\mathrm{HP}}$ (%)	4.55***	.18	1.13	1.35	.41
k_p^{HP} (%)	3.25***	.61	1.16	1.85	.95

Note.—The entry $\Delta_{p,t}x$ (respectively, $\Delta_{p,p+2}x$, $\Delta_{p-2,p}x$) denotes the percentage change of variable x from peak to trough (respectively, from peak to peak +2 years and peak −2 years to peak), where x denotes either GDP per capita (y) or the HP ($\lambda = 6.25$) cyclical component of credit per capita (k^{HP}). A severe (mild) recession is a recession associated with a peak-to-trough output loss in the top (bottom) three deciles of the distribution. All statistics are averages over the full sample of 1,736 country/year observations, except the peak-to-trough statistics, which are averages over the 161 complete business cycles only (1,333 country/year observations).

* The difference between financial (severe) recessions and other (mild) recessions is statistically significant at the 10 percent threshold.

** The difference between financial (severe) recessions and other (mild) recessions is statistically significant at the 5 percent threshold.

*** The difference between financial (severe) recessions and other (mild) recessions is statistically significant at the 1 percent threshold.

performing loans, bank capital losses, and ultimately banking crises. Indeed, should it be the case, we would observe that banking crises tend to break out after recessions start, once bank capital losses have accumulated. In our sample, we find the opposite: in 32 out of 41 financial recessions, the banking crisis breaks out either in the year before or in the very first year of the recession. Table 1 also shows that the credit gap falls by more during financial recessions (−3.80 percent) than during other recessions (−2.06 percent). We find similar differences with the growth rate of credit (not reported); during financial recessions, credit hardly increases, while during other recessions, it goes up by 4.16 percent (the difference is statistically significant at the 10 percent threshold). Together, these results suggest not only that financial recessions are accompanied with a credit crunch but also that, in the case of financial recessions, it is the severity of the recession that results from financial distress, and not the other way around.

Fact 3: Financial recessions follow credit booms. Banking crises and, a fortiori, financial recessions do not hit at random (Gorton 1988) but instead break out in the midst of a credit boom (Schularick and Taylor 2012). The bottom of table 1 illustrates this point. It shows that credit grows significantly faster before financial recessions than before other recessions (4.55 percent against 0.18 percent). Credit is also 3.25 percent above its trend in the peak year that precedes a financial recession, whereas it is only 0.61 percent above trend in the case of other recessions. All these differences are statistically significant at the 1 percent threshold. By contrast, there is no such difference between mild and severe recessions (last two columns of table 1). In other words, only financial recessions are associated with credit booms, which suggests that a country's recent history of credit growth helps predict a financial recession.

To test this idea formally we build on Schularick and Taylor (2012) and estimate the probability that a recession starts in a given country in a given year as a function of the country's lagged real credit growth and check whether credit indeed matters. To control for potential omitted variables and reverse causality, we also include five lags of real GDP growth in the model. The results are reported in table 2. For comparison, the first column replicates Schularick and Taylor's logit model for banking crises; the other three columns correspond to our model for all types of recessions, as well as for financial and severe recessions separately. The results essentially confirm the statistics in table 1. The sum of the coefficients on lagged credit growth is positive and significant in the case of financial recessions only. Credit helps predict financial recessions, but it is of no use to predict other types of recessions, be they severe or not. We conclude that financial recessions are indeed special, that they are credit booms gone bust. This result is important because

TABLE 2
PREDICTION OF RECESSIONS

	BANKING CRISIS: Schularick and Taylor Model 15	RECESSION		
		All	Financial	Severe
Sum of lag credit coefficients	19.32***	.21	15.39***	7.49
Observations	1,272	1,272	1,272	880
Countries	14	14	14	12
Pseudo R^2	.123	.053	.127	.152

NOTE.—All models include country fixed effects and the lagged growth rates of real credit and real GDP per capita (the coefficients are reported in table 1 in sec. C.1 of the supplementary appendix). The first column is a replication of Schularick and Taylor (2012, table 6, model 15); the dependent variable is a dummy equal to one in the country/year where a banking crisis breaks out. In the other models, the dependent variable is a dummy equal to one in the country/year where a recession starts. In the last column, as in table 1, a severe recession is a recession associated with a peak-to-trough growth rate of real GDP per capita in the bottom three deciles of the distribution; Denmark and Germany are excluded because of the fixed effects.

* Statistically significant at the 10 percent threshold.
** Statistically significant at the 5 percent threshold.
*** Statistically significant at the 1 percent threshold.

it establishes the role of credit as an endogenous source of financial instability, and not merely as an amplifier of exogenous shocks.

III. The Model

We consider a closed economy populated with one representative risk-averse household, one representative risk-neutral competitive firm, and a mass one of heterogeneous, risk-neutral, and competitive banks.

A. *The Representative Firm*

The representative firm lives for one period. It produces a homogeneous good that can be either consumed or invested, by means of capital, k_t, and labor, h_t, according to a constant returns to scale technology represented by a production function $z_tF(k_t, h_t)$ that satisfies standard Inada conditions. The level of total factor productivity (TFP), z_t, is assumed to follow an exogenous AR(1) process of the form

$$\log z_t = \rho_z \log z_{t-1} + \varepsilon_t, \tag{1}$$

where $|\rho_z| < 1$ and ε_t is a normally distributed innovation with zero mean and standard deviation σ_z. Variations in productivity are the only source of aggregate uncertainty, and ε_t is realized at the beginning of period t, before the firm decides on its production plan. Capital, k_t, depreciates at rate $\delta \in (0, 1)$. The firm is born with no resources and must borrow k_t from the banks at a gross corporate loan rate R_t at the beginning of the period. The corporate loan is repaid at the end of the period. The firm also rents labor services from the household at rate w_t. The production plan is decided so as to maximize profit,

$$\pi_t = z_t F(k_t, h_t) + (1-\delta)k_t - R_t k_t - w_t h_t. \tag{2}$$

B. *The Representative Household*

The economy is populated by an infinitely lived representative household that has preferences over consumption, c_t, represented by the utility function

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} u(c_{t+\tau}), \tag{3}$$

where $u(c_t)$ satisfies the usual regularity conditions (i.e., $u'(c) > 0$, $u''(c) < 0$, $u'(0) = \infty$, $u'(\infty) = 0$), $\beta \in (0, 1)$ is the psychological discount factor, and $\mathbb{E}_t(\cdot)$ denotes the expectation operator, which is taken over $\{\varepsilon_{t+\tau+1}\}_{\tau=0}^{+\infty}$. For expositional convenience, we assume that the household supplies inelastically one unit of labor; we will relax this assumption in the quantitative analysis of the model (see Sec. IV.A). The household enters period t with individual assets a_t (hereafter A_t will denote aggregate assets). Since there is no friction between the household and the banks, the composition of a_t is indeterminate, and a_t can be thought of either as bank deposit or as bank equity. To ease exposition, we will refer to it as bank deposit. Likewise, the associated gross return, r_t, will be referred to as the gross return on deposits. Following the macrofinance literature (e.g., Gertler and Kiyotaki 2010; Gertler and Karadi 2011), we assume that there exist frictions between the household and the firm that prevent the household from financing the firm directly. The household earns unit wage w_t from supplying her labor, receives profits π_t from the firm, and gets a lump-sum transfer χ_t that corresponds to banks' financial intermediation costs (see the next section). She therefore decides on her consumption/saving decisions maximizing her utility (3) subject to the budget constraint

$$c_t + a_{t+1} = r_t a_t + w_t + \pi_t + \chi_t. \tag{4}$$

C. *The Banking Sector*

The banking sector is at the core of the model and plays a nontrivial role because of two specific features. First, banks are heterogeneous with respect to their intermediation technology: some banks are more efficient than others, which potentially gives rise to an interbank market. Second, the banking sector is subject to both asymmetric information and moral hazard problems, which impair the functioning of the interbank market. Overall, banks engage in two types of activities. They perform traditional retail banking operations, which consist in collecting deposits from the household and in lending the funds to the firm. Banks also perform wholesale banking operations by issuing interbank claims, through which they reallocate assets among themselves—typically from the least efficient to the most efficient banks.

1. Banks

There is a continuum of one-period, risk-neutral, competitive banks indexed by p that each raise deposits a_t from the household at the end of period $t-1$. The banks that operate in period t are born at the end of period $t-1$ and die at the end of period t.[4] When they are born, banks are all identical and raise the same amount of deposits. They then draw a random, bank-specific, intermediation skill and become heterogeneous; by convention, bank p draws skill p. The p's are distributed over the interval [0, 1] with cumulative distribution $\mu(p)$, satisfying $\mu(0) = 0$, $\mu(1) = 1$, and $\mu'(p) > 0$. Bank p must pay an intermediation cost $(1-p)R_t$ per unit of corporate loan at the end of the period, so that the effective gross return of the loan is pR_t. These costs are rebated to the household in the form of lump-sum transfers that amount to χ_t, so that there are no deadweight losses in the economy. These intermediation costs reflect the prospection, screening, and monitoring costs that banks may face either as loan originators or as loan servicers.[5]

As an outside option, banks have the possibility to invest into a project that yields a constant gross return γ, and that is worth more than just letting the good depreciate, that is, $\gamma \geq 1-\delta$. To fix ideas, we will refer to this outside option as "storage," although it could lend itself to alternative interpretations, for example, in terms of subprime lending (see

[4] We will assume in a moment that banks are heterogeneous and that their types are private information. The assumption of one-period living banks is made to preserve this asymmetry of information over time. An alternative—but equivalent—approach would be to allow banks to live infinitely and, in order to rule out potential reputation effects, to assume that the types are randomly drawn afresh every period.

[5] The way we model these costs is not critical, in the sense that we would find qualitatively similar results (i) if they were dead weights (see Boissay, Collard, and Smets 2013), (ii) if they were paid at the beginning of the period, (iii) if p were the fraction of the loan that the firm repays to bank p, or (iv) if p were the probability that the firm reimburses bank p (see Boissay 2011). We do not introduce these features in the model so as to stay the closest possible to the textbook neoclassical model, where all deposits are channeled to firms and firms do not default.

sec. D.2 in the supplementary appendix available online). As will become clear later, this storage technology will play an important role in the working of the banking sector, but the assumption that the return on the storage technology, γ, is constant will not be critical (see fn. 7).

Bank heterogeneity gives rise to an intraperiodic interbank market, where the least efficient banks lend to the most efficient ones at gross rate ρ_t. In equilibrium this rate must be lower than the corporate loan rate, R_t; otherwise no bank would lend to the firm and the Inada conditions on the firm's production function would imply that R_t is infinite. Similarly, the interbank rate ρ_t must be bigger than the return on storage, γ; otherwise no bank would lend to other banks. It follows that in equilibrium it must be that $\gamma < R_t$, which means that storage is an inefficient technology. Banks take the rates ρ_t and R_t as given. Given these rates, bank p decides whether, and how much, it borrows or lends. By convention, we will call the banks that supply funds on the interbank market the "lenders" and those that borrow the "borrowers." Let ϕ_t be the—endogenous and publicly observable—amount borrowed on the interbank market per unit of deposit by a borrower p, with $\phi_t \geq 0$. Since ϕ_t is the ratio of market funding to traditional funding, we will refer to it as banks' "market funding ratio."[6] Bank p's gross unit return on deposits is equal to $pR_t(1+\phi_t) - \rho_t\phi_t$ when it borrows ϕ_t (per unit of deposit) from other banks at cost ρ_t and lends $1+\phi_t$ (per unit of deposit) to the firm for return pR_t. It is equal to ρ_t when, instead, bank p decides to lend to other banks. Denoting the gross return on deposits by $r_t(p)$, one therefore gets

$$r_t(p) \equiv \max\{pR_t(1+\phi_t) - \rho_t\phi_t, \rho_t\}, \tag{5}$$

and bank p chooses to be a borrower when

$$pR_t(1+\phi_t) - \rho_t\phi_t \geq \rho_t \Leftrightarrow p \geq \overline{p}_t \equiv \frac{\rho_t}{R_t}. \tag{6}$$

Inequality (6) is the participation constraint of bank p to the interbank market as a borrower rather than as a lender. It pins down the type of the marginal bank $\overline{p}_t$ that is indifferent between the two options. Banks with $p < \overline{p}_t$ delegate financial intermediation to more efficient banks with $p \geq \overline{p}_t$. In a frictionless world, all banks with $p < 1$ would lend to the most efficient bank, so that $\overline{p}_t = 1$. The bank $p = 1$ would then have an infinite market funding ratio ($\phi_t \to +\infty$) and the economy would reach the first-best allocation. Two frictions on the interbank market prevent the econ-

[6] Note that ϕ_t cannot be interpreted as a leverage ratio because (i) market funding is only one of the two components of bank liabilities—the other one being deposits—and (ii) the bank liability structure is indeterminate.

omy from reaching the first-best allocation: moral hazard and asymmetric information.

2. Moral Hazard

We assume that the proceeds of the storage technology are not traceable and cannot be seized by creditors. As a consequence, interbank loan contracts are not enforceable and banks may renege on their interbank debt by walking away from the lenders. A bank that walks away with $(1 + \phi_t)a_t$ and invests in the storage technology gets $\gamma(1 + \theta\phi_t)a_t$ as payoff, where $\theta \in [0, 1]$ captures the cost of walking away from the debt (the higher θ, the lower this cost). Following the current practice (e.g., Hart 1995; Burkart and Ellingsen 2004), we refer to such an opportunistic behavior as "diversion." From a corporate finance literature viewpoint (e.g., Tirole 2006), this is a standard moral hazard problem: (i) the gain from diversion increases with ϕ_t, and the opportunity cost of diversion increases (ii) with bank efficiency p and (iii) with the spread between the corporate loan rate (R_t) and the return on storage (γ). Points i and ii imply that efficient banks with "skin in the game" are less inclined to walk away than highly leveraged and inefficient banks. Point iii is similar to point ii, but in the "time-series" (as opposed to the "cross-sectional") dimension. This last feature, which implies that banks are more inclined to walk away when the return on corporate loans goes down, is consistent with recent empirical evidence that banks tend to take more risks when interest rates are low (see, e.g., Maddaloni and Peydró 2011; Jiménez et al. 2014).[7]

By limiting the borrowing capacity of the most efficient bank ($p = 1$), moral hazard will give less efficient banks room to borrow and generate a positive wedge between the corporate loan rate (R_t) and the deposit rate (r_t). But this friction alone will not be enough to induce market freezes; for this, uncertainty about the quality—and therefore some selection—of borrowers must also be assumed (see Boissay 2011).

