American Finance Association

Demand-Deposit Contracts and the Probability of Bank Runs
Author(s): Itay Goldstein and Ady Pauzner
Source: *The Journal of Finance*, Vol. 60, No. 3 (Jun., 2005), pp. 1293-1327
Published by: Blackwell Publishing for the American Finance Association
Stable URL: http://www.jstor.org/stable/3694927
Accessed: 23/11/2009 12:22

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/action/showPublisher?publisherCode=black.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Blackwell Publishing and *American Finance Association* are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Finance.*

http://www.jstor.org

THE JOURNAL OF FINANCE • VOL. LX, NO. 3 • JUNE 2005

Demand–Deposit Contracts and the Probability of Bank Runs

ITAY GOLDSTEIN and ADY PAUZNER*

ABSTRACT

Diamond and Dybvig (1983) show that while demand–deposit contracts let banks provide liquidity, they expose them to panic-based bank runs. However, their model does not provide tools to derive the probability of the bank-run equilibrium, and thus cannot determine whether banks increase welfare overall. We study a modified model in which the fundamentals determine which equilibrium occurs. This lets us compute the ex ante probability of panic-based bank runs and relate it to the contract. We find conditions under which banks increase welfare overall and construct a demand–deposit contract that trades off the benefits from liquidity against the costs of runs.

ONE OF THE MOST IMPORTANT ROLES PERFORMED BY banks is the creation of liquid claims on illiquid assets. This is often done by offering demand–deposit contracts. Such contracts give investors who might have early liquidity needs the option to withdraw their deposits, and thereby enable them to participate in profitable long-term investments. Since each bank deals with many investors, it can respond to their idiosyncratic liquidity shocks and thereby provide liquidity insurance.

The advantage of demand–deposit contracts is accompanied by a considerable drawback: The maturity mismatch between assets and liabilities makes banks inherently unstable by exposing them to the possibility of panic-based bank runs. Such runs occur when investors rush to withdraw their deposits, believing that other depositors are going to do so and that the bank will fail. As a result, the bank is forced to liquidate its long-term investments at a loss and indeed

*Goldstein is from the Wharton School, the University of Pennsylvania; Pauzner is from the Eitan Berglas School of Economics, Tel Aviv University. We thank Elhanan Helpman for numerous conversations, and an anonymous referee for detailed and constructive suggestions. We also thank Joshua Aizenman, Sudipto Bhattacharya, Eddie Dekel, Joseph Djivre, David Frankel, Simon Gervais, Zvi Hercowitz, Leonardo Leiderman, Nissan Liviatan, Stephen Morris, Assaf Razin, Hyun Song Shin, Ernst-Ludwig von Thadden, Dani Tsiddon, Oved Yosha, seminar participants at the Bank of Israel, Duke University, Hebrew University, Indian Statistical Institute (New-Delhi), London School of Economics, Northwestern University, Princeton University, Tel Aviv University, University of California at San Diego, University of Pennsylvania, and participants in The European Summer Symposium (Gerzensee (2000)), and the CFS conference on 'Understanding Liquidity Risk' (Frankfurt (2000)).

fails. Bank runs have been and remain today an acute issue and a major factor in shaping banking regulation all over the world.[1]

In a seminal paper, Diamond and Dybvig (1983, henceforth D&D) provide a coherent model of demand–deposit contracts (see also Bryant (1980)). Their model has two equilibria. In the "good" equilibrium, only investors who face liquidity shocks (impatient investors) demand early withdrawal. They receive more than the liquidation value of the long-term asset, at the expense of the patient investors, who wait till maturity and receive less than the full long-term return. This transfer of wealth constitutes welfare-improving risk sharing. The "bad" equilibrium, however, involves a bank run in which all investors, including the patient investors, demand early withdrawal. As a result, the bank fails, and welfare is lower than what could be obtained without banks, that is, in the autarkic allocation.

A difficulty with the D&D model is that it does not provide tools to predict which equilibrium occurs or how likely each equilibrium is. Since in one equilibrium banks increase welfare and in the other equilibrium they decrease welfare, a central question is left open: whether it is desirable, ex ante, that banks will emerge as liquidity providers.

In this paper, we address this difficulty. We study a modification of the D&D model, in which the fundamentals of the economy uniquely determine whether a bank run occurs. This lets us compute the probability of a bank run and relate it to the short-term payment offered by the demand–deposit contract. We characterize the optimal short-term payment that takes into account the endogenous probability of a bank run, and find a condition, under which demand–deposit contracts increase welfare relative to the autarkic allocation.

To obtain the unique equilibrium, we modify the D&D framework by assuming that the fundamentals of the economy are stochastic. Moreover, investors do not have common knowledge of the realization of the fundamentals, but rather obtain slightly noisy private signals. In many cases, the assumption that investors observe noisy signals is more realistic than the assumption that they all share the very same information and opinions. We show that the modified model has a unique Bayesian equilibrium, in which a bank run occurs if and only if the fundamentals are below some critical value.

It is important to stress that even though the fundamentals uniquely determine whether a bank run occurs, runs in our model are still panic based, that is, driven by bad expectations. In most scenarios, each investor wants to take the action she believes that others take, that is, she demands early withdrawal just because she fears others would. The key point, however, is that the beliefs of investors are uniquely determined by the realization of the fundamentals.

[1] While Europe and the United States have experienced a large number of bank runs in the 19th century and the first half of the 20th century, many emerging markets have had severe banking problems in recent years. For extensive studies, see Lindgren, Garcia, and Saal (1996), Demirguc-Kunt, Detragiache, and Gupta (2000), and Martinez-Peria and Schmukler (2001). Moreover, as Gorton and Winton (2003) note in a recent survey on financial intermediation, even in countries that did not experience bank runs recently, the attempt to avoid them is at the root of government policies such as deposit insurance and capital requirements.

In other words, the fundamentals do not determine agents' actions directly, but rather serve as a device that coordinates agents' beliefs on a particular outcome. Thus, our model provides testable predictions that reconcile two seemingly contradictory views: that bank runs occur following negative real shocks,[2] and that bank runs result from coordination failures,[3] in the sense that they occur even when the economic environment is sufficiently strong that depositors would not have run had they thought other depositors would not run.

The method we use to obtain a unique equilibrium is related to Carlsson and van Damme (1993) and Morris and Shin (1998). They show that the introduction of noisy signals to multiple-equilibria games may lead to a unique equilibrium. The proof of uniqueness in these papers, however, builds crucially on the assumption of *global strategic complementarities* between agents' actions: that an agent's incentive to take an action is monotonically increasing with the number of other agents who take the same action. This property is not satisfied in standard bank-run models. The reason is that in a bank-run model, an agent's incentive to withdraw early is highest not when all agents do so, but rather when the number of agents demanding withdrawal reaches the level at which the bank just goes bankrupt (see Figure 2). Yet, we still have *one-sided strategic complementarities*: As long as the number of agents who withdraw is small enough that waiting is preferred to withdrawing, the relative incentive to withdrawing increases with the number of agents who do so. We thus develop a new proof technique that extends the uniqueness result to situations with only one-sided strategic complementarities.[4]

Having established the uniqueness of equilibrium, we can compute the ex ante probability of a bank run. We find that it increases continuously in the degree of risk sharing embodied in the banking contract: A bank that offers a higher short-term payment becomes more vulnerable to runs. If the promised short-term payment is at the autarkic level (i.e., it equals the liquidation value of the long-term asset), only efficient bank runs occur. That is, runs occur only if the expected long-term return of the asset is so low that agents are better off if the asset is liquidated early. But if the short-term payment is set above the autarkic level, nonefficient bank runs will occur with positive probability, and the bank will sometimes be forced to liquidate the asset, even though the expected long-term return is high.

The main question we ask is, then, whether demand–deposit contracts are still desirable when their destabilizing effect is considered. That is, whether

[2] See Gorton (1988), Demirguc-Kunt and Detragiache (1998), and Kaminsky and Reinhart (1999).

[3] See Kindleberger (1978) for the classical form of this view. See also Radelet and Sachs (1998) and Krugman (2000) for descriptions of recent international financial crises.

[4] In parallel work, Rochet and Vives (2003), who study the issue of lender of last resort, apply the Carlsson and van Damme technique to a model of bank runs. However, they avoid the technical problem that we face in this paper by assuming a payoff structure with global strategic complementarities. This is achieved by assuming that investors cannot deposit money in banks directly, but rather only through intermediaries (fund managers). Moreover, it is assumed that these intermediaries have objectives that are different than those of their investors and do satisfy global strategic complementarities.

the short-term payment offered by banks should be set above its autarkic level. On the one hand, increasing the short-term payment above its autarkic level generates risk sharing. The benefit from risk sharing is of first-order significance, since there is a considerable difference between the expected payments to patient and impatient agents. On the other hand, there are two costs associated with increasing the short-term payment: The first cost is an increase in the probability of liquidation of the long-term investment, due to the increased probability of bank runs. This effect is of second order because when the promised short-term payment is close to the liquidation value of the asset, bank runs occur only when the fundamentals are very bad, and thus liquidation causes little harm. The second cost pertains to the range where liquidation is efficient: Because of the sequential-service constraint faced by banks, an increase in the short-term payment causes some agents not to get any payment in case of a run. This increased variance in agents' payoffs makes runs costly even when liquidation is efficient. However, this effect is small if the range of fundamentals where liquidation is efficient is not too large. To sum up, banks in our model are viable, provided that maintaining the underlying investment till maturity is usually efficient.

