Foundations and Trends® in Finance
Vol. 10, No. 2 (2015) 113–180
© 2015 I. Goldstein and A. Razin
DOI: 10.1561/0500000049

now
the essence of knowledge

Three Branches of Theories of Financial Crises

Itay Goldstein
University of Pennsylvania, the Wharton School
itayg@wharton.upenn.edu

Assaf Razin
Tel Aviv University, Eitan Berglas School of Economics
razin@post.tau.ac.il

Contents

Abstract

In this monograph, we review three branches of theoretical literature on financial crises. The first deals with banking crises originating from coordination failures among bank creditors. The second deals with frictions in credit and interbank markets due to problems of moral hazard and adverse selection. The third deals with currency crises. We discuss the evolutions of these branches in the literature, and how they have been integrated recently to explain the turmoil in the world economy during the East Asian crises and in the last few years. We discuss the relation of the models to the empirical evidence and their ability to guide policies to avoid or mitigate future crises.

I. Goldstein and A. Razin. *Three Branches of Theories of Financial Crises.* Foundations and Trends® in Finance, vol. 10, no. 2, pp. 113–180, 2015.

DOI: 10.1561/0500000049.

1

Introduction

Financial and monetary systems are designed to improve the efficiency of real activity and resource allocation. Many empirical studies in financial economics provide evidence that financial development and economic growth and efficiency are connected; see, for example, Levine [1997] and Rajan and Zingales [1998]. In theory, financial institutions and markets enable the efficient transmission of resources from savers to the best investment opportunities. In addition, they also provide risk sharing possibilities so that investors can take more risk and advance the economy. Finally, they enable aggregation of information that provides guidance for more efficient investment decisions. Relatedly, monetary arrangements, such as the European Monetary Union (EMU), are created to facilitate free trade and financial transactions among countries, thereby improving real efficiency.

A financial crisis — marked, for example, by the failure of banks, the sharp decrease in credit and trade, and/or the collapse of an exchange rate regime — causes extreme disruption of the normal functions of financial and monetary systems, thereby hurting the efficiency of the economy. Unfortunately, financial crises have happened frequently throughout history and, despite constant attempts to eliminate them,

it seems unlikely that they will disappear in the future. Clearly, the last decade has been characterized by great turmoil in the world's financial systems. The meltdown of leading financial institutions in the US and Europe, the sharp decrease in lending and trading activities, and the ongoing challenge in the European Monetary Union exhibit ingredients from several types of financial crises in recent history: banking crises, credit and market freezes, and currency crises.[1]

Understanding the different types of financial crises and the connection between them poses a challenge for academics, policymakers, and practitioners. Are crises caused by problems in the economy or are they creating the problems? Are crises inevitable for economies that wish to maintain a high level of financial development? Can we think of an optimal mix of regulations that will achieve financial development without much exposure to crises? Or, are the crises themselves sometimes a result of regulation and intervention in financial markets? Most financial economists will probably agree that crises are related to panics and externalities and that some policy is needed to reduce their frequency and severity. But how big is the problem and how extensive should intervention be? Ongoing research is critical to gain a better understanding of the origins of crises and the optimal response to them.

Over the years, many theories have been developed to explain financial crises and guide policymakers in trying to prevent and mitigate them. In this monograph, we review models from three different branches of literature that have been developed more or less in parallel: banking crises and panics, credit frictions and market freezes, and currency crises. At a later stage, mainly following the East Asian crisis in the late 1990s, these literatures have become more integrated as the events in the real world proved that the different types of crises can occur simultaneously and amplify each other in different ways. Our monograph is not meant to be a comprehensive survey of the financial-crises literature. The literature is too big to be meaningfully covered in full in one survey. In fact, there is no consensus on what this literature includes as different people have different views on what constitutes a

[1]Many authors provide detailed descriptions of the events of the recent crisis. See, for example, Brunnermeier [2009] and Gorton [2010].

financial crisis. Instead, we attempt to present basic frameworks linked to the broad topic of financial crises and describe some of the ways in which they influenced the literature and relate to recent events. We also address some of the policy challenges and shed light on them using the analytical tools at hand. We hope that this survey will be helpful in highlighting the basic underlying forces that have been studied in the literature for over three decades in a simple and transparent way, and will be an easy and accessible source for the many economists who are now interested in exploring the topic of financial crises following the events of the last few years.

In Section 2, we review the literature on banking crises and panics. This literature is perhaps most directly linked to the concept of crises. Banks are known to finance long-term assets with short-term liabilities. One advantage of this arrangement is that it enables banks to provide risk sharing to investors who might face early liquidity needs. However, this also exposes the bank to the risk of a bank run, whereby many creditors decide to withdraw their money early. The key problem is that of a coordination failure, which stands at the root of the fragility of banking systems: when more depositors withdraw their money from a bank the bank is more likely to fail, and other depositors have a stronger incentive to withdraw. These strategic complementarities lead to either multiple equilibria or abrupt regime shifts, and support the view held by many economists that crises are sudden and unexpected events that have an element of panic [see Friedman and Schwartz, 1963, Kindleberger, 1978]. In this section, we describe the theoretical underpinnings behind bank runs and the lessons for policy analysis.

Banking systems have been plagued with bank runs throughout history; see, for example, Calomiris and Gorton [1991]. Policy lessons from the early 20th century led governments to insure banks, which substantially reduced the likelihood of bank runs. However, runs are still a prominent phenomenon behind financial crises. In East Asian and Latin American countries, many runs occurred in the last two decades. In the recent turmoil, a classic type of bank run was seen in the United Kingdom (UK) at Northern Rock Bank [see Shin, 2009] when investors were lining up in the street to withdraw money from

their accounts. There are many other examples of runs in the financial system as a whole. The repo market, in which investment banks get short-term financing, was subject to a run according to Gorton and Metrick [2012]. This led to the failure of leading financial institutions, such as Bear Stearns and Lehman Brothers. One can think of the credit squeeze in the repo market as a coordination failure among providers of capital, who refused to roll over credit, expecting deterioration in the value of collateral and in the ability of borrowers to pay due to the refusal of other lenders to roll over credit. This is similar to the models of bank runs caused by coordination problems that we review in this section. Others documented runs in money-market funds and in the asset-backed-commercial-paper market [see for example, Schmidt et al., 2015, Covitz et al., 2013, Schroth et al., 2014], which were in distress during the recent crisis.

While Section 2 emphasizes fragility of financial institutions due to coordination failures by their creditors, we review models that analyze frictions in loans extended by financial institutions and other lenders in Section 3. Broadly speaking, these are models of credit frictions and market freezes. Traditionally, the literature on this topic has developed without addressing crises per se, but more recently the basic mechanisms have increasingly been mentioned in connection to major events during financial crises. This literature highlights two key problems that create frictions in the flow of credit and trade. One problem is that of moral hazard. If a borrower has the ability to divert resources at the expense of the creditor, then creditors will be reluctant to lend to borrowers. Hence, for credit to flow efficiently from the creditor to the borrower, it is crucial that the borrower maintains "skin in the game", that is, that he has enough at stake in the success of the project, and so does not have a strong incentive to divert resources. This creates a limit on credit, which can be amplified when economic conditions worsen, leading to a crisis. Another problem is that of adverse selection. Looking at financial markets and credit markets, many are puzzled by the fact that they freeze despite the presence of gains from trade. Adverse selection generated by asymmetric information is a powerful force that can generate a freeze. In the presence of asymmetric information, traders

are reluctant to trade as they are concerned that they are getting a "lemon". Again, this may lead to a crisis if asymmetric information is very extreme.

There is ample empirical evidence that shows the importance of the kind of credit frictions as described in this section. For example, Gan [2007a,b] documents reduced lending and firm investment as firms' collateral value and banks' capital deteriorated following the collapse of the Japanese real estate market in the early 1990s. In the period leading to the recent crisis, Chaney et al. [2012] find that increased real estate values for companies were related to increases in firm borrowing and investing. It is generally not difficult to link such forces to the events of the recent crisis. The credit freeze that followed the financial meltdown of 2008, in which financial institutions were reluctant to lend money to operating firms, and the ensuing freeze in the flow of funds between financial institutions in the interbank markets both seem to be related to the amplification of economic shocks due to the frictions in credit provision, brought on by the principal-agent models that we review here. As economic conditions deteriorated, borrowers found themselves with less "skin in the game", and so lenders refused to provide credit to them. This, in turn, worsened the economic conditions of borrowers, amplifying the initial shock. Similarly, the potential increase in asymmetric information that followed the collapse of Lehman Brothers in 2008 may have contributed to a total market freeze, where investors were reluctant to trade in assets with each other due to the heightened uncertainty about the value of assets they trade.

Overall, the models of Sections 2 and 3 show fragility on both sides of the balance sheet of a financial institution. It seems that both types of fragility have been at work in recent crises, as we mention above. Importantly, such fragilities can reinforce each other. For example, creditors of a financial institution are more likely to panic and run when problems of moral hazard and asymmetric information reduce the value of its assets or make it more uncertain. A small problem on the asset side of a financial institution that would not be usually called a 'crisis' might then worsen due the mechanisms highlighted in the two sections, and turn into a crisis. This is how the models described in Section 3

are connected to the traditional crises literature in Section 2, and now all are used to describe parts of the system of interdependent forces that lead to the pronounced outcomes we see around times of financial crises. We elaborate more on this in Section 3.

Another literature that evolved independently is focused on currency crises. Traditionally, these were viewed as a separate phenomenon, unrelated to banking crises, but more recently the literatures have moved towards each other. In Section 4, we review models of currency crises. Many currency crises, such as the early 1970s breakdown of the Bretton Woods global system, originate from the desire of governments to maintain a fixed exchange rate regime, which is inconsistent with other policy goals such as free capital flows and flexible monetary policy. This might lead to the sudden collapse of the regime. Like in the bank-run literature, coordination failures play an important role here. When the central bank tries to maintain a fixed exchange rate regime, it might decide to abandon it under pressure from speculators. Speculators then find themselves in a coordination problem, and they attack the regime if and only if they believe others will do so too. In such coordination failures, the event of a currency crisis becomes a self-fulfilling belief. This is also similar to debt crises, where the government may decide to default under pressure from creditors. Then, creditors are facing a coordination problem, and they liquidate their bond holdings if and only if they expect that others will liquidate their claims. Consequently, a debt crisis becomes a self-fulfilling expectation.

Such models are highly relevant to the current situation in the European Monetary Union. In the basis of the theory of currency crises is the famous international-finance trilemma, according to which a country can choose only two of three policy goals: free international capital flows (benefitting international risk sharing), monetary autonomy (the ability to employ monetary policy tools to stabilize inflation and output fluctuations), and the stability of the exchange rate (bringing about a reduction in transaction costs associated with trade and investment). Countries in the Euro zone now realize that in their attempt to achieve the first and third goal, they have given up on the second goal, and therefore have a limited ability to absorb the shocks in

economic activity and maintain their national debts. This is triggered by the global financial crisis. Coordination problems among investors and currency speculators aggravate this situation, and may have an important effect on whether individual countries in Europe are forced to default and/or leave the monetary union.

While the traditional literature on currency crises focused on the government alone, we review the 'third-generation' models of currency crises in Section 4.3. These models essentially connect models of banking crises and credit frictions (reviewed in Sections 2 and 3) with traditional models of currency crises (reviewed in Subsections 4.1 and 4.2). Such models were motivated by the East Asian crises of the late 1990s in which financial institutions and exchange rate regimes collapsed together, demonstrating the linkages between governments and financial institutions that can expose the system to further fragility. This is again relevant for the current situation in Europe as banks and governments are intertwined, and the fragility of the system depends to a large extent on the connections between them. We elaborate on this in Section 4.3.

2

Banking Crises and Panics

Depository institutions are inherently unstable because they have a mismatch in the maturity structure between their assets and liabilities. In particular, they finance long-term investments with short-term deposits. This puts banks at risk of bank runs because when many depositors demand their money in the short term, banks will have to liquidate long-term investments at a loss, leading to their failure. This can lead to a self-fulfilling prophecy because depositors believe that they are better off withdrawing their money if they believe others will do so as well, therefore the mere belief that a bank run will occur causes a bank run to occur.

Diamond and Dybvig [1983][1] provide a classic framework capturing this phenomenon. They first provide a rationale for why banks expose themselves to this maturity mismatch. In their model, agents might suffer idiosyncratic short-term liquidity needs. By offering demand-deposit contracts, banks enable short-term consumers to enjoy the fruits of long-term investments. Banks rely on the fact that only a forecastable fraction of agents will need to consume early, and thus offer a contract

[1] Another important paper on the topic from that period is Bryant [1980].

that transfers consumption from the long-term consumers to the short-term consumers. This way, banks enable risk sharing among agents who *ex ante* do not know whether they will have early liquidity needs or not. However, the contract may also lead to a catastrophic bank run.

2.1 Diamond-Dybvig economy

We now provide a formal description of an economy based on Diamond and Dybvig [1983]. The version here follows Goldstein and Pauzner [2005], which enables us to talk about equilibrium selection and policy implications.

