Journal of Economic Studies

Banking on experiments?

Martin Dufwenberg

Article information:

Banking on experiments?

Martin Dufwenberg
Department of Economics, University of Arizona, Tucson, Arizona, USA; Department of Decision Sciences and IGIER, Bocconi University, Milan, Italy and Department of Economics, University of Gothenburg, Gothenburg, Sweden

Abstract
Purpose – How can laboratory experiments help us understand banking crises, including the usefulness of various policy responses? After giving a concise introduction to the field of experimental economics more generally, the author attempts to provide answers. The paper aims to discuss this issue.
Design/methodology/approach – The author discusses methodology and surveys relevant work.
Findings – History is often too complicated to be meaningfully revamped or modified in the lab, for purposes of insight-by-analogy. But as people argue about how to understand financial history, they bring ideas to the table. It is possible and useful to test the empirical relevance of these ideas in lab experiments.
Originality/value – The paper pioneers broad discussion of how lab experiments may shed light on banking crises.
Keywords Banking crises, Bank runs, Insolvency, Lab experiments
Paper type Research paper

1. Introduction
How can laboratory experiments help us understand banking crises, and the usefulness of various policy tools in this connection? After giving a concise introduction to the field of experimental economics (EE) more generally, I attempt to provide answers. I discuss methodological issues and survey relevant work.

Section 2 sets the stage: I first discuss the substantive issues – banks, crises, policy – that motivate me, then introduce EE, and finally describe the fields of behavioral finance and behavioral economics since some of the topics to follow draw on notions from these areas. Section 3 discusses how experiments can inform understanding of banking crises: I identify themes that can or cannot be addressed, and discuss in some depth two main tracks that emerge as relevant. Section 4 sums up and draws policy lessons.

I focus mainly on banking crises rather than other aspects of banking, and mainly on lab experiments rather than randomized controlled trials in the field. These are choices of focus only[1]. I neither mean to signal that issues left out are unimportant nor that experiments might not be useful in their connection.

2. Background
2.1 Crises, banks, policy
The field of finance is concerned with value and economic exchange in regards to streams of income over time[2]. How will financial assets be priced? Who lends to whom on what terms?

The term financial crisis is used differently by various scholars; see Mishkin (1992) for an interesting discussion. His preferred definition requires that an unexpected change in some economic variable skews real investment incentives away from what

JEL Classification — C92, G21

Journal of Economic Studies
Vol. 42 No. 6, 2015
pp. 943-971

0144-3585
DOI 10.1108/JES-08-2015-0145

would be most productive, because of adverse selection or moral hazard. Many take a less structured approach, equating a crisis with large unexpected losses of value. Examples include exchange rate collapses, stock market crashes, bursting asset price bubbles, sovereign government defaults, and bank failures. The word "crisis" suggests that such episodes often involve major disruptions in people's lives, as savings evaporate or layoffs occur (say by a distressed bank or government) or, if falling demand and failing businesses reinforce each other, a recession or depression ensues. Avoiding financial crises may be an important policy objective if these episodes are considered grim.

This study is concerned with banking crises, where banks enter financial distress. Following Calomiris (2008), this may happen for two different reasons:

Banking crisis reason no. 1: Bank runs.
Banking crisis reason no. 2: Insolvency.

Bank runs occur when suddenly so many depositors decide to withdraw their savings that a bank finds itself illiquid. Bank runs are sometimes called banking panics (especially if all banks face runs). Insolvency occurs when the value of a bank's assets fall enough that it cannot meet its obligations, for example, if a real estate market collapses and people default on mortgage loans.

Scholars differ in their outlook on whether bank runs or fundamental shocks that cause insolvency are more important (and also whether one may cause the other). Calomiris (2008) explains how the idea that bank runs are key gained popularity through Friedman and Schwartz' (1963) classic book *A Monetary History of the United States*. His own view, based on a careful scrutiny of bank balance sheets, is that shocks causing insolvency are more important. Calomiris does not argue that runs do not happen, only that they are not associated with great social costs[3]. However, many other economists seem to attribute a much larger role to runs as being socially costly and important for understanding banking crises. For example, Bernanke (1983) argues that runs played an important and detrimental role causing economic losses during the great depression. More recently, witness the remarks of Larry Summers (2008) in his Arthur M. Okun Lecture "Learning from and Responding to Financial Crisis" delivered at Yale (check after 43 minutes). Lack of consensus is reflected also in the (FCIC/Angelides commission's) Financial Crisis Inquiry Report (2011), which reached partly split conclusions (along partisan lines).

The explanations of banking crises are important because the optimal policy response (e.g. whether or not to use deposit insurance) differs depending on the underlying reason. The controversy among scholars how best to explain then offers a nice starting point when considering whether and how a novel approach may inform the debate!

2.2 Experimental economics (EE)

EE differs from other subfields of economics like finance, public economics, labor, or social choice by being characterized by its method rather than its topic. This study belongs to macro and finance by its focus on banking crises, and to EE as it concerns using experiments. Experimental approaches to banking crises are not that common. I suspect some of my readers will be banking scholars without experience of EE. They may find it useful to get a concise introduction to EE before I approach the banking crisis theme. The goal of this section is to provide that. I start with a specific example of EE-research.

2.2.1 The Bubbles Example. Many historical episodes involve dramatic price hikes of some commodity or asset, followed by a sudden collapse. Perhaps most famous is the Dutch "tulipmania," where prices of tulip bulbs reached levels several times the yearly salary of an average wage earner, and then suddenly crashed in February 1637[4]. More recently we have the development of the Nasdaq around 2000, and the US real estate market 2003-2010. The patterns are often seen to involve "bubbles," a term suggesting that prices exceed the expected present value of all future returns to a given asset (the "fundamental value"). Commentators, who view some asset price development as a bubble often use terms like hysteria, mania, panics, or (Alan Greenspan's) "irrational exuberance," as in the titles of Kindleberger's (2001) and Shiller's (2000) books.

It is, however, treacherous to evaluate the degree of madness of a market, as it is difficult to know what the fundamental value really should be. In fact, some economists call to question the bubble/crash terminology altogether, arguing that what at first glance seem hysterical may on closer scrutiny have sensible fundamental explanations (say, changes in discount rates or profit expectations; see, e.g. Garber, 2000).

The situation is imperfectly understood and hotly debated. Are bubbles common or rare? The issue is policy relevant. Bubbles and crashes may generate huge shifts of wealth, and may affect how investments are made. If investment occurs in industries valued above fundamental value, then it may not always be the most lucrative investment projects that are embarked on. Whether bubbles occur moreover matters for evaluating policy, say the effect of turnover taxes in financial markets (so called "Tobin taxes") or the effect of monetary or fiscal policy.

Experiments offer a novel way to illuminate the issue. In the lab, one can construct a market with trading rules that share key features with a stock exchange, but where the dividend structure is controlled by the experimenter such that a fundamental value can be calculated. That value can be compared to the prices generated in the experimental market, so that insights on the bubbles vs fundamentals issue can be drawn by analogy.

The pioneering contribution along these lines is Smith *et al.* (1988). A large-related literature has followed. I will describe (part of) of a design used by Dufwenberg *et al.* (DLM, 2005), which is comparable.

DLM study experimental markets each involving six traders (students, recruited at Virginia Polytechnic Institute, where the experiment was run). Trade took place through continuous-time trading according to rules akin to those in stock markets, so-called "double-auction" rules, in which presumptive sellers place and revise asks, while buyers place and revise bids, and anyone can accept the lowest ask/highest bid at any time. A market involved ten periods. The financial assets traded had a life corresponding to those ten periods. In each period trading was allowed for two minutes. After each period, whoever owned the financial asset being traded was entitled to a stochastic dividend, which was either 0 or 20 US cent each with probability 1/2. This was publicly made known to everyone.

An asset's fundamental value can be calculated looking at the expected value of dividends to come (abstracting away from risk premia). With one period left, an asset should be worth 10 cents ($=1/2\times 0$ cents $+ 1/2\times 20$ cents); with two periods left, it should be worth 20 cent ($=2\times(1/2\times 0$ cents $+ 1/2\times 20$ cents)); with three periods left, it should be worth 30 cent, etc. The trading prices in the experiment can be compared to these fundamental values.

The experiment involved interaction over four rounds, each round corresponding to a ten-period market as described above. The spirit of the results is illustrated in Figure 1, based on data from a selected somewhat typical session (DLM's "Session 4"). Note that the horizontal axis re-starts the period count after 10, reflecting the four rounds. For now, focus only on the market of round 1, i.e., the leftmost ten-period block in the figure. The straight line that descends linearly from 100 (cents) to 10 indicates fundamental value. The other line indicates average price observed in the market (DLM's "Session 4").

One can note a distinct bubble. Trades initially occur on average slightly below fundamental values, but soon exceed them, eventually substantially. It seems natural to draw the conclusion that bubbles may be common in financial markets since they appeared in this design (beware, however, that this conclusion is subject to caveats when the markets are repeated so that traders gain experience; this topic will be discussed at some length in Section 3.3).

2.2.2 Treatments, causality, and theory testing. Experiments often involve treatments: multiple versions of a design, typically differing just in a single aspect, so that one can isolate the effect of that aspect. To exemplify, use *The Bubbles Example*: suppose someone were interested in exploring the effect of trader fatigue on bubble formation. That researcher may complement the DLM design with an additional treatment identical in every respect except that the session in question were run late at night.

Control offered by treatments is a main selling point of experiments, but having treatments should not be taken as the defining characteristic and in fact not all experiments have treatments. *The Bubbles Example* can illustrate. No change of design-feature was described (except the thought-up one of the previous paragraph), yet something useful was produced, namely, actual price data to be compared to theoretical fundamental values.

It makes methodological sense to distinguish between testing causal hypotheses (via treatments) and testing a theoretical prediction (by providing a condition such that the theory could apply). Of course, in practise, causal hypotheses are often derived through a formal theory (via comparative statics) and researchers end up testing theory in designs with many treatments[5].

2.2.3 Goals and topics. Almost 20 years old, the *Handbook of Experimental Economics* (Kagel and Roth, 1995) remains a good guide to its topic. In the introductory

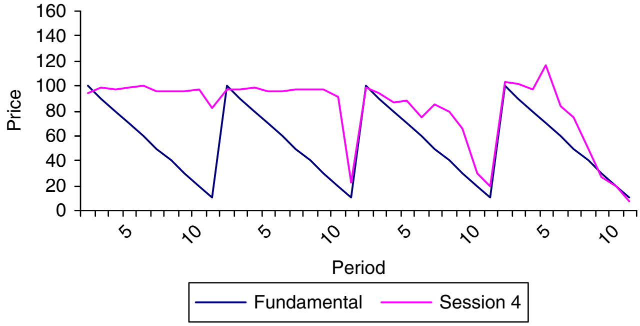

Figure 1.
Observed mean prices, alongside fundamental value

chapter Al Roth explains how experiments are run for different reasons (largely reflecting why empirical studies are done more generally):

Reason no. 1: Speaking to theorists.
Reason no. 2: Searching for facts.
Reason no. 3: Whispering in the ears of princes.

A propos no. 1: economic theorists tell stories about how thought-up worlds work. This may be valuable per se, but the scholarship gains relevance if it is empirically relevant. Experiments can evaluate that and *The Bubbles Example* illustrates: let the theoretical story be that market prices reflect fundamental values. The experiment tests that proposition.