Finally, we analyze the degree of risk sharing provided by the demand–deposit contract under the optimal short-term payment. We show that, since this contract must trade off the benefit from risk sharing against the cost of bank runs, it does not exploit all the potential gains from risk sharing.

The literature on banks and bank runs that emerged from the D&D model is vast, and cannot be fully covered here. We refer the interested reader to an excellent recent survey by Gorton and Winton (2003). Below, we review a few papers that are more closely related to our paper. The main novelty in our paper relative to these papers (and the broader literature) is that it *derives the probability of panic-based bank runs and relates it to the parameters of the banking contract.* This lets us study whether banks increase welfare even when the endogenous probability of panic-based runs they generate is accounted for.

Cooper and Ross (1998) and Peck and Shell (2003) analyze models in which the probability of bank runs plays a role in determining the viability of banks. However, unlike in our model, this probability is exogenous and unaffected by the form of the banking contract. Postlewaite and Vives (1987), Goldfajn and Valdes (1997), and Allen and Gale (1998) study models with an endogenous probability of bank runs. However, bank runs in these models are never panic based since, whenever agents run, they would do so even if others did not. If panic-based runs are possible in these models, they are eliminated by the assumption that agents always coordinate on the Pareto-optimal equilibrium. Temzelides (1997) employs an evolutionary model for equilibrium selection. His model, however, does not deal with the relation between the probability of runs and the banking contract.

A number of papers study models in which only a few agents receive information about the prospects of the bank. Jacklin and Bhattacharya (1988) (see also Alonso (1996), Loewy (1998), and Bougheas (1999)) explain bank runs as an equilibrium phenomenon, but again do not deal with panic-based runs. Chari

and Jagannathan (1988) and Chen (1999) study panic-based bank runs, but of a different kind: Runs that occur when uninformed agents interpret the fact that others run as an indication that fundamentals are bad.

While our paper focuses on demand–deposit contracts, two recent papers, Green and Lin (2003) and Peck and Shell (2003), inspired in part by Wallace (1988, 1990), study more flexible contracts that allow the bank to condition the payment to each depositor on the number of agents who claimed early withdrawal before her. Their goal is to find whether bank runs can be eliminated when such contracts are allowed. As the authors themselves note, however, such sophisticated contracts are not observed in practice. This may be due to a moral hazard problem that such contracts might generate: A flexible contract might enable the bank to lie about the circumstances and pay less to investors (whereas under demand–deposit contracts, the bank has to either pay a fixed payment or go bankrupt). In our paper, we focus on the simpler contracts that are observed in practice, and analyze the interdependence between the banking contract and the probability of bank runs, an issue that is not analyzed by Green and Lin or by Peck and Shell.

The remainder of this paper is organized as follows. Section I presents the basic framework without private signals. In Section II we introduce private signals into the model and obtain a unique equilibrium. Section III studies the relationship between the level of liquidity provided by the bank and the likelihood of runs, analyzes the optimal liquidity level, and inquires whether banks increase welfare relative to autarky. Concluding remarks appear in Section IV. Proofs are relegated to an appendix.

I. The Basic Framework

A. The Economy

There are three periods (0, 1, 2), one good, and a continuum [0, 1] of agents. Each agent is born in period 0 with an endowment of one unit. Consumption occurs only in period 1 or 2 (c_1 and c_2 denote an agent's consumption levels). Each agent can be of two types: With probability λ the agent is impatient and with probability $1 - \lambda$ she is patient. Agents' types are i.i.d.; we assume no aggregate uncertainty.[5] Agents learn their types (which are their private information) at the beginning of period 1. Impatient agents can consume only in period 1. They obtain utility of $u(c_1)$. Patient agents can consume at either period; their utility is $u(c_1 + c_2)$. Function u is twice continuously differentiable, increasing, and for any $c \geq 1$ has a relative risk-aversion coefficient, $-cu''(c)/u'(c)$, greater than 1. Without loss of generality, we assume that $u(0) = 0$.[6]

Agents have access to a productive technology that yields a higher expected return in the long run. For each unit of input in period 0, the technology

[5] Throughout this paper, we make the common assumption that with a continuum of i.i.d. random variables, the empirical mean equals the expectations with probability 1 (see Judd (1985)).

[6] Note that any vNM utility function, which is well defined at 0 (i.e., $u(0) \neq -\infty$), can be transformed into an equivalent utility function that satisfies $u(0) = 0$.

generates one unit of output if liquidated in period 1. If liquidated in period 2, the technology yields R units of output with probability $p(\theta)$, or 0 units with probability $1 - p(\theta)$. Here, θ is the state of the economy. It is drawn from a uniform distribution on [0, 1], and is unknown to agents before period 2. We assume that $p(\theta)$ is strictly increasing in θ. It also satisfies $E_\theta[p(\theta)]u(R) > u(1)$, so that for patient agents the expected long-run return is superior to the short-run return.

B. Risk Sharing

In autarky, impatient agents consume one unit in period 1, whereas patient agents consume R units in period 2 with probability $p(\theta)$. Because of the high coefficient of risk aversion, a transfer of consumption from patient agents to impatient agents could be beneficial, ex ante, to all agents, although it would necessitate the early liquidation of long-term investments. A social planner who can verify agents' types, once realized, would set the period 1 consumption level c_1 of the impatient agents so as to maximize an agent's ex ante expected welfare, $\lambda u(c_1) + (1-\lambda)u(\frac{1-\lambda c_1}{1-\lambda}R)E_\theta[p(\theta)]$. Here, λc_1 units of investment are liquidated in period 1 to satisfy the consumption needs of impatient agents. As a result, in period 2, each of the patient agents consumes $\frac{1-\lambda c_1}{1-\lambda}R$ with probability $p(\theta)$. This yields the following first-order condition that determines c_1 (c_1^{FB} denotes the first-best c_1):

$$u'(c_1^{\mathrm{FB}}) = Ru'\left(\frac{1-\lambda c_1^{\mathrm{FB}}}{1-\lambda}R\right)E_\theta[p(\theta)]. \tag{1}$$

This condition equates the benefit and cost from the early liquidation of the marginal unit of investment. The LHS is the marginal benefit to impatient agents, while the RHS is the marginal cost borne by the patient agents. At $c_1 = 1$, the marginal benefit is greater than the marginal cost: $1 \cdot u'(1) > R \cdot u'(R) \cdot E_\theta[p(\theta)]$. This is because $cu'(c)$ is a decreasing function of c (recall that the coefficient of relative risk aversion is more than 1), and because $E_\theta[p(\theta)] < 1$. Since the marginal benefit is decreasing in c_1 and the marginal cost is increasing, we must have $c_1^{\mathrm{FB}} > 1$. Thus, at the optimum, there is risk sharing: a transfer of wealth from patient agents to impatient agents.

C. Banks

The above analysis presumed that agents' types were observable. When types are private information, the payments to agents cannot be made contingent on their types. As D&D show, in such an environment, banks can enable risk sharing by offering a *demand–deposit contract*. Such a contract takes the following form: Each agent deposits her endowment in the bank in period 0. If she demands withdrawal in period 1, she is promised a fixed payment of $r_1 > 1$. If she waits until period 2, she receives a stochastic payoff of $\tilde{r}_2$, which is the proceeds

Table I
Ex Post Payments to Agents

Withdrawal in Period	$n < 1/r_1$	$n \geq 1/r_1$
1	r_1	$\begin{cases} r_1 \text{ with probability } \frac{1}{nr_1} \\ 0 \text{ with probability } 1 - \frac{1}{nr_1} \end{cases}$
2	$\begin{cases} \frac{(1-nr_1)}{1-n} R \text{ with probability } p(\theta) \\ 0 \quad \text{with probability } 1 - p(\theta) \end{cases}$	0

of the nonliquidated investments divided by the number of remaining depositors. In period 1, the bank must follow a *sequential service constraint*: It pays r_1 to agents until it runs out of resources. The consequent payments to agents are depicted in Table I (n denotes the proportion of agents demanding early withdrawal).[7]

Assume that the economy has a banking sector with free entry, and that all banks have access to the same investment technology. Since banks make no profits, they offer the same contract as the one that would be offered by a single bank that maximizes the welfare of agents.[8] Suppose the bank sets r_1 at c_1^{FB}. If only impatient agents demand early withdrawal, the expected utility of patient agents is $E_\theta[p(\theta)] \cdot u(\frac{1-\lambda r_1}{1-\lambda} R)$. As long as this is more than $u(r_1)$, there is an equilibrium in which, indeed, only impatient agents demand early withdrawal. In this equilibrium, the first-best allocation is obtained. However, as D&D point out, the demand–deposit contract makes the bank vulnerable to runs. There is a second equilibrium in which *all* agents demand early withdrawal. When they do so, period 1 payment is r_1 with probability $1/r_1$ and period 2 payment is 0, so that it is indeed optimal for agents to demand early withdrawal. This equilibrium is inferior to the autarkic regime.[9]

In determining the optimal short-term payment, it is important to know how likely each equilibrium is. D&D derive the optimal short-term payment under the implicit assumption that the "good" equilibrium is always selected. Consequently, their optimal r_1 is c_1^{FB}. This approach has two drawbacks. First, the contract is not optimal if the probability of bank runs is not negligible. It is

[7] Our payment structure assumes that there is no sequential service constraint in period 2. As a result, in case of a run ($n \neq \lambda$), agents who did not run in period 1 become "residual claimants" in period 2. This assumption slightly deviates from the typical deposit contract and is the same as in D&D.