There are three periods (0, 1, 2), one good, and a continuum [0, 1] of agents. Each agent is born in period 0 with an endowment of one unit. Consumption occurs only in period 1 or 2 (c_1 and c_2 denote an agent's consumption levels). Each agent can be one of two types: (1) with probability λ the agent is impatient and (2) with probability $1-\lambda$ she is patient. Agents' types are independent and identically distributed; we assume no aggregate uncertainty. Agents learn their types (which are their private information) at the beginning of period 1. Impatient agents can consume only in period 1, and they obtain a utility of $u(c_1)$. Patient agents can consume in both period 1 and 2, and their utility is $u(c_1 + c_2)$. Function u is twice continuously differentiable, increasing, and for any $c \geq 1$ has a relative risk-aversion coefficient, $-cu''(c)/u'(c)$, greater than 1. Without loss of generality, we assume that $u(0) = 0$.[2]

Agents have access to a productive technology that yields a higher expected return in the long run. For each unit of input in period 0, the technology generates one unit of output if liquidated in period 1. If liquidated in period 2, the technology yields R units of output with probability $p(\theta)$, or 0 units with probability $1 - p(\theta)$. Here, θ is the state of the economy. It is drawn from a uniform distribution on [0, 1], and is unknown to agents before period 2. We assume that $p(\theta)$ is

[2]Note that any von Neumann-Morgenstern utility function, which is well defined at 0 (that is, $u(0) \neq -\infty$), can be transformed into an equivalent utility function that satisfies $u(0) = 0$.

strictly increasing in θ. It also satisfies

$$E_\theta[p(\theta)]u(R) > u(1).$$

In autarky, impatient agents consume one unit in period 1, whereas patient agents consume R units in period 2 with probability $p(\theta)$. A transfer of consumption from patient agents to impatient ones could be beneficial, *ex-ante*, to all agents; although it would necessitate the early liquidation of long-term investments. A social planner who can verify agents' types, once realized, would set the period-1 consumption level c_1 of the impatient agents so as to maximize an agent's ex-ante expected welfare as

$$\lambda u(c_1) + (1-\lambda)u\left(\frac{1-\lambda c_1}{1-\lambda}R\right)E_\theta[p(\theta)].$$

Here, $\lambda\, c_1$ units of investment are liquidated in period 1 to satisfy the consumption needs of impatient agents. As a result, in period 2, each one of the patient agents consumes an amount of $\frac{1-\lambda c_1}{1-\lambda}R$ with probability $p(\theta)$.

The first best period-1 consumption c_1^{FB} is set to maximize this *ex-ante* expected welfare. It can be shown that $c_1^{FB} > 1$, i.e., the consumption available in period 1 to impatient consumers exceeds the endowment. Hence, at the first best allocation, there is risk sharing, which is achieved via maturity transformation: a transfer of wealth from patient agents to impatient ones.

Without a social planner, risk sharing can be achieved via a banking sector. Suppose agents put their money in the bank at date 0 in exchange for a demand-deposit contract that enables them to withdraw at date 1 or at date 2. The bank can set the payoff to early withdrawal r_1 at the first best level of consumption, c_1^{FB}. If only impatient agents demand early withdrawal, the expected utility of patient agents is $E_\theta[p(\theta)] \cdot u(\frac{1-\lambda r_1}{1-\lambda}R)$. As long as this is more than the utility from withdrawing early $u(r_1)$, there is an equilibrium in which, indeed, only impatient agents demand early withdrawal. In this equilibrium, the first best allocation is obtained.

However, as Diamond and Dybvig point out, the demand-deposit contract makes the bank vulnerable to runs. There is a second equilibrium in which *all* agents demand early withdrawal. When they do so,

period-1 payment is now r_1 with probability $1/r_1$, and period-2 payment is 0; so that it is indeed optimal for agents to demand early withdrawal. This equilibrium is evidently inferior to the autarkic regime. The reason for multiplicity of equilibria, as in other models featuring multiple equilibria in economics and finance, is the strategic complementarities among agents: it is optimal for them to run if they think that others are going to run.

Table 2.1 describes the payments expected by agents when they withdraw at period 1 and at period 2 as a function of the proportion n of agents (between 0 and 1) who decide to withdraw at period 1. Looking at the table, it is easy to see that, under the assumptions discussed earlier, there is an equilibrium with no run ($n = 0$) and an equilibrium with a run ($n = 1$).

Many perceive the multiplicity of equilibria to be a strength of the model, since it seems to capture the fragility of banks and the element of surprise in financial crises. However, it poses two major problems for researchers and policymakers. First, the model provides no prediction as to when a bank run is more likely to occur. This

Table 2.1: *Ex-post* payments to agents in a model of bank runs.

Period	$n < 1/r_1$	$n \geq 1/r_1$
1	r_1	$\begin{cases} r_1 & \text{prob } \frac{1}{nr_1} \\ 0 & \text{prob } 1 - \frac{1}{nr_1} \end{cases}$
2	$\begin{cases} \frac{(1-nr_1)}{1-n}R & \text{prob } p(\theta) \\ 0 & \text{prob } 1 - p(\theta) \end{cases}$	0

The table is based on Goldstein and Pauzner [2005] and describes the payments agents expect to get when demanding their money at period 1 vs. period 2. Here, n is the proportion of agents who demand their money at period 1; r_1 is the promised return to agents at period 1; R is the return that the bank's asset yields at period 2 in case it is successful, and $p(\theta)$ is the probability it will be successful.

is in contrast to the vast empirical evidence that financial crises are linked to various variables that capture the strength of fundamentals of the banking system (see, for example, Gorton [1988] and Demirguc-Kunt and Detragiache [1998]; for a recent review, see Goldstein [2012]). Second, policy analysis becomes quite difficult with multiple equilibria. If a policy measure is intended to reduce the likelihood of bank runs but also has other costs, then assessing the desirability of this policy measure becomes impossible if the likelihood of bank runs cannot be pinned down (with or without the policy measure in place).

2.2 Heterogeneous signals and unique equilibrium

The global-games literature offers a solution to the problems mentioned earlier while still maintaining the element of panic in financial crises. The literature was pioneered by Carlsson and van Damme [1993], and then applied to financial crises in the context of currency attacks by Morris and Shin [1998]. It assumes that agents observe noisy signals of the fundamentals of the economy, which leads to a unique equilibrium where the fundamentals uniquely determine whether a crisis will occur or not. Goldstein and Pauzner [2005] build on this literature in the context of bank runs and derive a unique equilibrium.

Technically, the proof of uniqueness in Goldstein and Pauzner [2005] is quite different from the proof employed in the rest of the global-games literature. This is due to the nature of payoffs in the bank run model, which violates a central assumption in the global-games framework. Specifically, in traditional global-games models, an agent's incentive to take a certain action monotonically increases in the proportion of other agents taking this action.[3] As one can see in Table 2.1, this does not hold in the bank run model since in the region where the bank is bankrupt, the net benefit from running decreases when more people run. Goldstein and Pauzner [2005] overcome this problem and nevertheless show uniqueness under some conditions. For the purpose of our review, we will not get into these complexities here, but rather

[3]This property is referred to as "Global Strategic Complementarities".

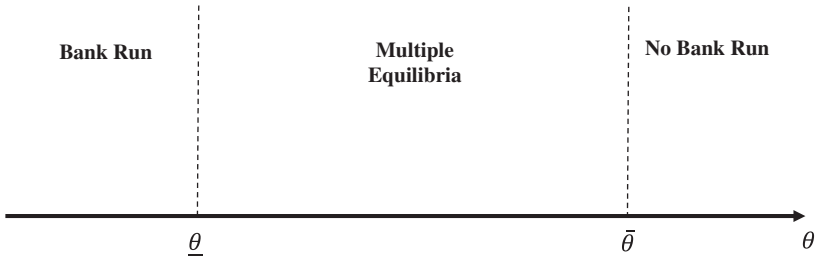

Figure 2.1: Bank runs with common knowledge. Tripartite classification of the fundamentals (from Goldstein [2012], based on Morris and Shin [1998] and Goldstein and Pauzner [2005]).

just briefly describe the intuition behind the traditional global-games framework and how it generates a unique equilibrium.

If the realization of the fundamental θ is common knowledge to agents before they make their choice to run or not, the model of Goldstein and Pauzner [2005] generates three regions of the fundamentals. These are depicted in Figure 2.1. Below a threshold $\underline{\theta}$, there is a unique equilibrium where all depositors — patient and impatient — run on the bank and demand early withdrawal. Here, the fundamentals are so low that the bank is insolvent and will fail no matter what other depositors do. Hence, each depositor undoubtedly finds it profitable to withdraw. Above a threshold $\bar{\theta}$, there is a unique equilibrium where patient depositors do not withdraw.[4] Here, the fundamentals are so high that the bank can survive and pay its liabilities even if all depositors demand early withdrawal. Hence, they choose not to withdraw. Between $\underline{\theta}$ and $\bar{\theta}$, there are multiple equilibria. Either everyone runs and the bank fails, or only impatient agents withdraw and the bank remains solvent. There are strategic complementarities, since depositors benefit from the run if, and only if, other depositors run, and therefore there are two possible equilibria.

However, introducing noise in speculators' information about the fundamental θ, such that every depositor gets a signal composed of the

[4]This upper dominance region is obtained with an additional assumption introduced by Goldstein and Pauzner [2005].

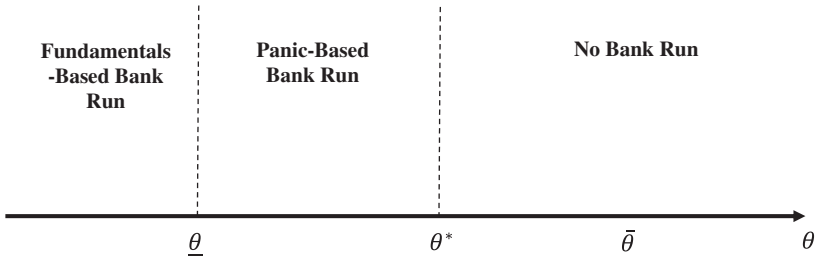

Figure 2.2: Equilibrium outcomes in a bank-run model with non-common knowledge (from Goldstein [2012], based on Morris and Shin [1998] and Goldstein and Pauzner [2005]).

true fundamental θ plus independent and identically distributed noise, changes the predictions of the model dramatically (even if the noise is very small). The new predictions are depicted in Figure 2.2. Now, the intermediate region between $\underline{\theta}$ and $\bar{\theta}$ is split into two sub-regions: below θ^*, a bank run occurs and the bank fails, whereas above it, there is no run and the bank remains solvent.[5]

This result can be best understood by applying the logic of a backward induction.[6] Due to the noise in patient depositors' information about θ and due to the presence of strategic complementarities, their decisions about whether to withdraw will be affected by what the signal conveys about other depositors' signals. Hence, a depositor observing a signal slightly below $\bar{\theta}$ knows that many other depositors may have observed signals above $\bar{\theta}$ and therefore choose not to run. Taking this into account, this depositor also chooses not to run. Then we know that depositors who receive signals just below $\bar{\theta}$ do not run on the bank. Applying the same logic, depositors who receive even lower signals also choose not to run. This logic can be repeated again and again, establishing a boundary well below $\bar{\theta}$, above which depositors do not run on

[5]This sharp outcome is obtained when the noise in the signal approaches zero. For larger noise, the transition from run to no-run will not be so abrupt. Rather there will be a range of partial run. This does not matter for the qualitative message of the theory.

[6]Strictly speaking, this intuition holds for the traditional global-games framework where global strategic complementarities hold. The intuition in the bank-run model of Goldstein and Pauzner [2005] is more involved.

the bank. The same logic can then be repeated from the other direction, establishing a boundary well above $\underline{\theta}$, below which depositors do run on the bank. The mathematical proof shows that the two boundaries coincide at a unique θ^*, such that all depositors run below θ^*, and do not run above θ^*.

Figure 2.2 shows that in the range between $\underline{\theta}$ and $\bar{\theta}$, the level of the fundamentals now perfectly predicts whether or not a crisis occurs. In particular, a crisis will surely occur below θ^*. We refer to crises in this range as panic-based because a crisis in this range is not necessitated by the fundamentals; it occurs only because agents think it will occur, and in that sense it is self-fulfilling. However, the occurrence of a self-fulfilling crisis here is uniquely pinned down by the fundamentals. So, in this sense, the panic-based approach and the fundamental-based approach are not inconsistent with each other. The key is that the fundamentals uniquely determine agents' expectations about whether a crisis will occur or not, and in doing this, they indirectly determine whether a crisis occurs. Agents' self-fulfilling beliefs amplify the effect of fundamentals on the economy. Similarly, even though the fundamentals between θ^* and $\bar{\theta}$ could support a crisis, it does not occur since agents' expectations are coordinated on the no-crisis outcome. Note that crises below $\underline{\theta}$ can be thought of as purely fundamental-based because here the fundamentals are so bad that crises must occur no matter what agents think other agents are going to do; that is, it is their dominant action to run.

The global-games approach produces empirical predictions that are consistent with the vast empirical literature that links the occurrence of a crisis to fundamental variables. A classic reference in this context is Gorton [1988]. Studying the national banking era in the US between 1863 and 1914, he shows that crises were depositors' responses to an increase in perceived risk. He demonstrates that crises occurred whenever key variables that are linked to the probability of recession reached a critical value. The most important variable is the liabilities of failed firms. He also shows effects of other variables, such as the production of pig iron, which he uses as a proxy for consumption. When the perceived risk of recession based on these variables becomes high,

depositors believe that their deposits in banks, which have claims in firms, become too risky. Hence, they demand early withdrawal, leading, in aggregate, to mass withdrawals.

Similar evidence is obtained in international studies that attempt to understand what brings down a whole banking sector. A banking crisis in this literature is manifested by large withdrawals out of the banking system leading to bank closures, government's aid to banks, or suspension of convertibility. For example, Demirguc-Kunt and Detragiache [1998] conduct an international study to understand the determinants of banking crises in a sample of developing and developed economies in the late 20th century (1980–1994). They again find that a number of variables connected to the fundamental state of the economy are related to the occurrence of crises. The key predictors in their study are: low gross domestic product (GDP) growth (reflecting declining economic activity that reduces the value of banks' assets), high real interest rates and inflation (inducing banks to offer higher deposit rates while the rates on their loans are fixed given that they are mostly long-term loans), and high levels of outstanding credit (making the banking system fragile).

Other authors have shown that bank-specific variables have an effect on the withdrawals from specific banks. One example is Schumacher [2000], who conducts her study around the runs on Argentine banks following the devaluation of the Mexican currency in December 1994. The devaluation in Mexico was of significance to the Argentine banks because it led to speculation that Argentina would also have to devalue its currency. Schumacher conjectures that depositors' runs were triggered by information they had about the ability of banks to survive the currency collapse, and that according to this information they transferred money from banks they considered "bad" to banks they considered "good". She finds evidence in support of this conjecture. Similarly, Martinez-Peria and Schmukler [2001] analyze the behavior of depositors in Argentina, Chile, and Mexico over two decades in the late 20th century, and show that depositors' behavior is affected by banks' risk characteristics. Deposits decrease and interest rates rise in banks with low ratio of capital to assets, low return on assets, high

level of non-performing loans, and high ratio of expenditures to assets. Finally, Calomiris and Mason [2003] study banking crises during the Great Depression. They show that bank specific variables — such as leverage, asset risk, and liquidity — affect the likelihood of failure, and so do variables that capture the local or regional economic situation.