No. 2: experiments sometimes just document interesting data (possibly to be addressed by future theory). *The Bubbles Example* illustrates, if its interpretation is tweaked. Suppose one did not refer to any theory, that one were merely curious about what patterns of trade emerge. The experiment produces data. Another, perhaps more interesting, example would be an experiment recording treatment effects (e.g. the trader-fatigue design mentioned earlier).

No. 3: the idea would be to offer policy advice. *The Bubbles Example* illustrates; recall the comments on potential relevance for evaluating turnover taxes. That angle is somewhat indirect though. To get a more clearly relevant example, imagine a new treatment, including a laboratory turnover tax, to be compared to the treatment previously described.

What topics did experimental economists explore? In the *Handbook*'s introductory chapter, Roth also gives a nice overview of the early history of EE (≈1940-1970). Contributions concerned three areas: first, individual decisions; second, game theory; third, market experiments. This tri-part classification works for organizing research to this day[6]. The third area was already illustrated, through *The Bubbles Example*. Examples related to second area come in Section 3.2, as bank run experiments belong to the category.

2.2.4 Vernon Smith's Nobel. Space constraints and desire to focus imply that I will not survey EE in much more detail. However, I would be remiss unless I commented on the scholarship of Vernon Smith, which landed him the Nobel Memorial Prize in 2002 (shared with Daniel Kahneman):

A basic tenet of much economic scholarship is that one may understand markets with many buyers and sellers such that price and quantity is given by the intersection of demand and supply curves. That outcome is often "efficient," meaning that no potential gains from trade go unexploited. It is important to assess whether or not this story is empirically relevant.

Smith's (1962) classic study of double-auction markets, in which buyers and sellers are active posting public bids and asks just like in a stock exchange, tackles that task. Trade took place in a commodity for which Smith "induced" valuations/production costs. Buyers were told that if at the end of a trading session they owned a unit, they would get paid a certain amount of money. Sellers were told that if they produced and sold a unit they would have to pay a certain cost. Controlling valuations/costs this way, the classical market predictions become testable.

The data matched the theory well. Later research has shown that requirements to get efficient outcomes are even less restrictive than theory would suggest, as the markets work well even if the number of participants is small, say two or three buyers or sellers. Smith and others have also performed research on how a wide variety of other market forms (e.g. monopoly markets, posted offer markets, duopolistic competion, auctions) work. More often than not, the outcomes are rather efficient, although in that regard the double auction is hardly surpassed[7].

Recall *The Bubbles Example*. The design involved double-auction trading rules. In light of what is known about the efficiency properties of such trading rules, with induced values, it is remarkable that bubbles may occur when experimental financial assets are traded.

2.3 Two "behavioral" fields

I briefly introduce two other fields that have some bearing on the banking discussion to follow.

2.3.1 Behavioral finance. "[…] uses insights from psychology to understand how human behavior influences the decisions of individuals and professional investors, markets and managers." The quote is from a recent textbook by Ackert and Deaves (2010, p. xxvi). My presentation of the field to follow draws in large part on the survey by Barberis and Thaler (B&T, 2003)[8].

Classical finance is built on the Efficient Market Hypothesis (EMH), according to which market prices track fundamental values (= the present value of expected returns). Texts motivating behavioral finance often describe data at odds with the EMH. For example, B&T (Section 2.3.1) describe how two companies "agreed to merge their interests on a 60:40 basis" (and so face proportional returns in the future) and yet did not exhibit a 60:40 price-of-equity ratio. Another case where the EMH is violated is *The Bubbles Example* (2.2).

B&T argue that there are two "building blocks" to behavioral finance:

Building block no. 1: Limits to arbitrage.
Building block no. 2: Psychology.

A common defense of the EMH is that if markets were not efficient rational investors would make profit through risk-less arbitrage, in the process making markets efficient. Building block no. 1 points out a variety of reasons why engaging in arbitrage may be difficult, or impossible. For example, arbitrage may require short-selling, which may be costly or not allowed. Or if there are "noise traders" in the market, people whose decisions cannot be described as rational (cf. De Long *et al.*, 1990), then it can be a rational response to not let fundamental value govern trade. *The Bubbles Example* illustrates: if a rational investor anticipates a bubble rising then, despite that fundamental value is lower than current prices, it would be rational to buy and not to sell. The idea would be to "ride-the bubble," and get out just before it bursts. This may feed the bubble further.

Building block no. 2 becomes relevant if one rejects the EMH. What comes instead? A literature attempts to describe how aspects of psychology shape investor strategy and pricing. B&T cover notions of overconfidence, wishful thinking, base rate neglect, gambler's fallacy, belief anchoring, availability heuristics, reference dependent utility, loss aversion, ambiguity aversion, sample size neglect, and the hot hand phenomenon[9]. I will describe only the last two notions, which have direct bearing on things to come (3.2). Take it from B&T (p. 1065):

> Sample size neglect means that in cases where people do not initially know the data-generating process, they will tend to infer it too quickly on the basis of too few data points. For instance, they will come to believe that a financial analyst with four good stock picks is talented because four successes are not representative of a bad or mediocre analyst. It also generates a "hot hand" phenomenon, whereby sports fans become convinced that a basketball player who has made three shots in a row is on a hot streak and will score again,

> even though there is no evidence of a hot hand in the data (Gilovich *et al.*, 1985). This belief that even small samples will reflect the properties of the parent population is sometimes known as the "law of small numbers" (Rabin, 2002).

These phenomena bear on how to understand *The Bubbles Example*. If traders see prices rise, they may simply believe the pattern will continue, so it seems wise to buy, not to sell[10],[11]. We will return to related topics, connecting them directly to some banking issues, in Section 3.2.

2.3.2 Behavioral economics. To define this field, reuse Ackert and Deaves's definition of behavioral finance (cited above), except that the goal is to understand economic outcomes more broadly. Viewed this way, behavioral finance is a special case of behavioral economics. However, it makes sense to present the fields separately, because techniques tend to differ, behavioral economics being the more theory-driven discipline. Drawing on game theory, behavioral economists have developed models of bounded rationality, of various forms of human motivation (say altruism, fairness, reciprocity, emotions, or social status), and of psychological biases. They like to argue that, and show how, nuanced notions of psychology shapes economic outcomes in important ways, and they use experiments to test the empirical relevance of the psychological stories that they tell.

The field is huge. To comprehensively survey it is beyond my scope. For entries to parts of the literature, try Camerer (2008), or look up the textbook Cartwright (2014). The one area I wish to highlight here, because it plays a role below, is evidence and theory of reciprocity, the inclination to be kind in return to kindness and to punish unkindness. I postpone giving relevant references until Section 3.3, where a banking topic where reciprocity matters is brought up.

3. Experiments on banking

3.1 Roadmap

There is controversy regarding how to explain, and avoid, banking crises. Experimental economists bring new tools to the discussion. This subsection attempts to identify areas where to dig; the following two subsection then explore what's been done.

To get an anchor for the discussion, consider how – in a briefing titled "Banking Crisis Yesterday & Today" – Calomiris (2009a, cf. 2009b) sums up his views:

> This brief survey of the history of banking crises traces unusual bank fragility to risk-inviting microeconomic rules of the banking game established by governments, the most important of which have been rules that subsidize risk. Other destabilizing rules include limits on bank entry and the failure to establish a proper lender of last resort. The subprime crisis exemplifies the historical pattern all too well. Government subsidization of risky mortgages in the U.S. accelerated markedly in the years prior to the crisis. That along with prudential regulatory failures to prevent excessive risk taking allowed the mortgage risk binge of 2003-2007 to produce a worldwide financial collapse. As the U.S. gears up to respond to the subprime crisis with regulatory reforms, history suggests important lessons.

The passage reflects two ideas: first (between the lines), insolvency is a more important source of banking crises than bank runs (cf. 2.1). Second, the reason is bad government policy causing behavior changes (a form of moral hazard).

I think it is fair to say that many economists would somewhat dispute the account. For example, Bernanke (1983) argues that bank runs "were clearly important part of the banking problems" during the great depression (p. 260). More recently, in his 2008 Okun Lecture, Larry Summers, who helped engineer much policy, while not thinking about or engaging in discussion geared directly to Calomiris, stressed the importance of bank runs for understanding banking crises and discussed the need for government intervention in this connection rather than focussed on the problems that government activity may have caused. My impression is that Summers would want to add riders to Calomiris' account.

Who is right? I will not say. I consider whether experiments can illuminate the debate. It does not take much reflection to see why one may hope this could be done. Counterfactual histories are difficult to observe! To test the empirical validity of Calomiris' claim it would be useful to compare the existing record with what would have happened had government pursued other policies. It is impossible to change history, but one wonders: can lab experiments help?

3.1.1 Cul-de-sac. My answer, as regards getting direct evidence, is that this may prove difficult or impossible. Reflect on what it would take. The key problem, described in the previous paragraph, is lack of data regarding counterfactual circumstances. Lab experiments would seem to have a shot at providing that, as a virtue of the lab is that one can compare treatments. However, the real world may prove too complex to allow direct insights-by-analogy that way. History involved a very complex game, with bank managers, their employees with varying incentives, their customers with their lives and trade-offs and deposit decisions, and government with all its people involved. Anyone who reads the Financial Crisis Inquiry Report (2011) will see this clearly. It is hard to imagine a design that would shed light on the empirical relevance of claims such as Calomiris' through an insight-by-analogy approach. That's the bad news.

3.1.2 Ways to go. Now the good news: experiments may inform the debate regarding how to understand and react to banking crises more indirectly, by evaluating the empirical relevance of various ideas that economists bring up when they engage in debate. That is, rather than attempt to evaluate historical events by creating grand-scale parallels, experiments can zoom in on isolated building blocks that economists argue matter and evaluate the relevance of those.

Section 3.2 considers the idea that bank runs cause crises. Economists have traditionally given substance to that via models where bank runs make sense. In particular, in a classic contribution Diamond and Dybvig (D&D, 1983) model a bank run as the occurrence of a "bad equilibrium" in a game admitting two equilibria. The other equilibrium is "good," and involves no run. What if players always coordinate on that one? That would undermine the relevance of the D&D model. Lab experiments can help illuminate the issue. In the lab one may create a D&D world, and explore the conditions under which its bank run equilibrium is viable. Whatever findings obtain, it may affect the confidence with which scholars can argue that the D&D story for explaining bank runs is relevant.

Section 3.3 considers the insolvency explanation of banking crises. Calomiris claims that banking crises largely reflect insolvency due to government policies that stimulated excessive risk taking. Is that plausible? I will discuss ways to evaluate that.

3.2 Bank runs

With 7000 + Google Scholar citations, Diamond and Dybvig's (D&D, 1983) model of bank runs is a classic. Calomiris and Gorton (1991, p. 120) summarize the outlook that motivated the work:

> [In the early 80s], theoretical work on banks and banking panics was aimed at addressing the following questions: How can bank debt contracts be optimal if such contracts lead to banking panics? Why would privately issued circulating bank debt be used to finance

nonmarketable assets if this combination leads to socially costly panics? Posed in this way, explaining panics was extremely difficult.