[8] This equivalence follows from the fact that there are no externalities among different banks, and thus the contract that one bank offers to its investors does not affect the payoffs to agents who invest in another bank. (We assume an agent cannot invest in more than one bank.)

[9] If the incentive compatibility condition $E_\theta[p(\theta)]u(\frac{1-\lambda r_1}{1-\lambda} R) \geq u(r_1)$ does not hold for $r_1 = c_1^{FB}$, the bank can set r_1 at the highest level that satisfies the condition, and again we will have two equilibria. In our model, the incentive compatibility condition holds for $r_1 = c_1^{FB}$ as long as $E_\theta[p(\theta)]$ is large enough. The original D&D model is a special case of our model, where $E_\theta[p(\theta)] = 1$, and thus the condition always holds.

not even obvious that risk sharing is desirable in that case. Second, the computation of the banking contract presumes away any possible relation between the amount of liquidity provided by the banking contract and the likelihood of a bank run. If such a relation exists, the optimal r_1 may not be c_1^{FB}. These drawbacks are resolved in the next section, where we modify the model so as to obtain firmer predictions.

II. Agents with Private Signals: Unique Equilibrium

We now modify the model by assuming that, at the beginning of period 1, each agent receives a private signal regarding the fundamentals of the economy. (A second modification that concerns the technology is introduced later.) As we show below, these signals force agents to coordinate their actions: They run on the bank when the fundamentals are in one range and select the "good" equilibrium in another range. (Abusing both English and decision theory, we will sometimes refer to demanding early withdrawal as "running on the bank.") This enables us to determine the probability of a bank run for any *given* short-term payment. Knowing how this probability is affected by the amount of risk sharing provided by the contract, we then revert to period 0 and find the optimal short-term payment.

Specifically, we assume that state θ is realized at the beginning of period 1. At this point, θ is not publicly revealed. Rather, each agent i obtains a signal $\theta_i = \theta + \varepsilon_i$, where ε_i are small error terms that are independently and uniformly distributed over $[-\varepsilon, \varepsilon]$. An agent's signal can be thought of as her private information, or as her private opinion regarding the prospects of the long-term return on the investment project. Note that while each agent has different information, none has an advantage in terms of the quality of the signal.

The introduction of private signals changes the results considerably. A patient agent's decision whether to run on the bank depends now on her signal. The effect of the signal is twofold. The signal provides information regarding the expected period 2 payment: The higher the signal, the higher is the posterior probability attributed by the agent to the event that the long-term return is going to be R (rather than 0), and the lower the incentive to run on the bank. In addition, an agent's signal provides information about other agents' signals, which allows an inference regarding their actions. Observing a high signal makes the agent believe that other agents obtained high signals as well. Consequently, she attributes a low likelihood to the possibility of a bank run. This makes her incentive to run even smaller.

We start by analyzing the events in period 1, assuming that the banking contract that was chosen in period 0 offers r_1 to agents demanding withdrawal in period 1, and that all agents have chosen to deposit their endowments in the bank. (Clearly, r_1 must be at least 1 but less than $\min\{1/\lambda, R\}$.) While all impatient agents demand early withdrawal, patient agents need to compare the expected payoffs from going to the bank in period 1 or 2. The ex post payoff of a patient agent from these two options depends on both θ and the proportion

n of agents demanding early withdrawal (see Table I). Since the agent's signal gives her (partial) information regarding both θ and n, it affects the calculation of her expected payoffs. Thus, her action depends on her signal.

We assume that there are ranges of extremely good or extremely bad fundamentals, in which a patient agent's best action is independent of her belief concerning other patient agents' behavior. As we show in the sequel, the mere existence of these extreme regions, no matter how small they are, ignites a contagion effect that leads to a unique outcome for any realization of the fundamentals. Moreover, the probability of a bank run does not depend on the exact specification of the two regions.

We start with the lower range. When the fundamentals are very bad (θ very low), the probability of default is very high, and thus the expected utility from waiting until period 2 is lower than that of withdrawing in period 1—even if all patient agents were to wait ($n = \lambda$). If, given her signal, a patient agent is sure that this is the case, her best action is to run, regardless of her belief about the behavior of the other agents. More precisely, we denote by $\underline{\theta}(r_1)$ the value of θ for which $u(r_1) = p(\theta)u(\frac{1-\lambda r_1}{1-\lambda}R)$, and refer to the interval $[0, \underline{\theta}(r_1))$ as the *lower dominance region*. Since the difference between an agent's signal and the true θ is no more than ε, we know that she demands early withdrawal if she observes a signal $\theta_i < \underline{\theta}(r_1) - \varepsilon$. We assume that for any $r_1 \geq 1$ there are feasible values of θ for which all agents receive signals that assure them that θ is in the lower dominance region. Since $\underline{\theta}$ is increasing in r_1, the condition that guarantees this for any $r_1 \geq 1$ is $\underline{\theta}(1) > 2\varepsilon$, or equivalently $p^{-1}(\frac{u(1)}{u(R)}) > 2\varepsilon$.[10] In most of our analysis ε is taken to be arbitrarily close to 0. Thus, it will be sufficient to assume $p^{-1}(\frac{u(1)}{u(R)}) > 0$.

Similarly, we assume an *upper dominance region* of parameters: a range $(\bar{\theta}, 1]$ in which no patient agent demands early withdrawal. To this end, we need to modify the investment technology available to the bank. Instead of assuming that the short-term return is fixed at 1, we assume that it equals 1 in the range $[0, \bar{\theta}]$ and equals R in the range $(\bar{\theta}, 1]$. We also assume that $p(\theta) = 1$ in this range. (Except for the upper dominance region, $p(\theta)$ is strictly increasing.) The interpretation of these assumptions is that when the fundamentals are extremely high, such that the long-term return is obtained with certainty, the short-term return improves as well.[11] When a patient agent knows that the fundamentals are in the region $(\bar{\theta}, 1]$, she does not run, whatever her belief regarding the behavior of other agents is. Why? Since the short-term return from a single investment unit exceeds the maximal possible value of r_1 (which is $\min\{1/\lambda, R\}$), there is no need to liquidate more than one unit of investment in order to pay one agent in period 1. As a result, the payment to agents who

[10] This sufficient condition ensures that when $\theta = 0$, all patient agents observe signals below ε. These signals assure them that the fundamentals are below $2\varepsilon < \underline{\theta}(r_1)$, and thus they must decide to run.

[11] This specification of the short-term return is the simplest that guarantees the existence of an upper dominance region. A potentially more natural assumption, that the short-term return increases gradually over [0, 1], would lead to the same results but complicate the computations.

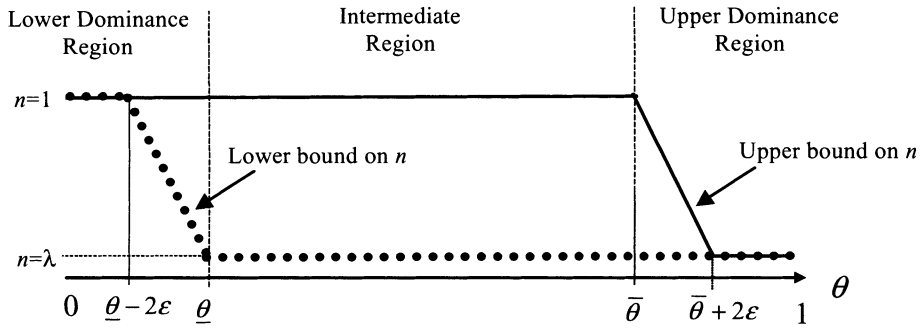

Figure 1. Direct implications of the dominance regions on agents' behavior.

withdraw in period 2 is guaranteed.[12] As in the case of the lower dominance region, we assume that $\bar{\theta} < 1 - 2\varepsilon$ (in most of our analysis ε is taken to be arbitrarily close to 0 and $\bar{\theta}$ arbitrarily close to 1).

An alternative assumption that generates an upper dominance region is that there exists an external large agent, who would be willing to buy the bank and pay its liabilities if she knew for sure that the long-run return was very high. This agent need not be a governmental institute; it can be a private agent, since she can be sure of making a large profit. Note however, that while such an assumption is very plausible if we think of a single bank (or country) being subject to a panic run, it is less plausible if we think of our bank as representing the world economy, when all sources of liquidity are already exhausted. (Our assumption about technology is not subject to this critique.)

Note that our model can be analyzed even if we do not assume the existence of an upper dominance region. In spite of the fact that in this case there are multiple equilibria, several equilibrium selection criteria (refinements) show that the more reasonable equilibrium is the same as the unique equilibrium that we obtain when we assume the upper dominance region. We discuss these refinements in Appendix B.