3. Asymmetric Information

Banks' intermediation skills (the p's) are privately known, and lenders can neither observe them ex ante nor verify them ex post. Lenders therefore ignore borrowers' private incentives to divert funds. As a result, the

[7] Note that this feature does not require that the return on storage be constant. Besides, our assumption that the return on storage is constant is not critical in our model. What is key, though, is that this return is less sensitive to the business cycle than the corporate loan rate. Should γ be strictly proportional to R_t, then any fall in R_t would leave the spread (ratio) between R_t and γ unchanged, and incentives would not be affected. In sec. D.3 of the supplementary appendix, we show that our results still hold when the return on storage is allowed to comove positively—but less than one-for-one—with the corporate loan rate.

loan contracts signed on the interbank market are the same for all banks and neither the market funding ratio (ϕ_t) nor the interbank rate (ρ_t) depends on p.[8]

Lenders want to deter borrowers from diverting. They can do so by limiting the quantity of funds that borrowers can borrow so that even the most inefficient banks with $p < \overline{p}_t$—those that should be lending—have no interest in demanding a loan and diverting it:

$$\gamma(1 + \theta\phi_t) \leq \rho_t. \tag{7}$$

This incentive compatibility constraint sets a limit for ϕ_t, which we interpret as lenders' funding tolerance, that is, the market funding ratio above which a bank refuses to lend. The program of bank $p \geq \overline{p}_t$ thus consists in maximizing its return on deposits $r_t(p)$ (see eq. [5]) with respect to market funding ϕ_t, subject to constraint (7). Proposition 1 follows; the proofs of the propositions are reported in the Appendix.

Proposition 1 (Optimal market funding ratio). At the optimum, the incentive compatibility constraint binds and borrowing banks exhaust their borrowing capacity. This borrowing capacity, $\phi_t = (\rho_t - \gamma)/\gamma\theta$, increases with ρ_t.

The positive relationship between market funding (ϕ_t) and the market funding cost (ρ_t) is a critical feature of the model. When the interbank rate ρ_t increases, banks have more incentives to lend, and only those with a high p keep demanding a loan. Since private incentives to divert are lower for high-p banks, lenders tolerate a higher market funding ratio (ϕ_t goes up). In this sense, a rise in the interbank rate has a positive selection effect on borrowers. Symmetrically, a fall in this rate has a detrimental effect on incentives. In the limit case in which $\rho_t = \gamma$, no borrower can commit herself to repay, and demand is null ($\phi_t = 0$).

4. Interbank Market

The equilibrium of the interbank market is characterized by the gross interbank rate ρ_t that clears the market. We look for an equilibrium in which the interbank rate exceeds the return on storage ($\rho_t > \gamma$) so that trade takes place ($\phi_t > 0$). Since a mass $\mu(\overline{p}_t)$ of banks lend and the com-

[8] To see this, consider a menu of debt contracts $\{\rho_t(p), \phi_t(\tilde{p})\}_{\tilde{p}\in[0,1]}$ intended for the borrowers of type $\tilde{p}$'s, and notice that lenders' arbitrage across these contracts requires that $\rho_t(\tilde{p}) = \rho_t$ for all $\tilde{p} \in [0, 1]$. It is easy to see that such a menu of contracts cannot be revealing because any borrower p (i.e., with $pR_t > \rho_t$) claiming being of type $\tilde{p}$ would make profit $r_t(\tilde{p}|p) = pR_t + (pR_t - \rho_t)\phi_t(\tilde{p})$ and pick the contract with the highest $\phi_t(\tilde{p})$, independent of its type. It is equally easy to see that there is no revealing menu of equity contracts either. Indeed, consider a menu of equity contracts $\{\eta_t(\tilde{p}), \phi_t(\tilde{p})\}_{\tilde{p}\in[0,1]}$, where $\eta_t(\tilde{p})$ would be the share of retained earnings. Then the net profit of bank p would be $\eta_t(\tilde{p})[1 + \phi_t(\tilde{p})]pR_t$, and in equilibrium, this bank would pick the contract that yields the highest $\eta_t(\tilde{p})[1 + \phi_t(\tilde{p})]$, independently of its own p.

plement mass $1 - \mu(\overline{p}_t)$ of banks borrow ϕ_t per unit of deposit, the market clears when (using relation [6], proposition 1, and the fact that $a_t = A_t$ in the general equilibrium)

$$\underbrace{A_t\mu\left(\frac{\rho_t}{R_t}\right)}_{\text{Supply}} = \underbrace{\overbrace{A_t\left[1 - \mu\left(\frac{\rho_t}{R_t}\right)\right]}^{\text{Extensive Margin}} \quad \overbrace{\frac{\rho_t - \gamma}{\gamma\theta}}^{\text{Intensive Margin}}}_{\text{Demand}}$$

$$\Leftrightarrow R_t = \Psi(\rho_t) \equiv \frac{\rho_t}{\mu^{-1}\left(\dfrac{\rho_t - \gamma}{\rho_t - \gamma(1-\theta)}\right)}. \tag{8}$$

The aggregate supply increases monotonically with the interbank rate, ρ_t, while aggregate demand is driven by two opposite forces. On the one hand, aggregate demand decreases with the interbank rate because fewer borrowers demand funds when this rate goes up; this is the "extensive margin" effect. On the other hand, a rise in the interbank rate also has a positive effect on aggregate demand because, following the rise, each borrower is able to borrow more; this is the "intensive margin" effect. At the aggregate level, this latter effect more than offsets the extensive margin effect when there are many borrowers, that is, when ρ_t is small enough. It follows that the aggregate demand curve bends backward, increasing with the interbank rate ρ_t for small values of ρ_t (see fig. 2).

One can check that the function $\Psi(\rho_t)$ is strictly convex, goes to infinity as ρ_t approaches γ, is greater than R_t when ρ_t approaches R_t, and reaches a minimum for some value $\rho_t = \overline{\rho} > \gamma$. Hence there exists a threshold $\overline{R} \equiv \Psi(\overline{\rho})$ for the corporate loan rate R_t below which there is no equilibrium with trade and above which there are two equilibria with trade. This threshold is the minimum corporate loan rate that is necessary for banks to accept to lend to each other in equilibrium. Figure 2 illustrates this point and depicts the shifts in aggregate supply and demand as R_t falls from R_H (equilibrium E) to R_L (equilibrium C), with $R_L < \overline{R} < R_H$.

Following the fall in the corporate loan rate, the supply curve shifts to the right while the demand curve shifts to the left. Given the initial equilibrium rate $\rho_t = \rho_E$, demand falls below supply. Market clearing then requires a decrease in the interbank rate, which results in more banks demanding funds (extensive margin effect). But since the banks that switch from the supply to the demand side are less efficient and have a relatively higher private incentive to divert cash, lenders require borrowers to reduce their demand for loan. By construction, this intensive margin effect is the strongest when $R_t < \overline{R}$. It follows that aggregate demand decreases and excess supply goes further up. This process feeds itself and goes on until the market freezes, at point C, where by convention we set $\rho_t = \gamma$.

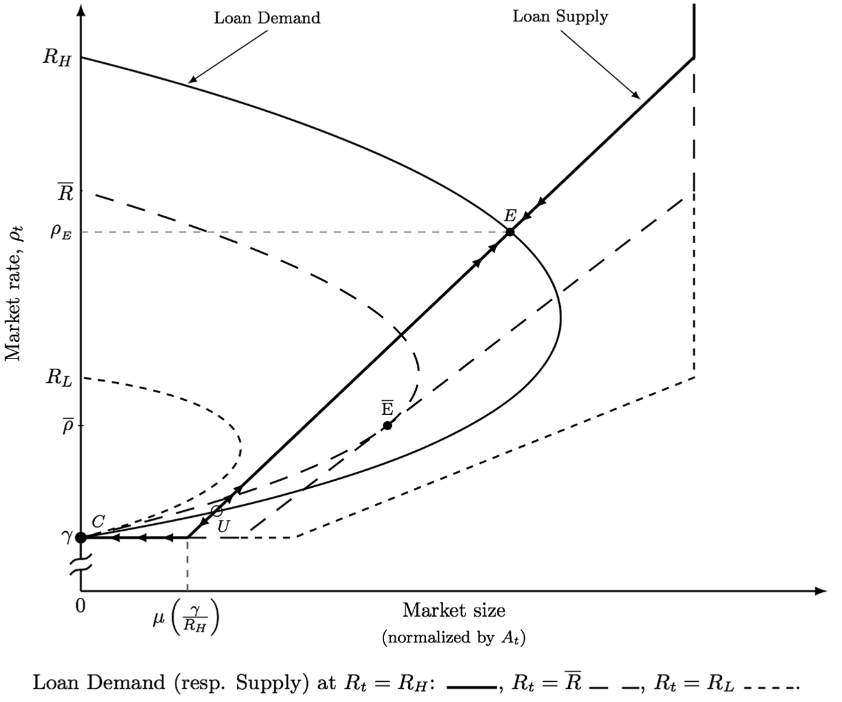


FIG. 2.—Interbank market clearing

When $R_t \geq \overline{R}$, the aggregate supply and demand curves intersect at points U and E. In this case it seems reasonable to restrict attention to E and to rule out U as not tatônnement stable (see Mas-Colell, Whinston, and Green 1995, chap. 17.H). At this point net aggregate demand is a decreasing function of the interbank rate, and following any small perturbation to ρ_t, a standard price tatônnement process would push the economy away from U. Similar reasoning shows that point E is a stable equilibrium. Since in this equilibrium trade takes place, we will refer to E as *normal times.*

In equilibrium C, in contrast, aggregate loan demand and supply are both equal to zero, and no trade takes place. In this case, there exists a cutoff $\overline{p}_t = \gamma/R_t$ such that banks with $p < \overline{p}_t$ store, banks with $p > \overline{p}_t$ are better off lending to the firm, and bank $\overline{p}_t$ is indifferent. We will refer to such no-trade equilibrium as a *banking crisis.*[9] This equilibrium always ex-

[9] The shift from E to C is akin to a regime switch, in which the economy experiences a large drop in output. Regime switching can be thought of as a way to endogenize rare exogenous shocks, or disaster risk (e.g., see Gourio 2012), or to generate nonlinear dynamics akin to those obtained in catastrophe theories (e.g., see Varian [1979] for an introduction).

ists, irrespective of R_t, is stable, and therefore coexists with E whenever E exists. Because it is Pareto-dominated by E, though, we rule it out by assuming that banks always coordinate on E.[10]

To complete the description of the banking sector we derive the sector's return on deposit using relations (5) and (8):

$$r_t \equiv \int_0^1 r_t(p)\mathrm{d}\mu(p)$$
$$= \begin{cases} R_t \displaystyle\int_{\overline{p}_t}^1 p\, \frac{\mathrm{d}\mu(p)}{1-\mu(\overline{p}_t)} & \text{if an equilibrium with trade exists} \\ R_t\left[\dfrac{\gamma}{R_t}\mu\left(\dfrac{\gamma}{R_t}\right) + \displaystyle\int_{\gamma/R_t}^1 p\,\mathrm{d}\mu(p)\right] & \text{otherwise.} \end{cases} \tag{9}$$

Relation (9) reflects the fact that, in normal times, the least efficient banks delegate financial intermediation to $1 - \mu(\overline{p}_t)$ efficient banks, each of which lends to the firm a multiple $1 + \phi_t = [1 - \mu(\overline{p}_t)]^{-1}$ of their initial assets at effective return pR_t. In the no-trade equilibrium, in contrast, banks do not delegate intermediation, and a mass $\mu(\gamma/R_t)$ of banks use the storage technology.

5. Aggregate Supply of Corporate Loans

The supply of corporate loans depends on whether or not the interbank market operates. When interbank trades are possible, all bank assets are channeled to the firm, and the supply of corporate loans is A_t. When the interbank market is frozen, banks with $p < \gamma/R_t$ use the storage technology and the aggregate supply of corporate loans is $[1 - \mu(\gamma/R_t)]A_t$. Denoting the aggregate supply of corporate loan ℓ_t, one gets

[10] Two comments are in order. First, note that when $R_t \geq \overline{R}$, we could possibly generate sunspot equilibria by randomizing between outcomes C, U, and E, with a sunspot variable coordinating the beliefs of the banks. This would neither complicate the numerical resolution of the model nor change in a material way its qualitative properties. But, from a quantitative viewpoint, sunspot equilibria would not be consistent with fact 3 that banking crises do not hit at random. For a discussion on the selection of the Pareto-dominant equilibrium in games with multiple Pareto-rankable Nash equilibria, see Cooper et al. (1990). Second, also note that rationing equilibria (with excess aggregate supply) are ruled out by bank competition. Indeed, in a rationing equilibrium, some inefficient banks would be rationed; and since these banks have only storage as a relevant outside investment opportunity, they would always be better off by deviating from the rationing equilibrium loan contract. Hence, our model differs from the standard Stiglitz and Weiss (1981) adverse selection model with monopolistic banks also in this respect.

$$\ell_t = \begin{cases} A_t & \text{if an equilibrium with trade exists} \\ \left[1 - \mu\left(\dfrac{\gamma}{R_t}\right)\right] A_t & \text{otherwise.} \end{cases} \tag{10}$$

During crises, banks inefficiently channel funds away from the firm into storage—a phenomenon we will refer to as a "credit crunch." Proposition 2 follows.

Proposition 2 (Interbank market freeze and credit crunch). An interbank market freeze is accompanied by a credit crunch.

D. Decentralized General Equilibrium

A general equilibrium of the economy is defined as follows.

Definition 1 (Competitive general equilibrium). A competitive general equilibrium is a sequence of prices $\mathcal{P}_t \equiv \{R_{t+i}, r_{t+i}, \rho_{t+i}, w_{t+i}\}_{i=0}^{\infty}$ and a sequence of quantities $\mathcal{Q}_t \equiv \{c_{t+i}, y_{t+i}, k_{t+i}, h_{t+i}, a_{t+i}, A_{t+i}, \ell_{t+i}\}_{i=0}^{\infty}$ such that (i) for a given sequence of prices, $\mathcal{P}_t$, the sequence of quantities, $\mathcal{Q}_t$, solves the optimization problems of the agents; (ii) $a_{t+i} = A_{t+i}$ for $i = 0, \ldots, \infty$; and (iii) for a sequence of quantities, $\mathcal{Q}_t$, the sequence of prices, $\mathcal{P}_t$, clears the markets; the market-clearing condition in the interbank market is given by equation (8), and the market-clearing conditions of the labor, good, and corporate loan markets are given by $h_{t+i} = 1$, $y_{t+i} = c_{t+i} + A_{t+1+i} - (1-\delta)A_{t+i}$, and $\ell_{t+i} = k_{t+i}$, respectively, for $i = 0, \ldots, \infty$.