D&D depict banking as game with one good equilibrium where people keep their money in the bank and let them grow with a profit, and a bad equilibrium where everyone panics and withdraws deposits so that banks run out of cash and collapse. Readers satisfied with that account may fast forward a few paragraphs; for others I illustrate D&D's model through an example that highlights key features and intuitions[12].

Consider a society with n individuals each of whom has 1 unit of money. There are three periods: 0, 1, 2. Technology permits investments made in period 0 to grow to $R > 1$ units in period 2. However, if an invested unit is pulled back in period 1, the value stays at 1 (it takes two periods for projects to bear fruit). All individuals make the investment in period 0 but a proportion $t < 1$ of "need money tomorrow" and are forced to pull their investments back at period 1; t is known by all beforehand although in period 0 no one knows who will be forced to pull back early. Let $u(w)$ be the utility of getting w. Each individual gets expected utility:

$$t \times u(1) + (1-t) \times u(R) \quad (1)$$

Individuals are risk averse: $t \times u(1) + (1-t) \times u(R) < u(t \times 1 + (1-t) \times R)$, for all $R > 1$. There is scope for a bank to improve the outcome by acting as an intermediary between the individuals and the technology. Instead of investing directly in the technology, individuals become bank customers. Each deposits his unit in period 0, then chooses when to withdraw funds. By judicious choice of offered returns, the bank may provide valuable insurance.

I illustrate under the assumptions that $n = 6$, $t = 1/2$, $R = 4$, and $u(w) = 2\text{-}2/w$ for all $w > 0$[13]. Using (1), one sees that without a bank each individual would get expected utility:

$$1/2 \times \left(2-2/1\right) + 1/2 \times \left(2-2/4\right) = 3/4$$

Suppose a bank offers a return of $1 + r$ if a withdrawal is made in period 1 and a return of $(1-r) \times 4$ if a withdrawal is made in period 2, so the bank breaks even when three-out-of-six customers withdraw early[14]. A customer's expected utility in period 0 is:

$$1/2 \times \left(2-2/(1+r)\right) + 1/2 \times \left(2-2/((1-r) \times 4)\right)$$

If, for example, $r = 1/3$ the expression becomes:

$$1/2 \times \left(2-2/\left(1+1/3\right)\right) + 1/2 \times \left(2-2/\left(\left(1-1/3\right) \times 4\right)\right) =$$
$$= \left(1-3/4\right) + \left(1-3/8\right) = 7/8$$

Since $7/8 > 3/4$, consumers are better off with the bank[15].

The analysis so far assumed that only the three-out-of-six customers who need the money in period 1 withdraw early. The described pattern is a Nash equilibrium (meaning that each customer optimizes, given the behavior of others). The three-out-of-six customers who need money in period 1 obviously do what's in their best interest. So do the other three, who each get $(1-r) \times 4 = (1-1/3) \times 4 = 8/3$, which is more than those who withdraw in period 1 get.

However – and here comes a key insight – the game has a second, bad, equilibrium in which all customers withdraw their money in period 1, whether they need the money or not. To see this, consider a customer who does not need money in period 1, but who believes the other five are all withdrawing then. Since 6×1 units were deposited in period 0, five withdrawals of the others in period 1, if served, would amount to $5 \times (1+r) = 5 \times (1+1/3) = 20/3$. Since $20/3 > 6$, there would not be enough money in the bank to serve these five early withdrawers. Consequently, there will also be no money left for our customer if he waits till period 2. His best response is to withdraw in period 1 too.

This "bank run equilibrium" is inefficient since the investments never yield the desired return; no units get quadrupled. In fact, it is worse than just keeping the money in the mattress (not even investing in the first place), as in the bank run four-out-of-six customers get $1+r = 1+1/3 = 4/3$, one customer gets 2/3, and one customer gets nothing [16]; in expectation they each get 1, but they are risk averse and would have preferred to get 1 for sure.

Note the key role played by $r > 0$. On the one hand, it is through the ability to offer a period 1 positive return to early withdrawers that the bank acquires its raison-d'être, the ability to offer insurance to risk-averse individuals. On the other hand, once a sufficiently high $r > 0$ is in place, the possibility of a bank run comes alive.

3.2.1 R as in insolvency?. D&D's bank run equilibrium can potentially explain how inefficient outcomes arise merely because economic agents come to coordinate on an inferior equilibrium, rather than because of some other adverse economic shocks. As indicated before, the empirical relevance is contested (e.g. Calomiris ,and Mason, 1997, 2003; cf. Section 2.1). However, to this day it is nevertheless often mentioned as highly relevant. I previously cited Bernanke's (1983) influential article which gives an important role to bank runs during the Great Depression, and Summers (2008) says that "this Diamond-Dybvig bank run metaphor has been very powerful in influencing financial policy in a wide range of situations,"[17] which echoes his earlier (2000) remarks (in a Richard T. Ely Lecture) that while the likelihood of bank runs "is driven and determined by the extent of fundamental weaknesses" it is also an issue of "bank-run psychology" (p. 7)[18]. Another example: Paul Krugman explained in his *New York Times* column that he is "a Diamond and Dybvig guy" (January 11, 2010).

In the D&D model, R is the returns offered by investments. It is useful to point out how to think of the link between R and the bank runs vs insolvency debate. In D&D's model, R is a constant. In the real world, R is surely not constant; returns depend in complicated ways on the state of the macro-economy, and on lending practices, entrepreneur behavior, credit ratings, and government policy. R not being a constant is what the insolvency aspects of banking crises are all about! For example, when Calomiris blames government for the financial crisis, one way to think of this is that policy affected incentives so that behavior changes caused R to dip, a form of moral hazard. The D&D model, with R constant, abstracts away from all that, which allows explaining how a banking crisis could arise for reasons that have nothing to do with insolvency[19].

3.2.2 Experiments. Lab experiments can inform the debate by evaluating the arguments on which debaters build when they discuss how crises work. The D&D story is a key part, and lab experiments may evaluate its empirical relevance under various conditions. Several rather recent studies have done this. They involve games that in structure resemble the D&D model, although details differ. For the most part I will not attempt to describe these details, but will rather concentrate on main insights. I finally say something about angles that have not been addressed but which it would seem interesting to explore.

3.2.3 Madiès (2006). A first finding is that most of the time neither of D&D's equilibria occurs. Some subjects withdraw in period 1, some do not. This is true even as the design allows subjects to gain experience, playing D&D games 30 times. Similar observations apply to most of the other studies to be discussed.

Madiès explores treatments that give subjects "more time to think before making further decisions" (p. 1854), interpreted as "suspension of deposit convertibility," which D&D show may be helpful for avoiding runs. The interpretation seems questionable as in D&D's theory, but not in the design, suspension entails reaching period 2 (where investment gains accrue) with denied opportunity of prior withdrawals. Anyway, Madiès' design feature reduces period 1 withdrawals, suggesting that in some situations a cooling off period or bank holiday may reduce panics. One may well imagine, however, that closing banks can possibly have the opposite effect, that is to increase panics[20].

Another set of results concern "deposit insurance" (guarantees of a certain level of return in case of a run). Madiès explores two "partial insurance" (75 and 25 percent) schemes. He reports that neither works, concluding that "only a total deposit coverage is effective in preventing bank runs" (p. 1855). One should note that Madiès scheme differs from D&D's in nature. D&D prove (their section V) that in their model deposit insurance eliminates runs. Their scheme is self-financed, with a form of ex post depositor taxation based on the number of withdrawals[21]. Madiès design, by contrast, involves outright transfers in case of many withdrawals and so does seem to not furnish a clear test of D&D's result. At the same time, his structure may resemble some deposit insurance schemes used in practice more than D&D's scheme does.

3.2.4 Garratt and Keister (G&K, 2009). G&K's experimental study have a 2×2 treatment structure. The first treatment variable concerns whether or not a random subset of players are forced to withdraw early[22]. G&K interpret this feature, the presence or absence of forced withdrawals, as a "proxy for macroeconomic conditions" (p. 301). The second treatment variable concerns how many opportunities subjects have to withdraw during period 1, one or three (where after each opportunity subjects learn how many others withdrew); G&K say this "adds a realistic feature of banking: depositors have a period of time during which they can choose to withdraw their funds, and they are able to observe some information about actions of other depositors, for example, by noticing if a line is forming outside the bank" (p. 301). Across treatments, both D&D equilibria are always viable.

Regarding the treatments with multiple withdrawal opportunities during period 1, initially the withdrawal rates were the same as in the treatment with a single withdrawal opportunity, however, exposure to the occasional bank runs that occurred early on had a greater effect leading to more voluntary period 1 withdrawals later on. G&K write: "the ability of people to coordinate on the payoff-dominant equilibrium [without a bank run] is sensitive to the presence of aggregate uncertainty about fundamental withdrawal demand, even when this uncertainty alone poses little or no threat to the solvency of the bank" and the random forced withdrawals "mimic the type of uncertainty that is likely to be present under unfavorable macroeconomic conditions or in times of financial distress" (p. 311). It is interesting to compare this finding to Summers' (2000) hypothesis that "bank runs [...] are not driven by sunspots" but rather "the extent of fundamental weaknesses" (p. 7)[23]. G&K's design explores related ideas, but falls short of being able to evaluate the following: Suppose macroeconomic conditions go sour, and yet a particular D&D style banking game has its payoffs unchanged (unlike G&K's games, where payoffs change alongside). Could it be that the tendency toward the bank run equilibrium is nevertheless enhanced? Put differently, G&K's forced withdrawals not only add an element of uncertainty but also change the riskiness of not withdrawing as the expected number of early withdrawers change; the design does not disentangle these effects.

As regards the treatments with multiple withdrawal opportunities during period 1, more actual period 1 withdrawals appear. G&K interpret this to "suggest that in countries where people have a history of exposure to financial crises, withdrawal behavior might depend on […] the information flow regarding the withdrawal activity of others" (p. 311).

3.2.5 Schotter and Yorulmazer (S&Y, 2009). S&Y are more interested in the dynamics of runs than in whether or not runs occur. They explore a design that differs from the previous two in that they consider cases where the bank is, at some point, for sure insolvent. Their games have a four-period structure, and in many cases all equilibria involve runs before period 4. The design includes treatments manipulating what subjects learn about others' withdrawals, and whether or not "insiders" know about degree-of-solvency. S&Y also explore treatments with partial deposit insurance.

One treatment concerning information about others' withdrawals is akin to G&K's triple-withdrawal opportunity; the results are comparable in that multiple opportunities to learn about others' withdrawals increases withdrawals.

Another result, counter to Madiès', is that partial deposit insurance "can help diminish the severity of bank runs" (p. 217). The (transfer) form of the insurance is similar to Madiès'.

S&Y also report results related to notions of transparency and insiders. One intriguing example is that the presence of insiders "mitigates the severity" of runs (p. 217). The reason seems related to the design feature (not seen in Madiès or G&K) that some banks are insolvent, so depositors may interpret failure of insiders to withdraw as evidence that their bank is solvent. Conceivably this supports the idea that insider trading can be useful (although I would suspect there are many counter-arguments to that idea which are not addressed by S&Y's design).