The two dominance regions are just extreme ranges of the fundamentals at which agents' behavior is known. This is illustrated in Figure 1. The dotted line represents a lower bound on n, implied by the lower dominance region. This line is constructed as follows. The agents who definitely demand early withdrawal are all the impatient agents, plus the patient agents who get signals below the threshold level $\underline{\theta}(r_1) - \varepsilon$. Thus, when $\theta < \underline{\theta}(r_1) - 2\varepsilon$, all patient agents get signals below $\underline{\theta}(r_1) - \varepsilon$ and n must be 1. When $\theta > \underline{\theta}(r_1)$, no patient agent gets a signal below $\underline{\theta}(r_1) - \varepsilon$ and no patient agent must run. As a result, the lower bound on n in this range is only λ. Since the distribution of signal errors is uniform, as θ grows from $\underline{\theta}(r_1) - 2\varepsilon$ to $\underline{\theta}(r_1)$, the proportion of patient agents observing signals below $\underline{\theta}(r_1) - \varepsilon$ decreases linearly at the rate $\frac{1-\lambda}{2\varepsilon}$. The solid line is the upper bound, implied by the upper dominance region. It is constructed

[12] More formally, when $\theta > \bar{\theta}$, an agent who demands early withdrawal receives r_1, whereas an agent who waits receives $(R - nr_1)/(1 - n)$ (which must be higher than r_1 because r_1 is less than R).

in a similar way, using the fact that patient agents do not run if they observe a signal above $\bar{\theta} + \varepsilon$.

Because the two dominance regions represent very unlikely scenarios, where fundamentals are so extreme that they determine uniquely what agents do, their existence gives little *direct* information regarding agents' behavior. The two bounds can be far apart, generating a large intermediate region in which an agent's optimal strategy depends on her beliefs regarding other agents' actions. However, the beliefs of agents in the intermediate region are not arbitrary. Since agents observe only noisy signals of the fundamentals, they do not exactly know the signals that other agents observed. Thus, in the choice of the equilibrium action at a given signal, an agent must take into account the equilibrium actions at nearby signals. Again, these actions depend on the equilibrium actions taken at further signals, and so on. Eventually, the equilibrium must be consistent with the (known) behavior at the dominance regions. Thus, our information structure places stringent restrictions on the structure of the equilibrium strategies and beliefs.

Theorem 1 says that the model with noisy signals has a unique equilibrium. A patient agent's action is uniquely determined by her signal: She demands early withdrawal if and only if her signal is below a certain threshold.

THEOREM 1: *The model has a unique equilibrium in which patient agents run if they observe a signal below threshold* $\theta^*(r_1)$ *and do not run above.*[13]

We can apply Theorem 1 to compute the proportion of agents who run at every realization of the fundamentals. Since there is a continuum of agents, we can define a deterministic function, $n(\theta, \theta')$, that specifies the proportion of agents who run when the fundamentals are θ and all agents run at signals below θ' and do not run at signals above θ'. Then, in equilibrium, the proportion of agents who run at each level of the fundamentals is given by $n(\theta, \theta^*) = \lambda + (1 - \lambda) \cdot \text{prob}[\varepsilon_i < \theta^* - \theta]$. It is 1 below $\theta^* - \varepsilon$, since there all patient agents observe signals below θ^*, and it is λ above $\theta^* + \varepsilon$, since there all patient agents observe signals above θ^*. Because both the fundamentals and the noise are uniformly distributed, $n(\theta, \theta^*)$ decreases linearly between $\theta^* - \varepsilon$ and $\theta^* + \varepsilon$. We thus have:

COROLLARY 1: *Given* r_1, *the proportion of agents demanding early withdrawal depends only on the fundamentals. It is given by:*

$$n(\theta, \theta^*(r_1)) = \begin{cases} 1 & \text{if } \theta \leq \theta^*(r_1) - \varepsilon \\ \lambda + (1 - \lambda)\left(\dfrac{1}{2} + \dfrac{\theta^*(r_1) - \theta}{2\varepsilon}\right) & \text{if } \theta^*(r_1) - \varepsilon \leq \theta \\ & \quad \leq \theta^*(r_1) + \varepsilon. \\ \lambda & \text{if } \theta \geq \theta^*(r_1) + \varepsilon \end{cases} \tag{2}$$

[13] We thank the referee for pointing out a difficulty with the generality of the original version of the proof.

Importantly, although the realization of θ uniquely determines how many patient agents run on the bank, most run episodes—those that occur in the intermediate region—are still driven by bad expectations. Since running on the bank is not a dominant action in this region, the reason patient agents do run is that they believe others will do so. Because they are driven by bad expectations, we refer to bank runs in the intermediate region as *panic-based runs*. Thus, the fundamentals serve as a coordination device for the expectations of agents, and thereby indirectly determine how many agents run on the bank. The crucial point is that this coordination device is not just a sunspot, but rather a payoff-relevant variable. This fact, and the existence of dominance regions, forces a unique outcome; in contrast to sunspots, there can be no equilibrium in which agents ignore their signals.

In the remainder of the section, we discuss the novelty of our uniqueness result and the intuition behind it.

The usual argument that shows that with noisy signals there is a unique equilibrium (see Carlsson and van Damme (1993) and Morris and Shin (1998)) builds on the property of *global strategic complementarities*: An agent's incentive to take an action is higher when more other agents take that action. This property does not hold in our model, since a patient agent's incentive to run is highest when $n = 1/r_1$, rather than when $n = 1$. This is a general feature of standard bank-run models: once the bank is already bankrupt, if more agents run, the probability of being served in the first period decreases, while the second-period payment remains null; thus, the incentive to run decreases. Specifically, a patient agent's utility differential, between withdrawing in period 2 versus period 1, is given by (see Table I):

$$v(\theta, n) = \begin{cases} p(\theta)u\left(\dfrac{1 - nr_1}{1 - n}R\right) - u(r_1) & \text{if } \dfrac{1}{r_1} \geq n \geq \lambda \\ 0 - \dfrac{1}{nr_1}u(r_1) & \text{if } 1 \geq n \geq \dfrac{1}{r_1} \end{cases} \tag{3}$$

Figure 2 illustrates this function for a given state θ. Global strategic complementarities require that v be always decreasing in n. While this does not hold in our setting, we do have *one-sided strategic complementarities*: v is monotonically decreasing whenever it is positive. We thus employ a new technical approach that uses this property to show the uniqueness of equilibrium. Our proof has two parts. In the first part, we restrict attention to *threshold equilibria*: equilibria in which all patient agents run if their signal is below some common threshold and do not run above. We show that there exists exactly one such equilibrium. The second part shows that any equilibrium must be a threshold equilibrium. We now explain the intuition for the first part of the proof. (This part requires even less than one-sided strategic complementarities. Single crossing, that is, v crossing 0 only once, suffices.) The intuition for the second part (which requires the stronger property) is more complicated—the interested reader is referred to the proof itself.

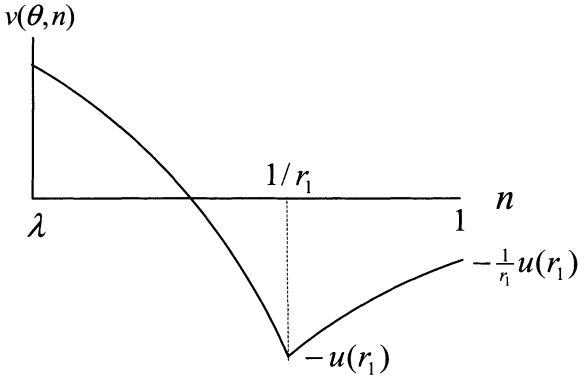

Figure 2. The net incentive to withdraw in period 2 versus period 1.

Assume that all patient agents run below the common threshold θ', and consider a patient agent who obtained signal θ_i. Let $\Delta^{r_1}(\theta_i, \theta')$ denote her expected utility differential, between withdrawing in period 2 versus period 1. The agent prefers to run (wait) if this difference is negative (positive). To compute $\Delta^{r_1}(\theta_i, \theta')$, note that since both state θ and error terms ε_i are uniformly distributed, the agent's posterior distribution of θ is uniformly distributed over $[\theta_i - \varepsilon, \theta_i + \varepsilon]$. Thus, $\Delta^{r_1}(\theta_i, \theta')$ is simply the average of $v(\theta, n(\theta, \theta'))$ over this range. (Recall that $n(\theta, \theta')$ is the proportion of agents who run at state θ. It is computed as in equation (2).) More precisely,

$$\Delta^{r_1}(\theta_i, \theta') = \frac{1}{2\varepsilon} \int_{\theta=\theta_i-\varepsilon}^{\theta_i+\varepsilon} v(\theta, n(\theta, \theta'))\, d\theta. \tag{4}$$

In a threshold equilibrium, a patient agent prefers to run if her signal is below the threshold and prefers to wait if her signal is above it. By continuity, she is indifferent at the threshold itself. Thus, an equilibrium with threshold θ' exists only if $\Delta^{r_1}(\theta', \theta') = 0$. This holds at exactly one point θ^*. To see why, note that $\Delta^{r_1}(\theta', \theta')$ is negative for $\theta' < \underline{\theta} - \varepsilon$ (because of the lower dominance region), positive for $\theta' > \bar{\theta} + \varepsilon$ (because of the upper dominance region), continuous, and increasing in θ' (when both the private signal and the threshold strategy increase by the same amount, an agent's belief about how many other agents withdraw is unchanged, but the return from waiting is higher). Thus, θ^* is the only candidate for a threshold equilibrium.