In the general equilibrium, the market-clearing condition on the corporate loan market can be expressed as

$$f_k^{-1}\left(\frac{R_t + \delta - 1}{z_t}\right) = \begin{cases} A_t & \text{if an equilibrium with trade exists} \quad \text{(a)} \\ A_t - \mu\left(\dfrac{\gamma}{R_t}\right) A_t & \text{otherwise,} \quad \text{(b)} \end{cases} \tag{11}$$

where $f_k(k_t) \equiv \partial F(k_t, 1)/\partial k_t$. Relation (11) characterizes the equilibrium corporate loan rate R_t as a function of the two aggregate state variables of the model, A_t (savings) and z_t (productivity). Together with figure 2, it also points to the two-way relationship that exists between the interbank loan and the corporate loan markets, as R_t both affects and is affected by whether or not the interbank market operates.

The corporate loan market equilibrium must therefore be solved sequentially. We first solve (11) for R_t assuming that an interbank market equilibrium exists (using [11a]) and check that $R_t \geq \overline{R}$. In the negative, the interbank market equilibrium with trade cannot emerge (the interbank market freezes), and banks reduce their supply of loans to the firm. The corporate loan rate, R_t, then solves (11b). Note that, following the credit crunch, R_t may well be above $\overline{R}$; but this is not relevant for the existence of interbank trades because, in the absence of coordination failures, the only rate that matters in this respect is the one that would prevail in the equilibrium with trade, if it existed. (Remember that the interbank market freezes if and only if an equilibrium with trade cannot be sustained in the first place.) It follows that condition $R_t < \overline{R}$ is neither necessary nor sufficient for the interbank market to freeze. Proposition 3 follows.

Proposition 3 (Interbank loan market freeze). An equilibrium with interbank trade exists if and only if

$$A_t \leq \overline{A}_t \equiv f_k^{-1}((\overline{R} + \delta - 1)/z_t) \Leftrightarrow z_t \geq \overline{z}_t \equiv (\overline{R} + \delta - 1)/f_k(A_t);$$

otherwise, the interbank loan market freezes.

The threshold $\overline{A}_t$ is the maximum quantity of assets that the banking sector can allocate efficiently, that is, without the storage technology being used. Above this threshold counterparty fears are so prevalent that no trade takes place in the interbank market. In the rest of the paper we will refer to $\overline{A}_t$ as banks' "absorption capacity." It increases with the level of TFP $(\partial\overline{A}_t/\partial z_t > 0)$ because following an increase in TFP the equilibrium corporate loan rate goes up and banks' opportunity cost of diversion rises, which alleviates the moral hazard problem. Symmetrically, there is a threshold $\overline{z}_t$ for the level of TFP below which the interbank market freezes; this threshold increases with banks' assets $(\partial\overline{z}_t/\partial A_t > 0)$. Market freezes thus result either from the overaccumulation of assets by the household (when, given z_t, A_t rises beyond $\overline{A}_t$) or from an adverse productivity shock that reduces banks' absorption capacity (when, given A_t, z_t falls below $\overline{z}_t$), or from a combination of both endogenous and exogenous factors. To illustrate this point, figure 3 reports a stylized representation of the household's optimal asset accumulation rules for three different levels of z_t: low (z_L), average (z_A), and high (z_H).

For a given z_t, the optimal saving rule is discontinuous in $A_t = \overline{A}_t$. The upper branch (with $A_t \leq \overline{A}_t$) corresponds to normal times, where the interbank market functions well and all available assets in the economy are used to finance the firm. The lower branch (with $A_t > \overline{A}_t$) corresponds to crisis times. During a crisis, banks reduce their supply of loans to the firm, the firm reduces its demand for labor, wages drop, and the household dissaves so as to sustain her consumption; hence the discontinuity in the decision rule when $A_t = \overline{A}_t$.

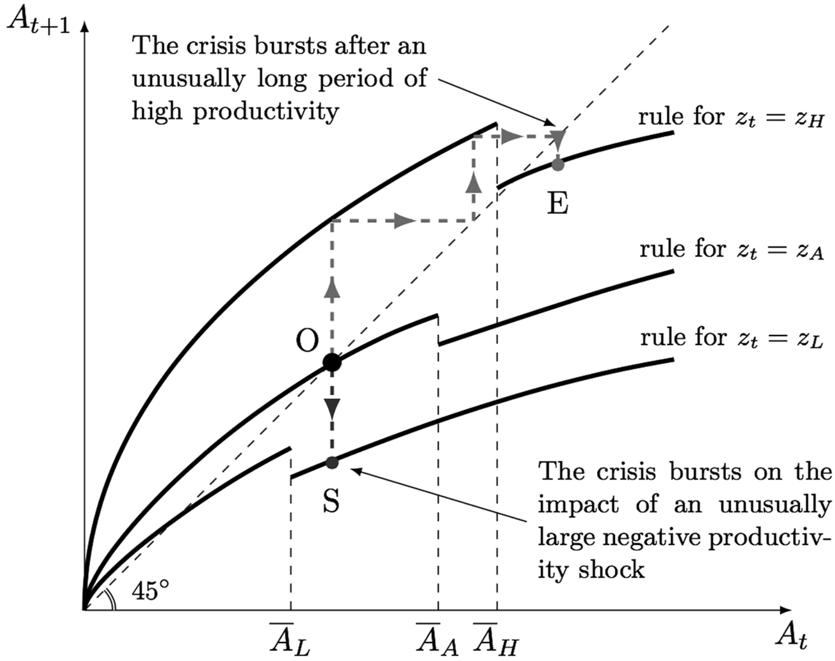


Fig. 3.—Stylized representation of the optimal decision rules

The figure focuses on two polar examples of crises: S and E. The crisis in S is triggered by a large, exogenous, adverse shock that makes productivity fall from z_A to z_L. Starting from the average steady state O, the shock instantaneously reduces banks' absorption capacity below the current level of assets, from $\overline{A}_A$ to $\overline{A}_L$, and the crisis bursts on impact. This crisis is purely driven by the bad shock and the amplification mechanism that the shock initiates. This is the channel usually depicted in the literature. The crisis in E is of a different nature. Starting from the average steady state O, the economy experiences an unusually long period of high productivity, z_H. During this period, banks' absorption capacity is high (at $\overline{A}_H$), but the household eventually accumulates assets beyond it (point E).[11] Points S and E provide useful, albeit extreme, examples of crises. In effect, crises in our quantitative version of the model will be a mix of these

[11] Notice that, in theory, this second type of crisis does not require any shock. To emphasize this point, we present in sec. D.1 of the supplementary appendix a purely deterministic version of the model—i.e., with a constant level of TFP—and show that the model can generate endogenous deterministic boom-bust cycles.

two polar cases. We will dub "shock-driven" the crises (i) associated with an unusually low level of TFP at the time they burst and (ii) not preceded by an unusually high level of assets—or credit, as in *S*. Similarly, we will dub "credit boom–driven" the crises (i) preceded by an unusually high level of assets—or credit—and (ii) not associated with an unusually low level of TFP at the time they burst, as in *E*.

The above discussion highlights that different combinations of z_t and A_t may prompt crises. For example, some crises follow credit booms, but a priori, not all crises do; some credit booms lead to crises, but not all credit booms do. Therefore, if one relies on either z_t or A_t to predict crises, one may generate many type I and type II errors. By construction, the best predictor of a crisis is given by the model probability that a crisis breaks out in the next period (see proposition 4).

PROPOSITION 4 (Model probability of a crisis). Given the state of the economy (z_t, A_t) at the end of period t, the probability that a banking crisis breaks out in period $t + 1$ is

$$\mathbb{P}(z_{t+1} < \bar{z}_{t+1} | z_t, A_t) = \Phi(\log \bar{z}_{t+1} - \rho_z \log z_t),$$

where $\Phi(\cdot)$ denotes the normal cumulative density function. Given z_t, $\mathbb{P}(z_{t+1} < \bar{z}_{t+1} | z_t, A_t)$ increases with A_t.

This probability is fully consistent with general equilibrium effects and agents' rational expectations. It increases as the economy approaches its absorption capacity. In Section V.B, we analyze the predictive properties of z_t and A_t within a calibrated version of the model, use the model probability of a crisis to construct early warnings of banking crises, and discuss the performance of these early warnings.

IV. Quantitative Analysis

The aim of this section is to investigate the quantitative properties of the model. To bring our model closer to the workhorse macroeconomic model, we extend it in two directions: we allow for endogenous labor supply decisions, and we introduce technological progress that is consistent with the observed long-run productivity growth in advanced economies. These features are standard and do not change the qualitative properties of the model. The equations characterizing the general equilibrium of the economy are reported in the Appendix.

A. *Calibration of the Model*

Technology is represented by a constant returns to scale production function of the form $z_t F(k_t, h_t) \equiv z_t k_t^{\alpha} (\Psi_t h_t)^{1-\alpha}$ with $\alpha \in (0, 1)$, where the term Ψ_t captures technological progress, is exogenous, and grows at the constant gross rate $\psi > 1$ (i.e., $\Psi_t = \Psi_0\psi^t$, with $\Psi_0 > 0$). The household is endowed with preferences over consumption, c_t, and hours worked, h_t, represented by a Greenwood, Hercowitz, and Huffman (1988) utility function,

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} \tilde{\beta}^{\tau} \frac{1}{1-\sigma} \left(c_{t+\tau} - \vartheta \Psi_{t+\tau} \frac{h_{t+\tau}^{1+\nu}}{1+\nu} \right)^{1-\sigma},$$

where $\nu \geq 0$ is the inverse Frisch labor supply elasticity and ϑ is the labor disutility parameter. The term $\tilde{\beta}$ denotes the discount factor in the economy not deflated for growth. The presence of the technological progress term in the utility function ensures the existence of a balanced growth path (for another application, see, e.g., Jaimovich and Rebelo [2009]). We chose this specification for a practical reason because it yields a closed form for the banking sector's (detrended) absorption capacity that greatly simplifies the numerical solution of the equilibrium,[12]

$$\overline{A}_t = \Gamma z_t^{(1+\nu)/\nu(1-\alpha)} \quad \text{with } \Gamma \equiv \left(\frac{1-\alpha}{\vartheta} \right)^{1/\nu} \left(\frac{\alpha}{\overline{R} + \delta - 1} \right)^{(\nu+\alpha)/\nu(1-\alpha)}. \tag{12}$$

The model is calibrated on a yearly basis to be consistent with the facts presented in Section II. The calibration, reported in table 3, is most standard. We set $\tilde{\beta}$ so that the household discounts the future at a 3 percent rate per annum in the economy deflated for growth. We set $\nu = 0.5$ so that the labor supply elasticity is equal to 2. The labor disutility parameter ϑ is such that the household would supply one unit of labor in a deterministic version of the model. The risk aversion parameter σ is set to 4.5, which lies within the range of estimated values. The capital elasticity in the production function α is set to 0.3, and we assume that capital depreciates at a 10 percent rate per annum ($\delta = 0.1$). To calibrate the data-generating process of TFP, we first back out a model-consistent series of the logarithm of TFP, for the US economy over the period 1950–2008,

$$\log(\text{TFP}_t) = \log(y_t) - \alpha \log(k_t) - (1-\alpha)\log(h_t),$$

where y_t is real GDP as reported in the National Income and Product Accounts (NIPA), and h_t is total annual hours worked as reported by the Conference Board Total Economy Database. The series of physical capital k_t is constructed by feeding the law of motion of capital in the model with the real investment series from the NIPA. The initial capital stock, k_{1950}, is set such that the economy is on a balanced growth path, implying

[12] As is well known, in this class of utility functions, wealth effects do not affect labor supply decisions. From a quantitative point of view, the main implication of the absence of wealth effects is that hours worked are particularly sensitive to the business cycle, which leads to larger boom-bust episodes. In sec. D.7 of the supplementary appendix, we present simulations of a version of the model featuring wealth effects in labor supply decisions.

TABLE 3
CALIBRATION

Parameter	Variable	Value
Discount factor (deflated for growth)	$\beta \equiv \tilde{\beta}\psi^{-\sigma}$	.970
Inverse of Frisch elasticity	ν	.500
Labor disutility	ϑ	.945
Risk aversion	σ	4.500
Capital elasticity	α	.300
Capital depreciation rate	δ	.100
Growth factor	ψ	1.012
Standard deviation of productivity shock	σ_z	.013
Persistence of productivity shock	ρ_z	.890
Bank distribution; $\mu(p) = p^{\lambda}$	λ	26
Diversion cost	θ	.085
Storage technology	γ	.952

that $k_{1950} = (k/y)y_{1950}$, where k/y denotes the capital output ratio of the economy implied by our model ($k/y = \alpha\beta/[\psi - \beta(1-\delta)] = 2.1$). Finally, we fit a linear trend to the (log) TFP series and use the deviations from this trend to estimate the AR(1) process for TFP in equation (1); we obtain $\rho_z = 0.89$ and $\sigma_z = 1.3$ percent as estimates.[13]

The remaining parameters pertain to the banking sector and include the return on storage γ, the diversion technology parameter, θ, and the distribution of banks $\mu(\cdot)$. For tractability reasons we assume that $\mu(p) = p^{\lambda}$, with $\lambda \in \mathbb{R}^{+}$. The parameters of the banking sector are calibrated jointly so that (i) the spread between the real corporate loan rate and the implicit real risk-free rate equals 1.7 percent, (ii) the real corporate loan rate equals 4.4 percent, and (iii) a financial recession occurs on average every 42 years (fact 1). For statistics i and ii, we use the real lending rate on midsize business loans for the United States between 1990 and 2011, as reported in the US Federal Reserve Bank's Survey of Terms of Business Lending and the real Federal Fund rate. We obtain $\gamma = 0.952$, $\lambda = 26$, and $\theta = 0.085$. On the basis of this calibration, the model generates an average interbank loan rate of 0.90 percent and an implied threshold for the real corporate loan rate of 3.20 percent (i.e., $\overline{R} = 1.032$). Note that we do not calibrate the model to replicate a priori any of facts 2 (financial recessions are deeper and longer) and 3 (financial recessions follow credit booms).

We solve the model numerically using a collocation method, allowing for discontinuities in the asset accumulation decision rule at the points at which the economy switches regime, that is, when $A_t = \overline{A}_t$. (Details on the numerical solution method are provided in the Appendix.) Given the importance of parameters σ, ν, ρ_z, σ_z, γ, θ, and λ, for the dynamics

[13] In sec. C.2 of the supplementary appendix, we investigate the robustness of these estimates to changes in the measurement of TFP.