3.2.6 Kiss et al. (KRR, 2012, 2014a, b, c). KRR (2012) have a design in which two subjects act as depositors, interacting with each other and with a third computer simulated early withdrawer depositor. Their two treatment variables concerns whether withdrawals are simultaneous or sequential with depositors informed of preceding decisions, and the degree of deposit insurance applied to payoffs (none, partial, or full; again, a transfer rather than a financed scheme). The "sequential setup decreases significantly the likelihood of bank runs" and "the main contribution" is that if "decisions are not simultaneous but sequential […] deposit insurance decreases the likelihood of bank runs, but the effects of full and partial insurance are not significantly different" (p. 1654). In related studies, Kiss *et al.* (2014a, b, c) study further variations on the information theme: three depositors act in sequence but the information about preceding decisions is altered. Kiss *et al.* (2014a) find, for example, that runs are much less frequent whenever depositor 2 is informed of depositor 1's decision (whatever it may be). Kiss *et al.* (2014b, c) explore, respectively, gender differences and cognitive ability in such settings. One finding is that men and women seem equally likely to panic.

On the one hand, there may be limits to how much general insight one may take away from these results since the imposed structure, with an exogenously given queue of depositors making their decisions, is so special and stylized. (e.g. how much sense does it really make, in general rather than in a queue, that depositors may observe others committing not to withdraw?) On the other hand, KRR's studies usefully highlight and alert us to how the prevalence of runs may depend on details concerning observability among the depositors[24]. Their finding in this regard indicate that the topic may deserve more scrutiny, including how to endogenize the timing of the order in which depositors move.

3.2.7 Arifovic et al. (AJX, 2013), Arifovic and Jiang (A&J, 2014). By varying parameters in the D&D model, one can influence how many period 1 withdrawals it takes to make a depositor who is not forced to withdraw in period 1 nevertheless want to do so. For example, consider my example above, with $n=6$, $t=1/2$, $R=4$, $r=1/3$. Simple calculations show that a customer who believes that exactly three others withdraw in period 1 (the three customers forced to do so, plus one more) would be indifferent between withdrawing or not. If instead $r>1/3$ ($r<1/3$) he would prefer (not) to withdraw. It seems very plausible that changes of that sort, say changes that make r lower, make a bank *r*un less likely. AJX report, essentially, that this intuition is supported by their data. Their data can be organized into three coordination parameter regions, depending on whether runs predictably do or do not occur, plus an intermediate indeterminacy region where the outcomes vary widely across different games.

D&D pointed out that, in principle, equilibrium selection may depend on a "sunspot variable," some phenomenon that everyone observes but which is unrelated to the structure of the banking game per se. Based on AJX's results, A&J conjecture that "if a sunspot variable is introduced to the bank run game, its power as a coordination device is likely to be weak if the coordination parameter lies in the run or non-run region, but strong if the parameter is in the indeterminacy region" (p. 3). They present a design where the sunspot takes the form of a randomly generated announced forecast how many depositors will withdraw funds, and report experimental support for their conjecture.

3.2.8 Klos and Sträter (K&S, 2012). Coordination games, of which bank run games are a special case, have multiple equilibria. The "global games" approach (Carlsson and van Damme, 1993) shows that, in many cases, if one incorporates uncertainty about a payoff parameter (think of R in a D&D game), and assume players get private signals of this parameter, then there may be a unique equilibrium (where behavior may change with the strength of the signal). See Dasgupta (2004), Goldstein and Pauzner (2005), and Rochet and Vives (2004) for applications to bank runs. There is an experimental literature on global games (see e.g. Heinemann *et al.*, 2004 study of currency attacks) and K&S (2012) take first steps testing global games-related ideas in experimental bank run games. They compare the predictive power of global games theory vs so-called level-k theory, declaring the latter winner (see Crawford *et al.*, 2013, for an introduction to level-k theory, and note their remarks in section 7 concerning bank runs).

3.2.9 Contagion. The studies discussed so far concern runs in given banks. When there are several banks around there is an issue of bank run contagion in which runs at one bank, when observed by depositors in other banks, induce runs there. Two recent, very similar in structure, experiments by Brown *et al.* (2012) (BTV) and Chakravarty *et al.* (CFK, 2013) are aimed at understanding the mechanics of bank runs in such setting. Their designs involve two banks; the customers of one make their decision to withdraw or not before those of the other bank. Some of the parameters of the environment (analogous, say, to my R above) are subject to stochastic shocks (which are small enough to not change the D&D style multiple equilibria feature of the games, but which move basins of attraction somewhat like in AJX). Treatments concern whether or not the realizations of the shocks that concern the first bank are correlated with the shocks that apply to the second bank. The question is, to cite BTV (p. 3) "whether bank runs are only contagious when there are economic linkages between the banks." Both studies report evidence that linkages/correlation fosters bank run contagion. Without linkages/correlation, the two studies reach somewhat different conclusions. BTV do not observe contagion, while CFK do (though to a more limited extent than with linkages/correlation).

In an intriguing study which the author views as an extension of this theme, Dijk (2015) argues that contagion would work via "emotional impact," somehow triggering withdrawal decisions. Referencing BTV's null result when bank strengths are not correlated, he suggests that "simply observing the outcome of other subjects playing a game in the lab does not have the same emotional impact as seeing a financial crisis unfold when your own life savings are on the line, and that therefore the emotional channel through which psychological contagion works would not be activated." He attempts to "take a more direct approach of first inducing an emotional state" and argues that the "main emotion of interest is fear, as this seems to be the most plausible candidate to explain changes in behaviour during a panic" (all quotes, p. 4). However, he also explores and compares with the role sadness and happiness. His chosen emotion induction procedure is to have his experimental subjects write short essays about a time when they experienced the target emotion, before playing one of BTV's bank run games. Dijk reports that "the presence of fear significantly increases the likelihood of withdrawal" and that women are "significantly more likely to withdraw then men, but only when fear is present" (p. 1).

3.2.10 Things to do. I offered critical remarks about various aspects of the studies I surveyed. Those comments implicitly suggest issues where more clarifying or corroborating evidence may be called for, and I will not say more about that here but rather bring up new themes.

D&D inspired a huge follow-up literature. For example, while D&D restrict attention to demand-deposit contracts, Green and Lin (2003) show that other forms may rule out equilibrium runs. It is beyond the scope of this paper report to summarize the literature but I would like to note that, surely, there must be exciting ideas developed that cry out for experimental testing.

D&D's analysis starts with deposits already in the bank. There must have been an earlier period where the deposit decisions were made. Suppose that stage is explicitly added to the model. As van Damme (1994, p. 22) notes, it would seem that no bank run equilibrium could be sustained, as presumably the best response would be to not deposit in the first place[25]. And if a deposit is made this might signal the intent to play the good equilibrium without a bank run, which may make others tag along. It may be fruitful to test the empirical relevance of these ideas in an experiment, and I'm thinking about doing such research myself.

Finally, I note two limitations that apply to most papers I surveyed. First, D&D's model embodies two key components: first, it explains why banks are needed and formed (to provide a form of insurance); and second, it explains bank runs in terms of coordination on an inefficient equilibrium. Much of the beauty of D&D's work lies in how first and second are inter-linked (via $r > 0$; cf. my earlier remarks). No such link is examined in experimental studies though; in fact, first, is not addressed at all[26]. Second, what are the key psychological factors involved in bank runs? Some would say "fear" or "panic." Recall Summers' (2000) remark on "bank run psychology." If this is modeled by D&D, it is done somewhat soporifically, as an inefficient equilibrium. With the exception of the paper by Dijk (2015) cited in the previous "contagion" section, the experimental designs follow suit; the psychology of fear or panic is not really addressed.

3.3 Insolvency

What about the relevance of the insolvency story? In this section I will approach this issue from two different angles that concern, respectively, misprediction and moral hazard.

3.3.1 Misprediction. As regards the 2007 + crisis, Calomiris points out that "[R] easonable, forward-looking estimates of risk were ignored" (quote from Calomiris, 2009c, p. 3; compare the longer analysis in Calomiris, 2009b, pp. 13-21)[27]. Part of the argument is that despite the writing on the wall, various market participants benefited from excessive loan origination. However, many participants were hurt as the market collapsed! It is puzzling that this could happen with such force, if there was writing on the wall. One key to explaining this would be if people were affected by sample size neglect and the hot hand phenomenon (recall 2.3).

In Section 2.2, *The Bubbles Example* was presented chiefly to illustrate general themes discussed there. However, bubbles experiments may be relevant also for evaluating the empirical relevance of sample size neglect & the hot hand phenomenon which may make traders oblivious that a crash is imminent. With those goggles on, the presentation now picks up where *The Bubbles Example* ended. Prepare for a roller-coaster ride!

Recall the observation that bubbles appeared in DLM's design, echoing earlier findings of Smith *et al.* and others (King *et al.*, 1993; Peterson, 1993; Van Boening *et al.*, 1993; Porter and Smith, 1995; Caginalp *et al.*, 1998, 2000, 2001; Fisher and Kelly, 2000; Lei *et al.*, 2001; Noussair *et al.*, 2001). It may seem natural to draw the conclusion, by analogy, that bubbles may be common in naturally occurring markets as well. However, a closer looks gives reason for pause. As seen at the time they published their paper, bubble and crash pricing patterns appear under a variety of circumstances, with one exception. If subjects have gained experience through participation in at least two preceding markets, then bubbles are significantly abated. Does this mean that the support from experiments for the view that bubbles and crashes occur is weak? That's not clear. One could argue that the studies discussed so far did not settle the issue. In nearly all experiments either all or none of the market participants were experienced. By contrast, in naturally occurring markets, there is often a mixture of experienced and inexperienced traders. It is hard to know which result applies. Therefore, it would seem desirable to run experiments where the subjects in a market differ as regards their level of experience.

DLM report results from a lab experiment intended to shed light on the issue. Recall (from 2.2) that a market design consisted of ten periods. Subjects actually interacted repeatedly in three such "market rounds," and gained experience (note the distinction between a round and a period. A round consists of ten periods). Then a fourth round was played, in which some (two or four, by now experienced) traders were replaced by inexperienced traders (that up till that point have solved crosswords, rather than interacted in any market). This fourth round is crucial. By comparing its data with data from rounds one and three, one can evaluate whether a mixed experience setting more resembles one with inexperienced or experienced traders.

Looking back at Figure 1, this time focussing also on the data from rounds three and four data, one sees that bubbles are common in the first round but abate by the third round. The bubbles do not return in the fourth round, where they resemble the prices of the third round (where all traders are arguably experienced) and do not resemble the prices of the first round (where all traders are inexperienced). DLM actually ran ten sessions and Figure 2 shows the data averaged across all of these, which underscores the pattern reported so far (this claim is supported also through formal statistical test, reported by DLM).

DLM thus support the proposition that prices are fundamental, rather than bubbly, when market participants are heterogeneous as regards experience level. They sum it up: the "results may shift the burden of proof somewhat between those who believe in the madness of the market and the market fundamentalists. [The] results speak in favor of the latter position" (pp. 1735-1736).

DLM turns on its head the conclusion of *The Bubbles Example* (2.1). However the story does not end here. Hussam *et al.* (HPS, 2008) respond to DLM in a way that to an extent brings back the message of *The Bubbles Example*. HPS report data from a design which allows subjects to gain experience through repeat interaction. They "impose a large increase in liquidity and dividend uncertainty to shock the environment of experienced subjects who have converged to equilibrium, and this environment rekindles a bubble" (p. 924); "when important elements in the underlying market environment change for experienced subjects, a bubble can reignite" (p. 937). The burden of proof, alluded to by DLM as cited above, shifts back […].