In order to show that it is indeed an equilibrium, we need to show that $\Delta^{r_1}(\theta_i, \theta^*)$ is negative for $\theta_i < \theta^*$ and positive for $\theta_i > \theta^*$ (recall that it is zero at $\theta_i = \theta^*$). In a model with global strategic complementarities, these properties would be easy to show: A higher signal indicates that the fundamentals are better and that fewer agents withdraw early—both would increase the gains from waiting. Since we have only partial strategic complementarities, however, the effect of more withdrawals becomes ambiguous, as sometimes the incentive to

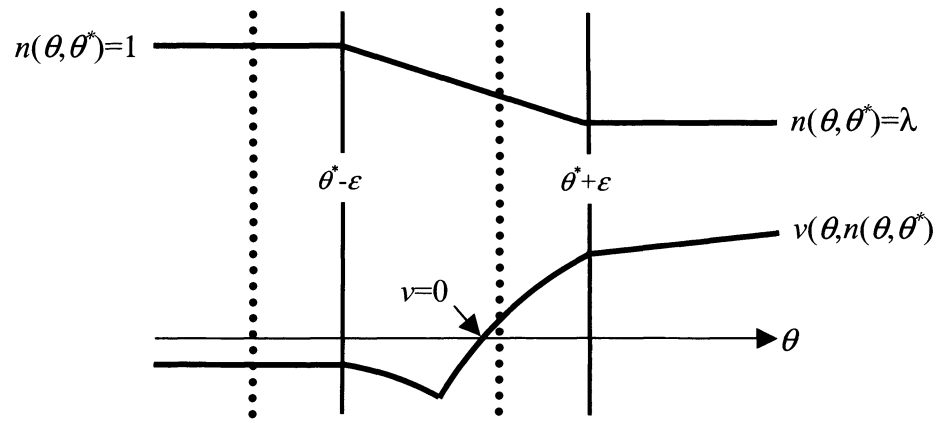

Figure 3. Functions $n(\theta, \theta^*)$ and $v(\theta, n(\theta, \theta^*))$.

run decreases when more agents do so. The intuition in our model is thus somewhat more subtle and relies on the single crossing property. Since $v(\theta, n(\theta, \theta'))$ crosses zero only once and since the posterior average of v, given signal $\theta_i = \theta^*$, is 0 (recall that an agent who observes θ^* is indifferent), observing a signal θ_i below θ^* shifts probability from positive values of v to negative values (recall that the distribution of noise is uniform). Thus, $\Delta^{r_1}(\theta_i, \theta^*) < \Delta^{r_1}(\theta^*, \theta^*) = 0$. This means that a patient agent with signal $\theta_i < \theta^*$ prefers to run. Similarly, for signals above θ^*, $\Delta^{r_1}(\theta_i, \theta^*) > \Delta^{r_1}(\theta^*, \theta^*) = 0$, which means that the agent prefers to wait.

To see this more clearly, consider Figure 3. The top graph depicts the proportion $n(\theta, \theta^*)$ of agents who run at state θ, while the bottom graph depicts the corresponding v. To compute $\Delta^{r_1}(\theta_i, \theta^*)$, we need to integrate v over the range $[\theta_i - \varepsilon, \theta_i + \varepsilon]$. We know that $\Delta^{r_1}(\theta^*, \theta^*)$, that is, the integral of v between the two solid vertical lines, equals 0. Consider now the integral of v over the range $[\theta_i - \varepsilon, \theta_i + \varepsilon]$ for some $\theta_i < \theta^*$ (i.e., between the two dotted lines). Relative to the integral between the solid lines, we take away a positive part and add a negative part. Thus, $\Delta^{r_1}(\theta_i, \theta^*)$ must be negative. (Note that v crossing zero only once ensures that the part taken away is positive, even if the right dotted line is to the left of the point where v crosses 0.) This means that a patient agent who observes a signal $\theta_i < \theta^*$ prefers to run. A similar argument shows that an agent who observes a signal $\theta_i > \theta^*$ prefers to stay.

To sum up the discussion of the proof, we reflect on the importance of two assumptions. The first is the uniform distribution of the fundamentals. This assumption does not limit the generality of the model in any crucial way, since there are no restrictions on the function $p(\theta)$, which relates the state θ to the probability of obtaining the return R in the long term. Thus, any distribution of the probability of obtaining R is allowed. Moreover, in analyzing the case of small noise ($\varepsilon \to 0$), this assumption can be dropped. The second assumption is the uniform distribution of noise. This assumption is important in deriving a unique equilibrium when the model does not have global strategic

complementarities. However, Morris and Shin (2003a), who discuss our result, show that if one restricts attention to monotone strategies, a unique equilibrium exists for a broader class of distributions.

III. The Demand–Deposit Contract and the Viability of Banks

Having established the existence of a unique equilibrium, we now study how the likelihood of runs depends on the promised short-term payment, r_1. We then analyze the optimal r_1, taking this dependence into consideration. We show that when $\underline{\theta}(1)$ is not too high, that is, when the lower dominance region is not too large, banks are viable: Demand–deposit contracts can increase welfare relative to the autarkic allocation. Nevertheless, under the optimal r_1, the demand deposit contract does not exploit all the potential gains from risk sharing. To simplify the exposition, we focus in this section on the case where ε and $1-\bar{\theta}$ are very close to 0. As in Section II, $\bar{\theta} < 1 - 2\varepsilon$ must hold. (All our results are proved for nonvanishing ε and $1-\bar{\theta}$, as long as they are below a certain bound.)

We first compute the threshold signal $\theta^*(r_1)$. A patient agent with signal $\theta^*(r_1)$ must be indifferent between withdrawing in period 1 or 2. That agent's posterior distribution of θ is uniform over the interval $[\theta^*(r_1) - \varepsilon, \theta^*(r_1) + \varepsilon]$. Moreover, she believes that the proportion of agents who run, as a function of θ, is $n(\theta, \theta^*(r_1))$ (see equation (2)). Thus, her posterior distribution of n is uniform over $[\lambda, 1]$. At the limit, the resulting indifference condition is $\int_{n=\lambda}^{1/r_1} u(r_1) + \int_{n=1/r_1}^{1} \frac{1}{nr_1} u(r_1) = \int_{n=\lambda}^{1/r_1} p(\theta^*) \cdot u(\frac{1-r_1 n}{1-n} R)$.[14] Solving for θ^*, we obtain:

$$\lim_{\varepsilon \to 0} \theta^*(r_1) = p^{-1}\left(\frac{u(r_1)(1 - \lambda r_1 + \ln(r_1)}{r_1 \int_{n=\lambda}^{1/r_1} u\left(\frac{1-r_1 n}{1-n} R\right)} \right). \tag{5}$$

Having characterized the unique equilibrium, we are now equipped with the tools to study the effect of the banking contract on the likelihood of runs. Theorem 2 says that when r_1 is larger, patient agents run in a larger set of signals. This means that the banking system becomes more vulnerable to bank runs when it offers more risk sharing. The intuition is simple: If the payment in period 1 is increased and the payment in period 2 is decreased, the incentive of patient agents to withdraw in period 1 is higher.[15] Note that this incentive is

[14] In this condition, the value of θ decreases from $\theta^* + \varepsilon$ to $\theta^* - \varepsilon$, as n increases from λ to 1. However, since ε approaches 0, θ is approximately θ^*. Note also that this implicit definition is correct as long as the resulting θ^* is below $\bar{\theta}$. This holds as long as r_1 is not too large; otherwise, θ^* would be close to $\bar{\theta}$. (Below, we show that when $\bar{\theta}$ is close to 1, the bank chooses r_1 sufficiently small so that there is a nontrivial region where agents do not run; i.e., such that θ^* is below $\bar{\theta}$.)

[15] There is an additional effect of increasing r_1, which operates in the opposite direction. In the range in which the bank does not have enough resources to pay all agents who demand early withdrawal, an increase in r_1 increases the probability that an agent who demands early withdrawal will not get any payment. This increased uncertainty over the period 1 payment reduces the incentive to run on the bank. As we show in the proof of Theorem 2, this effect must be weaker than the other effects if ε is not too large.

further increased since, knowing that other agents are more likely to withdraw in period 1, the agent assigns a higher probability to the event of a bank run.

THEOREM 2: *$\theta^*(r_1)$ is increasing in r_1.*

Knowing how r_1 affects the behavior of agents in period 1, we can revert to period 0 and compute the optimal r_1. The bank chooses r_1 to maximize the ex ante expected utility of a representative agent, which is given by

$$\lim_{\substack{\bar{\theta}\to 1\\ \varepsilon\to 0}} \mathrm{EU}(r_1) = \int_0^{\theta^*(r_1)} \frac{1}{r_1} u(r_1)\, d\theta + \int_{\theta^*(r_1)}^{1} \lambda \cdot u(r_1) + (1-\lambda)\cdot p(\theta)\cdot u\left(\frac{1-\lambda r_1}{1-\lambda} R\right) d\theta. \tag{6}$$

The ex ante expected utility depends on the payments under all possible values of θ. In the range below θ^*, there is a run. All investments are liquidated in period 1, and agents of both types receive r_1 with probability $1/r_1$. In the range above θ^*, there is no run. Impatient agents (λ) receive r_1 (in period 1), and patient agents ($1-\lambda$) wait till period 2 and receive $\frac{1-\lambda r_1}{1-\lambda} R$ with probability $p(\theta)$.

The computation of the optimal r_1 is different than that of Section I. The main difference is that now the bank needs to consider the effect that an increase in r_1 has on the expected costs from bank runs. This raises the question whether demand–deposit contracts, which pay impatient agents more than the liquidation value of their investments, are still desirable when the destabilizing effect of setting r_1 above 1 is taken into account. That is, whether provision of liquidity via demand–deposit contracts is desirable even when the cost of bank runs that result from this type of contract is considered. Theorem 3 gives a positive answer under the condition that $\underline{\theta}(1)$, the lower dominance region at $r_1 = 1$, is not too large. The exact bound is given in the proof (Appendix A).