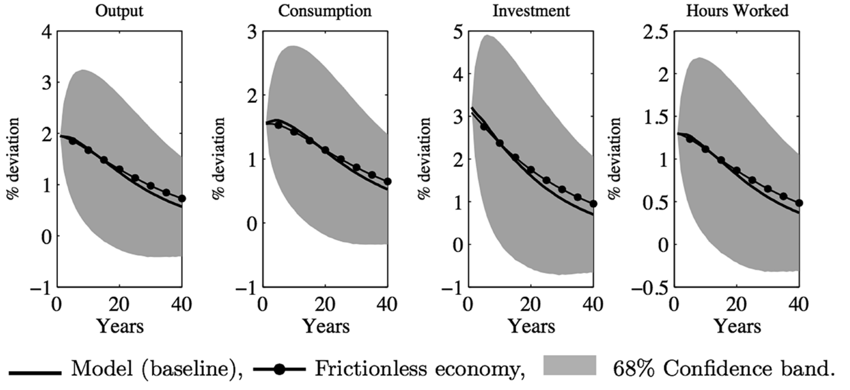


FIG. 4.—Impulse response to a one standard deviation technology shock. The impulse responses are computed with Koop et al.'s (1996) methodology, using 100,000 Monte Carlo simulations of the model. The confidence band is drawn around the baseline model.

of the model, we will discuss the sensitivity of the quantitative results to changes in their values in Section V.C.

B. *The Economy in Normal Times*

We analyze the response of the economy to a positive one standard deviation productivity shock, when the economy is initially at the steady state associated with the long-run average of z_t. We compute the impulse response functions as in Koop, Pesaran, and Potter (1996) so as to take nonlinearities into account (for the computational details, see sec. B.3 of the supplementary appendix). Figure 4 compares the dynamics in our model (thick plain line) with those in a frictionless model (round markers), which is obtained by setting either $\gamma = 0$ (no moral hazard), $\theta = 0$ (no moral hazard), or $\lambda = +\infty$ (all banks have $p = 1$).[14] This frictionless version of our model corresponds to a standard real business cycle (RBC) model.

The responses of output, hours worked, consumption, and investment are essentially the same in the two models. Our model exhibits only slightly more amplification than the RBC model because of a financial accelerator mechanism that tends to magnify the effects of the shock. Following the shock the corporate loan rate rises, which relaxes banks' borrowing constraints in the interbank market. As the aggregate demand for interbank loans increases, the equilibrium interbank loan rate goes up and inefficient banks switch from the demand to the supply side of the market, so that the moral hazard problem recedes; hence the—mild—financial accelerator.

[14] In all these cases savings are channeled to the most efficient bank $p = 1$, either as an equilibrium result or because $p = 1$ is the only type of bank in the economy.

Importantly, the economy does not experience any banking crisis after the shock. Even though the corporate loan rate eventually falls below its steady-state level as the household accumulates assets, at no point does it fall below $\overline{R}$ (not shown). We find that the model also behaves like a standard financial accelerator model following a negative one standard deviation productivity shock (not shown). Banking crises in the model are indeed rare events that occur under specific conditions, as we show in the next section.

C. *Typical Boom-Bust*

This section describes the typical conditions under which financial recessions occur in the model. As already discussed in Section III.D banking crises can be driven by exogenous productivity shocks, by the household's asset accumulation behavior, or by both. Which factor typically leads to such events is therefore a quantitative question, which we settle by simulating the calibrated version of the model.

Starting from the average steady state, we simulate the model over 500,000 periods and identify the years of recessions and crises. We define recessions as the periods between business cycle peaks and troughs, for which the drop in output is such that, overall, recessions occur 11.29 percent of the time, as in the data. A recession is then said to be a financial recession when a banking crisis breaks out between the peak that precedes it and the trough that ends it. In the simulations, all banking crises are associated with a recession, and all financial recessions begin the same year as the crisis they are associated with.

Then, we consider all the paths of the variables of the model—in deviation from their trend—30 years before and 10 years after the start of each recession and compute the median of the distribution of those paths for each variable. We refer to this median path as the "typical path" to recessions.

Figure 4 reports the typical path to financial recessions. The left panel describes the endogenous dynamics of banks' assets (A_t) as well as banks' absorption capacity ($\overline{A}_t$) along the typical path; the right panel describes the exogenous dynamics of TFP (z_t) and its innovations (ε_t). By convention, period 0 corresponds to the period when the financial recession (or the crisis) bursts.

The first important result that emerges from figure 5 is that the typical financial recession is not due to an unusually large, exogenous, negative shock. It is triggered by a 1.45-σ_z negative TFP shock (right panel, square markers), which is neither large nor unusual. Following the shock, TFP falls to 2.1 percent below its trend, which is not unusual either; in the simulations, this happens 17.2 percent of the time. The confidence bands around the typical path also indicate that TFP is sometimes above its trend at the time crises burst; in the simulations, this is the case for

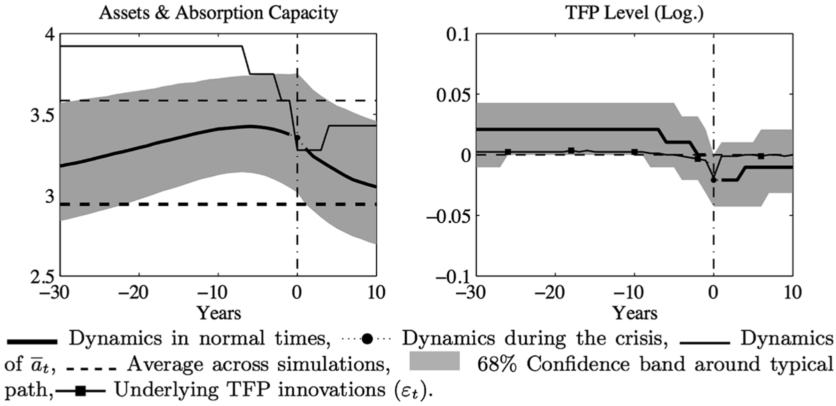


FIG. 5.—Typical path to financial recessions (I). For presentation purposes, the reported series have been deflated for the underlying exogenous growth trend.

13.5 percent of the crises. In those cases, bank assets are so large that even a small negative TFP shock suffices to trigger the crisis (the median size of this shock is 0.63 σ_z). In some—but very few—instances, the model even generates a crisis without a negative technology shock. These results suggest that the exogenous dynamics of technology alone are not enough to explain why financial recessions occur and why they are so rare. The left panel of figure 5 shows that banks' assets are disproportionately large at the time the crisis hits; they are 13.5 percent above their average steady state and 41.25 percent above the steady state associated with the current TFP level. That banks' assets grow so much oversize happens only 15.2 percent of the time in our simulations. From proposition 3, it is clear that, had banks not been so large in the first place, the fall in their absorption capacity would not have triggered the crisis.

This prompts the question why the banking sector can grow so large. Since in our model banks' assets are entirely financed by the household, banks' expansion results from an excessive accumulation of savings or, as Bernanke (2005) coined it, from a "saving glut."[15] In a neoclassical framework like ours, where savings are driven by the household's consumption smoothing behavior, a saving glut can develop only following an unusually and unexpectedly long string of positive technology shocks, as the right panel of figure 5 illustrates (square markers). To understand how things work, it is useful to decompose the precrisis period into two phases.

[15] Another way to obtain a saving glut would be to consider a two-country model, where the banking sector would be more developed (i.e., would have a lower θ) in one of the countries and would attract the developing countries' savings, as in, for instance, Caballero, Farhi, and Gourinchas (2008), Mendoza, Quadrini, and Ríos-Rull (2009), and Boissay (2011).

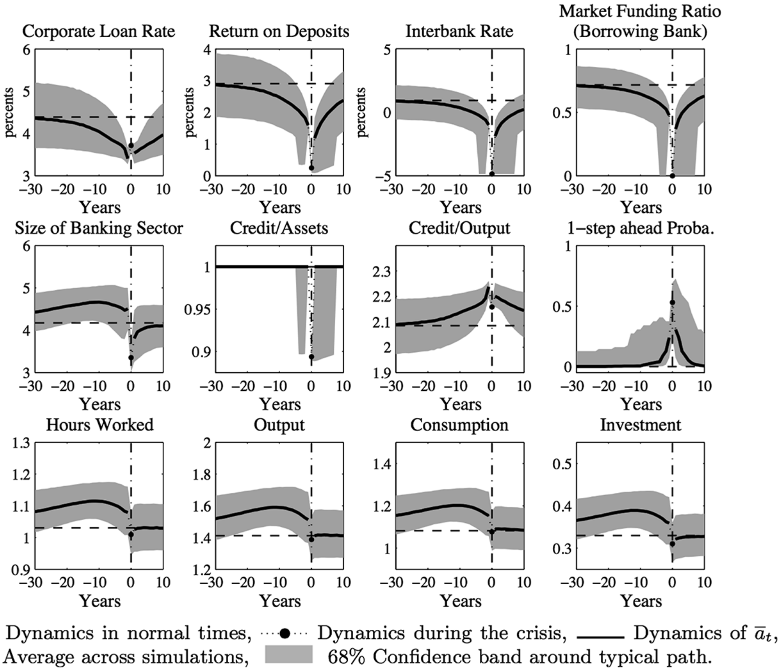


FIG. 6.—Typical path to financial recessions (II). The size of the banking sector corresponds to the sum (as opposed to consolidation) of all banks' assets and therefore includes interbank funding. Formally, we report $a_t + [1 - \mu(\overline{p}_t)]\phi_t a_t$. For presentation purposes, the reported series have been deflated for the underlying exogenous growth trend.

During the first phase, TFP rises above trend and remains above it for a long time (right panel, plain line). These productivity gains result from the sequence of positive TFP innovations, which are small but yet large enough to overcome the mean-reverting dynamics of TFP. These are good times, where the economy is far from its absorption capacity, the moral hazard problem is benign, crises are unlikely (see the middle-right panel of fig. 6), and it is optimal for the household to accumulate savings. This first phase lasts long enough to give the household the time to build up an unusually large stock of assets. This is the phase during which the saving glut develops and feeds a boom in credit.

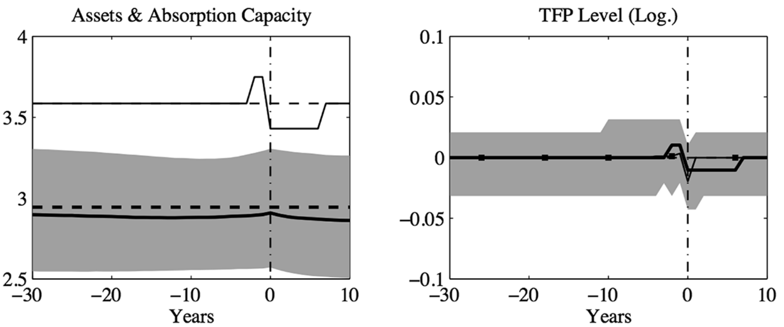
The second precrisis phase starts when TFP reverts back to its trend.[16] As productivity gains peter out, the economy gets closer to its absorption capacity, output reverts to trend, and the credit-to-output ratio rises. The equilibrium corporate and interbank loan rates both fall significantly below their average level, counterparty fears spread, and the crisis probabil-

[16] It should be clear that, given the positive 1.2 percent slope of the underlying productivity growth trend, productivity also increases during this second phase, but at a slower pace than its trend; we illustrate this point in sec. C.4 of the supplementary appendix, where we report the nondeflated typical path to financial recessions.

ity reaches 25 percent in the year before the crisis. The household's asset accumulation process during this precrisis period results from two countervailing forces. On the one hand, as the expected return on bank deposits diminishes, the household tends to save less. On the other hand, the household also realizes that a crisis is looming and that she will face large income losses if it materializes. To hedge against such a bad outcome and smooth her consumption profile, the household is willing to accumulate more. All in all, the stock of assets goes down, but consumption smoothing slows down the pace of disaccumulation. In effect, the household does not dissave fast enough to avoid the crisis because she does not internalize that, by maintaining her savings high, she makes the crisis more likely, an externality that we dub "the saving glut externality." (We discuss the effects of this externality in the next section.)

In this context of financial fragility, the 1.45-σ_z negative TFP shock in period 0 is enough to bring the banking sector down. As it breaks out, the crisis manifests itself by a sudden reversal of the economy, with the size of the banking sector falling by 28.2 percent and the credit-to-assets ratio falling by around 11 percent. The banking crisis ends a long period of economic expansion in which output, investment, hours worked, and consumption were more than 10 percent above their respective trends (see fig. 6). During the crisis, all these aggregates return back to trend.

The dynamics described in figures 5 and 6 are specific to financial recessions and are very different from those of other recessions. To illustrate this point, we compare the typical path toward financial recessions in figure 5, with the typical path toward other recessions in figure 7. In the run-up to nonfinancial recessions, bank assets and productivity are on their respective trends; there is no boom in credit. Moreover, during those recessions, bank assets remain stable, around their trend; the



Dynamics in normal times, Dynamics during the crisis, Dynamics of $\bar{a}_t$, Average across simulations, 68% Confidence band around typical path, Underlying TFP innovations (ε_t).

FIG. 7.—Typical path to other recessions. For presentation purposes, the reported series have been deflated for the underlying exogenous growth trend.

banking sector does not collapse. This sharply contrasts with financial recessions. In fact, financial and nonfinancial recessions have only one feature in common: they are both triggered by a rather mild, negative TFP shock (a 1.45-σ_z shock in the case of financial recessions and a 1.5-σ_z shock in the case of other recessions). These findings indicate that the economy responds much more aggressively to the very same shock when it is far off its average steady state than when it is close to it. As a complement to figures 5 and 7, we also compare in section C.6 of the supplementary appendix the ergodic distributions of TFP levels at the beginning of financial recessions and at the beginning of other recessions. One can hardly see a difference between the two. By contrast, when we do the same comparison for bank assets, we find that bank assets are distributed—clearly—above their trend at the beginning of financial recessions only.

Altogether, these results confirm the critical role of credit booms as an endogenous source of financial instability in our model.

D. Inspecting the Mechanisms

There is nothing systematic about the prevalence of credit boom–driven crises. This is a result that can be partly explained by the asymmetric effects of the household's consumption smoothing behavior on financial stability. In bad times TFP is relatively low and the household dissaves, which by lowering the TFP threshold $\bar{z}_t$ makes crises less likely (remember that $\partial \bar{z}_t / \partial A_t > 0$). Hence, in bad times the dynamics of asset accumulation tend to stabilize the banking sector. By contrast, in good times TFP is relatively high and the household accumulates assets, which by raising $\bar{z}_t$ makes crises more likely. In this case the dynamics of savings tend to destabilize the financial sector. As already discussed, though, asset accumulation by the household would not be as destabilizing if the household internalized its adverse effects on financial stability. Therefore, we would argue that another important reason why most crises follow credit booms in the model is the presence of the saving glut externality.