I will leave the bubbles topic here, noting that research in the area is ongoing and new studies keep shifting the burden of proof in one or the other direction. I will highlight just one interesting recent direction, namely, studies that attempt to reach a deeper psychological understanding of how bubbles are produced, using brain imaging techniques (Smith *et al.*, 2013), emotion induction procedures (Andrade *et al.*, 2014), or facereading software (Breaban and Noussair, 2014)[28].

3.3.2 Moral hazard. Recall the remarks in Section 3.2, about "R as in insolvency." In the D&D model R is constant, effectively ruling out moral hazard as a source of insolvency. However, one may consider banking games in which (some counterpart of) R is not constant but rather depends on the decisions of economic agents. Theoretical

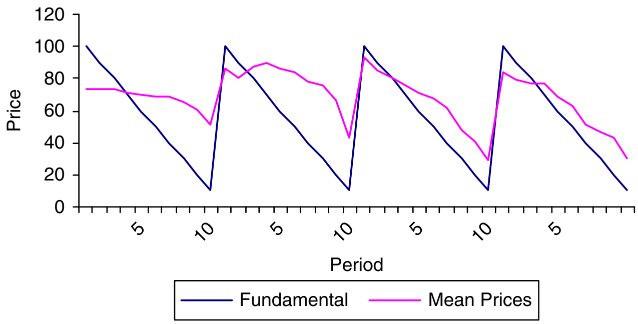

Figure 2.
Observed mean prices, and fundamental value, across all sessions

examination may generate valuable insights, and subsequent experimental testing can explore empirical relevance. This would be an area where not much research seems to have been done[29].

To give a flavor of how relevant work may look, I will discuss a specific new idea. It concerns deposit insurance. First note that this policy tool is relevant for both crises reasons – bank runs and insolvency – on which I have focussed. D&D show theoretically how deposit insurance can help avoid bank runs (and some experiments discussed in Section 3.2 explored related issues). As regards insolvency, however, the argument has been made that deposit insurance might be detrimental rather than helpful. Calomiris (1990, 2008) and interprets historical records to indicate that deposit insurance often caused insolvency by inviting excessive risk taking. See also Demirgüç-Kunt and Detragiache (2002).

How can this insolvency connection be understood? Dufwenberg and Rietzke (2015) explore the issue theoretically, formulating a model with bankers, depositor, and an institution providing deposit insurance (say the FDIC). If customers deposit, then bankers can invest/give loans. Without deposit insurance, bankers are residual claimants once deposits are covered; bankers go bankrupt if investments go sour or if loans that they issue are not repaid. With deposit insurance, bankers are still residual claimants, but now customers' losses are covered by the FDIC. Compare the last two sentences, and reflect on how the monetary payoff to the bankers is the same in the two cases. If bankers maximize profit the outcome would seem to be the same for the bank, with or without insurance, although the customers' payoffs vary greatly.

The model sketched so far does thus not furnish a compelling reason why deposit insurance creates enhanced incentives for excessive risk taking[30]. But this conclusion changes if one takes a relationship aspect between bankers and customers into account by applying reciprocity theory (Dufwenberg and Kirchsteiger, 2004; cf. Rabin, 1993; Falk and Fischbacher, 2006). Customers who deposit will be viewed by the bank as kind, as deposits are what allow investment/lending. Without insurance, bankers will want to be kind in return, so they hold back on risk taking[31]. With deposit insurance, by contrast, excessive risk taking is no longer unkind to the customers (it may be unkind to whoever foots the insurance payments, but, according to reciprocity theory, bankers do not care if that party were not kind to them).

Is this plausible? Should not recent bankers' behavior during the crisis be interpreted it in terms of greed rather than in terms of a desire for reciprocation? I note that several reasonable responses are available: First, much experimental evidence suggests that many people are not exclusively selfish, and often motivated by reciprocity (see e.g. Fehr and Gächter, 2000 and Sobel, 2005 for some relevant discussion of the evidence and theoretical issues), so why not bankers?[32] Second, the reciprocity story gets the predictions right while, as I argued, sticking with traditional assumptions makes the empirical evidence look puzzling. Third, in the reciprocity story, with deposit insurance, in equilibrium bankers act as if they cared only about profit, which seems consistent with the historical record according to Calomiris.

However, the point to be made here does not concern arm-chair reflection on the plausibility of a piece of theory. Dufwenberg and Rietzke (2015) tell a consistent story that provide a link between deposit insurance and insolvency. It is natural to wonder about the empirical relevance of that account. One way to address that topic is to create a laboratory environment which as regards strategic structure and payoff consequences matches the theory, and then to test it experimentally. David Rietzke and I are planning to do this.

4. Conclusions

Reasonable people disagree on how to understand financial crises in general and banking crises in particular, and what to do to avoid or mitigate these events. How would history have changed had government policy, lending standards, credit rating practices, et cetera been different? One cannot go back in time and check. However, one can run lab experiments under controlled conditions. Can this be useful?

I presented bad news and good news. History is probably too complicated to be meaningfully revamped or modified in the lab, for purposes of insight-by-analogy. But as people argue about how to understand and shape financial history, they bring ideas to the table. It may be possible and useful to test the empirical relevance of these ideas in lab experiments.

The debate on how to understand banking crises may be seen as centered on a bank runs vs insolvency issue. The bank run experiments, surveyed in section III-A, explore the empirical relevance of the story told by Diamond and Dybvig's (D&D) model, which abstracts away from insolvency issues. The results are a mixed bag. It neither says that D&D's depiction of bank runs seems irrelevant nor that it is super compelling. There is moreover room for related work that may produce more robust insights.

What about the relevance of the insolvency story? In Section 3.2 I approached this issue from two angles. First, I discussed the notions (developed by psychologists and emphasized in behavioral finance) of sample size neglect and the hot hand phenomenon, which may help explain why individuals might believe prices-will-keep-rising while neglecting important information about excessive risk taking in the lending industry. Experiments on bubbles in financial markets furnish relevant test beds, and on balance somewhat support this idea. Second, I discussed moral hazard in banking, the use of models to examine that, and the possibility of running-related experiments.

What about the external validity of experimental work? That is, what is the extent to which insights that hold in a lab environment have bearing on how to understand the banking world in which one is ultimately interested? I expressed a caveat that lab experiments hardly allow recreating history for insight-by-analogy purposes. Beyond that, in many cases I would not worry too much. For example, lab experiments with students should form adequate test beds for exploring the empirical relevance of sample size neglect and the hot hand effect among bankers, borrowers, and government officials. People are people! Lab experiments are also useful for testing theory, since by design one can often well represent the situation a theory depicts. Thus, for example, testing the empirical relevance of D&D's logic in the lab makes sense[33].

How can lab experiments inform policy? In order to do well it is useful to first diagnose how the world works. Therefore, exploring, e.g. the bank runs vs insolvency reasons for banking crises must have some policy relevance at least indirectly. As regards specific tools, consider deposit insurance. According to D&D, such schemes may be helpful for avoiding bank run equilibria. Some of the experiments tested their effect in D&D designs, but the schemes used (conditional transfers rather than funded schemes) do not map well to what D&D have in their theory which makes it somewhat unclear what is tested. Besides, the results were pointing in different direction, for example, concerning the usefulness of various degrees of partial insurance.

When reading these results, one must not forget that deposit insurance schemes may have poor properties for reasons that are not addressed in the bank run experiments.

Take it from Calomiris (2008, p. 350; compare, e.g. Calomiris, 1990; Demirgüç-Kunt and Detragiache, 2002):

> Empirical studies […] conclude that deposit insurance and other policies that protect banks from market discipline, intended as a cure for instability, have become the single greatest source of banking instability. […] Deposit insurance removes depositors' incentives to monitor and discipline banks, and frees bankers to take imprudent risk.

D&D style bank run games, abstracting away from insolvency and moral hazard, do not address that but in Section 3.2 I discussed the possibility of coming up with new games and experimental designs that do. Not much work has been done but I see the area as one where future research effort may be fruitful. I sketched a specific example, whereby reciprocity theory may prove useful for explaining a reason why deposit insurance may stimulate excessive risk taking by banks, and the potential role of lab experiments in this connection.

Next consider capital requirements (cash reserve ratios) that governments impose on banks. Are they useful? What level is best? Can lab experiments help evaluate that? I have not found any experiment done on the topic. Calomiris (2009c, p. 10) suggest that a regulatory rule of the following sort may be useful:

> [V]ary capital and liquidity requirements over time in response to changes in macroeconomic and financial system circumstances. For example, during booms, minimum capital would be set higher, especially if a boom were occurring in which asset prices and credit were rising rapidly. Raising capital requirements on banks would discourage a protracted bubble from forming and create a larger equity cushion for banks if a bubble should burst.

His comments about "bubbles" suggest ways that asset market experiments of the sort I have covered could be useful in this connection. Finally, *mutatis mutandis*, introducing capital requirements to a D&D game influences its payoffs[34]. It is natural to wonder what the empirical relevance may be and whether experiments could shed light on that.

An important feature of recent financial crises in many places (e.g. USA, Ireland, Spain, Eastern Europe) was a credit boom and bust, rooted mainly in the housing/mortgage market. Ideas circulate regarding how these episodes happen, say when lending is collateralized, and regulation has emerged (countercyclical capital requirements, loan-to-value caps, etc.) aimed at smoothing fluctuations. Lab experiments may help understand causes and effects. For example, bubbles experiments (cf. Section 3.3) may be relevant for understanding how households might buy into overheated housing markets. Perhaps experimental designs could explore particular relevant market features (say that housing is a consumption as well as an investment good, that substantial leverage is possible for investors, and various lending restrictions) may (or may not) lead to stronger bubble activity than in "ordinary" asset markets.

Another prominent feature of recent crises was the freeze (i.e. collapse of trade) of wholesale funding markets and the spillover of liquidity shocks to asset markets. As a result, tighter liquidity regulations were introduced in the Third Basel Accord, and regulators push for money market reforms (centralized counterparties rather than OTC contracts, higher quality collateral for repo transactions). See Brunnermeier and Pedersen (2009) for some relevant-related theory, possibly usefully testable in experiments. Perhaps experiments can further help understand under which conditions market freezes may be more likely to occur, for example, whether secured markets are more subject to freezes than unsecured markets. Or does the concentration of counterparties (say having a few large broker-dealer banks) exacerbate or mitigate the possibility of a freeze. Experiments could maybe also shed light on which institutional solutions (lender of last resort, central counterparties) may make money markets more resilient in the face of idiosyncratic/aggregate liquidity shocks[35].

In preparing this text I found a lot of inspiration in the work of Charles Calomiris, so it seems fitting to present his what-should not-have-been-done (in the USA, 2003-2007) and what-should-be-done lists (from Calomiris, 2009a, c), as regards avoiding banking crises:

What should not have been done: "Political pressures from Congress on the government-sponsored enterprises (GSEs) Fannie Mae and Freddie Mac, to promote "affordable housing" by investing in high-risk subprime mortgages; Lending subsidies for housing finance via the Federal Home Loan Bank System to its member institutions; Federal Housing Administration subsidization of extremely high-mortgage leverage and risk; Government and GSE mortgage foreclosure mitigation protocols that were developed in the late 1990s and early 2000s to reduce the costs to borrowers of failing to meet debt service requirements on mortgages, which further promoted risky mortgages; and 2006 legislation enacted to encourage ratings agencies to relax standards for subprime securitizations."