THEOREM 3: *If $\underline{\theta}(1)$ is not too large, the optimal r_1 must be larger than* 1.

The intuition is as follows: Increasing r_1 slightly above 1 enables risk sharing among agents. The gain from risk sharing is of first-order significance, since the difference between the expected payment to patient agents and the payment to impatient agents is maximal at $r_1 = 1$. On the other hand, increasing r_1 above 1 is costly because of two effects. First, it widens the range in which bank runs occur and the investment is liquidated slightly beyond $\underline{\theta}(1)$. Second, in the range $[0, \underline{\theta}(1))$ (where runs always happen), it makes runs costly. This is because setting r_1 above 1 causes some agents not to get any payment (because of the sequential service constraint). The first effect is of second order. This is because at $\underline{\theta}(1)$ liquidation causes almost no harm since the utility from the liquidation value of the investment is close to the expected utility from the long-term value. The second effect is small provided that the range $[0, \underline{\theta}(1))$ is not too

large. Thus, overall, when $\underline{\theta}(1)$ is not too large, increasing r_1 above 1 is always optimal. Note that in the range $[0, \underline{\theta}(1))$, the return on the long-term technology is so low, such that early liquidation is efficient. Thus, the interpretation of the result is that if the range of fundamentals where liquidation is efficient is not too large, banks are viable.

Because the optimal level of r_1 is above 1, panic-based bank runs occur at the optimum. This can be seen by comparing $\theta^*(r_1)$ with $\underline{\theta}(r_1)$, and noting that the first is larger than the second when r_1 is above 1. Thus, under the optimal r_1, the demand–deposit contract achieves higher welfare than that reached under autarky, but is still inferior to the first-best allocation, as it generates panic-based bank runs.

Having shown that demand–deposit contracts improve welfare, we now analyze the forces that determine the optimal r_1, that is, the optimal level of liquidity that is provided by the banking contract. First, we note that the optimal r_1 must be lower than $\min\{1/\lambda, R\}$: If r_1 were larger, a bank run would always occur, and ex ante welfare would be lower than in the case of $r_1 = 1$. In fact, the optimal r_1 must be such that $\theta^*(r_1) < \bar{\theta}$, for the exact same reason. Thus, we must have an interior solution for r_1. The first-order condition that determines the optimal r_1 is (recall that ε and $1 - \bar{\theta}$ approach 0):

$$\begin{aligned}
&\lambda \int_{\theta^*(r_1)}^{1} \left[u'(r_1) - p(\theta) \cdot R \cdot u' \left(\frac{1 - \lambda r_1}{1 - \lambda} R \right) \right] d\theta \\
&= \frac{\partial \theta^*(r_1)}{\partial r_1} \left[\left(\lambda u(r_1) + (1 - \lambda) p(\theta^*(r_1)) u \left(\frac{1 - \lambda r_1}{1 - \lambda} R \right) \right) - \frac{1}{r_1} u(r_1) \right] \\
&\quad + \int_0^{\theta^*(r_1)} \left[\frac{u(r_1) - r_1 u'(r_1)}{r_1^2} \right] d\theta. \qquad (7)
\end{aligned}$$

The LHS is the marginal gain from better risk sharing, due to increasing r_1. The RHS is the marginal cost that results from the destabilizing effect of increasing r_1. The first term captures the increased probability of bank runs. The second term captures the increased cost of bank runs: When r_1 is higher, the bank's resources (one unit) are divided among fewer agents, implying a higher level of risk ex ante. Theorem 4 says that the optimal r_1 in our model is smaller than c_1^{FB}. The intuition is simple: c_1^{FB} is calculated to maximize the gain from risk sharing while ignoring the possibility of bank runs. In our model, the cost of bank runs is taken into consideration: Since a higher r_1 increases both the probability of bank runs and the welfare loss that results from bank runs, the optimal r_1 is lower.

THEOREM 4: *The optimal r_1 is lower than c_1^{FB}.*

Theorem 4 implies that the optimal short-term payment does not exploit all the potential gains from risk sharing. The bank can increase the gains from risk sharing by increasing r_1 above its optimal level. However, because of the increased costs of bank runs, the bank chooses not to do this. Thus, in the model

with noisy signals, the optimal contract must trade off risk sharing versus the costs of bank runs. This point cannot be addressed in the original D&D model.

IV. Concluding Remarks

We study a model of bank runs based on D&D's framework. While their model has multiple equilibria, ours has a unique equilibrium in which a run occurs if and only if the fundamentals of the economy are below some threshold level. Nonetheless, there are panic-based runs: Runs that occur when the fundamentals are good enough that agents would not run had they believed that others would not.

Knowing when runs occur, we compute their probability. We find that this probability depends on the contract offered by the bank: Banks become more vulnerable to runs when they offer more risk sharing. However, even when this destabilizing effect is taken into account, banks still increase welfare by offering demand deposit contracts (provided that the range of fundamentals where liquidation is efficient is not too large). We characterize the optimal short-term payment in the banking contract and show that this payment does not exploit all possible gains from risk sharing, since doing so would result in too many bank runs.

In the remainder of this section, we discuss two possible directions in which our results can be applied.

The first direction is related to policy analysis. One of the main features of our model is the endogenous determination of the probability of bank runs. This probability is a key element in assessing policies that are intended to prevent runs, or in comparing demand deposits to alternative types of contracts. For example, two policy measures that are often mentioned in the context of bank runs are suspension of convertibility and deposit insurance. Clearly, if such policy measures come without a cost, they are desirable. However, these measures do have costs. Suspension of convertibility might prevent consumption from agents who face early liquidity needs. Deposit insurance generates a moral hazard problem: Since a single bank does not bear the full cost of an eventual run, each bank sets a too high short-term payment and, consequently, the probability of runs is above the social optimum. Being able to derive an endogenous probability of bank runs, our model can be easily applied to assess the desirability of such policy measures or specify under which circumstances they are welfare improving. This analysis cannot be conducted in a model with multiple equilibria, since in such a model the probability of a bank run is unknown and thus the expected welfare cannot be calculated. We leave this analysis to future research.

Another direction is related to the proof of the unique equilibrium. The novelty of our proof technique is that it can be applied also to settings in which the strategic complementarities are not global. Such settings are not uncommon in finance. One immediate example is a debt rollover problem, in which a firm faces many creditors that need to decide whether to roll over the debt or

not. The corresponding payoff structure is very similar to that of our bank-run model, and thus exhibits only one-sided strategic complementarities. Another example is an investment in an IPO. Here, a critical mass of investors might be needed to make the firm viable. Thus, initially agents face strategic complementarities: their incentive to invest increases with the number of agents who invest. However, beyond the critical mass, the price increases with the number of investors, and this makes the investment less profitable for an individual investor. This type of logic is not specific to IPOs, and can apply also to any investment in a young firm or an emerging market. Finally, consider the case of a run on a financial market (see, e.g., Bernardo and Welch (2004) and Morris and Shin (2003b)). Here, investors rush to sell an asset and cause a collapse in its price. Strategic complementarities may exist if there is an opportunity to sell before the price fully collapses. However, they might not be global, as in some range the opposite effect may prevail: When more investors sell, the price of the asset decreases, and the incentive of an individual investor to sell decreases.

Appendix A: Proofs

Proof of Theorem 1: The proof is divided into three parts. We start with defining the notation used in the proof, including some preliminary results. We then restrict attention to threshold equilibria and show that there exists exactly one such equilibrium. Finally, we show that any equilibrium must be a threshold equilibrium. Proofs of intermediate lemmas appear at the end.

A. Notation and Preliminary Results

A (mixed) strategy for agent i is a measurable function $s_i : [0-\varepsilon, 1+\varepsilon] \to [0, 1]$ that indicates the probability that the agent, if patient, withdraws early (runs) given her signal θ_i. A strategy profile is denoted by $\{s_i\}_{i\in[0,1]}$. A given strategy profile generates a random variable $\tilde{n}(\theta)$ that represents the proportion of agents demanding early withdrawal at state θ. We define $\tilde{n}(\theta)$ by its cumulative distribution function:

$$F_\theta(n) = \text{prob}[\tilde{n}(\theta) \le n] = \underset{\{\varepsilon_i : i\in[0,1]\}}{\text{prob}} \left[\lambda + (1-\lambda) \cdot \int_{i=0}^{1} s_i(\theta+\varepsilon_i)\, di \le n \right]. \quad \text{(A1)}$$

In some cases, it is convenient to use the inverse CDF:

$$n^x(\theta) = \inf\{n : F_\theta(n) \ge x\}. \quad \text{(A2)}$$

Let $\Delta^{r1}(\theta_i, \tilde{n}(\cdot))$ denote a patient agent's expected difference in utility from withdrawing in period 2 rather than in period 1, when she observes signal θ_i and holds belief $\tilde{n}(\cdot)$ regarding the proportion of agents who run at each state θ. (Note that since we have a continuum of agents, $\tilde{n}(\theta)$ does not depend on the action of the agent herself.) When the agent observes signal θ_i, her posterior distribution of θ is uniform over the interval $[\theta_i - \varepsilon, \theta_i + \varepsilon]$. For each θ

in this range, the agent's expected payoff differential is the expectations, over the one-dimensional random variable $\tilde{n}(\theta)$, of $v(\theta, \tilde{n}(\theta))$ (see definition of v in equation (3)).[16] We thus have[17]

$$\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) = \frac{1}{2\varepsilon} \int_{\theta=\theta_i-\varepsilon}^{\theta_i+\varepsilon} \mathrm{E}_n[v(\theta, \tilde{n}(\theta))]\, d\theta$$
$$\equiv \frac{1}{2\varepsilon} \int_{\theta=\theta_i-\varepsilon}^{\theta_i+\varepsilon} \left(\int_{n=\lambda}^{1} v(\theta, n)\, dF_\theta(n) \right) d\theta. \tag{A3}$$

In case all patient agents have the same strategy (i.e., the same function from signals to actions), the proportion of agents who run at each state θ is deterministic. To ease the exposition, we then treat $\tilde{n}(\theta)$ as a number between λ and 1, rather than as a degenerate random variable, and write $n(\theta)$ instead of $\tilde{n}(\theta)$. In this case, $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot))$ reduces to

$$\Delta^{r_1}(\theta_i, n(\cdot)) = \frac{1}{2\varepsilon} \int_{\theta=\theta_i-\varepsilon}^{\theta_i+\varepsilon} v(\theta, n(\theta))\, d\theta. \tag{A4}$$

Also, recall that if all patient agents have the same threshold strategy, we denote $n(\theta)$ as $n(\theta, \theta')$, where θ' is the common threshold—see explicit expression in equation (2).