To shed more light on those mechanisms, we compare in figure 8 the typical paths of assets and TFP in our baseline economy with the typical paths in two "counterfactual" economies: one without the saving glut externality and another with a constant saving rate.[17] Our intention is to

[17] The economy without the saving glut externality corresponds to a hypothetical economy in which the household internalizes—all else equal—the adverse effects of her saving decisions on the efficiency of the banking sector and on the overall cost of financial intermediation. But she does not realize that those costs are eventually rebated to her. (Hence, the no saving glut allocation is not constrained efficient; we analyze the constrained efficient allocation later, in Sec. VI.) We provide a formal description of the underlying optimization problem in sec. B.1.3 of the supplementary appendix. In the case of the constant saving rate economy, the saving rate is set at the average steady-state saving rate of the baseline economy (i.e., at 24 percent). We describe the constant saving rate version of the model in sec. D.9 of the supplementary appendix.

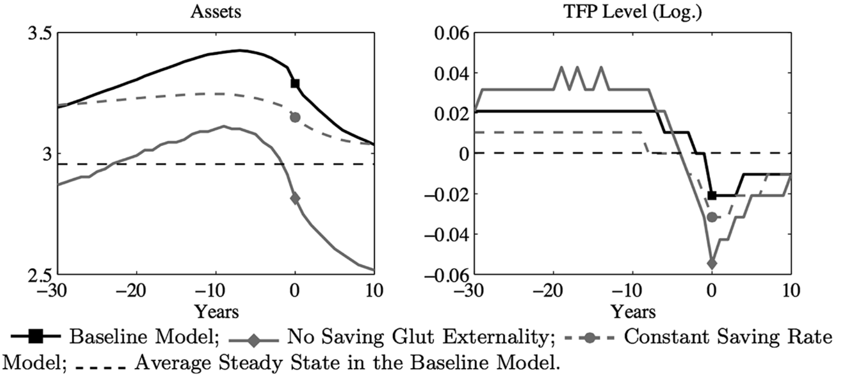


FIG. 8.—Typical paths to financial recessions without smoothing or externality. The marker indicates the value of the variable at the period the crisis breaks out. For presentation purposes, the reported series have been deflated for the underlying exogenous growth trend.

show that in these two economies the typical crisis is not credit boom–driven; that is, either it is not preceded by an unusually high level of assets (or credit) or it is associated with an unusually low level of TFP at the time it bursts. If this is the case, then we would conclude that the externality and consumption smoothing are both key ingredients to generate credit boom–driven crises.

The first result that emerges from figure 8 is that in the constant saving rate version of the model (gray dashed line, circle markers) the typical banking crisis is not due to a boom in credit. As the crisis breaks out, TFP is significantly—3.2 percent—below trend, which happens 9 percent of the time in the simulations, whereas bank assets are at their average steady-state level. In this case, the dynamics of TFP can almost drive, on their own, financial recessions. Hence we would argue that consumption smoothing is a necessary ingredient to generate credit boom–driven crises. But again, it is not sufficient. The typical path of TFP in the economy without the saving glut externality (gray plain line, diamond markers) indeed shows that in this economy too the crisis occurs when the level of TFP is unusually low. The reason is that the household, who now internalizes the adverse effects of her savings, not only saves overall less but also hastily dissaves as soon as the likelihood of a crisis increases—precisely to avoid that the crisis materializes. To trigger the crisis the negative TFP shocks must therefore surprise the household and be relatively large; here TFP must fall to 5.5 percent below trend, which happens only 1.3 percent of the time in the simulations.

We conclude that both the consumption smoothing behavior and the saving glut externality are necessary to generate credit boom–driven crises in the calibrated version of the model. Consumption smoothing is important because it causes the development of the saving glut in the first precrisis phase and creates the conditions under which even small shocks matter; the saving glut externality is equally important because it prevents the saving glut from resorbing when the economy is on the brink of a crisis.

V. Quantitative Implications

A. *Financial Recession Statistics*

The aim of this section is to assess the ability of our model to account for the three stylized facts documented in Section II.[18] To do so, we simulate a 500,000-period time series and calculate the same statistics as in table 1; the results are reported in table 4.

By construction, the model generates financial recessions with a frequency of 2.34 percent (fact 1). These recessions feature a 9.69 percent drop in output from peak to trough and last, on average, 1.84 years. This is more than in other recessions, which are about 5 months shorter and where output falls by about half as much (4.58 percent). These differences have the same order of magnitude as in the actual data. Thus, the model replicates fact 2 that financial recessions are deeper and last longer than other recessions. However, the magnitude of the average financial recession is 40 percent larger than in the data, where output falls by 6.84 percent from peak to trough. Three features of the model contribute to these excessively large output losses. First, financial recessions are accompanied by a severe credit crunch, with credit falling by 9.55 percent from peak to trough (this is more than twice as much as what is observed in the data). Second, the firm is assumed to refinance its capital through bank loans only and to roll over those loans every period; it does not retain earnings and does not issue bonds or outside equity. As a consequence, capital is not predetermined (only the household's stock of assets is), and the fall in credit implies an equally large fall in capital. Finally, the absence of wealth effects in labor supply decisions, and therefore of leisure smoothing behavior, makes hours very sensitive to fluctuations in wages. As capital falls, labor productivity and labor demand tank, wages go down, and the household reduces her supply of labor, which amplifies the initial fall in output; hence the too severe recession. In section D.6 of the supplementary appendix, we consider a version of the model with a lower labor Frisch elasticity, equal to 0.5 ($\nu = 2$) instead of 2 ($\nu = 0.5$), and recalibrate the labor disutility parameter (ϑ) and the

[18] In Sec. IV.B we showed that financial frictions in our model hardly matter in normal times, which implies that our model has essentially the same business cycle properties (second-order moments) as the textbook RBC model; see sec. C.8 of the supplementary appendix.

TABLE 4
Statistics on Recessions in the Model

	Financial	Other	All	Severe	Mild
Events	11,724	44,730	56,454	18,818	18,818
Frequency (%)	2.34	8.95	11.29	3.76	3.76
Duration (years)	1.84	1.34	1.44	1.95	1.04
Magnitude ($\Delta_{p,t}y$, %)	−9.69	−3.24	−4.58	−8.28	−2.13
Credit crunch:					
$\Delta_{p,t}k^{\mathrm{HP}}$ (%)	−9.55	.23	−1.80	−5.86	.20
$\Delta_{p,p+2}k^{\mathrm{HP}}$ (%)	−4.95	.11	−.94	−2.99	.11
Credit boom:					
$\Delta_{p-2,p}k^{\mathrm{HP}}$ (%)	3.55	.13	.84	2.27	.08
k_p^{HP} (%)	3.72	.07	.83	2.34	.01

Note.—The entry $\Delta_{p,t}x$ (respectively, $\Delta_{p,p+2}x$, $\Delta_{p-2,p}x$) denotes the percentage change of variable x from peak to trough (from peak to peak +2 years and peak −2 years to peak), where x denotes either output (y) or the HP ($\lambda = 6.25$) cyclical component of credit (k^{HP}). Statistics are based on a 500,000 time period simulation.

financial block of the model (λ, γ, and θ) accordingly. Since hours worked are less sensitive to the business cycle in that case, financial recessions are not as severe as in the baseline: from peak to trough, output falls by 4.82 percent, and credit falls by 7.67 percent. So, this version of the model matches fact 2 better than our baseline. However, the credit booms that precede financial recessions are smaller: in the 2 years that precede the financial recession, credit goes up by 2.90 percent, against 3.55 percent in the baseline and 4.55 percent in the data. Therefore, with regard to fact 3, the model with a low labor Frisch elasticity performs worse than our baseline. We find very similar results when we introduce wealth effects in the household's labor supply decisions (see sec. D.7 of the supplementary appendix).

The last two rows of table 4 show statistics on the evolution of credit before recessions. Credit increases in the 2 years that precede financial recessions, whereas it is stable before other recessions. Moreover, at the time financial recessions start, credit is 3.72 percent above its Hodrick-Prescott (HP) trend, whereas it is essentially on its trend in the case of other recessions.

As a last exercise to assess the quantitative properties of our model, we compare in figure 9 the dynamics of output and credit gaps around recessions, in the model and in the data. In both cases, the gaps correspond to the HP-filtered series; for the actual data, the figure is the same as figure 1. The results confirm that in the model, financial recessions (panel *a*) are deeper and longer than other recessions (panel *b*) and are associated with a boom-bust cycle in credit, whereas other recessions are not. Hence, the model can reproduce fact 2 and fact 3. In addition, figure 9 also shows that the dynamics of output and credit gaps observed in the data lie within the 68 percent confidence band of the model. We

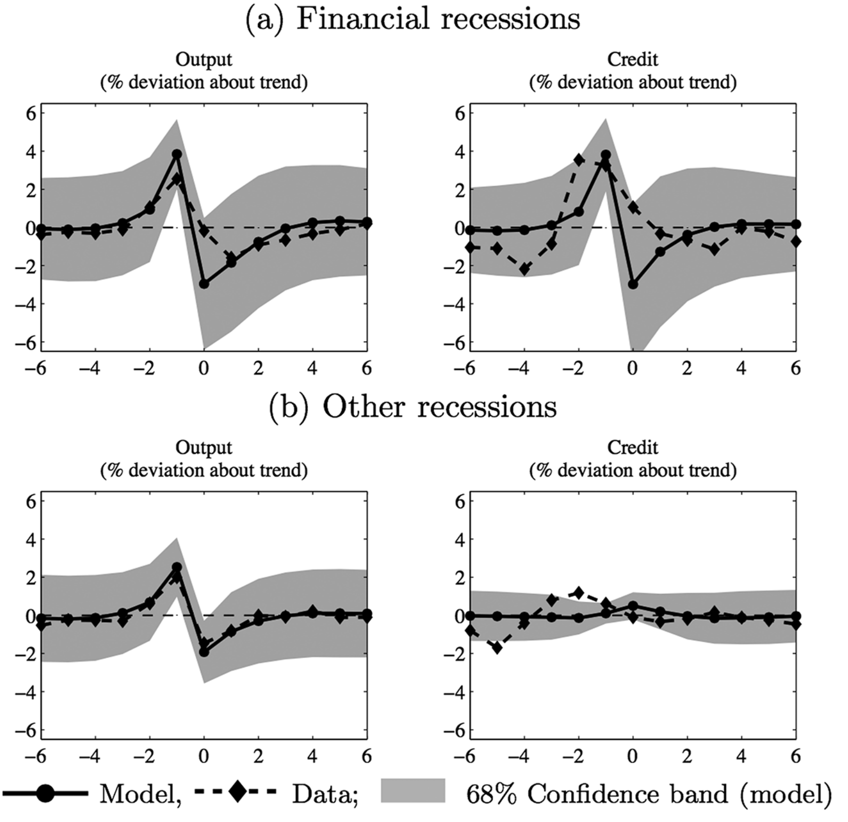


FIG. 9.—Dynamics of output and credit gaps around recessions. Average dynamics of the Hodrick-Prescott ($\lambda = 6.25$) cyclical component of (log) output and credit six periods before and after the beginning of a recession (period 0) with and without a banking crisis. To be consistent, we treat the simulated series of output and credit as we treat the actual data.

therefore conclude that the quantitative properties of our model seem overall consistent with Schularik and Taylor's (2012) data.

B. Crisis Prediction

The probability of falling into a crisis (see proposition 4) can be used to construct early warnings in the context of our model, but it is not observable in the real world. Our aim in this section is to look into whether this model probability can be backed out from variables that are directly observable in the data, to construct early warnings based on this crisis probability, and to discuss the properties of those early warnings (type I and type II errors).

In the model three variables are key to understanding the dynamics to crises: the level of bank deposit (or credit), TFP, and the credit-to-output ratio. (Instead of the credit-to-output ratio, one could consider the corporate loan rate, since given our production function the latter is a simple transformation of the former.) We simulate the model over 500,000 time periods and then regress the one-period-ahead crisis probability at period t on these variables at period t. The results are reported in table 5.[19]

We start by including the levels of credit ($k_t = A_t$ before the crisis) and TFP (z_t) separately and find—as expected—that the fit is relatively poor. TFP, for example, is significant but explains only 2 percent of the variations in the crisis probability (regression 1). In regression 2 with credit, the R^2 is higher (53 percent), which reflects the endogenous nature of the crises in the model. But it is still below that of regressions 3 and 4, which feature both credit supply and demand factors; in regression 4, with the credit-to-output ratio, the R^2 goes up to 73 percent.

Next, we construct early warnings of crises based on the crisis probabilities. By convention, a warning is issued whenever the model or fitted probability of a crisis is above 12.75 percent. This cutoff value is chosen so that the warnings derived from the model probability (our benchmark) are statistically correct at the 5 percent threshold (i.e., so that the type II error is less than 5 percent). In this case, the model issues almost two and a half times as many warnings as there are crises in the simulations (28,586 warnings for 11,724 crises; see tables 4 and 5) but signals two-thirds of the crises (the percentage of type I errors is 33.77 percent). Among the regression-based warnings, only those derived from the regression with the credit-to-output ratio (regression 4) perform as well, with type I and type II errors of 33.81 percent and 5.98 percent, respectively. One caveat of regression 4, though, is that the dependent variable—the model probability of a crisis—is not observable and has no clear counterpart in the data. So, as a last experiment, we check whether the credit-to-output ratio remains a good predictor when we instead use as the dependent variable a crisis dummy equal to one when a crisis breaks out in period t (and to zero otherwise) and estimate a logit regression (regression 5). In this case, to be consistent, we regress the crisis dummy on the one-period-lagged credit-to-output ratio. We find that the warn-

[19] Note that, unlike the data, the model cannot disentangle credit from physical capital. In the model, capital is therefore as good a predictor of financial recessions as credit. This prompts the question whether this holds also in the data. In sec. C.1 of the supplementary appendix (table 2), we present estimates of the same logit regressions as in table 2 using capital instead of credit as a regressor. We do find that capital is a good predictor of financial recessions. However, when credit is added into the regression, capital is not statistically significant at the 5 percent threshold anymore; only credit is. Capital being positively correlated with credit, one may argue that, in the data, only the credit-financed part of the capital stock helps predict financial recessions.