It is important to note again that even though in the global-games approach the occurrence of crises is determined by fundamentals, this approach still maintains the flavor of panic or self-fulfilling beliefs that emerges from the Diamond-Dybvig model. An alternative line of models describes banking crises as a result of bad fundamentals only. See, for example, Jacklin and Bhattacharya [1988], Chari and Jagannathan [1988] and Allen and Gale [1998]. As we will discuss in the next subsection, the distinction between fundamental-based and panic-based crises is important for policy analysis.

An important question is how to provide empirical validation for the existence of panic and self-fulfilling beliefs in real-world crises. In the past, authors interpreted the evidence of the link between fundamentals and crises to go against theories of panic and self-fulfilling beliefs (for example, Gorton [1988]), but given the results of the global-games literature described here, this conclusion is clearly flawed. Two recent papers attempt to identify the role of panic and strategic complementarities more directly. Chen et al. [2010] identify the effect of strategic complementarities in outflows from mutual funds by showing that the sensitivity of outflows to bad performance is stronger in funds that exhibit stronger strategic complementarities. Hertzberg et al. [2012] use a natural experiment from Argentina to show that the release of public information makes banks react to information they already had, essentially because they expect other banks to react to it. The use of such methodologies in more traditional crises datasets can prove useful for our understanding of the role that strategic complementarities and panic may have in such crises.[7]

Another appealing feature of the global-games solution is that the equilibrium in the global-games model captures the notion of strategic

[7]See Goldstein [2012] for a review of the empirical literature and a discussion of strategies to identify strategic complementarities.

risk. Depositors who observe signals near the threshold where the bank fails, who ultimately determine the likelihood of a run, are not sure about how many people are going to run and whether the bank will fail. This strategic risk is of course very realistic; although it is missing from the multiple-equilibria framework where in equilibrium agents know for sure how many people run and whether the bank will survive.

Another advantage of pinning down a unique equilibrium is that it enables the researcher to compute the probability of a run and relate it to the terms of the banking contract.[8] Questions to ask are: do demand deposit contracts improve welfare even when their destabilizing consequences are taken into account? And how will they be designed in light of their effect on fragility? Goldstein and Pauzner [2005] show that banks become more vulnerable to bank runs when they offer a higher level of risk sharing. That is, the threshold θ^*, below which a run happens, is an increasing function of the short-term payment offered to depositors r_1. However, even when this destabilizing effect is taken into account, banks still increase welfare by offering demand deposit contracts, provided that the range of fundamentals where liquidation is efficient is not too large. Characterizing the short-term payment in the banking contract chosen by banks that takes into account the probability of a run, Goldstein and Pauzner show that this payment does not exploit all possible gains from risk sharing since doing so would result in too many bank runs. Still, in equilibrium, panic-based runs occur as a result of coordination failures among bank depositors. This leaves room for government policy to improve overall welfare.[9]

2.3 A basis for micro policy analysis

One of the basic policy remedies to reduce the loss from panic-based runs is the introduction of deposit insurance by the government. This

[8] Cooper and Ross [1998] study the relation between the banking contract and the probability of bank runs in a model where the probability of bank runs is exogenous.

[9] Note that the Goldstein–Pauzner model only focuses on demand deposit contracts and asks whether they improve welfare and how much risk sharing they should provide. Outside the global-games framework, there are papers that study a wider variety of contracts; for example, Green and Lin [2003], Peck and Shell [2003], and Ennis and Keister [2009].

idea goes back to Diamond and Dybvig [1983], where the government promises to collect taxes and provide liquidity (or bailout money) to the bank in case the bank faces financial distress (that is, when the number of agents demanding early withdrawal n exceeds the number of impatient agents λ).

In the context of the model described earlier, with deposit insurance, patient agents know that if they wait, they will receive the promised return independently of the number of agents who run. Hence, panic-based runs are prevented as patient agents withdraw their deposits only when this is their dominant action; that is, when θ is below $\underline{\theta}(r_1)$, rather than below the higher threshold $\theta^*(r_1)$. Then, in many cases, federal deposit insurance deters bank runs without the need to exercise the liquidity enhancing power. Extending the context of the model, Keister [2015] highlights another benefit of deposit insurance: it helps providing a better allocation of resources by equating the marginal utility that agents derive from private consumption and public-good consumption. That is, when bank runs occur, private consumption decreases, generating a gap between the marginal utility of private consumption and that of public-good consumption. So, with bailouts, the government can reduce the public good and increase private consumption to correct the distortion.

However, deposit insurance also has a drawback, like any insurance it creates moral hazard: when the bank designs the optimal contract, it does not internalize the cost of the taxes that might be required to pay the insurance. Thus, the bank has an incentive to over-exploit the deposit insurance by setting r_1 higher than the socially optimal level. This drawback of deposit insurance is consistent with the critique made by Calomiris [1990] that "today's financial intermediaries can maintain higher leverage and attract depositors more easily by offering higher rates of return with virtually no risk of default". In the context of the model, this is costly as it increases the lower threshold $\underline{\theta}(r_1)$, below which crises occur even without a coordination failure.

The framework developed earlier enables one to compare the benefits and costs of deposit insurance, and provide policy recommendations regarding the optimal amount and design of this insurance. As

mentioned above, the unique equilibrium coming out of the global-games framework enables the researcher to pin down the likelihood of a crisis and analyze the effect of deposit insurance on it. One can then compare the benefit of deposit insurance due to the reduction in the probability of panic-based runs to the cost of the moral hazard that leads to an increase in government expenditure to help banks and potentially an increase in the probability of fundamental-based runs. In a recent paper, Allen et al. [2014] use the global-games framework to conduct such analysis of optimal deposit insurance policy.[10]

While deposit insurance was enacted for banks and was effective in reducing the likelihood of traditional bank runs, there are many sectors of the financial system — such as money market funds and repo markets — that are uninsured, and in which massive runs have occurred in recent years.[11] The institutional details of these parts of the financial system are quite different from those of traditional commercial banks. For example, in the repo market, a lender is secured by collateral, and so the run of other lenders does not have a direct effect on how much he can get back from the borrower. Considering market forces, strategic complementarities can however still emerge: a fire sale of assets used as collateral, due to a systemic run, can lead to the reduction in the value of collateral, making quick withdrawal the optimal action for each lender (as he may be able to avoid the effect of the expected decrease in the value of collateral). A recent paper by Martin et al. [2014] adapts the traditional theories of runs to these new settings to discuss when runs will emerge and when they will not. Overall, given the proven fragility of these parts of the financial system, there is room to consider optimal insurance and regulation for them in light of the tradeoffs described here.

[10]Keister [2015] analyzes the optimal deposit insurance policy without employing the global-games methodology (and thus without pinning down the probability of a crisis) by checking the effect that the policy has on the range of fundamentals where a run may occur.

[11]See, for example, Gorton and Metrick [2012], Schmidt et al. [2015], Covitz et al. [2013], and Schroth et al. [2014].

2.4 Why debt contracts? The reasons behind bank fragility

One of the main criticisms of the Diamond and Dybvig [1983] model was raised by Jacklin [1987]. He showed that demand-deposit contracts offered by banks are not necessary for achieving the optimal allocation. Instead, this can be achieved in a market, where patient agents and impatient agents trade after finding out their types. In such a market solution, fragility does not arise. Hence, one can say that demand deposit contracts and bank runs will only occur in a model with limitations on trading.

Such limitations on trading may be quite pertinent in the real world. Indeed, many households stay away from financial markets for various reasons. Diamond [1997] formally analyzes the implications of such limitations. He studies an environment where some agents have access to financial markets and others do not. While the reason for this feature is not modeled, one can think of it as a result of differences in sophistication across investors. In such a model, banks will naturally emerge as part of the optimal solution, and so will the fragility and runs they bring with them.

However, the question of why banks are so highly leveraged and why they expose themselves to such high fragility remains open and deserves attention. A line of literature led by Calomiris and Kahn [1991] and Diamond and Rajan [2001] argues that bank fragility generated by high leverage is desirable in a framework with an agency problem between bank managers and outsiders. In such a framework, managers might act in their own private benefit. The fact that depositors can run at the sign of bad news disciplines the managers and prevents them from expropriating value from outsiders.

Gorton and Pennacchi [1990] argue for a different reason for debt contracts in banks: the fact that such contracts are not very sensitive to information makes them desirable for uninformed investors who might need to sell securities to informed investors. In this sense, ignorance is bliss: debt contracts do not incentivize information production and thus attract investors who do not want to expose themselves to information asymmetry.

However, various explanations for bank debt have a hard time explaining why banks are so much more leveraged than non-financial firms. For example, monitoring needs are not so different between banks and non-financial firms. This observation led Admati and Hellwig [2013] and others to argue that there is a strong element of moral hazard in bank leverage. They argue that one thing that distinguishes banks from most non-financial firms is the fact that their debt is guaranteed either in the form of explicit guarantees, such as deposit insurance, or implicit guarantees, such as expected bailouts for large and interconnected financial institutions. Knowing that their debt is guaranteed by the government implies that banks can borrow more cheaply, and incentivizes them to have excessive leverage as the cost is imposed on the government.[12]

The policy conclusion that is advanced by Admati and Hellwig [2013] is that banks should be subject to much higher capital requirements than they are today; that is, they should be required to be financed by much more equity. This is because the government cannot credibly commit to abandon the guarantees, and so the only way to undo the distortion that such guarantees generate for banks' capital structure is to require them to hold more equity.

On the other hand, others argue that banks are very different from non-financial firms and that the very nature of their business involves liquidity creation, which amounts to having a lot of debt on their balance sheets. For example, the type of liquidity creation described by Gorton and Pennacchi [1990] inherently involves the creation of securities that are not very sensitive to information. As this is one role of banks, it means that the nature of their business involves having debt, and this differentiates them very clearly from non-financial firms. More recently, different authors have extended this line of thinking and state that investors not only demand liquid assets but actually demand safe assets, and this demand is satisfied by banks (see Stein [2012] and a recent survey by Perotti and Golec [2015]). DeAngelo and Stulz [2015] provide a model that rationalizes banks' choice of high leverage in a

[12] Another government subsidy to debt is the tax subsidy, but this is common to both financial firms and non-financial firms.

world with demand for liquidity and without any moral-hazard or tax distortions.

The debate about bank capital is likely to continue in the coming years. The academic literature has an important role in informing this debate. The argument in the literature that banks are inherently different from non-financial firms due to their liquidity-creation role is indeed convincing, and so it is probably not very informative to compare the leverage ratios of the two. However, the literature has not made a very compelling case for why banks require such high leverage ratios as they have now to perform the role of liquidity creation. Even if they need to have debt to perform this role, it does not preclude raising equity as well. There is probably a significant moral-hazard component in the current leverage ratios, and increasing capital requirements to some extent is therefore warranted. Ultimately, setting the numbers for capital requirements has to rely on more quantitative work based on the microfoundations described here and new ones that hopefully will be developed in the future.

Importantly, work on bank capital should also connect to the policy issues discussed in the previous subsection in order to understand optimal guarantees by the government in light of the tradeoff between panic reduction and moral hazard that might increase fundamental fragility. In some sense, the analysis of the onset of a run, using the global-games framework as described in Section 2.2, is not so much affected by the source of the vulnerability. The global-games analysis, which helps us pin down the likelihood of a crisis can be employed in interaction with the analysis of the sources of debt and fragility to get the full picture.

2.5 Contagion and systemic risk

An important reason for concern with banking crises is that they spread across banks, leading many to fail at the same time, and hence creating systemic risk. Much has been written on contagion of banking crises, highlighting the different sources for spillovers and coordination among banks. Allen and Gale [2000b] and Lagunoff and Schreft [2001] show how contagion arises due to bank inter-linkages. Banks

facing idiosyncratic liquidity needs insure each other and so provide efficient risk sharing. However, this creates links across banks, leading to spillover of shocks and contagion of crises. Dasgupta [2004] extends their model, using the global-games framework described earlier to analyze the optimal insurance contracts among banks, taking into account their undesirable implications for contagion. In Goldstein and Pauzner [2004], contagion is generated due to a common pool of investors investing in different banks. The failure of one bank leads investors to lose wealth and become more risk averse. As a consequence, they are more likely to run on the other bank. Kyle and Xiong [2001] and Kodres and Pritsker [2002] analyze related models in which contagion across assets is generated by the portfolio rebalancing made by investors who hold the different assets.

Some authors analyze contagion as a result of transmission of information. In these models, a crisis in one market/bank reveals some information about the fundamentals in the other, and may thus induce a crisis in the other market/bank as well (examples include King and Wadhwani [1990] and Chen [1999]). Calvo and Mendoza [2000] suggest that the high cost of gathering information on each and every market may induce rational contagion. Recently, Oh [2013] analyzed a model of contagion in which investors learn about other investors' types and points out that this can be a source of contagion.

Another source of systemic risk is the "too big to fail" problem. Banks that become too big pose a large threat to the economy if they fail, and so governments are willing to provide a bail out to prevent this from happening. This, in turn, generates disincentives so that the bank will take on excessive risk knowing that the consequences will be borne by the taxpayer. Similarly, the government might be particularly concerned about the possibility of several banks failing together due to the adverse implications this might have on the economy. Therefore, the government will bail out banks only when many of them are about to fail. As pointed out by Acharya and Yorulmazer [2007] and Farhi and Tirole [2012], this might provide incentives for banks to choose correlated risks *ex ante*, which leads to correlated failures and destabilizes the system as a whole.

3

Credit Frictions and Market Freezes

In the models of financial-institution failures that we discussed, the returns on assets and loans held by the bank were generally assumed to be exogenous, and the focus was on the behavior of depositors or creditors of the banks. However, problems in the financial sector often arise from the other side of the balance sheet. The quality of loans provided by the banks is determined in equilibrium by the behavior of the bank and the behavior of its borrowers. Moral hazard between the bank and its borrowers and between the bank and its lenders affects the amount of bank lending and its return. This can lead to frictions in the flow of credit in the economy.