What should be done: "regulatory taxes and reforms of resolution processes that would discourage too-big-to-fail protection of large, complex banks; macro prudential regulatory authority to gauge overall risk in the financial system and structure dynamic capital and liquidity requirements accordingly; elimination of leverage subsidies in housing; rules to encourage OTC clearing in clearinghouses; disclosure standards for OTC market participants; improvements in the measurement of regulatory risk that would include market-based measures; changes in the use of rating agencies' opinions to discourage grade inflation; eliminating regulatory limits on the concentration of ownership in banks."

Might lab experiments be relevant for shedding light on the usefulness of the policies indicated? If I led my readers to consider that question natural, then I would be happy.

Acknowledgments

This paper grew out of a lecture the author prepared for an internal seminar on banking organized by the Norwegian Ministry of Finance, as input to a report they prepared on financial markets (Oslo, September 24, 2012; Financial Markets Report, 2012). Previous versions have been posted as working papers at the University of Arizona (Department of Economics WP 13-08) and at Bocconi University (IGIER WP 534). I have benefited from very helpful comments by a referee as well as from Martin Brown, Price Fishback, Rod Garratt, Francesco Guala, Michael Kirchler, Hubert Kiss, Johanna Möllerström, Asaf Plan, Stefanie Ramirez, David Rietzke, Alec Smith, Morten Søberg, Stefan Trautmann, Tiemen Woutersen, and the seminars participants in Oslo as well as at the Universities of Arizona and Gothenburg.

Notes

1. I am grateful to Martin Brown who encouraged me to emphasize that experiments may be useful for understanding financial intermediation beyond questions of crises. Why are particular arrangements (say relationship banking, threat of exclusion, collateral, demandable deposits, term deposits, commitment savings products, or microfinance arrangements useful, say for overcoming information asymmetries in lending or provision of liquidity? Lab, and field, experiments may be helpful to shed light on these issues. A collection of relevant studies include Karlan (2005), Fehr and Zehnder (2009), Brown and Serra-Garcia (2014), Brown and Zehnder (2007), Giné *et al.* (2010), Flatnes (2014), Ashraf *et al.* (2006).

2. See Ross (2008) and Wang (2008) for more discussion of the field.
3. As regards a banking panic originating in New York in 1907, Calomiris writes that "failures and losses were not much higher than in normal times. As the crisis worsened, banks suspended covertibility until uncertainty of the incidence of the shock had been resolved" (p. 348). Calomiris ,and Mason (1997, p. 863) report that the Chicago bank panic of 1932 "did not produce significant social cost in terms of failures among solvent banks" and their (2003) paper bring home similar conclusions for other panics. See also the detailed descriptions of banking panics of the "gilded age" and the "great depression," reported by Wicker (1996, 2000) in a partly comparable spirit.
4. See Dash (1999), Chancellor (1999), Garber (2000), and Kindleberger (2001).
5. See Guala (2005) for further discussion of the methodology of experimental testing.
6. For a different breakdown, consider the list of chapters of the 1995 *Handbook*, indicating key streams focus up till then: Public Goods (author: J. Ledyard), Coordination Problems (J. Ochs), Bargaining Experiments (A. Roth), Industrial Organization (C. Holt), Experimental Asset Markets (S. Sunder), Auctions (J. Kagel), Individual Decision Making (C. Camerer). A second volume, again edited by Kagel and Roth, is in the making; its chapters indicate streams of post-1995 focus (information from Al Roth's web page): Introduction: The Last Ten+Years (A. Roth), Political Economy (T. Palfrey), Voluntary Giving to Public Goods (L. Vesterlund), Learning & the Economics of Small Decisions (I. Erev and E. Haruvy), Field Experiments: Psychology & Economics in the Field (author TBD), Neuroeconomics (C. Camerer, J. Cohen, E. Fehr, P. Glimcher, D. Laibson), Other-Regarding Preferences (D. Cooper and J. Kagel), Auctions (J. Kagel and D. Levin), Macroeconomics (J. Duffy), Market Design (A. Roth).
7. Charlie Holt's chapter in the *Handbook*, cited in footnote 6, is a good guide to much of the literature.
8. For a shorter/dictionary description see Bloomfield (2008), for another review see Hirshleifer (2014).
9. Much work builds on insights reached by Daniel Kahneman (who shared the 2002 Nobel Memorial Prize with Vernon Smith) and Amos Tversky (who died in 1996). See the edited volume Kahneman and Tversky (2000).
10. B&T's phrasing "do not initially know the data-generating process" suggests that they may have in mind settings where the process is not only unknown to a decision maker but also independent of his decisions. Of course, in *The Bubbles Example*, the second assumption does not hold. Nevertheless, the underlying psychology may be the same. The idea is furthermore reminiscent of Hyman Minsky's (1975, 1992) "Financial Instability Hypothesis."
11. I should also note a caveat: the validity of Gilovich *et al.*'s classic results presumes that the belief of a hot streak in basketball shooting is mistaken, so that there is no actual hot hand effect. A recent paper by Miller and Sanjurjo (2014), involving a field experiment as well as data re-analysis, questions that presumption.
12. For a more elaborate account, still condensed and more concise than D&D's full presentation, I recommend the presentation of van Damme (1994).
13. D&D work with an infinite number of customers, but the intuition is the same. Readers familiar with risk aversion concepts will notice that the utility function exhibits a constant relative risk aversion (CRRA) coefficient equal to 2; it has the properties of $u(w) = w^{1-\alpha}/(1-\alpha)$, with $\alpha = 2$. The essence of the example would go through for any $\alpha > 1$. CRRA is a frequent assumption in theoretical models on bank runs, see, e.g. (2) in Ennis and Keister (2009) to have just one more example.

14. With $n = 6$ the bank gets deposits of 6×1 in period 0. It pays out $3 \times (1 + r)$ in period 1 and is then left with $3 \times (1 - r)$ which by period 2 grows to $3 \times (1 - r) \times 4$, which is shared equally by its three remaining customers.

15. 7/8 is in fact the best outcome possible as $r = 1/3$ is the optimal level of r. To see this, take the derivative with respect to r of the expected utility expression $1/2 \times (2 - 2/(1 + r)) + 1/2 \times (2 - 2/((1 - r) \times 4)$, set it equal to 0, and solve for r.

16. D&D assume that whenever multiple customers simultaneously attempt to withdraw funds, a "sequential service constraint" applies such that it is randomly determined who gets stipulated return ($1 + r = 4/3$, in our example), up until when the bank runs out of funds. Note that $4 \times (4/3) + (2/3) + 0 = 18/3 = 6 =$ the total period 0 deposits.

17. Check after 49 minutes. Summers exemplifies with his own advice to Bill Clinton in 1995 (Mexican crisis).

18. Several empirical studies have explored to what extent fundamentally weaker banks are more likely to be affected by bank runs than fundamentally healthy ones, using old and new data. See, e.g. Davison and Ramirez (2014), De Graeve and Karas (2014), Iyer and Puri (2012), Starr and Yilmaz (2007).

19. I focus on bank runs reflecting coordination on a bad equilibrium but note that some work highlights how runs may in some settings reflect problems with fundamentals, and possibly then also be useful. See, e.g. Calomiris and Gorton (1991) or Allen and Gale (1998) on how runs may reflect adverse news about the macroeconomy, and Gu (2011) on how uncertainty about the quality of a bank's portfolio may produce runs due to herding effects.

20. Even announcing support for a bank may trigger a panic. See Shin (2009) for an interesting account (and suggested explanation of) the events that lead to a bank run at the UK bank Northern Rock: "On September 13, 2007, the BBC's evening television news broadcast first broke the news that Northern Rock had sought the Bank of England's support. The next morning, the Bank of England announced that it would provide emergency liquidity support. It was only after that announcement – that is, after the central bank had announced its intervention to support the bank – that retail depositors started queuing outside the branch offices" (pp. 101-102).

21. See also Wallace (1988) who questions the feasibility of such a policy with sequential service of withdrawals.

22. G&K's subjects first played a few games without forced withdrawals followed by a few games with forced withdrawals, so the treatment comparison is not across different sessions but rather "within subjects." In the second phase, the number of forced withdrawal was stochastic, unlike in the version of the D&D model that I presented in Section 3.2, where t was known by all. However, in any case the games have one good and one bad equilibrium.

23. See also Gorton (1988) who explores related issued looking at historical data.

24. KRR (2012, pp. 1652-1653) discuss many studies documenting historical bank run episodes where depositors' withdrawal decisions were influenced by observations of others: Sprague (1910), Wicker (1996), Kelly and Ó Grada (2000), Bruner and Carr (2007), Starr and Yilmaz (2007), Iyer and Puri (2012).

25. See also Peck and Shell (2000, pp. 106-107) for a related discussion incorporating also the effect of sunspots.

26. And keep in mind that D&D's story is not the only one for why banks, and the contracts they offer, exist. For example, while D&D show how demandable debt can provide insurance to risk averse depositors, Calomiris and Kahn (1991) show how such contracts may provide useful incentives for monitoring which loans are issued.

27. The Financial Crisis Inquiry Report (2011, p. xvii) echoes this: "Despite the expressed view of many on Wall Street and in Washington that the crisis could not have been foreseen or avoided, there were warning signs. The tragedy was that they were ignored or discounted. There was an explosion in risky subprime lending and securitization, an unsustainable rise in housing prices, widespread reports of egregious and predatory lending practices, dramatic increases in household mortgage debt, and exponential growth in financial firms' trading activities, unregulated derivatives, and short-term "repo" lending markets, among many other red flags. Yet there was pervasive permissiveness; little meaningful action was taken to quell the threats in a timely manner."
28. An incomplete list of other recent experimental studies that explore bubbles includes Haruvy and Noussair (2006), Haruvy *et al.* (2007), Lei and Vesely (2007), Stöckl *et al.* (2010), Oechssler *et al.* (2011), Kirchler *et al.* (2012), Corgnet *et al.* (2013, 2015), Baghestanian and Walker (2014), Bao *et al.* (2014), Eckel and Füllbrunn (2015).
29. There are experiments on moral hazard more generally from which insights may perhaps be drawn (e.g. Charness and Dufwenberg, 2006). However, in this report I restrict attention to studies that focus explicitly on banking.
30. Should one object that deposit insurance reduces depositors' incentives to monitor banks? Arguably, if depositors can monitor and constrain bankers' choices then the FDIC would presumably also have that ability. The FDIC, an organization run by professionals, would seem better positioned to monitor than would private depositors who are likely to form a fragmented and hard-to-organize group.
31. The sentence is written as if the risk level is chosen by the bank. In general, that may depend of borrower behavior, which may be influenced by other aspects of the banking game (e.g. information sharing among lenders or the long-term prospects of the banking relation). See Capra *et al.* (2005), Brown and Zehnder (2007, 2010) and Trautmann and Vlahu (2012) for some interesting experimental studies on related topics.
32. Cohn *et al.* (2014) experimentally explore bank employees' "honesty," a quality which in their design involves non-selfish behavior, which many exhibit. Cohn *et al.*'s main result is that on balance bank employees' honesty withers if they are reminded of their occupational role, but even then selfishness is certainly not universal.
33. Moreover, while lab experiments are usually conducted with students they can be replicated in more representative populations. Cornée *et al.* (2012), e.g. replicated Brown and Zehnder (2007) with real bankers.
34. With reference to my D&D account (3.2), assume that for each unit deposited with the bank it has to hold cash reserves $k \leqslant 1$, and if an invested unit is pulled back in period, its value is $y \leqslant 1$. With n depositors, the amount available for period 1 withdrawals is $n \times (k + (1-k) \times y)$. Section 3.2 considered the case $y = 1$, where $n \times (k + (1-k) \times y) = n$, independent of k. If $y < 1$ then $n \times (k + (1-k) \times y) < n$, and the payoffs of various strategy profiles now depend on k.
35. I owe the substance of the last two paragraphs to Martin Brown.