TABLE 5
PREDICTION OF BANKING CRISES

	MODEL PROBABILITY (Benchmark)	PROBABILITY REGRESSIONS z_t (1)	A_t (2)	(A_t, z_t) (3)	k_t/y_t (4)	LOGIT: k_t/y_t (5)
R^2	...	.02	.53	.68	.73	.38
F-test (%)	...	.12	.12	.00	.12	.00
Type I (%)	33.77	100.00	70.95	55.44	33.81	37.27
Type II (%)	4.26	.00	4.24	5.85	5.98	5.07
Warnings	28,586	0	22,893	32,129	35,275	30,674

NOTE.—Regressions 1–4 are linear regressions using a logistic transformation of the period t model probability that a crisis breaks out in $t + 1$ as the dependent variable. Regression 5 is a logit regression using a dummy equal to one when a crisis breaks out in $t + 1$ (and zero otherwise) as the dependent variable. All models include period t independent variables (in logarithm) and are conditional on not being in a crisis in period t. The regression sample includes 471,746 observations and 11,724 crises. Type I is the probability that the model fails to issue a warning. Type II is the probability that the model mistakenly issues a warning. A warning is issued as soon as the—model or fitted—crisis probability is above 12.75 percent, which threshold is chosen so that the type II error is less than 5 percent in the case of model-based warnings (first column).

ings derived from this regression have the same properties as those derived from the model probability. Together, these findings suggest that it is reasonable to look at the credit-to-output ratio to predict crises.

C. *Sensitivity Analysis*

The aim of this section is to evaluate the sensitivity of the model to changes in its parameters and to get a sense of which parameters are critical to replicate the three stylized facts of Section II. We run several experiments and calculate the same statistics on recessions as in tables 1 and 4, with a focus on financial and nonfinancial recessions. The results are reported in table 6.

The model replicates fact 2 and fact 3 throughout all the experiments. Financial recessions are always preceded by a credit boom, whereas other recessions are not; and they are deeper and longer than other recessions because, when they burst, output falls from far above its trend (see fig. 6). These properties are closely linked to the household's consumption smoothing behavior and the saving glut externality (see Sec. IV.D). So they are to a large extent immune to changes in parameter values. By contrast, but not surprisingly, changes in parameters do affect the probability of financial recessions and therefore the ability of the model to replicate fact 1.

As a first experiment, we consider a version of the model in which the financial intermediation cost (χ_t) is not rebated to the household,

TABLE 6
Statistics on Recessions in the Model: Sensitivity Analysis

	Panel A									
	Baseline		Deadweight Losses		$\sigma = 10$		$\nu = .25$		$\theta = .15$	
	FR	OR	FR	OR	FR	OR	FR	OR	FR	OR
Events	11,724	44,730	11,944	46,470	21,555	44,745	16,903	45,812	37,458	39,121
Frequency (%)	2.34	8.95	2.39	9.29	4.31	8.95	3.38	9.16	7.49	7.82
Duration (years)	1.84	1.34	2.09	1.33	1.66	1.34	2.07	1.35	1.58	1.32
Magnitude ($\Delta_{p,t}y$, %)	−9.69	−3.24	−10.32	−3.08	−8.55	−3.18	−12.77	−3.97	−6.44	−3.15
Credit crunch:										
$\Delta_{p,t}k^{\mathrm{HP}}$ (%)	−9.55	.23	−9.68	.31	−9.47	−.06	−10.03	.15	−6.15	−.33
$\Delta_{p,p+2}k^{\mathrm{HP}}$ (%)	−4.95	.11	−4.91	.07	−5.71	.06	−5.69	.12	−3.98	−.08
Credit boom:										
$\Delta_{p-2,p}k^{\mathrm{HP}}$ (%)	3.55	.13	3.48	.24	4.50	.08	4.04	.20	3.32	.05
k_p^{HP} (%)	3.72	.07	3.68	.12	4.51	.06	4.14	.11	3.24	.09

Panel B

	$\gamma = .955$		$\lambda = 20$		$\sigma_z = .025$		$\rho_z = .44$		Alternative TFP	
	FR	OR	FR	OR	FR	OR	FR	OR	FR	OR
Events	19,671	40,857	37,187	37,435	30,103	49,246	21,181	56,030	11,667	58,621
Frequency (%)	3.93	8.17	7.44	7.49	6.02	9.85	4.24	11.21	2.33	11.72
Duration (years)	1.73	1.32	1.63	1.36	1.96	1.45	1.51	1.18	1.42	1.19
Magnitude ($\Delta_{p,t}y$, %)	−9.09	−3.14	−9.47	−3.17	−14.37	−6.71	−14.76	−7.10	−8.12	−2.29
Credit crunch:										
$\Delta_{p,t}k^{\mathrm{HP}}$ (%)	−9.51	.11	−11.54	−.41	−9.92	−.40	−10.29	.92	−9.87	.40
$\Delta_{p,p+2}k^{\mathrm{HP}}$ (%)	−5.33	.17	−7.00	.28	−6.27	−.12	−4.40	.18	−3.59	.10
Credit boom:										
$\Delta_{p-2,p}k^{\mathrm{HP}}$ (%)	3.98	.07	5.66	−.26	4.90	.37	2.67	.31	2.03	.16
k_p^{HP} (%)	4.08	.04	5.62	−.10	4.80	.31	3.16	.03	2.54	.03

Note.—FR = financial recessions; OR = other recessions. The entry $\Delta_{p,t}x$ (respectively, $\Delta_{p,p+2}x$, $\Delta_{p-2,p}x$) denotes the percentage change of variable x from peak to trough (from peak to peak +2 years and peak −2 years to peak), where x denotes either output (y) or the HP ($\lambda = 6.25$) cyclical component of credit (k^{HP}). Statistics are based on a 500,000 time period simulation. The baseline column is reproduced from table 4. The column "deadweight losses" corresponds to a version of the model in which the intermediation costs, χ_t, are not rebated to the household (in this case we still use the baseline calibration, as in table 3). Column alternative TFP corresponds to a process estimated on the HP-filtered logarithm of TFP; in this case, $\rho_z = 0.44$ and $\sigma_z = 0.0105$, parameter γ is recalibrated in order to match the crisis probability observed in the data, and since our model is not designed to accommodate a smooth TFP growth trend, we set $\psi = 1$.

keeping our calibration unchanged; the results are reported in the columns "deadweight losses" (third and fourth columns in panel A). Since these costs are particularly large during crises, the household's income tends to fall by more during financial recessions in this version of the model than it does in the baseline model. As a result, financial recessions are slightly larger and longer than in the baseline; they are also slightly more frequent.

Next, we analyze the effects of household's consumption and leisure smoothing behaviors (parameters σ and ν). We run two experiments. First, we raise σ from 4.5 (baseline) to 10. Compared to the baseline, the household is more willing to smooth consumption over time. For a given sequence of positive TFP shocks, savings grow faster and credit booms are bigger. For instance, before the start of the recession, credit is 4.51 percent above trend, compared to 3.72 percent in the baseline economy. This makes the banking sector more sensitive to small adverse shocks. Since smaller shocks may trigger a crisis, the probability of a financial recession is higher than in the baseline (4.31 percent vs. 2.34 percent). In the second experiment, we raise the labor supply elasticity from 2 (baseline) to 4 ($\nu = 0.25$). Compared to the baseline, labor supply and income are more procyclical. The household saves faster, and the frequency of financial recessions goes up to 3.38 percent. Moreover, financial recessions, which are magnified by the larger response of hours worked, are more severe and last longer than in the baseline.

In the next set of experiments, we analyze the sensitivity of the results to financial frictions (parameters θ, γ, and λ). We first reduce the cost of diversion and raise θ from 0.1 (baseline) to 0.15. This aggravates the moral hazard problem between banks, reduces the absorption capacity of the economy (through its effect on $\overline{R}$), and increases the probability of a financial recession. The latter jumps from 2.34 percent in the baseline to 7.49 percent; also, financial recessions are milder and shorter. We find qualitatively similar results when we raise the return on storage, γ, from 0.952 (baseline) to 0.955 (first and second columns in panel B) or when we lower λ from 26 (baseline) to 20. In the latter case, the banking sector is overall less efficient than in the baseline; there are more inefficient banks. Since those banks have more incentives to divert cash, the moral hazard problem is more stringent, counterparty fears increase, and the frequency of financial recessions goes up to 7.44 percent.

Finally, we analyze the role of uncertainty. We start by raising σ_z from 0.013 (baseline) to 0.025, leaving ρ_z unchanged. The consequences are straightforward: the household accumulates more assets for consumption smoothing motives, the corporate loan rate decreases with respect to the baseline, and the financial sector is more fragile. Financial recessions are more frequent, and the credit booms that precede them are much bigger than in the baseline. We then cut the persistence of the shock by a half to $\rho_z = 0.44$ (instead of 0.89 in the baseline), leaving σ_z unchanged. Since TFP reverts faster back to its trend in this case, the periods of high TFP are shorter compared to the baseline. This has two opposite effects on the frequency of crises. On the one hand, the household accumulates assets faster and savings are more sensitive to shocks, which implies that credit booms faster than in the baseline. On the other hand, the household does not have as much time to accumulate assets; so credit booms tend to be smaller than in the baseline. Given our calibration, the first effect dominates, and financial recessions are more frequent. However, credit is only 3.16 percent above trend before financial recessions, against 3.72 percent in the baseline.

As a last experiment, we reestimate the AR(1) process of TFP using the HP-filtered series for z_t as a way to control for any potential change in the productivity growth trend. We find that the technology shock is less persistent ($\rho_z = 0.44$) and less volatile ($\sigma_z = 0.0105$) in this case than in the baseline. For comparison purposes, we recalibrate (reduce) the return on storage, γ, to obtain the same frequency of financial recessions as in the data. In this version of the model, financial recessions are marginally milder and shorter. Since σ_z is smaller, the household also accumulates less assets for consumption smoothing, and the credit booms that precede crises are smaller than in the baseline.

VI. Welfare Cost of Financial Frictions and Policy Intervention

The aim of this section is to evaluate the welfare costs of financial frictions, and thereby of the existence of financial recessions. Along the lines of Lucas (1987), we measure these welfare costs as the permanent percentage difference in consumption that would make the household indifferent between the decentralized equilibrium allocation and the first-best allocation, which would prevail in a frictionless world (here, the RBC model). This measure takes into account the cost of lower consumption in the transition to the first-best allocation. The result is reported in the first column in table 7. We find that, all in all, financial frictions reduce consumption permanently by 0.54 percent.

These welfare losses can be decomposed into two types. The first type of losses is imputable to financial frictions per se, to institutional deficiencies (limited contract enforceability and asymmetric information). In our model, these frictions have a negative effect on welfare because, by reducing banks' borrowing capacity, they restrict the set of feasible allocations. The second type of losses is imputable to the externalities, which the institutional deficiencies induce. Indeed, the household does not internalize the effects of her saving behavior on interest rates and on the likelihood of a crisis ("saving glut externality"). She does not inter-

TABLE 7
Welfare Costs: Percentage of Permanent Consumption

Financial Frictions DEA → FBA	Externalities DEA → CEA	Deficient Institutions CEA → FBA	Financial Underdevelopment NIM → DEA
.54	.15	.38	2.56

Note.—FBA = first-best allocation; DEA = decentralized equilibrium allocation; CEA = constrained efficient allocation; NIM = equilibrium allocation in a model with no interbank market. The welfare associated to an allocation j is measured by $\mathcal{W}^j(\mathcal{C}^j,\mathcal{H}^j)=\mathbb{E}\sum_{t=0}^{\infty}\beta^t U(c_t^j,h_t^j)$, where $\mathcal{C}^j\equiv\{c_t^j\}_{t=0}^{\infty}$ and $\mathcal{H}^j\equiv\{h_t^j\}_{t=0}^{\infty}$, with $j\in$ {FBA, CEA, DEA, NIM}. The table reports the percentage increase in permanent consumption ω when the household moves from allocation i to allocation j ($i\to j$); this measure explicitly takes into account the cost of the transition from i to j. The term ω solves $\mathcal{W}^i((1+\omega)\mathcal{C}^i,\mathcal{H}^i)=\mathcal{W}^j(\mathcal{C}^j,\mathcal{H}^j)$ and measures the welfare cost of allocation i with respect to allocation j. The results are based on 10,000 Monte Carlo simulations of the model, where the initial conditions are drawn from the ergodic distribution of allocation i.

nalize either that the financial intermediation costs are eventually rebated to her (“rebate externality”). As a result, among the feasible allocations, she does not pick the Pareto-optimal one.

To disentangle the two types of welfare losses, we devise the optimal allocation of a benevolent central planner, who faces the same set of feasible allocations as the private household, but who is not subject to the externalities. (Our approach is similar to those in Bianchi and Mendoza [2010] and Bianchi [2011].) We will refer to this allocation as the “constrained efficient allocation.” The welfare cost of institutional deficiencies will then be measured as the permanent percentage difference in consumption that makes the household indifferent between the constrained efficient allocation and the first-best allocation. Similarly, the welfare cost of externalities will be measured as the permanent percentage difference in consumption that makes the household indifferent between the decentralized equilibrium allocation and the constrained efficient allocation.

The benevolent central planner maximizes the household’s utility with respect to assets, consumption, labor, and capital, subject to the household’s resource constraints, and to the same set of allocations as the decentralized equilibrium supports. The only—but key—difference with respect to the decentralized equilibrium is that the central planner internalizes all the effects of her saving decisions. In particular, she understands that her savings, A_t, affects the equilibrium retail loan market interest rate, R_t (see relation [11]); that in turn R_t affects the equilibrium interbank market rate, ρ_t (see relation [8]); that together R_t and ρ_t determine banks’ efficiency, $\overline{p}_t$ (see relation [6]); and that bank efficiency ultimately affects the return on her savings, r_t (see relation [9]). She also understands that her savings decisions have an effect on the cost of financial intermediation, χ_t, and that this cost is rebated to her lump-sum. More formally, the constrained efficient allocation is defined as follows. (See also sec. B.1.2 in the supplementary appendix.)

Definition 2 (Recursive constrained efficient allocation). A recursive constrained efficient allocation is defined by a set of decision rules $\{c(A_t, z_t), h(A_t, z_t), k(A_t, z_t), A'(A_t, z_t)\}$ with the value function $V^{CP}(A_t, z_t)$ that solve the recursive optimization problem

$$V^{CP}(A_t, z_t) = \max_{A_{t+1}, c_t, k_t} \frac{1}{1-\sigma}\left(c_t - \vartheta \frac{h_t^{1+\nu}}{1+\nu}\right)^{1-\sigma} + \beta \mathbb{E}_t V^{CP}(A_{t+1}, z_{t+1})$$

subject to

$$A_{t+1} + c_t = z_t F(k_t, h_t) + (1-\delta)k_t + \gamma(A_t - k_t),$$

$$k_t = \begin{cases} A_t & \text{if } A_t \leq \overline{A}_t \\ \left[1 - \mu\left(\dfrac{\gamma}{z_t F_k(k_t, h_t) + 1 - \delta}\right)\right] A_t & \text{otherwise,} \end{cases}$$

where $\overline{A}_t$ is given by equation (12), h_t solves $\vartheta h_t^{\nu} = z_t F_h(k_t, h_t)$, and $F_x(k_t, h_t)$ denotes the first derivative of $F(k_t, h_t)$ with respect to $x \in \{k, h\}$.