Lemma A1 states a few properties of $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot))$. It says that $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot))$ is continuous in θ_i, and increases continuously in positive shifts of both the signal θ_i and the belief $\tilde{n}(\cdot)$. For the purpose of the lemma, denote $(\tilde{n} + a)(\theta) = \tilde{n}(\theta + a)$ for all θ:

LEMMA A1: *(i) Function $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot))$ is continuous in θ_i. (ii) Function $\Delta^{r_1}(\theta_i + a, (\tilde{n} + a)(\cdot))$ is continuous and nondecreasing in a. (iii) Function $\Delta^{r_1}(\theta_i + a, (\tilde{n} + a)(\cdot))$ is strictly increasing in a if $\theta_i + a < \bar{\theta} + \varepsilon$ and if $\tilde{n}(\theta) < 1/r_1$ with positive probability over $\theta \in [\theta_i + a - \varepsilon, \theta_i + a + \varepsilon]$.*

A Bayesian equilibrium is a measurable strategy profile $\{s_i\}_{i\in[0,1]}$, such that each patient agent chooses the best action at each signal, given the strategies of the other agents. Specifically, in equilibrium, a patient agent i chooses $s_i(\theta_i) = 1$ (withdraws early) if $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) < 0$, chooses $s_i(\theta_i) = 0$ (waits) if $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) > 0$, and may choose any $1 \geq s_i(\theta_i) \geq 0$ (is indifferent) if $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) = 0$. Note that, consequently, patient agents' strategies must be the same, except for signals θ_i at which $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) = 0$.

[16] While this definition applies only for $\theta < \bar{\theta}$, for brevity we use it also if $\theta > \bar{\theta}$. It is easy to check that all our arguments hold if the correct definition of v for that range is used.

[17] Note that the function Δ^{r_1} defined here is slightly different than the function defined in the intuition sketched in Section II (see equation (4)). This is because here we do not restrict attention only to threshold strategies.

B. There Exists a Unique Threshold Equilibrium

A threshold equilibrium with threshold signal θ^* exists if and only if, given that all other patient agents use threshold strategy θ^*, each patient agent finds it optimal to run when she observes a signal below θ^*, and to wait when she observes a signal above θ^*:

$$\Delta^{r_1}\big(\theta_i, n(\cdot, \theta^*)\big) < 0 \quad \forall \theta_i < \theta^*; \tag{A5}$$

$$\Delta^{r_1}\big(\theta_i, n(\cdot, \theta^*)\big) > 0 \quad \forall \theta_i > \theta^*; \tag{A6}$$

By continuity (Lemma A1 part (i)), a patient agent is indifferent when she observes θ^*:

$$\Delta^{r_1}(\theta^*, n(\cdot, \theta^*)) = 0. \tag{A7}$$

We first show that there is exactly one value of θ^* that satisfies (A7). By Lemma A1 part (ii), $\Delta^{r_1}(\theta^*, n(\cdot, \theta^*))$ is continuous in θ^*. By the existence of dominance regions, it is negative below $\underline{\theta}(r_1) - \varepsilon$ and positive above $\bar{\theta} + \varepsilon$. Thus, there exists some θ^* at which it equals 0. Moreover, this θ^* is unique. This is because, by part (iii) of Lemma A1, $\Delta^{r_1}(\theta^*, n(\cdot, \theta^*))$ is strictly increasing in θ^* as long as it is below $\bar{\theta} + \varepsilon$ (by the definition of $n(\theta, \theta^*)$ in equation (2), note that $n(\theta, \theta^*) < 1/r_1$ with positive probability over the range $\theta \in [\theta^* - \varepsilon, \theta^* + \varepsilon]$).

Thus, there is exactly one value of θ^*, which is a candidate for a threshold equilibrium. To establish that it is indeed an equilibrium, we need to show that given that (A7) holds, (A5) and (A6) must also hold. To prove (A5), we decompose the intervals $[\theta_i - \varepsilon, \theta_i + \varepsilon]$ and $[\theta^* - \varepsilon, \theta^* + \varepsilon]$, over which the integrals $\Delta^{r_1}(\theta_i, n(\cdot, \theta^*))$ and $\Delta^{r_1}(\theta^*, n(\cdot, \theta^*))$ are computed, respectively, into a (maybe empty) common part $c = [\theta_i - \varepsilon, \theta_i + \varepsilon] \cap [\theta^* - \varepsilon, \theta^* + \varepsilon]$, and two disjoint parts $d^i = [\theta_i - \varepsilon, \theta_i + \varepsilon] \setminus c$ and $d^* = [\theta^* - \varepsilon, \theta^* + \varepsilon] \setminus c$. Then, using (A4), we write

$$\Delta^{r_1}(\theta^*, n(\cdot, \theta^*)) = \frac{1}{2\varepsilon} \int_{\theta \in c} v(\theta, n(\cdot, \theta^*)) + \frac{1}{2\varepsilon} \int_{\theta \in d^*} v(\theta, n(\cdot, \theta^*)), \tag{A8}$$

$$\Delta^{r_1}\big(\theta_i, n(\cdot, \theta^*)\big) = \frac{1}{2\varepsilon} \int_{\theta \in c} v(\theta, n(\cdot, \theta^*)) + \frac{1}{2\varepsilon} \int_{\theta \in d^i} v(\theta, n(\cdot, \theta^*)). \tag{A9}$$

Analyzing (A8), we see that $\int_{\theta \in c} v(\theta, n(\cdot, \theta^*))$ is negative. This is because $\Delta^{r_1}(\theta^*, n(\cdot, \theta^*)) = 0$ (since (A7) holds); because the fundamentals in the range d^* are higher than in the range c (since $\theta_i < \theta^*$); and because in the interval $[\theta^* - \varepsilon, \theta^* + \varepsilon]$, $v(\theta, n(\cdot, \theta^*))$ is positive for high values of θ, negative for low values of θ, and crosses zero only once (see Figure 3). Moreover, by the definition of $n(\theta, \theta^*)$ (see equation (2)), $n(\theta, \theta^*)$ is always one over the interval d^i (which is below $\theta^* - \varepsilon$). Thus, $\int_{\theta \in d^i} v(\theta, n(\cdot, \theta^*))$ is also negative. Then, by (A9), $\Delta^{r_1}(\theta_i, n(\cdot, \theta^*))$ is negative. This proves that (A5) holds. The proof for A6 is analogous.

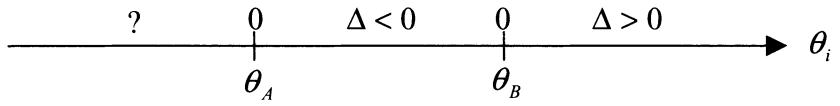

Figure A1. Function Δ in a (counterfactual) nonthreshold equilibrium.

C. Any Equilibrium Must Be a Threshold Equilibrium

Suppose that $\{s_i\}_{i\in[0,1]}$ is an equilibrium and that $\tilde{n}(\cdot)$ is the corresponding distribution of the proportion of agents who withdraw as a function of the state θ. Let θ_B be the highest signal at which patient agents do not strictly prefer to wait:

$$\theta_B = \sup\left\{\theta_i : \Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) \leq 0\right\}. \tag{A10}$$

By continuity (part (i) of Lemma A1), patient agents are indifferent at θ_B; that is, $\Delta^{r_1}(\theta_B, \tilde{n}(\cdot)) = 0$. By the existence of an upper dominance region, $\theta_B < 1 - \varepsilon$.

If we are not in a threshold equilibrium, there are signals below θ_B at which $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) \geq 0$. Let θ_A be their supremum:

$$\theta_A = \sup\left\{\theta_i < \theta_B : \Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) \geq 0\right\}. \tag{A11}$$

Again, by continuity, patient agents are indifferent at θ_A. Thus, we must have

$$\Delta^{r_1}(\theta_B, \tilde{n}(\cdot)) = \Delta^{r_1}(\theta_A, \tilde{n}(\cdot)) = 0. \tag{A12}$$

Figure A1 illustrates the (counterfactual) nonthreshold equilibrium. Essentially, patient agents do not run at signals above θ_B, run between θ_A and θ_B, and may or may not run below θ_A. Our proof goes on to show that this equilibrium cannot exist by showing that (A12) cannot hold. Since this is the most subtle part of the proof, we start by discussing a special case that illustrates the main intuition behind the general proof, and then continue with the complete proof that takes care of all the technical details.