The forces in this literature can be traced back to Stiglitz and Weiss [1981], who provide a basic rationale for the presence of credit rationing, which is a common phenomenon in financial crises. While basic economic theory suggests that in equilibrium prices adjust so that supply equals demand and no rationing arises, they show that this will not occur in the credit market because of the endogeneity of the quality of the loan. There are two key frictions. The first one is moral hazard: if borrowers are charged a very high cost for credit, they lose the incentive to increase the value of their projects, and are therefore less likely

capable of paying back. The second one is adverse selection: if interest rates are high, only borrowers with bad projects will attempt to get loans, and again the bank is unlikely to get the money back.

For these reasons, banks will ration credit, hampering the effectiveness of the financial system in providing capital to those who need it, and, in extreme cases, leading to a financial crisis, where credit drops dramatically. As mentioned in the introduction, credit rationing and credit freeze have been a very important part of the recent financial crisis, as lending to firms and households decreased sharply (see, for example, Ivashina and Scharfstein [2010]). In this section, we review basic theories of this kind.

3.1 Moral hazard

When an entrepreneur borrows money to finance a project, he can take actions that reduce the value of the project and increase his own private benefits. Hence, a lender needs to make sure that the entrepreneur has a large enough incentive to preserve (or improve) the quality of the project, which will enable him to repay the loan. A direct implication is that the entrepreneur has to have a large enough stake in the investment or he has to be able to secure the loan with collateral. These considerations limit the amount of credit available to firms. They can lead to amplification of shocks to fundamentals and ultimately to financial crises.

Holmstrom and Tirole [1997] provide a canonical representation of this mechanism. In their model, there is a continuum of entrepreneurs, with access to the same investment technology and different amounts of capital A. The distribution of assets across entrepreneurs is described by the cumulative distribution function $G(A)$. The investment required is I, so an entrepreneur needs to raise $I - A$ from outside investors. The gross return on the investment is either 0 or $R > 0$, and the probability of getting R instead of 0 depends on the type of project that the entrepreneur chooses. The possible projects are described in Table 3.1.

Table 3.1: Project outcomes in a model of moral hazard in the credit market.

Project	Good	Bad (low private benefit)	Bad (hihg private benefit)
Private benefit	0	b	B
Probability of success	P_H	P_L	P_L

The table is based on Holmstrom and Tirole [1997]. An entrepreneur can choose among three projects. The good project yields no private benefits and succeeds (that is, yields R) with probability p_H (otherwise, it fails and yields 0). There are two bad projects that succeed with probability p_L. They differ in the amount of private benefits they generate to the entrepreneur, which can be either b or B.

If the entrepreneur chooses a good project, the probability of a high return is p_H. On the other hand, if he chooses a bad project, the probability of a high return is only $p_L < p_H$. However, the entrepreneur may choose a bad project because a bad project provides him a non-pecuniary private benefit. The private benefit is either b or B, where $B > b$. So, if unconstrained, the entrepreneur will always choose a bad project with a private benefit of B over a bad project with a private benefit of b.

The rate of return demanded by outside investors is denoted by γ, which can either be fixed or coming from an upward sloping supply function $S(\gamma)$. The assumption is that only the good project is viable:

$$p_H R - \gamma I > 0 > p_L R - \gamma I + B.$$

That is, investing in the bad project generates a negative total surplus. Hence, for outside investors to put money in the firm, it is essential to make sure that the entrepreneur undertakes the good project. The incentive of the entrepreneur to choose the good project will depend on how much "skin in the game" he has. This means that the entrepreneur will need to keep enough ownership of the project so that he has a monetary incentive to make the "right" decision. A key implication is that it would be easier to provide external financing to entrepreneurs

with large assets A, since they are more likely to internalize the monetary benefit and choose the good project rather than enjoying the non-pecuniary private benefits of the bad project.

To see this, let us derive the solution of this basic model. Consider a contract where the entrepreneur invests his funds A together with an amount $I-A$ raised from an outside investor. Clearly, no one will receive any payment if the project fails and yields 0. The key is to determine how the entrepreneur and the outside investor split the return of the project in case it succeeds, yielding R. In general, one can denote the payment to the entrepreneur as R_f and the payment to the outside investor as R_u, such that $R_f + R_u = R$.

A necessary condition for outside investors to be willing to provide financing to the entrepreneur is that the entrepreneur has an incentive to choose the good project. Otherwise, the total net present value is negative, and the outside investor cannot break even. Hence, it is crucial that the entrepreneur benefits more from taking the good project than from taking the bad project. This implies:

$$p_H R_f \geq p_L R_f + B.$$

Denoting $\Delta p = p_H - p_L$, we get the incentive compatibility constraint:

$$R_f \geq B/\Delta p.$$

This implies that the maximum expected amount that can be promised to the outside investors — the pledgeable expected income — is:

$$p_H(R - B/\Delta p).$$

Hence, to satisfy the participation constraint of the outside investors, that is, making sure that they get a high enough expected income to at least break even, we need:

$$\gamma(I - A) \leq p_H(R - B/\Delta p).$$

This puts an endogenous financing constraint on the entrepreneur, which depends on how much internal capital A he has. Defining the threshold $\overline{A}(\gamma)$ as:

$$\overline{A}(\gamma) = I - p_H(R - B/\Delta p)/\gamma,$$

we get that only entrepreneurs with capital at or above $\overline{A}(\gamma)$ can raise external capital and invest in their projects. This is the classic credit rationing result, which goes back to Stiglitz and Weiss [1981]. The entrepreneur cannot get unlimited amounts of capital, since he needs to maintain a high enough stake in the project to satisfy his incentive compatibility constraint and the outside investors' participation constraint.

Holmstrom and Tirole go on to introduce financial intermediaries, who have the ability to monitor entrepreneurs.[1] The monitoring technology available to financial intermediaries is assumed to prevent the entrepreneur from taking a bad project with high non-pecuniary private benefit B, thereby reducing the opportunity cost that the entrepreneur incurs when taking the good project from B to b. Monitoring yields a private cost of c to the financial intermediary. Financial intermediaries themselves need to have an incentive to pay the monitoring cost and prevent entrepreneurs from enjoying high private benefits B. To achieve this, they need to put in their own capital. Hence, the amount of intermediary capital K_m available in the economy is going to be a key parameter in determining how much lending will occur.

An intermediary can help relax the financing constraint of the entrepreneur by monitoring him and reducing his incentive to take the bad project. This means that even entrepreneurs with a level of capital lower than the threshold $\overline{A}(\gamma)$ will be able to get financing assisted by the intermediaries. Denoting the return required by the intermediaries as β, where β is determined in equilibrium and is decreasing in the amount of capital K_m that is available in the financial-intermediary sector, the threshold $\underline{A}(\gamma, \beta)$ of entrepreneur's capital A above which the entrepreneur can raise capital via financial intermediaries and invest is:

$$\underline{A}(\gamma, \beta) = I - I_m(\beta) - p_H(R - (b + c)/\Delta p)/\gamma.$$

Here, $I_m(\beta)$ is the amount of capital provided by the financial intermediary, which is decreasing in the return β demanded by financial

[1] Strictly speaking, the financial intermediaries here are not necessarily intermediating between the outside investors and the entrepreneurs. Rather, they could provide a different type of financing that can relax financial constraints via monitoring.

intermediaries. Hence, the entrepreneur only needs to raise $I - I_m(\beta)$ directly from outside investors. At the same time, the entrepreneur can only promise them an expected payment of $p_H(R-(b+c)/\Delta p)$, so that the entrepreneur and the financial intermediary maintain incentives to pick the good project and monitor. This implies that only entrepreneurs with more internal capital than $\underline{A}(\gamma,\beta)$ (defined above) will be able to raise capital via the financial intermediary sector.

Figure 3.1 depicts the equilibrium outcomes with regard to which entrepreneurs will be financed and invest, depending on how much capital they have. We can see that entrepreneurs with little capital — that is, below $\underline{A}(\gamma,\beta)$ — cannot get financed and do not invest in their projects. Entrepreneurs with an intermediate level of capital — that is, between $\underline{A}(\gamma,\beta)$ and $\overline{A}(\gamma)$ — can get financed through financial intermediaries who assist them with their monitoring technology. Entrepreneurs with a high level of capital — that is, above $\overline{A}(\gamma)$ — can get financed directly by the outside investors without the monitoring of the financial intermediaries. Of course, a key condition for this figure to hold is that $\underline{A}(\gamma,\beta)$ is smaller than $\overline{A}(\gamma)$. This will happen when c and β are not too large. In such a case, the financial intermediary sector is efficient enough to be able to provide financing to entrepreneurs who are otherwise rationed in the credit market. If this condition does not hold, then the financial intermediary sector does not exist, and

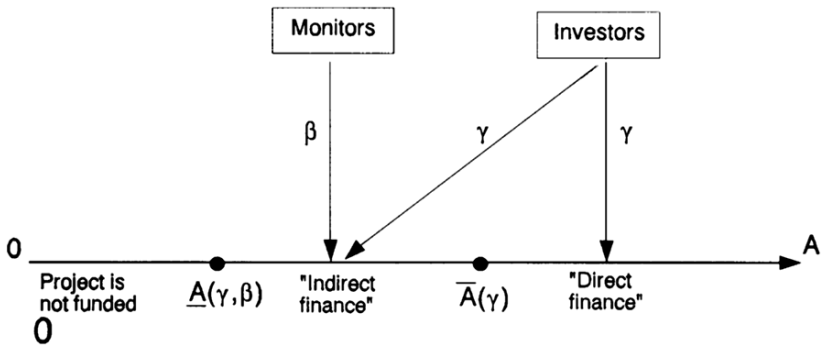

Figure 3.1: Direct and indirect financing of investment (from Holmstrom and Tirole [1997]).

entrepreneurs simply get financing if and only if their level of internal capital is above $\overline{A}(\gamma)$.

Overall, the model demonstrates the frictions in the provision of capital to real investments. Due to a moral hazard problem, entrepreneurs with profitable investment opportunities might not be able to finance them if they do not have enough capital already. Gan [2007a] and Chaney et al. [2012] provide micro-level evidence of the effect of this 'balance-sheet channel', where a decrease in the value of firms' assets reduces their ability to borrow and invest. For identification, they use exogenous shocks to the prices of real estate.

While such frictions always exist, they might be exacerbated. This leads to severe credit rationing, which can be referred to as a crisis. This literature was not developed to describe crises per se, but is increasingly used to describe phenomena around financial crises, given that during financial crises the flow of credit is interrupted or even freezes. For example, in this model, a negative aggregate shock in the economy — shifting the distribution of capital $G(A)$ to the left so that entrepreneurs have less capital on average — will be amplified as entrepreneurs having less wealth will face stricter financial constraints and will be less likely to raise external financing. Hence, there is an accelerator effect, whereby shocks to the economy are amplified: an initial loss of capital causes further losses due to the tightening of financial constraints, making entrepreneurs unable to make profitable investments.

Another form of the accelerator effect in this model operates via the financial intermediary sector, as a decrease in the capital K_m of the financial intermediary sector will also have an adverse effect on the real economy. This is because it leads to an increase in the equilibrium return β demanded by financial intermediaries, and to an increase in the threshold $\underline{A}(\gamma, \beta)$, above which middle-size entrepreneurs can get financed and invest. Hence, a decrease in financial intermediary capital will lead to contraction in real investment, specifically of middle-sized firms. Gan [2007b] provides evidence of this 'lending channel', where the decrease in the value of banks' assets reduces their ability to lend to firms, and hence reduces the ability of firms to borrow and invest. Note that this approach does not explain the source of a shock in the

financial system but rather sheds light on how shocks to the financial system can be amplified into large effects on the real economy.

Holmstrom and Tirole [1998] study a related setup and develop the implications for government policy. Recall that entrepreneurs need to keep sufficient ownership in the firms that they run (R_f needs to be sufficiently high), so that they take the good project rather than the bad project. This limits their ability to offer sufficient return to outside investors (R_u is limited). Therefore, in case of an adverse liquidity shock, they are limited in how much capital they can raise to keep running their projects and prevent welfare-reducing bankruptcy. This creates an incentive for holding liquid securities *ex ante* so that they can use them when they are hit by adverse shocks and are financially constrained. Holmstrom and Tirole [1998] show that, in case of aggregate uncertainty, the government can improve overall welfare by issuing government debt and supplementing the supply of liquid securities in the economy.

Other models of government intervention to alleviate problems in the provision of credit are based on externalities between lenders that might lead to a credit freeze. Such models include Bebchuk and Goldstein [2011] and Benmelech and Bergman [2012]. In Bebchuk and Goldstein [2011], network externalities across firms imply that the profitability of providing credit by one bank depends on the amount of credit provided by other banks. Then, a credit freeze might arise as a self-fulfilling belief. They discuss and compare various policy measures to alleviate the freeze. In Benmelech and Bergman [2012], the externalities originate from the endogenous value of collateral.

On the other hand, others have focused on periods of excessive lending or credit booms. They show that these can also be inefficient and might call for government intervention to tame the boom. Allen and Gale [2000a] present a model in which the possibility of risk shifting leads borrowers to borrow excessively, inflating the prices of the projects they invest in (this is most common in the real estate sector). When a bad shock occurs, the bubble bursts, and credit dries up. This seems to match the events leading to a typical crisis, with excessive lending and asset bubbles leading to a crash. Lorenzoni [2008] studies

a model of inefficient credit booms, where borrowers and lenders do not internalize the negative externalities that credit imposes on others. These externalities emerge due to fire sales that happen following a bad shock that forces many borrowers to liquidate their assets at once.

The papers described here so far all consider the provision of credit by financial institutions to the real economy. An important part of the financial system, however, is the network that connects different financial institutions through which they lend to and borrow from one another. Such a network was traditionally known in the interbank market, and analyzed in many papers such as Bhattacharya and Gale [1987] and Allen and Gale [2000b]. More recently, it has been quite prominent in markets like the repo market and the market for asset-backed commercial paper, which were the epicenter of the recent financial crisis.