The results of the welfare decomposition are reported in the second and third columns of table 7. We find that the externalities cost the household 0.15 percent of permanent consumption, which lies slightly above what is usually found in the literature; Bianchi and Mendoza (2010) and Bianchi (2011), for example, report values ranging from 0.05 percent to 0.135 percent. However, these losses are still small relative to those due to financial markets being deficient, which costs the household 0.38 percent of permanent consumption.

Finally, for the sake of completeness, we compute the welfare loss if interbank loans were not permitted. To do so we compare our baseline, decentralized economy with a hypothetical, financially underdeveloped economy with no interbank market. In line with Mendoza, Quadrini, and Ríos-Rull (2007), we find that financial underdevelopment has a sizable, detrimental effect on welfare. To be compensated for financial underdevelopment, the household should be given 2.56 percent of extra permanent consumption.

The welfare cost of financial frictions calls for policy intervention. One set of policies are macro-prudential policies that make private agents internalize the effects of their decisions on financial stability. In this respect, the saving glut externality and the rebate externality have very different implications. As we discussed in Section IV.D, the saving glut externality essentially kicks in when a crisis is imminent. This happens when, following a marginal increase in deposits, the crisis probability becomes economically relevant. In this case, the household tends to keep assets to hedge against the fall in consumption she will face if the crisis materializes (see fig. 6) and, by doing so, inadvertently precipitates the crisis. Since a crisis induces an instantaneous fall in welfare, the optimal policy then consists in giving the household an incentive to dissave fast to avoid the crisis, for example, by taxing the returns on savings. By contrast, the rebate externality always affects welfare, if a crisis is imminent or not. Were the household to internalize that the cost of financial intermediation is rebated to her, then she would understand that the return on assets is more than r_t and therefore accumulate more assets. (In normal times, for instance, the relevant return on assets is R_t, and not r_t.) Compared with the benevolent central planner, the household thus accumulates too little assets when the economy is far away from a crisis and too much assets when it is close to it. Accordingly, a welfare-improving policy would consist in subsidizing the marginal unit of deposit (or credit) when the probability of crisis is null and the banking sector can safely absorb it, and in taxing deposits—or reducing the subsidy—when a crisis is looming.[20] Of course, the profile and timing of such a policy are critical. In the run-up to a crisis, notably, taxes should not be raised too early, so as to let the economy reap the most benefits from high productivity; but they should not be raised too late, to give the household enough time to dissave and keep assets below banks' absorption capacity.

Our finding that the welfare cost of externalities is low relative to that of institutions being deficient, however, suggests that even an optimal macroprudential policy would have rather limited effects on welfare. Structural reforms aiming at improving financial contract enforceability and transparency in the banking sector may potentially have a much bigger impact.

VII. Concluding Remarks

We offered a simple quantitative macroeconomic model in which endogenous financial developments could turn the real business cycle around. The model features a nontrivial banking sector, in which banks are heterogeneous with respect to their intermediation skills. This heterogeneity gives rise to an interbank market. Moral hazard and asymmetric information in this market lead to infrequent market meltdowns, banking crises, credit crunches, and, ultimately, severe recessions. Our model has the potential to generate credit boom–driven crises, and we view it as a step toward understanding the interactions between financial and real busi-

[20] Notice that this policy implication heavily relies on our assumption that there is no deadweight loss in the economy. If the intermediation costs were not rebated, then there would be no underaccumulation of assets in normal times and therefore no need to subsidize deposits (or credit).

ness cycles. In this respect, the quantitative properties of the model are encouraging.

For tractability reasons we made some strong assumptions and presented a stylized model, whose narrative of banking crises may not be taken at face value. Indeed, we left out some aspects of the banking sector that would deserve more attention for future extensions. For instance, we assumed that the household cannot finance the firm directly and—by the same token—that the firm cannot issue bonds or equity. A proper investigation of the interactions between financial and real business cycles must take into account the substitution effects that exist between direct and intermediated finance (see De Fiore and Uhlig 2013). To keep matters simple we modeled a closed economy. Accordingly, the saving glut, which is key to the dynamics in our model, can be fed only by the domestic household. Opening the economy could have interesting insights as the saving glut would become global, thus offering the possibility to study how banking crises spread internationally. Finally, we ignored frictions between the banks and the household. Therefore, we could not pin down banks' optimal funding structure, which in turn prevented us from studying the dynamics of bank leverage. It would be interesting to model bank leverage explicitly so as to get richer dynamics and be in a position to discuss the effects of macro-prudential policies like changes in bank capital requirements. This and other extensions are left to future research.

Appendix

Solution Method

The decentralized equilibrium is solved using a collocation method. Solving the equilibrium boils down to finding the optimal decision rule $A_{t+1}(A_t; z_t) = \mathcal{G}(A_t; z_t)$ for all possible values of A_t and z_t.

We first construct a discrete representation of the distribution of the technology shock using the approach proposed by Tauchen and Hussey (1991). This yields a Markov chain representation of the technology shock with $z_t \in \{z_1, \ldots, z_{n_z}\}$ and a transition matrix $\Pi = (\pi_{ij})_{i,j=1}^{n_z}$ where $\pi_{ij} = \mathbb{P}(z_{t+1} = z_j | z_t = z_i)$. We use $n_z = 15$ and look for n_z decision rules $\mathcal{G}(A_t; z_i)$, for $i = 1, \ldots, n_z$. Since the economy can be in two regimes—normal and crisis times—and we know when it reaches one regime or the other (see proposition 3), we approximate the decision rules as

$$\mathcal{G}(A_t; z_i) \equiv \exp\left[\sum_{j=0}^{q} \xi_j^N(z_i) T_j(\varphi(A_t))\right] \mathbb{I}_{A_t \leq \overline{A}_i} + \exp\left[\sum_{j=0}^{q} \xi_j^C(z_i) T_j(\varphi(A_t))\right] \mathbb{I}_{A_t > \overline{A}_i},$$

where z_i denotes a particular level of the TFP in the grid and $\overline{A}_i = \overline{A}(z_i)$ is the corresponding absorption capacity (see Sec. IV.A). The term $T_j(\cdot)$ is a Chebychev polynomial of order j; $\xi_j^N(z_i)$ (respectively, $\xi_j^C(z_i)$) is the coefficient associated with this polynomial when the economy is in normal times (respectively, in a banking crisis) for the productivity level z_i; $\varphi(A_t)$ is a function that maps the level of assets into interval $(-1, 1)$; finally, $\mathbb{I}_{A_t \le \overline{A}_i}$ (respectively, $\mathbb{I}_{A_t > \overline{A}_i}$) is an indicator function that takes value one in normal (respectively, crisis) times and zero otherwise. The optimal decision rule $A_{t+1}(A_t; z_i)$ is given by the fixed-point solution to the Euler equation.

Our resolution algorithm proceeds as follows.

1. Choose a domain $[A_m, A_s]$ of approximation for A_t. We consider values that guarantee that the conditional steady state associated with each level of the technology shock can be reached. This led us to use $A_m = 0.5$ and $A_s = 8$.

2. Choose an order of approximation q, compute the $q + 1$ roots of the Chebychev polynomial of order $q + 1$ as

$$\zeta_k = \cos\left[\frac{(2k-1)\pi}{2(q+1)}\right] \quad \text{for } k = 1, \ldots, q+1,$$

and formulate initial guesses $\xi_j^{N(0)}(z_i)$ and $\xi_j^{C(0)}(z_i)$ for the coefficients of the Chebychev polynomial, for $i = 1, \ldots, n_z$ and $j = 1, \ldots, q$. We chose $q = 15$ so as to obtain an accurate approximation of the decision rule.

3. Define the (reciprocal of the) mapping function $\varphi(A_k)$, and compute A_k (for $k = 1, \ldots, q + 1$) as

$$A_k = \varphi^{-1}(\zeta_k) \equiv \begin{cases} \exp\left[\log(A_m) + (\zeta_k + 1)\dfrac{\log(\overline{A}_i) - \log(A_m)}{2}\right] & \text{for normal times} \\ \exp\left[\log(\overline{A}_i) + (\zeta_k + 1)\dfrac{\log(A_s) - \log(\overline{A}_i)}{2}\right] & \text{for crisis times} \end{cases}$$

for $k = 1, \ldots, q + 1$. Note that the grid of values for a depends fundamentally on the level of the technology shock as the threshold $\overline{A}_i$ depends on the level of productivity z_i (see Sec. IV.A):

$$\overline{A}_i = \Gamma z_i^{(1+\nu)/\nu(1-\alpha)}, \quad \text{with } \Gamma \equiv \left(\frac{1-\alpha}{\vartheta}\right)^{1/\nu}\left(\frac{\alpha}{\overline{R} + \delta - 1}\right)^{(\nu+\alpha)/\nu(1-\alpha)}.$$

4. Using the nth iteration on coefficients $\xi_j^{N(n)}(z_i)$ and $\xi_j^{C(n)}(z_i)$, compute for each level of A_k (with $k = 1, \ldots, q + 1$) the values of the decision rule:

$$\mathcal{G}^{(n)}(A_k; z_i) \equiv \exp\left[\sum_{j=0}^{q} \xi_j^{N(n)}(z_i) T_j(\zeta_k)\right]\mathbb{I}_{A_k \le \overline{A}_i} + \exp\left[\sum_{j=0}^{q} \xi_j^{C(n)}(z_i) T_j(\zeta_k)\right]\mathbb{I}_{A_k > \overline{A}_i}$$

as well as the control variables, in particular hours worked, $h_t^{(n)}(A_k, z_i)$, and household's income, $e_t^{(n)}(A_k, z_i) \equiv y_t^{(n)}(A_k, z_i) + (1-\delta)A_k$.

5. Use the decision rule $\mathcal{G}^{(n)}(A_k; z_i)$ to compute period $t + 2$'s level of assets $A_{t+2}^{(n)}(A_k; z_i, z_\ell) = \mathcal{G}^{(n)}(\mathcal{G}^{(n)}(A_k; z_i), z_\ell)$ as well as period $t + 1$'s hours worked $h_{t+1}^{(n)}(A_k; z_i, z_\ell)$, income $e_{t+1}^{(n)}(A_k; z_i, z_\ell)$, and interest rate $r_{t+1}^{(n)}(A_k; z_i, z_\ell)$, for all $k = 1, \ldots, q + 1$, $i = 1, \ldots, n_z$, and $\ell = 1, \ldots, n_z$.

6. Using the Euler equation and the budget constraint $c_t^{(n)}(A_k; z_i) = e_t(A_k; z_i) - \psi A_{t+1}^{(n)}(A_k; z_i)$, compute

$$\begin{aligned}\tilde{a}_{t+1}^{(n+1)}(A_k; z_i) \equiv \frac{1}{\psi}\Bigg(& e_t^{(n)}(A_k; z_i) - \vartheta \frac{h_t^{(n)}(A_k; z_i)^{1+\nu}}{1+\nu} \\ & - \Bigg\{\beta \sum_{\ell=1}^{n_z} \pi_{i\ell} r_{t+1}^{(n)}(A_k; z_i, z_\ell)\Bigg[e_{t+1}^{(n)}(A_k; z_i, z_\ell) \\ & - \psi A_{t+2}^{(n)}(A_k; z_i, z_\ell) - \vartheta \frac{h_{t+1}^{(n)}(A_k; z_i, z_\ell)^{1+\nu}}{1+\nu}\Bigg]^{-\sigma}\Bigg\}^{-1/\sigma}\end{aligned}$$

for each value of k (with $k = 1, \ldots, q + 1$) and each value of i (with $i = 1, \ldots, n_z$).

7. Project $\tilde{a}_{t+1}^{(n+1)}(A_k; z_i)$ on the Chebychev polynomials $T_j(\zeta_k)$ to obtain estimates $\xi_j^N(z_i)$ and $\xi_j^C(z_i)$. If the distance between $(\xi_j^N(z_i), \xi_j^C(z_i))$ and $(\xi_j^{N(n)}(z_i), \xi_j^{C(n)}(z_i))$ is close enough to zero in the sense that the norm of the difference is less than $\varepsilon \equiv 1\text{e}{-6}$, then stop. Otherwise, update the coefficients as

$$\begin{cases} \xi_j^{N(n+1)}(z_i) = (1-\omega)\xi_j^{N(n)}(z_i) + \omega \xi_j^N(z_i), \\ \xi_j^{C(n+1)}(z_i) = (1-\omega)\xi_j^{C(n)}(z_i) + \omega \xi_j^C(z_i), \end{cases}$$

where $\omega \in (0, 1)$ sets the speed at which we update the decision rule, and go back to step 4.

8. Once convergence is achieved, check that the residuals of the Euler equation are close enough to zero, in the sense that the intertemporal error an agent would make by using the approximation is small enough (see Judd 1992).