References

Ackert, L. and Deaves, R. (2010), *Behavioral Finance: Psychology, Decision-Making, and Markets*, Southwestern, Boston.

Allen, F. and Gale, D. (1998), "Optimal financial crises", *Journal of Finance*, Vol. 53, pp. 1245-1284.

Andrade, E., Odean, T. and Shengle, L. (2014), "Bubbling with excitement: an experiment", Mimeo.

Arifovic, J. and Jiang, J.H. (2014), "Do sunspots matter? Evidence from an experimental study of bank runs", Working Paper No. 2014-12, Bank of Canada, Vancouver.

Arifovic, J., Jiang, J.H. and Xu, Y. (2013), "Experimental evidence of bank runs as pure coordination failures", *Journal of Economic Dynamics and Control*, Vol. 37, pp. 2446-2465.

Ashraf, N., Karlan, D. and Yin, W. (2006), "Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines", *Quarterly Journal of Economics*, Vol. 121, pp. 635-672.

Baghestanian, S. and Walker, T.B. (2014), "Thar she blows again: reducing anchoring rekindles bubbles", SSRN Scholarly Paper ID 2456941, Rochester, NY.

Bao, T., Hommes, C. and Makarewicz, T. (2014), "Bubble formation and (in)efficient markets in learning-to-forecast and – optimise experiments", Mimeo.

Barberis, N. and Thaler, R. (2003), "A survey of behavioral finance", in Constantinides, G.M., Harris, M. and Stulz, R. (Eds), *Handbook of the Economics of Finance*, Elsevier.

Bernanke, B. (1983), "Nonmonetary effects of the financial crisis in the propagation of the great depression", *American Economic Review*, Vol. 73, pp. 257-276.

Bloomfield, R. (2008), "Behavioral finance", in Durlauf, S.N. and Blume, L.E. (Eds), *The New Palgrave Dictionary of Economics*, 2nd ed., Vol. 1, Palgrave Macmillan, London, pp. 438-444.

Breaban, A. and Noussair, C. (2014), "Emotional state and market behavior", Mimeo.

Brown, M. and Serra-Garcia, M. (2014), "The threat of exclusion and relational contracting", Discussion Paper No. 2010-2018, European Banking Center, Boston.

Brown, M. and Zehnder, C. (2007), "Credit reporting, relationship banking, and loan repayment", *Journal of Money, Credit & Banking*, Vol. 39, pp. 1883-1918.

Brown, M. and Zehnder, C. (2010), "The emergence of information sharing in credit markets", *Journal of Financial Intermediation*, Vol. 19, pp. 255-278.

Brown, M., Trautmann, S. and Vlahu, R. (2012), "Contagious bank runs: experimental evidence", Working Paper No. 363, De Nederlandsche Bank (DNB), Amsterdam.

Bruner, R. and Carr, S. (2007), *The Panic of 1907: Lessons Learned from the Markets Perfect Storm*, John Wiley & Sons, Hoboken, NJ.

Brunnermeier, M. and Pedersen, L.H. (2009), "Market liquidity and funding liquidity", *Review of Financial Studies*, Vol. 22, pp. 2201-2238.

Caginalp, G., Porter, D. and Smith, V. (1998), "Initial cash/asset ratio and asset prices: an experimental study", *Proceedings for the National Academy of Sciences*, *Vol. 95*, pp. 756-761.

Caginalp, G., Porter, D. and Smith, V. (2000), "Momentum and overreaction in experimental asset markets", *International Journal of Industrial Organization*, Vol. 18, pp. 187-204.

Caginalp, G., Porter, D. and Smith, V. (2001), "Financial bubbles, excess cash, momentum, and incomplete information", *Journal of Psychology & Financial Markets*, Vol. 2, pp. 80-99.

Calomiris, C. and Gorton, G. (1991), "The origin of banking panics: models, facts, and bank regulations", in Hubbard, R.G. (Ed.), *Financial Markets & Financial Crises*, University of Chicago Press, Chicago.

Calomiris, C. and Kahn, C. (1991), "The role of demandable debt in structuring optimal banking arrangements", *American Economic Review*, Vol. 81, pp. 497-513.

Calomiris, C. and Mason, J. (1997), "Contagion and bank failures during the great depression: the June 1932 Chicago banking panic", *American Economic Review*, Vol. 87, pp. 863-883.

Calomiris, C. and Mason, J. (2003), "Fundamentals, panics and bank distress during the great depression: the June 1932 Chicago banking panic", *American Economic Review*, Vol. 93, pp. 1615-1647.

Calomiris, C. (1990), "Is deposit insurance necessary? A historical perspective", *Journal of Economic History*, Vol. 50, pp. 283-295.

Calomiris, C. (2008), "Banking crises", in Durlauf, S.N. and Blume, L.E. (Eds), *The New Palgrave Dictionary of Economics*, 2nd ed., Vol. 1, Palgrave Macmillan, London, pp. 348-352.

Calomiris, C. (2009a), "Banking crisis yesterday and today", Briefing to the pew financial reform task force.

Calomiris, C. (2009b), "The subprime turmoil: what's old, what's new, and what's next", *Journal of Structured Finance*, Vol. 15, pp. 6-52.

Calomiris, C. (2009c), "Prudential bank regulation: what's broke and how to fix it", unpublished manuscript.

Camerer, C. (2008), "Behavioral game theory", in Durlauf, S.N. and Blume, L.E. (Eds), *The New Palgrave Dictionary of Economics*, 2nd ed., Vol. 1, Palgrave Macmillan, London, pp. 444-448.

Capra, M., Fernandez, M. and Ramirez-Comeig, I. (2005), "Moral hazard and collateral as screening device: empirical and experimental evidence", Mimeo.

Carlsson, H. and van Damme, E. (1993), "Global games and equilibrium selection", *Econometrica*, Vol. 61, pp. 989-1018.

Cartwright, E. (2014), *Behavioral Economics*, 2nd ed., Routledge, London.

Chakravarty, S., Fonseca, M. and Kaplan, T. (2013), "An experiment on the causes of bank run contagions", *European Economic Review*, forthcoming.

Chancellor, E. (1999), *Devil Take the Hindmost: A History of Financial Speculation*, Farrar, Straus & Giraux, New York, NY.

Charness, G. and Dufwenberg, M. (2006), "Promises & partnership", *Econometrica*, Vol. 74, pp. 1579-1601.

Cohn, A., Fehr, E. and Maréchal, M. (2014), "Business culture and dishonesty in the banking industry", *Nature*, Vol. 516, pp. 86-89.

Corgnet, B., Kujal, P. and Porter, D. (2013), "Reaction to public information in markets: how much does ambiguity matter?", *Economic Journal*, Vol. 123, pp. 699-737.

Corgnet, B., Hernán-González, R., Kujal, P. and Porter, D. (2015), "The effect of earned versus house money on price bubble formation in experimental asset markets", *Review of Finance*, Vol. 19, pp. 1455-1488.

Cornée, S., Masclet, D. and Thenet, G. (2012), "Credit relationships: evidence from experiments with real bankers", *Journal of Money, Credit and Banking*, Vol. 44, pp. 957-980.

Crawford, V., Costa-Gomes, M. and Iriberri, N. (2013), "Structural models of nonequilibrium strategic thinking: theory, evidence & applications", *Journal of Econ Literature*, Vol. 51, pp. 5-62.

Dasgupta, A. (2004), "Financial contagion through capital connections: a model of the origin and spread of bank panics", *Journal of the European Economic Association*, Vol. 2, pp. 1049-1084.

Dash, M. (1999), *Tulipomania: The Story of the World's Most Coveted Flower and the Extraordinary Passions it Aroused*, Crown Publishers, New York, NY.

Davison, L.K. and Ramirez, C.D. (2014), "Local banking panics of the 1920s: identification and determinants", *Journal of Monetary Economics*, Vol. 66, pp. 164-177.

De Graeve, F. and Karas, A. (2014), "Evaluating theories of bank runs with heterogeneity restrictions", *Journal of the European Economic Association*, Vol. 12, pp. 969-996.

De Long, B., Shleifer, A., Summers, L. and Waldmann, R. (1990), "Noise trader risk in financial markets", *Journal of Political Economy*, Vol. 98, pp. 703-738.

Demirgüç-Kunt, A. and Detragiache, E. (2002), "Does deposit insurance increase banking system stability? An empirical investigation", *Journal of Monetary Economics*, Vol. 49, pp. 1373-1406.

Diamond, D. and Dybvig, P. (1983), "Bank runs, deposit insurance and liquidity", *Journal of Political Economy*, Vol. 91, pp. 401-419.

Dijk, O. (2015), *Bank Run Psychology*, Mimeo, Radboud University Nijmegen, Nijmegen.

Dufwenberg, M. and Kirchsteiger, G. (2004), "A theory of sequential reciprocity", *Games & Economic Behavior*, Vol. 47, pp. 268-298.

Dufwenberg, M. and Rietzke, D. (2015), "Banking on reciprocity: deposit insurance and insolvency," Mimeo.

Dufwenberg, M., Lindqvist, T. and Moore, E. (2005), "Bubbles & experience: an experiment", *American Economic Review*, Vol. 95, pp. 1731-1736.

Eckel, C. and Füllbrunn, S. (2015), "Thar SHE blows? Gender, competition, and bubbles in experimental asset markets", *American Economic Review*, Vol. 105, pp. 906-926.

Ennis, H. and Keister, T. (2009), "Bank runs and institutions: the perils of intervention", *American Economic Review*, Vol. 99, pp. 1588-1607.

Falk, A. and Fischbacher, U. (2006), "A theory of reciprocity", *Games & Economic Behavior*, Vol. 54, pp. 293-315.

Fehr, E. and Gächter, S. (2000), "Fairness and retaliation: the economics of reciprocity", *Journal of Economic Perspecitves*, Vol. 14, pp. 159-191.

Fehr, E. and Zehnder, C. (2009), Reputation and credit market formation: how relational incentives and legal contract enforcement interact", Discussion Papaer No. 4351, University of Lausanne, Lausanne.