Moral hazard frictions as described by Holmstrom and Tirole [1997] and the related literature reviewed above are also prevalent in the lending relationships among financial institutions and have been used to explain the shocks in the financial sector. As opposed to the papers that show how a shock in the financial sector can have an amplified effect on the real economy, these papers can shed more light on what is going on within the financial sector. A recent example is Acharya and Viswanathan [2011] who use a moral-hazard rationale to explain how liquidity dries up in the financial system. The basis of their model is a different moral hazard story, which is based on risk shifting instead of the rent seeking in the Holmstrom and Tirole [1997] model developed earlier, but the consequences in terms of the financial accelerator are similar. That is, a negative shock to the capital of some borrowers within the financial system reduces their ability to borrow, and so leading to much lower transmission of capital within the financial system. Acharya et al. [2012] also use a moral-hazard model to study frictions in the interbank market. They add market power, which is held by large institutions and hurts the flow of capital across banks.

Models by Martin et al. [2014] and Kuong [2015] apply more directly to the repo market. They are also based on moral hazard frictions and employ them to explain the use of collateral in repo markets. They analyze circumstances under which the frictions in these markets can lead

to coordination failures that resemble runs. Di Maggio and Tahbaz-Salehi [2015] embed a moral-hazard friction in a model of networks, where the connections established by each bank are important for understanding the transmission of shocks. They study how the friction can propagate itself over the network via the connections across banks. They show that beyond the overall availability of collateral, a key aspect for understanding fragility is how this collateral is distributed over the network.

3.2 Implications for macroeconomic models

Financial accelerators in the spirit of Holmstrom and Tirole [1997] have been discussed in macroeconomic setups, showing how shocks to asset values can be amplified and become persistent in equilibrium. In many cases, the downside in macroeconomic models is that they do not provide rich microfoundations. Yet this is where the interaction between microeconomic models and macroeconomic models is important. Embedding the basic frictions and mechanisms in macroeconomic models enables researchers to move towards calibration and direct use of these models in policymaking. While there is still a lot of progress to be made before these models replace the traditional models used in central banks, a lot of progress has been made already.[2]

Bernanke and Gertler [1989] provide one of the first financial-accelerator models in macroeconomics, emphasizing that financial frictions amplify adverse shocks and that they are persistent. That is, a temporary shock depresses not only current but also future economic activity. The mechanism goes through the agency problem between borrowers and lenders as described earlier. A negative shock to the net worth of a borrower strengthens the agency problem between the borrower and potential lenders, which reduces lending and investment in equilibrium, and thus amplifies the initial shock.

Kiyotaki and Moore [1997] identify an important dynamic feedback mechanism that strengthens these forces. The reduction in future investments following a negative shock today lowers future economic

[2]For a survey of this literature, see Brunnermeier et al. [2013].

activity and will reduce future asset prices. But since this decline is anticipated, it is immediately reflected in a drop of current asset prices. As a result, current net worth of potential borrowers is reduced today, lowering the collateral value they can provide, and limiting their debt capacity even further. Then, investments decline and the demand for assets drops even more. As a consequence, prices decline further, eroding productive agents' net worth, and so on. This feedback loop can amplify shocks significantly.

A large body of recent work builds on the models of Bernanke and Gertler [1989] and Kiyotaki and Moore [1997] to analyze the effect of financial frictions due to credit constraints in macroeconomic settings. One important point in the data is the asymmetry between recessions and booms. Kocherlakota [2000] builds a model where credit cycles of the kind described by Kiyotaki and Moore [1997] are asymmetric: sharp downturns are followed by slow recoveries. Eisfeldt and Rampini [2006] develop a model where credit constraints are more binding in recessions, and they match the empirical regularity that capital reallocation is lower in downturns than in booms. Iacoviello [2005] adds nominal mortgage debt using real estate as collateral to evaluate the quantitative relevance of the Kiyotaki-Moore mechanism. Other authors have adopted similar frameworks to an international setting; for example Caballero and Krishnamurthy [2001] and Mendoza [2010], who study the dry-up of international capital inflows. We discuss this line of research more in the next section.

While in Kiyotaki and Moore [1997] credit is limited by the expected price of the collateral in the next period, other models emphasize the role of volatility. In Brunnermeier and Pedersen [2009], for example, borrowing capacity is limited due to the volatility of future prices. In Brunnermeier and Sannikov [2014], even productive entrepreneurs are concerned about hitting their solvency constraint in the future and consequently do not fully exploit their debt capacity. As volatility rises they cut back on borrowing by selling assets. This depresses prices further, leading to rich volatility dynamics.

A different angle on the role of credit frictions in the macro economy is provided by Eggertsson and Krugman [2012]. They study a model

with heterogeneous agents, where patient agents lend and impatient agents borrow subject to a collateral constraint. If, for some reason, the collateral requirement becomes tighter, impatient agents will have to go into a process of deleveraging, reducing the aggregate demand. This excess saving leads to a reduction in the natural interest rate, which might become negative. The nominal (policy) interest rate then hits the zero bound, putting the economy into a liquidity trap. Consequently, traditional monetary policy becomes impossible but fiscal policy regains some potency.

A key feature that is missing from these models is the role of financial intermediaries. The model of Holmstrom and Tirole [1997] gives banks a special role in their ability to monitor firms. Evidence from the recent crisis has shown the important role of financial intermediaries and the importance of distinguishing between loans extended to firms by banks and other types of corporate debt [see Adrian et al., 2013]. Adrian et al. [2014] demonstrate the importance of the conditions in the financial intermediary sector for determining asset prices in the economy. Recently, models by Gertler and Kiyotaki [2011], He and Krishnamurthy [2012, 2013], and Rampini and Viswanathan [2015] add a financial intermediary sector, in the spirit of Holmstrom and Tirole [1997], and analyze the dynamic interactions between this sector and the rest of the economy. Introducing this sector into macroeconomic models enables elaborate discussions on various policies conducted by governments during the recent crisis in an attempt to stimulate the economy via the financial intermediation sector. Such policies are discussed by Gertler and Kiyotaki [2011].

Still, while all these models are able to analyze how a financial shock interacts with the real economy, they do not go into detail about what happens in the financial sector to generate a sizable financial shock. Integrating the models described here with those of Section 2, in which sizable financial shocks are generated by panic of investors in the financial system, could help in this direction. This is particularly important in macroeconomic models that are trying to provide quantification and guide policymakers on the quantitative implications of various policy steps. In the recent crisis, it seems that frictions on both sides of the

balance sheet reinforced each other. The deterioration in the quality of the financial intermediaries' assets gave creditors of the intermediary a stronger incentive to run. This further reduced the value of bank assets, and so on. Recently, Gertler and Kiyotaki [2015] made this exact point. They build on the traditional macroeconomic model with moral hazard frictions in lending, and add to it fragility on the liability side due to potential runs. They analyze the extent to which runs further amplify the effects of shocks on the economy. In our view, more work is needed in this direction, perhaps utilizing the global-games framework that enables determination of the probability of a run and how it interacts with key parameters of the economy.

Alternatively, one could think of sizable shocks in the financial system being generated by the moral hazard or adverse selection problems in the networks that connects different financial institutions. We described the moral hazard models in the previous subsection and turn to a short description of the adverse selection models in the next one.

3.3 Asymmetric information

Another key factor behind the breakdown of financial markets, including credit markets, is adverse selection. The analysis of market breakdown due to asymmetric information and adverse selection goes back to Akerlof [1970]. He analyzed a market where sellers have private information about the quality of the assets that they are trying to sell. As a result, buyers are reluctant to buy the assets from them because they realize that the sale represents negative information about the asset. In extreme situations, when the only motivation to trade is based on information, this leads to a complete market breakdown: no transactions will happen in equilibrium. If there are other gains to trade between sellers and buyers, trade may still occur. However, the increase in the magnitude of asymmetric information, due to an increase in the share of informed agents or in the degree of underlying uncertainty, might reduce trade (or the efficiency of trade).

In the context of credit markets, Stiglitz and Weiss [1981] point to adverse selection as a potential reason for credit rationing: when lenders

do not know the quality of their borrowers, increasing the interest rate will only attract bad borrowers. This means that the interest rate cannot increase freely to clear the market, and we might get an equilibrium with credit rationing. Another canonical application in the context of finance is by Myers and Majluf [1984] who show that a firm, which is better informed about the value of its assets, will have a hard time raising equity for the purpose of financing new investment as outsiders will fear that the firm is raising equity because the value of its assets in place is low. They go on to show that debt may be a better security as it is less sensitive to the information about the value of the assets, leading to the famous pecking order theory of capital structure.

While these papers deal with provision of capital for non-financial firms, adverse selection seems to have been a big problem in the interactions of financial institutions during the recent crisis. Various institutions held so-called toxic assets, such as mortgage-backed securities or asset-backed commercial paper, and the market for exchanging these securities between institutions froze despite the likely presence of gains to trade. In Section 3.1, we described how moral hazard can contribute to the freeze of such a market. Essentially, moral hazard problems have the potential to eliminate the gains from trade and freeze transactions. With adverse selection, on the other hand, the gains from trade are clearly there, but the fact that one party knows more about the asset than the other might hamper trade. For example, if the seller is more informed than the buyer, then buyers might fear that only sellers of bad assets are willing to sell at existing prices, causing markets to break down. This could well explain what has happened in the interactions between financial institutions where mortgage-backed securities or asset-backed commercial paper could have exchanged hands.

Given this premise, several authors have written new theories of adverse selection, which are designed to shed more direct light on the freezing of markets in the last crisis. Glode et al. [2012] endogenize the choice of financial institutions on how much information to acquire about the assets. They describe the acquisition of information as an arms race in which each party chooses to become more informed about its assets in order to get a bargaining advantage against other parties.

Yet in equilibrium no such advantage is obtained, since all institutions end up being equally informed. This is akin to a prisoner's dilemma. Moreover, when uncertainty increases like in the run-up to the crisis, the expertise that is acquired becomes quite damaging and can lead to a complete market freeze.

Gorton and Ordonez [2014] provide a different angle on the incentive to produce information in an adverse selection model. They consider the repo market, in which complex securities are used as collateral for funding of financial institutions. In the spirit of Gorton and Pennacchi [1990], they show that the lack of information production supports the flow of credit in normal times. However, there are two reasons why this might change over time. First, after a long time with no information production, institutions become quite uninformed about the underlying asset, making information production more desirable. Second, a small change in uncertainty can make information production more desirable. This can lead to a sudden increase in information production, which can cause a breakdown in the repo market.

Malherbe [2014] is motivated by the debate about whether liquidity hoarding or adverse selection is most to blame for the market freeze in the crisis. He points out that these two are interrelated and can aggravate each other. When market participants expect adverse selection in the market, they have a greater incentive to hoard liquidity as this is a more effective way to satisfy their potential liquidity needs. Then, when more institutions are hoarding liquidity, it is known that their attempt to sell assets is not driven by a liquidity need but by bad information about their assets, and so the adverse selection problem is aggravated. This generates a feedback loop with potential for multiple equilibria: one with high liquidity hoarding and adverse selection and one with low levels of both. Other papers that study the interplay between adverse selection and liquidity hoarding include Bolton et al. [2011] and Heider et al. [2015].

The potential for self-fulfilling outcomes in models of adverse selection has been pointed out before in other settings. In models by Pagano [1989] and Dow [2004], uninformed traders have stronger incentive to participate in the market if they know that there are more uninformed

traders as they are then exposed to a lesser adverse selection problem. As a result, there is a coordination problem that can lead to sharp changes in market depth, resembling what we see in a financial crisis. Recently, Morris and Shin [2012] show that the amplification becomes even more severe when traders have different information about the extent of the adverse selection problem; that is, about how many informed traders are present. This leads to a contagious process, by which very small changes can lead to a market freeze.

The adverse selection problem was perceived to be a central issue in the freezing of financial markets in the recent crisis, and so several government programs were designed to alleviate adverse selection by, for example, removing toxic assets from the financial system and restoring flows of trade and credit as a result. Recent papers by Tirole [2012] and Philippon and Skreta [2012] perform theoretical analysis of such policy intervention in light of the problem of asymmetric information, trying to characterize what optimal intervention may look like.

Other solutions to the adverse selection problem may be in the form of market design. Interestingly, Glode and Opp [2015] show that the presence of intermediation chains can help alleviate adverse selection. The idea is that the asset being sold moves along the chain of intermediaries between the seller and the buyer. While the adverse selection between the seller and the buyer might be too large to support direct trade between them, the fact that adverse selection is not too large between any two adjacent participants in the chain implies that trade can be supported by the chain. As we discussed in Section 3.1, networks of this kind are common in modern financial markets. In that section, moral hazard issues were discussed and the network had the potential to propagate them. Here, from the point of view of adverse selection problems, it turns out that networks can contribute to stability.

4

Currency Crises

For a long time, the literature on currency crises has been developed independently from the literatures on runs and frictions in the financial sector described above. Empirically, however, there is a strong link between the different types of crises, which led to some integration of the different literatures. Indeed, historically, financial crises have often been marked with large disturbances in currency markets, which have spilled over to the financial sectors and the real economies of affected countries in various ways. In recent years, the deepening of the crisis in Europe is strongly associated with the attempt to maintain the common currency area, which also has implications for the financial sectors and real economies of countries in the Euro zone. In this section, we review the development of the theoretical literature on currency crises from its very beginning. We then discuss the connection of this literature to the literatures on banking panics and credit frictions.

In general, currency crises originate from the attempt of governments to maintain certain financial and monetary arrangements, most notably a fixed-exchange rate regime. Their goal is to stabilize the economy. At times, these arrangements become unstable, which leads to a

speculative attack on a fixed exchange rate regime and from there to a financial crisis.

The best way to understand the origins of currency crises is to think about the basic trilemma in international finance. A trilemma is a situation in which someone faces a choice among three options, each of which comes with some inevitable problems. In international finance, it stems from the fact that, in most nations, economic policy makers would like to achieve the following goals. First, make the country's economy open to international capital flows, because by doing so they let investors diversify their portfolios overseas and achieve risk sharing. They also benefit from the expertise brought to the country by foreign investors. Second, use monetary policy as a tool to help stabilize inflation, output, and the financial sector in the economy. This is achieved as the central bank can increase the money supply and reduce interest rates when the economy is depressed, or reduce money growth and raise interest rates when it is overheated. Moreover, it can serve as a lender of last resort in case of financial panic. Third, maintain stability in the exchange rate. This is because a volatile exchange rate, at times driven by speculation, can be a source of broader financial volatility, and makes it harder for households and businesses to trade in the world economy and for investors to plan for the future.