Proof of Proposition 1

The program of a borrowing bank writes

$$\max_{\phi_t} pR_t(1 + \phi_t) - \rho_t \phi_t$$

subject to

$$pR_t(1 + \phi_t) - \rho_t \phi_t \geq \rho_t,$$

$$\gamma(1 + \theta\phi_t) \leq \rho_t.$$

The participation constraint indicates that only banks with ability $p \geq \overline{p}_t \equiv \rho_t / R_t$ will borrow. We focus on this segment of the market. The problem simplifies to

$$\max_{\phi_t} pR_t(1+\phi_t) - \rho_t\phi_t$$

subject to

$$\gamma(1+\theta\phi_t) \leq \rho_t$$

for $p \geq \overline{p}_t$. Let us denote by ξ the Lagrange multiplier associated with the incentive constraint; the first-order conditions are then

$$pR_t - \rho_t = \gamma\theta\xi,$$

$$\xi(\rho_t - \gamma(1+\theta\phi_t)) = 0.$$

The result follows from ξ being strictly positive for all $p > \overline{p}_t$. QED

Proof of Proposition 3

The market-clearing condition (11a) in normal times together with the optimal demand for capital yields the normal times equilibrium corporate loan rate $R_t = z_t f_k(A_t) + 1 - \delta$. The interbank market freezes if and only if this normal time rate is below $\overline{R}$, that is, if and only if $z_t f_k(A_t) < \overline{R} + \delta - 1$, which can be restated as either $A_t \geq \overline{A}_t \equiv f_k^{-1}((\overline{R}+\delta-1)/z_t)$ or $z_t \leq \overline{z}_t \equiv (\overline{R}+\delta-1)/f_k(A_t)$. QED

Equations of the Model with Endogenous Labor Supply

$$y_t = z_t k_t^{\alpha} h_t^{1-\alpha} + (\gamma+\delta-1)(A_t - k_t), \tag{A1}$$

$$R_t = \alpha k_t^{-\nu(1-\alpha)/(\nu+\alpha)} z_t^{(1+\nu)/(\nu+\alpha)} \left(\frac{1-\alpha}{\vartheta}\right)^{(1-\alpha)/(\nu+\alpha)} + 1 - \delta, \tag{A2}$$

$$\left(c_t - \vartheta\frac{h_t^{1+\nu}}{1+\nu}\right)^{-\sigma} = \beta\mathbb{E}_t\left[\left(c_{t+1} - \vartheta\frac{h_{t+1}^{1+\nu}}{1+\nu}\right)^{-\sigma} r_{t+1}\right], \tag{A3}$$

$$h_t = \left[\frac{(1-\alpha)z_t}{\vartheta}\right]^{1/(\nu+\alpha)} k_t^{\alpha/(\nu+\alpha)}, \tag{A4}$$

$$\overline{A}_t \equiv [(1-\alpha)/\vartheta]^{1/\nu}[\alpha/(\overline{R}+\delta-1)]^{(\nu+\alpha)/\nu(1-\alpha)} z_t^{(1+\nu)/\nu(1-\alpha)}, \tag{A5}$$

$$i_t = \psi A_{t+1} - (1-\delta)A_t, \tag{A6}$$

$$y_t = c_t + i_t, \tag{A7}$$

$$\chi_t = (R_t - r_t)A_t - (R_t - \gamma)(A_t - k_t). \tag{A8}$$

If $A_t \leq \overline{A}_t$ (normal times):

$$k_t = A_t, \tag{A9a}$$

$$\frac{r_t}{R_t} = \int_{\overline{p}_t}^{1} p \frac{\mathrm{d}\mu(p)}{1 - \mu(\overline{p}_t)}, \tag{A10a}$$

$$\overline{p}_t = \frac{\rho_t}{R_t}, \tag{A11a}$$

$$R_t = \frac{\rho_t}{\mu^{-1}\left(\dfrac{\rho_t - \gamma}{\rho_t - (1-\theta)\gamma}\right)}, \quad \text{with } \rho_t > \overline{\rho}. \tag{A12a}$$

If $A_t > \overline{A}_t$ (crisis times):

$$k_t = A_t - \mu(\gamma/R_t)A_t, \tag{A9b}$$

$$\frac{r_t}{R_t} = \frac{\gamma}{R_t}\mu(\gamma/R_t) + \int_{\gamma/R_t}^{1} p \, \mathrm{d}\mu(p), \tag{A10b}$$

$$\overline{p}_t = \gamma/R_t, \tag{A11b}$$

$$\rho_t = \gamma. \tag{A12b}$$

A few comments are in order here. (i) During a crisis assets $A_t - k_t$ are stored for return γ. Since capital depreciates at rate δ, the value added of storage is $\gamma + \delta - 1$, as reflected in equation (1). (Remember that $\gamma + \delta - 1 > 0$.) (ii) The good market-clearing condition (7) is derived from Walras's law and the agents' budget constraints. Summing up the household budget constraint and the firm's profits, one gets

$$c_t + \psi A_{t+1} = z_t k_t^{\alpha} h_t^{1-\alpha} + (1-\delta)k_t + r_t A_t - R_t k_t + \chi_t,$$

which simplifies to

$$\begin{aligned} c_t + i_t &= y_t - (\gamma + \delta - 1)(A_t - k_t) + (1-\delta)(k_t - A_t) + r_t A_t - R_t k_t + \chi_t \Leftrightarrow \\ y_t + \chi_t &= c_t + i_t + (R_t - r_t)A_t - (R_t - \gamma)(A_t - k_t). \end{aligned}$$

It can be checked that

$$\chi_t = \int_{\overline{p}_t}^{1} (1-p)(1+\phi_t)R_t A_t \, \mathrm{d}\mu(p)$$

and that the resource identity reduces to $y_t = c_t + i_t$. In crisis times, by definition one has

$$\chi_t = \int_{\gamma/R_t}^{1} (1-p)R_t A_t \, \mathrm{d}\mu(p),$$

which simplifies to (using [A9b] and [A10b])

$$\begin{aligned}\chi_t &= [1-\mu(\gamma/R_t)]R_tA_t - A_t\int_{\gamma/R_t}^{1} pR_t\mathrm{d}\mu(p)\\ &= R_tk_t - r_tA_t + \gamma(k_t - A_t)\\ &= (R_t - r_t)A_t - (R_t-\gamma)(A_t-k_t).\end{aligned}$$

In normal times, by definition one has

$$\chi_t = \int_{\overline{p}_t}^{1}(1-p)(1+\phi_t)R_tA_t\ \mathrm{d}\mu(p),$$

which simplifies to (using [A9a] and [A10a] and the fact that in equilibrium $1 + \phi_t = 1/[1-\mu(\overline{p}_t)]$) $\chi_t = (R_t - r_t)A_t$.

References

Angeloni, Ignazio, and Ester Faia. 2013. "Capital Regulation and Monetary Policy with Fragile Banks." *J. Monetary Econ.* 60 (April): 311–24.

Bernanke, Ben. 2005. "The Global Saving Glut and the U.S. Current Account Deficit." Remarks, Sandridge Lecture, Virginia Assoc. Economists, Richmond, VA.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, vol. 1, edited by J. B. Taylor and M. Woodford. Amsterdam: Elsevier.

Bianchi, Javier. 2011. "Overborrowing and Systemic Externalities in the Business Cycle." *A.E.R.* 101 (December): 3400–3426.

Bianchi, Javier, and Enrique G. Mendoza. 2010. "Overborrowing, Financial Crises and 'Macroprudential' Policy." Working Paper no. 16091, NBER, Cambridge, MA.

Boissay, Frédéric. 2011. "Financial Imbalances and Financial Fragility." Working Paper no. 1317, European Central Bank, Frankfurt.

Boissay, Frédéric, Fabrice Collard, and Frank Smets. 2013. "Booms and Systemic Banking Crises." Working Paper no. 1514, European Central Bank, Frankfurt.

Borio, Claudio, and Mathias Drehmann. 2009. "Assessing the Risk of Banking Crises: Revisited." *Bank Internat. Settlements Q. Rev.* (March): 29–46.

Brunnermeier, Markus K., and Yuliy Sannikov. 2012. "A Macroeconomic Model with a Financial Sector." *A.E.R.* 104 (February): 379–421.

Burkart, Mike, and Tore Ellingsen. 2004. "In-Kind Finance: A Theory of Trade Credit." *A.E.R.* 94 (June): 569–90.

Caballero, Ricardo, Emmanuel Farhi, and Pierre-Olivier Gourinchas. 2008. "An Equilibrium Model of 'Global Imbalances' and Low Interest Rates." *A.E.R.* 98 (March): 358–93.

Christiano, Lawrence, Roberto Motto, and Massimo Rostagno. 2014. "Risk Shocks." *A.E.R.* 104 (January): 27–65.

Claessens, Stijn, Ayhan Kose, and Marco Terrones. 2008. "How Do Business and Financial Cycles Interact?" Working Paper no. WP/11/88, Internat. Monetary Fund, Washington, DC.

———. 2011. "What Happens during Recessions, Crunches and Busts?" Working Paper no. WP/08/274, Internat. Monetary Fund, Washington, DC.

Cooper, Russell, Douglas V. DeJong, Robert Forsythe, and Tom Ross. 1990. "Selection Criteria in Coordination Games: Some Experimental Results." *A.E.R.* 80 (March): 218–33.

De Fiore, Fiorella, and Harald Uhlig. 2013. "Corporate Debt Structure during the Financial Crisis." Manuscript, European Central Bank, Frankfurt.

Gertler, Mark, and Peter Karadi. 2011. "A Model of Unconventional Monetary Policy." *J. Monetary Econ.* 58 (January): 17–34.

Gertler, Mark, and Nobuhiro Kiyotaki. 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." In *Handbook of Monetary Economics*, vol. 3, edited by Benjamin M. Friedman and Michael Woodford, 547–99. Amsterdam: Elsevier.

———. 2013. "Banking, Liquidity and Bank Runs in an Infinite Horizon Economy." Manuscript, Princeton Univ.

Gorton, Gary B. 1988. "Banking Panics and Business Cycles." *Oxford Econ. Papers*, NS, 40 (December): 751–81.

———. 2010. Interview. *Region*, December.

———. 2012. "Some Reflections on the Recent Financial Crisis." Working Paper no. 18397, NBER, Cambridge, MA.

Gourio, François. 2012. "Disaster Risk and Business Cycles." *A.E.R.* 102 (October): 2734–66.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory Huffman. 1988. "Investment, Capacity Utilization, and the Real Business Cycle." *A.E.R.* 78 (June): 402–17.

Hahm, Joon-Ho, Hyun-Song Shin, and Kwanho Shin. 2011. "Non-core Bank Liabilities and Financial Vulnerability." Working Paper no. 18428, NBER, Cambridge, MA.

Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Oxford Univ. Press.

He, Zhiguo, and Arvind Krishnamurthy. 2012. "A Model of Capital and Crises." *Rev. Econ. Studies* 79 (April): 735–77.

Jaimovich, Nir, and Sergio Rebelo. 2009. "Can News about the Future Drive the Business Cycle?" *A.E.R.* 99 (September): 1097–1118.

Jermann, Urban J., and Vincenzo Quadrini. 2012. "Macroeconomic Effects of Financial Shocks." *A.E.R.* 102 (February): 238–71.

Jiménez, Gabriel, Steven Ongena, José-Luis Peydró, and Jesús Saurina. 2014. "Hazardous Times for Monetary Policy: What Do 23 Million Loans Say about the Impact of Monetary Policy on Credit Risk-Taking?" *Econometrica* 82 (March): 463–505.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2011. "Financial Crises, Credit Booms and External Imbalances: 140 Years of Lessons." *IMF Econ. Rev.* 59 (June): 340–78.

———. 2013. "When Credit Bites Back: Leverage, Business Cycles, and Crises." *J. Money, Credit, and Banking* 45 (December): 3–28.

Judd, Kenneth L. 1992. "Projection Methods for Solving Aggregate Growth Models." *J. Econ. Theory* 58 (December): 410–52.

Kaminsky, Graciela L., and Carmen M. Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance of Payments Problems." *A.E.R.* 89 (June): 473–500.

Koop, Gary, Hashem Pesaran, and Simon M. Potter. 1996. "Impulse Response Analysis in Nonlinear Multivariate Models." *J. Econometrics* 74 (September): 119–47.

Laeven, Luc, and Fabian Valencia. 2008. "Systemic Banking Crises: A New Database." Working Paper no. 08224, Internat. Monetary Fund, Washington, DC.

Lucas, Robert. 1987. *Models of Business Cycles*. New York: Blackwell.

Maddaloni, Angela, and José-Luis Peydró. 2011. "Bank Risk-Taking, Securitization, Supervision, and Low Interest Rates: Evidence from the Euro Area and US Lending Standards." *Rev. Financial Studies* 24 (June): 2121–65.

Mankiw, Gregory. 1986. "The Allocation of Credit and Financial Collapse." *Q.J.E.* 101 (August): 455–70.

Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. Oxford: Oxford Univ. Press.

Mendoza, Enrique G. 2010. "Sudden Stops, Financial Crises and Leverage." *A.E.R.* 100 (December): 1941–66.

Mendoza, Enrique G., Vincenzo Quadrini, and José-Víctor Ríos-Rull. 2007. "On the Welfare Implications of Financial Globalization without Financial Development." In *NBER International Seminar on Macroeconomics*, edited by Richard Clarida and Francesco Giavazzi. Chicago: Univ. Chicago Press.

———. 2009. "Financial Integration, Financial Development, and Global Imbalances." *J.P.E.* 117 (June): 371–416.

Mendoza, Enrique G., and Katherine A. Smith. 2005. "Quantitative Implications of a Debt-Deflation Theory of Sudden Stops and Asset Prices." *J. Internat. Econ.* 70 (September): 82–114.

Mendoza, Enrique G., and Marco Terrones. 2012. "An Anatomy of Credit Booms and Their Demise." Working Paper no. 670, Banco Central de Chile, Santiago.

Minsky, Hyman. 1977. "A Theory of Systemic Fragility." In *Financial Crises: Institutions and Markets in a Fragile Environment*, edited by Edward I. Altman and Arnold W. Sametz. New York: Wiley.

Perri, Fabrizio, and Vincenzo Quadrini. 2011. "International Recessions." Working Paper no. 17201, NBER, Cambridge, MA.

Reinhart, Carmen M., and Vincent Reinhart. 2008. "Capital Flow Bonanzas: An Encompassing View of the Past and Present." In *NBER International Seminar on Macroeconomics*, edited by Jeffrey Frankel and Christopher Pissarides. Chicago: Univ. Chicago Press.

Reinhart, Carmen M., and Kenneth Rogoff. 2009. *This Time Is Different—Eight Centuries of Financial Folly*. Princeton, NJ: Princeton Univ. Press.

———. 2013. "Banking Crises: An Equal Opportunity Menace." *J. Banking and Finance* 37 (November): 4557–73.

Schularick, Moritz, and Alan M. Taylor. 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *A.E.R.* 102 (April): 1029–61.

Shin, Hyun-Song. 2009. "Securitisation and Financial Stability." *Econ. J.* 119 (March): 309–32.

———. 2010. *Risk and Liquidity*. Clarendon Lecture in Finance. Oxford: Oxford Univ. Press.

Stiglitz, Joseph E., and Andrew Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *A.E.R.* 71 (June): 393–410.

Tauchen, George, and Robert Hussey. 1991. "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models." *Econometrica* 59 (March): 371–96.

Tirole, Jean. 2006. *The Theory of Corporate Finance*. Princeton, NJ: Princeton Univ. Press.

Uhlig, Harald. 2010. "A Model of a Systemic Bank Run." *J. Monetary Econ.* 57 (January): 78–96.

Varian, Hal. 1979. "Catastrophe Theory and the Business Cycle." *Econ. Inquiry* 17 (January): 14–28.