Financial Crisis Inquiry Report (2011), "National Commission on the causes of the financial & economic crisis in the USA, The Financial Crisis Inquiry Commission (FCIC), Chair, Phil Angelides, 10 Members, Washington, DC, available at: www.gpo.gov/fdsys/pkg/GPO-FCIC/pdf/GPO-FCIC.pdf

Financial Markets Report (2012), "Meld. St. 30 (2012-2013)", Report to the storting (white paper), Norwegian Ministry of Finance, Oslo.

Fisher, E. and Kelly, F. (2000), "Experimental foreign exchange markets", *Pacific Economic Review*, Vol. 5, pp. 365-387.

Flatnes, J.E. (2014), "The role of collateral in joint liability group lending: evidence from a framed field experiment in Tanzania", Mimeo, UC Davis, Davis, CA.

Friedman, M. and Schwartz, A. (1963), *A Monetary History of the United States*, Princeton University Press, Princeton, NJ, pp. 1867-1960.

Garber, P. (2000), *Famous First Bubbles: The Fundamentals of Early Manias*, MIT Press, Cambridge.

Garratt, R. and Keister, T. (2009), "Bank runs as coordination failures: an experimental study", *Journal of Economic Behavior & Organization*, Vol. 71, pp. 300-317.

Gilovich, T., Vallone, R. and Tversky, A. (1985), "The hot hand in basketball: on the misperception of random sequences", *Cognitive Psychology*, Vol. 17, pp. 295-314.

Giné, X., Jakiela, P., Karlan, D. and Morduch, J. (2010), "Microfinance games", *American Economic Journal: Applied Economics*, Vol. 2, pp. 60-95.

Goldstein, I. and Pauzner, A. (2005), "Demand-deposit contracts and the probability of bank runs", *Journal of Finance*, Vol. 60, pp. 1293-27.

Gorton, G. (1988), "Banking panics & business cycles", *Oxford Economic Papers*, Vol. 40, pp. 751-781.

Green, E. and Lin, P. (2003), "Implementing efficient allocations in a model of financial intermediation", *Journal of Economic Theory*, Vol. 109, pp. 1-23.

Guala, F. (2005), *The Methodology of Experimental Economics*, Cambridge University Press, Cambridge.

Haruvy, E. and Noussair, C. (2006), "The effect of short-selling on bubbles and crashes in experimental spot asset markets", *Journal of Finance*, Vol. 61, pp. 1119-1157.

Haruvy, E., Lahav, Y. and Noussair, C. (2007), "Traders' expectations in asset markets: experimental evidence", *American Economic Review*, Vol. 97, pp. 1901-1920.

Heinemann, F., Nagel, R. and Ockenfels, P. (2004), "The theory of global games on test: experimental analysis of coordination games with public and private information", *Econometrica*, Vol. 72, pp. 1583-1599.

Hirshleifer, D. (2014), "Behavioral Finance", Working Paper No. 59028, MPRA, Irvine, CA.

Hussam, R., Porter, D. and Smith, V. (2008), "Thar she blows: can bubbles be rekindled with experienced subjects?", *American Economic Review*, Vol. 98, pp. 924-937.

Iyer, R. and Puri, M. (2012), "Understanding bank runs: the importance of depositor-bank relationships and networks", *American Economic Review*, Vol. 102, pp. 1414-1445.

Kagel, J. and Roth, A. (Eds) (1995), *Handbook of Experimental Economics*, Princeton University Press, Princeton, NJ.

Kahneman, D. and Tversky, A. (Eds) (2000), *Choices, Values, and Frames*, Cambridge University Press, Cambridge, MA.

Karlan, D. (2005), "Using experimental economics to measure social capital and predict financial decisions", *American Economic Review*, Vol. 95, pp. 1688-1699.

Kelly, M. and Grada, C.Ó. (2000), "Market contagion: evidence from the panics of 1854 and 1857", *American Economic Review*, Vol. 90, pp. 1110-1124.

Kindleberger, C. (2001), *Manias, Panics & Crashes: A History of Financial Crisis*, Wiley, New York, NY.

King, R., Smith, V., Williams, A. and van Boening, M. (1993), "The robustness of bubbles and crashes in experimental stock markets", in Day, R. and Chen, P. (Eds), *Nonlinear Dynamics and Evolutionary Economics*, Oxford University Press, Oxford.

Kirchler, M., Huber, J. and Stöckl, T. (2012), "Thar she bursts: reducing confusion reduces bubbles", *American Economic Review*, Vol. 102, pp. 865-883.

Kiss, H.J., Rodríguez-Lara, I. and Rosa-Garcia, A. (2012), "On the effects of deposit insurance and observability on bank runs: an experimental study", *Journal of Money, Credit, & Banking*, Vol. 44, pp. 1651-1665.

Kiss, H.J., Rodríguez-Lara, I. and Rosa-Garcia, A. (2014a), "Do social networks promote or prevent bank runs?", *Journal of Economic Behavior and Organization*, Vol. 101, pp. 87-99.

Kiss, H.J., Rodríguez-Lara, I. and Rosa-Garcia, A. (2014b), "Do women panic more than men? An experimental study of financial decisions", *Journal of Behavioral and Experimental Economics*, Vol. 52, pp. 40-51.

Kiss, H.J., Rodriguez-Lara, I. and Rosa-Garcia, A. (2014c), "Think twice before running! Bank runs and cognitive abilities", Mimeo.

Klos, A. and Sträter, N. (2012), "How strongly do players react to changes in payoff parameters in an experimental bank run game?", Mimeo.

Krugman, P. (2010), "Too big to fail", *The Conscience of a Liberal*, January 11, *New York Times*, New York, NY.

Lei, V. and Vesely, F. (2007), "Market efficiency: evidence from a no-bubble asset market experiments", *Pacific Economic Review*, Vol. 14, pp. 246-258.

Lei, V., Noussair, C. and Plott, C. (2001), "Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs Actual irrationality", *Econometrica*, Vol. 69, pp. 831-859.

Madiès, P. (2006), "An experimental exploration of self-fulfilling banking panics: their occurrence, persistence, and prevention", *Journal of Business*, Vol. 79, pp. 1831-1866.

Miller, J. and Sanjurjo, A. (2014), "A cold shower for the hot hand fallacy", Working Paper No. 518, IGIER, Bocconi University.

Minsky, H. (1975), *John Maynard Keynes*, Columbia University, Annandale-on-Hudson, NY.

Minsky, H. (1992), "The financial instability hypothesis", Working Paper No. 74, The Jerome Levy Economics Institute of Bard College, May, Annandale-on-Hudson, NY.

Mishkin, F. (1992), "Anatomy of a financial crisis", *Journal of Evolutionary Economics*, Vol. 2, pp. 115-130.

Noussair, C., Robin, S. and Ruffieux, B. (2001), "Price bubbles in laboratory assets markets with constant fundamental values", *Experimental Economics*, Vol. 4, pp. 87-105.

Oechssler, J., Schmidt, C. and Schnedler, W. (2011), "On the ingredients for bubble formation: informed traders & communication", *Journal of Econic Dynamics & Control*, Vol. 35, pp. 1831-1851.

Peck, J. and Shell, K. (2000), "Equilibrium bank runs", *Journal of Political Economy*, Vol. 111, pp. 103-123.

Peterson, S. (1993), "Forecasting dynamics and convergence to market fundamentals: evidence from experimental asset markets", *Journal of Economic Behavior & Organization*, Vol. 22, pp. 269-284.

Porter, D. and Smith, V. (1995), "Futures contracting and dividend uncertainty in experimental asset markets", *Journal of Business*, Vol. 68, pp. 509-541.

Rabin, M. (1993), "Incorporating fairness into game theory & economics", *American Economic Review*, Vol. 83, pp. 1281-1302.

Rabin, M. (2002), "Inference by believers in the law of small numbers", *Quarterly Journal of Economics*, Vol. 117, pp. 775-816.

Rochet, J.-C. and Vives, X. (2004), "Coordination failures and the lender of last resort: was bagehot right after all?", *Journal of the European Economic Association*, Vol. 2, pp. 1116-1147.

Ross, S. (2008), "Finance", in Durlauf, S.N. and Blume, L.E. (Eds), *The New Palgrave Dictionary of Economics*, 2nd ed., Vol. 3, Palgrave Macmillan, London, pp. 314-333.

Schotter, A. and Yorulmazer, T. (2009), "On the dynamics and severity of bank runs: an experimental study", *Journal of Financial Intermediation*, Vol. 18, pp. 217-241.

Shiller, R. (2000), *Irrational Exuberance*, Princeton University Press.

Shin, H.S. (2009), "Reflections on Northern rock: the bank run that heralded the global financial crisis", *Journal of Economic Perspectives*, Vol. 23, pp. 101-120.

Smith, A., Lohrenz, T., King, J., Montague, R. and Camerer, C. (2013), "Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles", *Proceedings of the National Academy of Sciences*, Vol. 111, pp. 10503-10508.

Smith, V. (1962), "An experimental study of competitive market behavior", *Journal of Political Economy*, Vol. 70, pp. 111-137.

Smith, V., Suchanek, G. and Williams, A. (1988), "Bubbles, crashes and endogenous expectations in experimental spot asset markets", *Econometrica*, Vol. 56, pp. 1119-1151.

Sobel, J. (2005), "Interdependent preferences and reciprocity", *Journal of Economic Literature*, Vol. 43, pp. 392-436.

Sprague, O. (1910), *History Of Crises Under the National Banking System*, US Government Printing Office, Washington, DC.

Starr, M. and Yilmaz, R. (2007), "Bank runs in emerging-market economies: evidence from turkey's special finance houses", *Southern Economic Journal*, Vol. 73, pp. 1112-1132.

Stöckl, T., Huber, J. and Kirchler, M. (2010), "Bubble measures in experimental asset markets", *Experimental Economics*, Vol. 13, pp. 284-298.

Summers, L. (2000), "International financial crisis: causes, prevention, and cures", *American Economic Review*, Vol. 90, pp. 1-15.

Summers, L. (2008), *Learning from and Responding to Financial Crisis*, Arthur M. Okun Lecture delivered at Yale University, New Haven, CN, available at: http://wn.com/diamond_dybvig_model

Trautmann, S. and Vlahu, R. (2012), "Strategic loan defaults and coordination: an experimental analysis", *Journal of Banking & Finance*, Vol. 37, pp. 747-760.

van Boening, M., Williams, A. and LaMaster, S. (1993), "Price bubbles and crashes in experimental call markets", *Economics Letters*, Vol. 41, pp. 179-185.

van Damme, E. (1994), "Banking: a survey of recent microeconomic theory", *Oxford Review of Economic Policy*, Vol. 10, pp. 14-33.

Wallace, N. (1988), "Another attempt to explain an illiquid banking system: the diamond and dybvig model with sequential service taken seriously", *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 12, pp. 3-16.

Wang, J. (2008), "Finance (new developments)", in Durlauf, S.N. and Blume, L.E. (Eds), *The New Palgrave Dictionary of Economics*, 2nd ed., Vol. 3, Palgrave Macmillan, London, pp. 333-358.

Wicker, E. (1996), *The Banking Panics of the Great Depression*, Cambridge University Press, Cambridge, MA.

Wicker, E. (2000), *Banking Panics of the Gilded Age*, Cambridge University Press, Cambridge, MA.

Corresponding author
Professor Martin Dufwenberg can be contacted at: martind@eller.arizona.edu