The problem, however, is that a country can only achieve two of these three policy goals. In order to maintain a fixed exchange rate and capital mobility, the central bank loses its ability to control its policy instruments: the interest rate, or, equivalently, the monetary base. This is because, under free capital mobility, the interest rate becomes anchored to the world interest rate by the interest rate parity, and the monetary base is automatically adjusted to the pre-determined money demand. This is the case of individual members of the European Monetary Union. In order to keep control over the interest rate or, equivalently, the money supply, the central bank has to let the exchange rate float freely; as in the case of the US. If the central bank wishes to maintain both exchange rate stability and control over the monetary policy, the only way to do it is by imposing capital controls; as in the case of China.

Currency crises occur when the country is trying to maintain a fixed exchange rate regime with capital mobility, but faces conflicting policy needs, such as fiscal imbalances or a fragile financial sector, that need to be resolved by independent monetary policy. This leads to a shift in the regime from the first solution of the trilemma described earlier to the second one. The sudden depreciation in the exchange rate is often referred to as a currency crisis. It often has implications for the financial system as a whole and for the real economy, where agents were used to rely on a fixed exchange rate regime and often have to adjust to the change abruptly and unexpectedly. We elaborate more on this in Subsection 4.3. The theoretical currency-crises literature is broadly classified into three generations of models, which we now turn to describe in more detail.

4.1 First-generation models of currency crises

This branch of models, the so-called 'first generation models of currency attacks' was motivated by a series of events in which fixed exchange rate regimes collapsed following speculative attacks, for example, the early 1970s breakdown of the Bretton Woods global system.

The first paper here is the one by Krugman [1979].[1] He described a government that tries to maintain a fixed exchange rate regime but is subject to a constant loss of reserves, due to the need to monetize persistent government budget deficits. These two features of the policy are inconsistent with each other, and lead to an eventual attack on the reserves of the central bank, that culminates in a collapse of the fixed exchange rate regime. Flood and Garber [1984] extended and clarified the basic mechanism suggested by Krugman [1979], generating the formulation that is widely used since then.

Let us provide a simple description of this model. The model is based on the central bank's balance sheet. The asset-side of the central bank's balance sheet at time t is composed of domestic assets $B_{H,t}$, and the domestic-currency value of foreign assets $S_t B_{F,t}$, where S_t denotes

[1] The model by Krugman [1979] builds on an earlier paper by Salant and Henderson [1978] about a speculative attack on gold reserves.

the exchange rate, that is, the value of foreign currency in terms of domestic currency. The total assets have to equal the total liabilities of the central bank, which are, by definition, the monetary base, denoted as M_t.

In the model, the domestic assets grow in a fixed and exogenous rate due to fiscal imbalances:

$$\frac{B_{H,t} - B_{H,t-1}}{B_{H,t-1}} = \mu.$$

Because of perfect capital mobility, the domestic interest rate is determined through the interest rate parity as follows:

$$1 + i_t = (1 + i_t^*)\frac{S_{t+1}}{S_t},$$

where i_t denotes the domestic interest rate at time t and i_t^* denotes the foreign interest rate at time t. Finally, the supply of money, the monetary base, has to be equal to the demand for money, which is denoted as $L(i_t)$, a decreasing function of the domestic interest rate.

The inconsistency between a fixed exchange rate regime, $S_t = S_{t+1} = \bar{S}$, with capital mobility and the fiscal imbalances comes due to the fact that the domestic assets of the central bank keep growing but the total assets cannot change since the monetary base is pinned down by the demand for money, $L(i_t^*)$, which is determined by the foreign interest rate. Hence, the obligation of the central bank to keep financing the fiscal needs puts downward pressure on the domestic interest rate, which, in turn, puts upward pressure on the exchange rate. In order to prevent depreciation, the central bank has to intervene by reducing the inventory of foreign reserves. Overall, $\bar{S}B_{F,t}$ decreases by the same amount as $B_{H,t}$ increases, so the monetary base remains the same.

The problem is that this process cannot continue forever because the reserves of foreign currency have a lower bound. Eventually, the central bank will have to abandon the current solution of the trilemma — fixed exchange rate regime and perfect capital mobility — for another solution — flexible exchange rate with flexible monetary policy (that is, flexible monetary base or equivalently flexible domestic interest rate) and perfect capital mobility.

The question is what the critical level of domestic assets $B^*_{H,T} = B_{H,T}$ is and what the corresponding period of time T, at which the fixed-exchange rate regime collapses, is. As pointed out by Flood and Garber [1984], this happens when the shadow exchange rate — defined as the flexible exchange rate under the assumption that the central bank's foreign reserves reached their lower bound while the central bank keeps increasing the domestic assets to accommodate the fiscal needs — is equal to the pegged exchange rate. At this point, there will be a speculative attack on the currency, depleting the central bank's foreign reserves and forcing an immediate devaluation of the domestic currency.

This is depicted in Figure 4.1. The upper panel depicts the shadow exchange rate schedule and the pegged rate schedule as functions of domestic assets. Their intersection determines the level of domestic

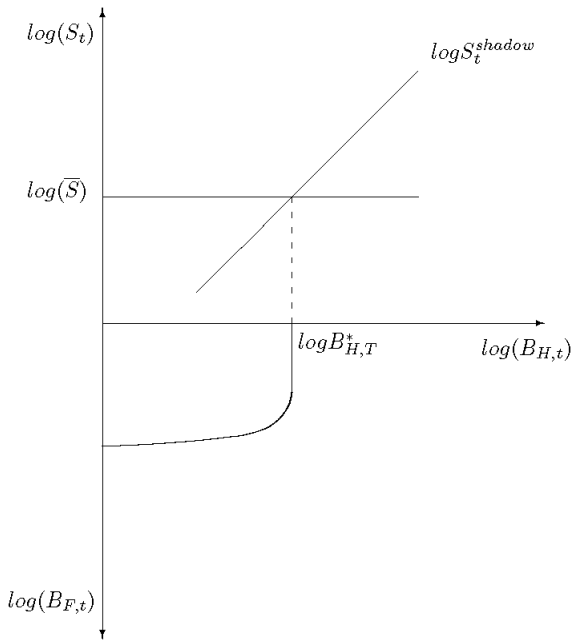

Figure 4.1: Currency-Regime switch (based on Krugman [1979] and Flood and Garber [1984]).

asset where the regimes switch occurs. The switch will not occur at a lower level of domestic assets because at that point there is no incentive to launch the speculative attack (it will yield a trading loss to the speculators). The switch will not occur at a higher level either because there is gain to be made at the intersection point, leading all speculators to attack at that point. The lower panel describes the path of foreign assets, which are quickly depleted at the point of the attack.

4.2 Second-generation models of currency crises

Following the collapse of the European Exchange Rate Mechanism (ERM) in the early 1990s, the so-called first-generation model of currency attacks did not seem suitable anymore to explain the ongoing crisis phenomenon. The events in Europe at that time featured governments actively making decisions between fighting the declining economic activity level and remaining in the exchange rate management system. Hence, there was a need for a model in which the government's choice is endogenized, rather than the first-generation models in which the exchange rate regime is essentially on 'automatic pilot'. This led to the development of the so-called 'second generation model of currency attacks,' pioneered by Obstfeld [1994, 1996].

In this line of models, the government/central bank is setting the policy endogenously, trying to maximize a well-specified objective function, without being able to fully commit to a given policy. An outcome of these models is that there are usually self-fulfilling multiple equilibria, where the expectation of a collapse of the fixed exchange rate regime leads to the government abandoning the regime. This feature seemed to capture the fact that crises where often unexpected. This is related to the Diamond and Dybvig [1983] model of bank runs described in Section 2, creating a link between these two literatures.

Obstfeld [1996] discusses various mechanisms that can create the multiplicity of equilibria in a currency-crisis model. Let us describe one mechanism, which is inspired by Barro and Gordon [1983]. Suppose that the government minimizes a loss function of the following type:

$$(y - y^*)^2 + \beta\varepsilon^2 + cI_{\varepsilon\neq 0}.$$

Here, y is the level of output, y^* is the target level of output, and ε is the rate of depreciation, which in the model is equal to the inflation rate. The government wants to minimize some combination of the rate of inflation and the distance from the target level of output. In addition, the third term is an index function, which says that there is a fixed cost in case the government deviates from the existing exchange rate. The interpretation is that the government is in a regime of zero depreciation (a fixed exchange rate regime), and deviating from it is costly.

Overall, when deciding on the rate of depreciation, the government has to weigh the costs against the benefit of depreciation. The costs are coming from the second and third terms above: there is a cost in operating the economy under inflation and there is a cost in deviating from the promise of a fixed exchange rate regime. The benefit in depreciation is that it enables reduction in the deviations from the target level of output. More precisely, creating inflation (which is equivalent to depreciation here) above the expected level serves to boost output.

This effect of inflation on output is coming from the Philips Curve. It is demonstrated in the following expression, specifying how output is determined:

$$y = \bar{y} + \alpha(\varepsilon - \varepsilon^e) - u.$$

Here, $\bar{y}$ is the natural output ($\bar{y} < y^*$, that is, the government sets an ambitious output target level to overcome distortions in the economy), u is a random shock, and ε^e is the expected level of depreciation/inflation that is set endogenously in the model by wage setters based on rational expectations. The idea is that an unexpected inflationary shock ($\varepsilon > \varepsilon^e$) boosts output by reducing real wages and increasing production.

Importantly, the government cannot commit to a fixed exchange rate. Otherwise, it would achieve minimum loss by committing to $\varepsilon = 0$. However, due to a lack of commitment, a sizable shock u will lead the government to depreciate and achieve the increase in output bearing the loss of credibility. Going back to the trilemma discussed above, a fixed exchange rate regime prevents the government from using monetary policy to boost output, and a large enough shock will cause the government to deviate from the fixed exchange rate regime.

It can be shown that this model generates multiplicity of equilibria. If wage setters coordinate on a high level of expected depreciation/inflation, then the government will validate this expectation with its policy by depreciating more often. If they coordinate on a low level of expected depreciation, then the government will have a weaker incentive to deviate from the fixed exchange rate regime. Hence, the expectation of depreciation becomes self-fulfilling.

Similarly, closer to the spirit of the Krugman [1979] model, one can describe mechanisms in which speculators can force the government to abandon an existing fixed-exchange rate regime by attacking its reserves and making the maintenance of the regime too costly. If many speculators attack, the government will lose a significant amount of reserves and will be more likely to abandon the regime. A speculative attack is profitable only if many speculators join it. Consequently, there is one equilibrium with a speculative attack and a collapse of the regime, and there is another equilibrium where these things do not happen.[2]

Similar mechanisms are likely at work in sovereign debt crises and, in particular, in the situation in Europe in the last few years. Speculators can attack government bonds, demanding higher rates due to expected sovereign-debt default, and creating an incentive for the central bank to abandon a currency regime and reduce the value of the debt or alternatively to default. This justifies the initial high rates. Such self-fulfilling debt crises have been studied in the literature. For example, Cole and Kehoe [2000] analyze the debt maturity structure under financial crises brought on by a loss of confidence in a government, which can arise within a dynamic, stochastic general equilibrium model.

As we discussed in Section 2, having a model of multiple equilibria creates an obstacle for policy analysis. Morris and Shin [1998] were the first to tackle the problem of multiplicity in the second-generation models of speculative attacks. They first express this model in an explicit game theoretic framework, where speculators are players having to

[2]Note that self-fulfilling speculative attacks can arise naturally from a first-generation model as demonstrated by Obstfeld [1986]. Hence, this is not the distinguishing feature of the second-generation models. Rather, the optimizing government is the distinguishing feature of the second-generation models.

make the decision of whether to attack the currency or not. Then, using the global-games methodology, pioneered by Carlsson and van Damme [1993], they are able to derive a unique equilibrium in which the fundamentals of the economy uniquely determine whether a crisis occurs or not. This is important since it enables one to ask questions about the effect of policy tools on the probability of a currency attack. The global-games methodology, relying on heterogeneous information across speculators, also brought to the forefront the issue of information in currency-attack episodes, leading to an analysis of the effect that transparency, signaling, and learning can have on such episodes [see, for example, Angeletos et al., 2006, Goldstein et al., 2011].[3]

4.3 Third-generation models of currency crises

In the late 1990s, a wave of crises hit the emerging economies in Asia, including Thailand, South Korea, Indonesia, Philippines, and Malaysia. A clear feature of these crises was the combination of the collapse of fixed exchange rate regimes, capital flows, financial institutions, and credit.[4] As a result, many researchers felt that the first two generations of models of currency crises were not sufficient for analyzing the events in Asia. There was a strong need to incorporate banking panics and credit frictions into these models. This led to extensive research on the interplay between currency crises and banking crises, sometimes referred to as the *twin crises*, and between currency crises and credit frictions.[5] Such models are often referred to as the 'third-generation models of currency crises'. In the context of this survey, it is important to note that such models bring together elements from the early currency crises literature described in Subsections 4.1 and 4.2 with

[3]For a broad review of the global–games methodology and its various applications, see Morris and Shin [2003]. Many followed the original developments, analyzing conditions under which the unique-equilibrium result fails to hold. See, for example, Angeletos and Werning [2006] and Hellwig et al. [2006].

[4]For a broad description of the events around the Asian Crisis and the importance of capital flows in conjunction with the collapse of the exchange rate see Radelet and Sachs [1998] and Calvo [1998].

[5]For empirical evidence on the twin crises, see Kaminsky and Reinhart [1999].

elements from the vast literature on banking panics and credit frictions described in Sections 2 and 3.

One of the first models to capture this joint problem was presented in Krugman [1999]. In his model, firms suffer from a currency mismatch between their assets and liabilities: their assets are denominated in domestic goods and their liabilities are denominated in foreign goods. Then, a real exchange rate depreciation increases the value of liabilities relative to assets, leading to deterioration in firms' balance sheets. Because of credit frictions, as in Holmstrom and Tirole [1997] described in Section 3, this deterioration in firms' balance sheets implies that they can borrow less and invest less. The novelty of Krugman's paper is that the decrease in investment validates the depreciation in a general equilibrium setup. This is because the decreased investment by foreigners in the domestic market implies that there will be a decrease in the aggregate demand for the local goods, relative to foreign goods (the Keynes–Ohlin "transfer problem" in international trade), leading to real depreciation. Hence, the system has multiple equilibria with high economic activity, appreciated exchange rate, and strong balance sheets in one equilibrium, and low economic activity, depreciated exchange rate, and weak balance sheets in the other equilibrium. Other models that extended and continued this line of research include: Aghion et al. [2001], Caballero and Krishnamurthy [2001], and Schneider and Tornell [2004]. The latter endogenizes the *ex-ante* currency mismatch between firms' assets and liabilities.

A different line of research links currency problems with the bank runs described in Section 2. Chang and Velasco [2001] and Goldstein [2005] model the vicious circle between bank runs and speculative attacks on the currency. On the one hand, the expected collapse of the currency worsens banks' prospects, as they have foreign liabilities and domestic assets, and thus generates bank runs. On the other hand, the collapse of the banks leads to capital outflows that deplete the reserves of the government, encouraging speculative attacks against the currency.

Accounting for the circular relationship between currency crises and banking crises is important for policy analysis, leading to some

surprising conclusions. For example, traditional banking models may advocate a lender-of-last-resort policy, or other expansionary policies, during a banking crisis to mitigate the bank-run problem. However, accounting for the circularity between bank runs and currency attacks, it is shown that such policies might backfire as they deplete the reserves available to the government. This makes a currency crisis more likely, which in turn might further hurt the banking sector that is subject to a currency mismatch problem.

As we mentioned in Section 4.2, there is a strong link between currency-crises models and sovereign-debt models (exemplified by the Cole and Kehoe [2000] framework). Hence, the models, reviewed in this subsection, that tie banking and credit problems with currency crises, can be very helpful in analyzing the connection between banking crises and sovereign-debt crises. This seems to be a very relevant and timely avenue for research given the current situation in Europe, where the faith of governments is intertwined with that of banks due to the various connections between banks' balance sheets and governments' balance sheets.

An interesting review of the current problems in Europe is provided by Shambaugh [2012]. He argues that Europe suffers from three types of crises that reinforce and amplify each other: a banking crisis, a sovereign debt crisis, and a real crisis due to a decline in growth and competitiveness. Concerning the banking crisis, banks in Europe face similar problems to those of banks in the US since the beginning of the financial crisis. Deterioration in the value of their assets and the mismatch between the maturity of assets and liabilities made them fragile, and led to failures and to government intervention. Coordinating banking supervision and intervention across the European countries seems important to alleviate problems in the banking sector in the future. Concerning the sovereign debt crisis, several countries in Europe face high yields on their sovereign debt, reflecting the expectation that they are at high risk of defaulting on their debt.

As Shambaugh [2012] emphasizes, the problems in Europe are amplified by the linkages between the different types of crisis. This is a theme that we emphasized throughout this survey. In Europe,

weaknesses in the banking sector increase doubts about the ability of governments to pay back their debt. This is because governments will attempt to support banks in order to prevent their collapse but on the way they might harm their own ability to pay back the debt, making the sovereign debt crisis more severe. Similarly, as the sovereign-debt problems become more severe, banks are also becoming weaker. This is because banks are holding a lot of government bonds and rely on government guarantees. Hence, there is a feedback loop between banking crises and sovereign debt crises in which they reinforce each other and make the overall problem more severe. This is very similar to the feedback loop between banking crises and currency crises, analyzed in the twin-crises literature described earlier. More research is needed to fully understand the amplification mechanisms operating between banking and sovereign debt crises, and this is necessary for the analysis of the European crisis. Such research should link to earlier works on the twin crises, inspired by the Asian crises of the late 1990s.

The problems in Europe are becoming even more severe due to the connections to the real economy. The fact that growth and competitiveness are in decline makes the problems for banks and sovereign debt even more serious because firms are less profitable, governments collect less taxes, and so on. Moreover, the problems in the financial sector and the austerity measures taken to help the government budget are making the recovery in the real economy more difficult. This is also where the fact that the European countries are in a monetary union plays a crucial role. Without a monetary union, individual countries could use monetary policy to address the problems but now they cannot. Given that other mechanisms for absorbing shocks that are present in other monetary unions are not very strong in Europe — for example, labor mobility, fiscal transfers, and a unified financial system — the existence of the monetary union may be making the problems worse.

Hence, over time, it seems possible that some members of the European monetary union will leave the common currency. While this seemed remote a few years ago, it is now mentioned more often, especially in the case of Greece. As the economic problems are becoming more severe, and the measures to address them are getting costlier,

the monetary union may not be sustainable in its current form. This is where the current banking and sovereign debt crisis may turn into a currency crisis. Again, the interconnectedness of different problems and crises is very clear and needs to be addressed in research on financial crises and in policies to address them.

4.4 Contagion of currency crises

Finally, in Section 2.5, we reviewed theories of contagion that are focused on the contagion of crises across different banks. The forceful transmission of crises across countries generated much literature on international financial contagion, which is very strongly related to the literature reviewed in Section 2.5. Kaminsky et al. [2003] provide a review of the theories behind this contagion.[6] They define contagion as an immediate reaction in one country to a crisis in another country.

As we wrote in Section 2.5, there are several theories that link contagion to fundamental explanations. The clearest one would be that there is common information about the different countries, and so the collapse in one country leads investors to withdraw from other countries; see, for example, Calvo and Mendoza [2000]. Models of the connections of portfolios across different countries also shed light on such international contagion; see, for example, Allen and Gale [2000b], Kodres and Pritsker [2002], Dasgupta [2004], and Goldstein and Pauzner [2004]. An explanation that is more directly related to currency depreciation is proposed by Gerlach and Smets [1995]. If two countries compete in export markets, the devaluation of one's currency hurts the competitiveness of the other, leading it to devalue the currency as well.

Empirical evidence has followed the theories of contagion. The common information explanation is greatly supported by the data. Several of the clearest examples of contagion involve countries that appear very similar. Examples include the contagion that spread across East Asia in the late 1990s and in Latin America in the early 1980s. A vast empirical literature provides evidence that trade links can account for contagion to some extent. These include Eichengreen et al. [1996] and Glick and

[6] For a broader review, see the collection of articles in Claessens and Forbes [2001].

Rose [1999]. Others have shown that financial linkages are also empirically important in explaining contagion. For example, Kaminsky et al. [2004] have shown that US-based mutual funds contribute to contagion by selling shares in one country when prices of shares decrease in another country. Caramazza et al. [2004], Kaminsky and Reinhart [2000] and Van Rijckeghem and Weder [2001] show similar results for common commercial banks.

5

Concluding Remarks

The global financial crisis that started in 2007, and has not been completely resolved, took much of the economic profession by surprise. Explaining the forces behind the crisis and coming up with suggestions for policymakers how to solve it and fix the system going forward have become top priorities for many economists, some of whom are new to the topic of financial crises.

As we argue in this monograph, many of the forces at play in the recent and current turmoil have been featured in the literature on financial crises for more than three decades now. Hence, it is important to go back to the main streams of this literature and summarize them to better understand the main forces behind crises, how they interact, how they apply to current and recent events, and what they imply for future policy. In this monograph, we attempted to achieve these goals.

The monograph covered three main streams of models of financial crises: (1) banking crises and panics, (2) credit frictions and market freezes, and (3) currency crises. The studies of these topics, as they are reviewed here, evolved almost in parallel for many years, until more recently they have been integrated to account for the connections between the different types of crises in real-world events. For each

one of these topics, we introduced a simple analytical framework that provides a formal description of the forces at work. We then reviewed the developments in the literature, describing the interactions between the forces, the implications for policy, and the connection to empirical evidence from the recent turmoil and before it.

While the survey presented here is extensive, it should not be viewed as a comprehensive survey of research on financial crises. The theoretical literature on financial crises is simply too large for such a comprehensive survey. We therefore focused on identifying three main streams and describing the insights they generate. We provided many references along the way that the interested reader can use to deepen his/her understanding of financial crises. In addition, we did not cover all types of theories on crises; while we mention sovereign debt crises and asset-market bubbles and crashes, there is certainly room to elaborate more on them.

As we mentioned earlier, a main benefit of this survey is that it puts together some of the basic insights on financial crises from a literature of more than three-decades old, so that people interested in studying the topic in light of recent events will have easier access to it and will know how current events are already reflected in existing literature. However, while we believe that the existing literature covers a lot of ground, there are still many open questions for future research. Along the way in this survey, we pointed out some of these open questions. We pointed out several times how, with regard to policy issues in particular, the tools reviewed here can be expanded and used to analyze optimal policies to avoid and solve crises. While there is research in this direction in the literature, it is still mostly in its early stages.

A major challenge in policy analysis going forward is to incorporate the frictions highlighted in this survey — coordination failures, incentive problems, and asymmetric information — into a macroeconomic model that can be calibrated and provides quantitative output for the optimal mix and magnitudes of policies. Some interesting work is being done in this direction in the context of credit frictions, which we reviewed here, but much less in different contexts. Developing such models is an important challenge for future research.

In addition, as the reader can observe, while there are many models discussing different forces, integrative models that combine the various forces are still in their early stages (although some exceptions have been reviewed here). This remains a major challenge to researchers going forward, since only with an integrative model, can one understand the relative contribution of different forces and the interaction between them. This is crucial for empirical work and for the design of policy to move forward. After all, as we pointed out many times in this article, the enormous force of financial crises seems to be coming from the interaction between different types of frictions that feed back into each other and reinforce each other.

Acknowledgements

We are grateful to the editor, Rick Green, for reading the manuscript carefully and giving us very insightful comments. Rick Green died a few days after the submission of the final version of this manuscript after a brief illness. His input on this manuscript was one of many examples of his dedicated service to the community of researchers in financial economics throughout his career. He will be missed.

References

V. Acharya and S. Viswanathan. Leverage, moral hazard, and liquidity. *Journal of Finance*, 66:99–138, 2011.

V. Acharya and T. Yorulmazer. Too many to fail — an analysis of time-inconsistency in bank closure policies. *Journal of Financial Intermediation*, 16:1–31, 2007.

V. Acharya, D. Gromb, and T. Yorulmazer. Imperfect competition in the interbank market for liquidity as a rationale for central banking. *American Economic Journal: Macroeconomics*, 4:184–217, 2012.

A. Admati and M. Hellwig. *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It*. Princeton University Press, 2013.

T. Adrian, P. Colla, and H. Song Shin. Which financial frictions? parsing the evidence from the financial crisis of 2007–09. *NBER Macroeconomics Annual (2012)*, 27:159–214, 2013.

T. Adrian, E. Etula, and T. Muir. Financial intermediaries and the cross-section of asset returns. *Journal of Finance*, 69:2557–2596, 2014.

P. Aghion, P. Bacchetta, and A. Banerjee. Currency crises and monetary policy in an economy with credit constraints. *European Economic Review*, 45:1121–1150, 2001.

G. A. Akerlof. The market for lemons: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84:488–500, 1970.

F. Allen and D. Gale. Optimal financial crises. *Journal of Finance*, 53:1245–1284, 1998.

F. Allen and D. Gale. Bubbles and crises. *Economic Journal*, 110:236–255, 2000a.

F. Allen and D. Gale. Financial contagion. *Journal of Political Economy*, 108:1–33, 2000b.

F. Allen, E. Carletti, I. Goldstein, and A. Leonello. Government guarantees and financial stability. Working Paper, 2014.

G.-M. Angeletos and I. Werning. Crises and prices: Information aggregation, multiplicity, and volatility. *American Economic Review*, 96:1720–1736, 2006.

G.-M. Angeletos, C. Hellwig, and A. Pavan. Signaling in a global game: Coordination and policy traps. *Journal of Political Economy*, 114:452–484, 2006.

R. Barro and D. Gordon. Rules, discretion and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12:101–121, 1983.

L. Bebchuk and I. Goldstein. Self-fulfilling credit market freezes. *Review of Financial Studies*, 24:3519–3555, 2011.

E. Benmelech and N. Bergman. Credit traps. *American Economic Review*, 104:3004–3032, 2012.

B. S. Bernanke and M. Gertler. Agency costs, net worth, and business fluctuations. *American Economic Review*, 79:14–31, 1989.

S. Bhattacharya and D. Gale. Preference shocks, liquidity and central bank policy. In W. Barnett and K. Singleton, editors, *New Approaches to Monetary Economics*, 1987.

P. Bolton, T. Santos, and J. Scheinkman. Outside and inside liquidity. *Quarterly Journal of Economics*, 126:259–321, 2011.

M. Brunnermeier. Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives*, 23:77–100, 2009.

M. Brunnermeier and L. Pedersen. Market liquidity and funding liquidity. *Review of Financial Studies*, 22:2201–2238, 2009.

M. Brunnermeier and Y. Sannikov. A macroeconomic model with a financial sector. *American Economic Review*, 104:379–421, 2014.

M. Brunnermeier, T. M. Eisenbach, and Y. Sannikov. Macroeconomics with financial frictions: A survey. In *Advances in Economics and Econometrics, Tenth World Congress of the Econometric Society*. Cambridge University Press, New York, 2013.

J. Bryant. A model of reserves, bank runs, and deposit insurance. *Journal of Banking and Finance*, 4:335–344, 1980.

R. Caballero and A. Krishnamurthy. International and domestic collateral constraints in a model of emerging market crises. *Journal of Monetary Economics*, 48:513–548, 2001.

C. Calomiris. Is deposit insurance necessary? A historical perspective. *Journal of Economic History*, 50:283–295, 1990.

C. Calomiris and G. Gorton. The origins of banking panics: Models, facts, and bank regulation. In G. Hubbard, editor, *Financial Markets and Financial Crises*. University of Chicago Press, 1991.

C. Calomiris and C. Kahn. The role of demandable debt in structuring optimal banking arrangements. *American Economic Review*, 81:497–513, 1991.

C. Calomiris and J. R. Mason. Fundamentals, panics, and bank distress during the depression. *American Economic Review*, 93:1615–1647, 2003.

G. Calvo. Capital flows and capital-market crises: The simple economics of sudden stops. *Journal of Applied Economics*, 1:35–54, 1998.

G. Calvo and E. Mendoza. Rational contagion and the globalization of securities markets. *Journal of International Economics*, 51:79–113, 2000.

F. Caramazza, L. Ricci, and R. Salgado. International financial contagion in currency crises. *Journal of International Money and Finance*, 23:51–70, 2004.

H. Carlsson and E. van Damme. Global games and equilibrium selection. *Econometrica*, 61:989–1018, 1993.

T. Chaney, D. Sraer, and D. Thesmar. The collateral channel: How real estate shocks affect corporate investment. *American Economic Review*, 102:2381–2409, 2012.

R. Chang and A. Velasco. A model of financial crises in emerging markets. *Quarterly Journal of Economics*, 116:489–517, 2001.

V. V. Chari and Ravi Jagannathan. Banking panics, information, and rational expectations equilibrium. *Journal of Finance*, 43:749–760, 1988.

Q. Chen, I. Goldstein, and W. Jiang. Payoff complementarities and financial fragility: evidence from mutual fund outflows. *Journal of Financial Economics*, 97:239–262, 2010.

Y. Chen. Banking panics: The role of the first-come, first-served rule and information externalities. *Journal of Political Economy*, 107:946–968, 1999.

S. Claessens and K. Forbes, editors. *International Financial Contagion*. Kluwer Academic Publishers, Norwell, MA, 2001.

H. L. Cole and T. Kehoe. Self-fulfilling debt crises. *Review of Economic Studies*, 67:91–116, 2000.

R. Cooper and T. W. Ross. Bank runs: Liquidity costs and investment distortions. *Journal of Monetary Economics*, 41:27–38, 1998.

D. Covitz, N. Liang, and G. Suarez. The evolution of a financial crisis: Collapse of the asset-backed commercial paper market. *Journal of Finance*, 68:815–848, 2013.

A. Dasgupta. Financial contagion through capital connections: A model of the origin and spread of bank panics. *Journal of the European Economic Association*, 2:1049–1084, 2004.

H. DeAngelo and R. Stulz. Liquid-claim production, risk management, and bank capital structure: Why high leverage is optimal for banks. *Journal of Financial Economics*, 116:219–236, 2015.

A. Demirguc-Kunt and E. Detragiache. The determinants of banking crises: Evidence from developed and developing countries. *IMF Staff Papers*, 45: 81–109, 1998.

M. Di Maggio and A. Tahbaz-Salehi. Financial intermediation networks. Working Paper, 2015.

D. W. Diamond. Liquidity, banks, and markets. *Journal of Political Economy*, 105:928–956, 1997.

D. W. Diamond and P. H. Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91:401–419, 1983.

D. W. Diamond and R. Rajan. Liquidity risk, liquidity creation and financial fragility: a theory of banking. *Journal of Political Economy*, 109:2431–2465, 2001.

J. Dow. Is liquidity self-fulfilling? *Journal of Business*, 77:895–908, 2004.

G. Eggertsson and P. Krugman. Debt, deleveraging, and the liquidity trap: A Fisher-Minsky-Koo approach. *Quarterly Journal of Economics*, 127:1469–1513, 2012.

B. Eichengreen, A. Rose, and C. Wyplosz. Contagious speculative attacks: First tests. *Scandinavian Journal of Economics*, 98:463–484, 1996.

A. L. Eisfeldt and A. Rampini. Capital reallocation and liquidity. *Journal of Monetary Economics*, 53:369–399, 2006.

H. Ennis and T. Keister. Run equilibria in the Green-Lin model of financial intermediation. *Journal of Economic Theory*, 144:1996–2020, 2009.

E. Farhi and J. Tirole. Collective moral hazard, maturity mismatch and systemic bailouts. *American Economic Review*, 102:60–93, 2012.

R. Flood and P. Garber. Collapsing exchange-rate regimes, some linear examples. *Journal of International Economics*, 17:1–13, 1984.

M. Friedman and A. J. Schwartz. *A Monetary History of the United States: 1867–1960*. Princeton University Press, 1963.

J. Gan. Collateral, debt capacity, and corporate investment: Evidence from a natural experiment. *Journal of Financial Economics*, 85:709–734, 2007a.

J. Gan. The real effects of asset market bubbles: Loan- and firm-level evidence of a lending channel. *Review of Financial Studies*, 21:2677–2703, 2007b.

S. Gerlach and F. Smets. Contagious speculative attacks. *European Journal of Political Economy*, 11:45–63, 1995.

M. Gertler and N. Kiyotaki. Financial intermediation and credit policy in business cycle analysis. In B. Friedman and M. Woodford, editors, *Handbook of Monetary Economics*. Elsevier, 2011.

M. Gertler and N. Kiyotaki. Banking, liquidity and bank runs in an infinite horizon economy. *American Economic Review*, 105:2011–2043, 2015.

R. Glick and A. Rose. Contagion and trade: Why are currency crises regional? *Journal of International Money and Finance*, 18:603–617, 1999.

V. Glode and C. Opp. Adverse selection and intermediation chains. Working Paper, 2015.

V. Glode, R. Green, and R. Lowery. Financial expertise as an arms race. *Journal of Finance*, 67:1723–1759, 2012.

I. Goldstein. Strategic complementarities and the twin crises. *Economic Journal*, 115:368–390, 2005.

I. Goldstein. Empirical literature on financial crises: Fundamentals vs. panic. In G. Caprio, editor, *The Evidence and Impact of Financial Globalization*. Elsevier, 2012.

I. Goldstein and A. Pauzner. Contagion of self-fulfilling financial crises due to diversification of investment portfolios. *Journal of Economic Theory*, 119: 151–183, 2004.

I. Goldstein and A. Pauzner. Demand deposit contracts and the probability of bank runs. *Journal of Finance*, 60:1293–1328, 2005.

I. Goldstein, E. Ozdenoren, and K. Yuan. Learning and complementarities in speculative attacks. *Review of Economic Studies*, 78:263–292, 2011.

G. Gorton. Banking panics and business cycles. *Oxford Economic Papers*, 40: 751–781, 1988.

G. Gorton. *Slapped by the Invisible Hand The Panic of 2007.* Oxford University Press, 2010.

G. Gorton and A. Metrick. Securitized banking and the run on repo. *Journal of Financial Economics*, 104:425–451, 2012.

G. Gorton and G. Ordonez. Collateral crises. *American Economic Review*, 104:343–378, 2014.

G. Gorton and G. Pennacchi. Financial intermediaries and liquidity creation. *Journal of Finance*, 45:49–71, 1990.

E. Green and P. Lin. Implementing efficient allocations in a model of financial intermediation. *Journal of Economic Theory*, 109:1–23, 2003.

Z. He and A. Krishnamurthy. A model of capital and crises. *Review of Economic Studies*, 79:735–777, 2012.

Z. He and A. Krishnamurthy. Intermediary asset pricing. *American Economic Review*, 103:732–770, 2013.

F. Heider, M. Hoerova, and C. Holthausen. Liquidity hoarding and interbank market spreads: The role of counterparty risk. *Journal of Financial Economics,* forthcoming, 2015.

C. Hellwig, A. Mukherji, and A. Tsyvinski. Self-fulfilling currency crises: The role of interest rates. *American Economic Review*, 96:1769–1787, 2006.

A. Hertzberg, J. M. Liberti, and D. Paravisini. Public information and coordination: evidence from a credit registry expansion. *Journal of Finance*, 66:379–412, 2012.

B. Holmstrom and J. Tirole. Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics*, 112:663–691, 1997.

B. Holmstrom and J. Tirole. Private and public supply of liquidity. *Journal of Political Economy*, 106:1–40, 1998.

M. Iacoviello. House price, borrowing constraints, and monetary policy in the business cycle. *American Economic Review*, 95:739–764, 2005.

V. Ivashina and D. Scharfstein. Bank lending during the financial crisis of 2008. *Journal of Financial Economics*, 97:319–338, 2010.

C. J. Jacklin. Demand deposits, trading restrictions, and risk-sharing. In E. Prescott and N. Wallace, editors, *Contractual Arrangements for Intertemporal Trade*. University of Minneapolis Press, 1987.

C. J. Jacklin and S. Bhattacharya. Distinguishing panics and information-based bank runs: Welfare and policy implications. *Journal of Political Economy*, 96:568–592, 1988.

G. L. Kaminsky and C. Reinhart. The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review*, 89:473–500, 1999.

G. L. Kaminsky and C. Reinhart. On crises, contagion, and confusion. *Journal of International Economics*, 51:145–168, 2000.

G. L. Kaminsky, C. M. Reinhart, and C. A. Vegh. The unholy trinity of financial contagion. *Journal of Economic Perspectives*, 17:51–74, 2003.

G. L. Kaminsky, R. Lyons, and S. L. Schmukler. Managers, investors, and crises: Mutual fund strategies in emerging markets. *Journal of International Economics*, 64:113–134, 2004.

T. Keister. Bailouts and financial fragility. *Review of Economic Studies,* forthcoming, 2015.

C. P. Kindleberger. Manias, panics, and crashes: A history of financial crises. Basic Books, 1978.

M. A. King and S. Wadhwani. Transmission of volatility between stock markets. *Review of Financial Studies*, 3:5–33, 1990.

N. Kiyotaki and J. Moore. Credit cycles. *Journal of Political Economy*, 105: 211–248, 1997.

N. R. Kocherlakota. Creating business cycles through credit constraints. *Federal Reserve Bank of Minneapolis Quarterly Review*, 24:2–10, 2000.

L. Kodres and M. Pritsker. A rational expectations model of financial contagion. *Journal of Finance*, 57:769–799, 2002.

P. R. Krugman. A model of balance-of-payments crises. *Journal of Money, Credit, and Banking*, 11:311–325, 1979.

P. R. Krugman. Balance sheets, the transfer problem, and financial crises. *International Tax and Public Finance*, 6:459–472, 1999.

J. C. F. Kuong. Self-fulfilling fire sales: Fragility of collateralised short-term debt markets. Working Paper, 2015.

A. S. Kyle and W. Xiong. Contagion as a wealth effect. *Journal of Finance*, 56:1401–1440, 2001.

R. Lagunoff and S. Schreft. A model of financial fragility. *Journal of Economic Theory*, 99:220–264, 2001.

R. Levine. Financial development and economic growth: Views and agenda. *Journal of Economic Literature*, 35:688–726, 1997.

G. Lorenzoni. Inefficient credit booms. *Review of Economic Studies*, 75:809–833, 2008.

F. Malherbe. Self-fulfilling liquidity dry-ups. *Journal of Finance*, 69:947–970, 2014.

A. Martin, D. Skeie, and E.-L. von Thadden. Repo runs. *Review of Financial Studies*, 27:957–989, 2014.

M. S. Martinez-Peria and S. L. Schmukler. Do depositors punish banks for bad behavior? market discipline, deposit insurance, and banking crises. *Journal of Finance*, 56:1029–1051, 2001.

E. G. Mendoza. Sudden stops, financial crises and leverage. *American Economic Review*, 100:1941–1966, 2010.

S. Morris and H. S. Shin. Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, 88:587–597, 1998.

S. Morris and H. S. Shin. Global games: Theory and applications. In M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, editors, *Advances in Economics and Econometrics*. Cambridge University Press, 2003.

S. Morris and H. S. Shin. Contagious adverse selection. *American Economic Journal: Macroeconomics*, 4:1–21, 2012.

S. Myers and N. Majluf. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13:187–221, 1984.

M. Obstfeld. Rational and self-fulfilling balance of payments crises. *American Economic Review*, 76:72–81, 1986.

M. Obstfeld. The logic of currency crises. *Cahiers Economiques et Monetaires*, 43:189–213, 1994.

M. Obstfeld. Models of currency crises with self-fulfilling features. *European Economic Review*, 40:1037–1047, 1996.

F. D. Oh. Contagion of a liquidity crisis between two firms. *Journal of Financial Economics*, 107:386–400, 2013.

M. Pagano. Trading volume and asset liquidity. *Quarterly Journal of Economics*, 104:255–274, 1989.

J. Peck and K. Shell. Equilibrium bank runs. *Journal of Political Economy*, 111:103–123, 2003.

E. Perotti and P. Golec. Safe asset demand: A review. Working Paper, 2015.

T. Philippon and V. Skreta. Optimal interventions in markets with adverse selection. *American Economic Review*, 102:1–30, 2012.

S. Radelet and J. Sachs. The east asian financial crisis: Diagnosis, remedies, prospects. *Brookings Papers on Economic Activity*, 1:1–74, 1998.

R. Rajan and L. Zingales. Financial dependence and growth. *American Economic Review*, 88:559–586, 1998.

A. Rampini and S. Viswanathan. Financial intermediary capital. Working Paper, 2015.

S. Salant and D. Henderson. Market anticipation of government policies and the price of gold. *Journal of Political Economy*, 86:627–648, 1978.

L. Schmidt, A. Timmermann, and R. Wermers. Runs on money market mutual funds. Working Paper, 2015.

M. Schneider and A. Tornell. Balance sheet effects, bailout guarantees and financial crises. *Review of Economic Studies*, 71:883–913, 2004.

E. Schroth, G. Suarez, and L. Taylor. Dynamic debt runs and financial fragility: Evidence from the 2007 abcp crisis. *Journal of Financial Economics*, 112:164–189, 2014.

L. Schumacher. Bank runs and currency run in a system without a safety net: Argentina and the 'tequila' shock. *Journal of Monetary Economics*, 46:257–277, 2000.

J. C. Shambaugh. The Euro's three crises. *Brookings Papers on Economic Activity,* Spring, 2012:157–231, 2012.

H. S. Shin. Reflections on Northern Rock: The bank run that heralded the global financial crisis. *Journal of Economic Perspectives*, 23:101–119, 2009.

J. Stein. Monetary policy as financial-stability regulation. *Quarterly Journal of Economics*, 127:57–95, 2012.

J. E. Stiglitz and A. Weiss. Credit rationing in markets with imperfect information. *American Economic Review*, 71:393–410, 1981.

J. Tirole. Overcoming adverse selection: How public intervention can restore market functioning. *American Economic Review*, 102:29–59, 2012.

C. Van Rijckeghem and B. Weder. Sources of contagion: Is it finance or trade. *Journal of International Economics*, 54:293–308, 2001.