INTUITION: Suppose that the distance between θ_B and θ_A is greater than 2ε and that the proportion of agents who run at each state θ is deterministic and denoted as $n(\theta)$ (rather than $\tilde{n}(\theta)$). From (A4), we know that $\Delta^{r_1}(\theta_B, n(\cdot))$ is an integral of the function $v(\theta, n(\theta))$ over the range $(\theta_B - \varepsilon, \theta_B + \varepsilon)$. We denote this range as d^B. Similarly, $\Delta^{r_1}(\theta_A, n(\cdot))$ is an integral of $v(\theta, n(\theta))$ over the range d^A.

To show that (A12) cannot hold, let us compare the integral of $v(\theta, n(\theta))$ over d^B with that over d^A. We can pair each point θ in d^B with a "mirror image" point $\overleftarrow{\theta}$ in d^A, such that as θ moves from the lower end of d^B to its upper end, $\overleftarrow{\theta}$ moves from the upper end of d^A to its lower end. From the behavior of agents illustrated in Figure A1, we easily see that: (1) All agents withdraw at the left-hand border of d^B and at the right-hand border of d^A. (2) As θ increases in the range d^B, we replace patient agents who always run with patient agents who never run, implying that $n(\theta)$ decreases (from 1 to λ) at the fastest feasible rate

over d^B. These two properties imply that n (the number of agents who withdraw) is higher in $\overleftarrow{\theta}$ than in θ. In a model with global strategic complementarities (where v is always decreasing in n), this result, together with the fact that $\overleftarrow{\theta}$ is always below θ, would be enough to establish that the integral of v over d^B yields a higher value than that over d^A, and thus that (A12) cannot hold. However, since, in our model, we do not have global strategic complementarities (as v reverses direction when $n > 1/r_1$), we need to use a more complicated argument that builds only on the property of one-sided strategic complementarities (v is monotonically decreasing whenever it is positive—see Figure 2).

Let us compare the distribution of n over d^B and its distribution over d^A. While over d^B,n decreases at the fastest feasible rate from 1 to λ, over $d^A n$ decreases more slowly. Thus, over d^A,n ranges from 1 to some $n^* > \lambda$, and each value of n in that range has more weight than its counterpart in d^B. In other words, the distribution over d^A results from moving all the weight that the distribution over d^B puts on values of n below n^* to values of n above n^*.

To conclude the argument, consider two cases. First, suppose that v is nonnegative at n^*. Then, since v is a monotone in n when it is nonnegative, when we move from d^B to d^A, we shift probability mass from high values of v to low values of v, implying that the integral of v over d^B is greater than that over d^A, and thus (A12) does not hold. Second, consider the case where v is negative at n^*. Then, since one-sided strategic complementarities imply single crossing (i.e., v crosses zero only once), v must always be negative over d^A, and thus the integral of v over d^A is negative. But then this must be less than the integral over d^B, which equals zero (by (A10))—so again (A12) does not hold.

Having seen the basic intuition as to why (A12) cannot hold, we now provide the formal proof that allows $\theta_B - \theta_A$ to be below 2ε and the proportion of agents who run at each state θ to be nondeterministic (denoted as $\tilde{n}(\theta)$). The proof proceeds in three steps.

Step 1: First note that θ_A is strictly below θ_B, that is, there is a nontrivial interval of signals below θ_B, where $\Delta^{r_1}(\theta_i, \tilde{n}(\cdot)) < 0$. This is because the derivative of Δ^{r_1} with respect to θ_i at the point θ_B is negative. (The derivative is given by $\mathrm{E}_n v(\theta_B + \varepsilon, \tilde{n}(\theta_B + \varepsilon)) - \mathrm{E}_n v(\theta_B - \varepsilon, \tilde{n}(\theta_B - \varepsilon))$. When state θ equals $\theta_B + \varepsilon$, all patient agents obtain signals above θ_B and thus none of them runs. Thus, $\tilde{n}(\theta_B + \varepsilon)$ is degenerate at $n = \lambda$. Now, for any $n \geq \lambda$, $v(\theta_B + \varepsilon, \lambda)$ is higher than $v(\theta_B - \varepsilon, n)$, since $v(\theta, n)$ is increasing in θ, and given θ, is maximized when $n = \lambda$.)

We now decompose the intervals over which the integrals $\Delta^{r_1}(\theta_A, \tilde{n}(\cdot))$ and $\Delta^{r_1}(\theta_B, \tilde{n}(\cdot))$ are computed into a (maybe empty) common part $c = (\theta_A - \varepsilon, \theta_A + \varepsilon) \cap (\theta_B - \varepsilon, \theta_B + \varepsilon)$, and two disjoint parts: $d^A = [\theta_A - \varepsilon, \theta_A + \varepsilon] \setminus c$ and $d^B = [\theta_B - \varepsilon, \theta_B + \varepsilon] \setminus c$. Denote the range d^B as $[\theta_1, \theta_2]$ and consider the "mirror image" transformation $\overleftarrow{\theta}$ (recall that as θ moves from the lower end of d^B to its upper end, $\overleftarrow{\theta}$ moves from the upper end of d^A to its lower end):

$$\overleftarrow{\theta} = \theta_A + \theta_B - \theta. \tag{A13}$$

Using (A3), we can now rewrite (A12) as

$$\int_{\theta=\theta_1}^{\theta_2} \mathrm{E}_n v(\theta, \tilde{n}(\theta)) = \int_{\theta=\theta_1}^{\theta_2} \mathrm{E}_n v(\overleftarrow{\theta}, \tilde{n}(\overleftarrow{\theta})). \tag{A14}$$

To interpret the LHS of (A14), we note that the proportion of agents who run at states $\theta \in d^B = [\theta_1, \theta_2]$ is deterministic. We denote it as $n(\theta)$, which is given by

$$n(\theta) = \lambda + (1-\lambda)(\theta_2 - \theta)/2\varepsilon \quad \text{for all } \theta \in d^B = [\theta_1, \theta_2]. \tag{A15}$$

To see why, note that when $\theta = \theta_2 (= \theta_B + \varepsilon)$, no patient agent runs and thus $n(\theta_2) = \lambda$. As we move from state θ_2 to lower states (while remaining in the interval d^B), we replace patient agents who get signals above θ_B and do not run, with patient agents who get signals between θ_A and θ_B and run. Thus, $n(\theta)$ increases at the rate of $(1-\lambda)/2\varepsilon$.

As for the RHS of (A14), we can write

$$\begin{aligned} \int_{\theta_1}^{\theta_2} \mathrm{E}_n v(\overleftarrow{\theta}, \tilde{n}(\overleftarrow{\theta}))\, d\theta &= \int_{\theta_1}^{\theta_2} \int_{x=0}^{1} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))\, dx\, d\theta \\ &= \int_{x=0}^{1} \int_{\theta_1}^{\theta_2} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))\, d\theta\, dx, \end{aligned} \tag{A16}$$

where $n^x(\theta)$ denotes the inverse CDF of $\tilde{n}(\theta)$ (see (A2)). Thus, to show that (A14) cannot hold, we go on to show in Step 3 that for each x,

$$\int_{\theta_1}^{\theta_2} v(\theta, n(\theta))\, d\theta > \int_{\theta_1}^{\theta_2} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))\, d\theta. \tag{A17}$$

But before, in Step 2, we derive a few intermediate results that are important for Step 3.

Step 2: We first show that $n(\theta)$ in the LHS of (A17) changes faster than $n^x(\overleftarrow{\theta})$ in the RHS:

LEMMA A2: $\left|\frac{\partial n^x(\overleftarrow{\theta})}{\partial \theta}\right| \leq \left|\frac{\partial n(\theta)}{\partial \theta}\right| = \frac{1-\lambda}{2\varepsilon}$.

This is based, as before, on the idea that $n(\theta)$ changes at the fastest feasible rate, $\frac{1-\lambda}{2\varepsilon}$, in $d^B = [\theta_1, \theta_2]$. The proof of the lemma takes care of the complexity that arises because $n^x(\theta)$ is not a standard deterministic behavior function, but rather the inverse CDF of $\tilde{n}(\theta)$. We can now collect a few more intermediate results:

Claim 1: For any $\theta \in [\theta_1, \theta_2]$, $\overleftarrow{\theta} < \theta_1$.

This follows from the definition of $\overleftarrow{\theta}$ in (A13).

Claim 2: If c is nonempty, then for any $\theta \in c = [\overleftarrow{\theta}_1, \theta_1]$, $n^x(\theta) \geq n(\theta_1)$.

This follows from the fact that as we move down from θ_1 to $\overleftarrow{\theta}_1$, we replace patient agents who get signals above θ_B and do not run, with patient agents who get signals below θ_A and may or may not run, implying that the whole support of $\tilde{n}(\theta)$ is above $n(\theta_1)$.

Claim 3: The LHS of (A17), $\int_{\theta=\theta_1}^{\theta_2} v(\theta, n(\theta))$, must be nonnegative.

When c is empty, this holds because $\int_{\theta=\theta_1}^{\theta_2} v(\theta, n(\theta))$ is equal to $\Delta^{r_1}(\theta_B, n(\theta))$, which equals zero. When c is nonempty, if $\int_{\theta=\theta_1}^{\theta_2} v(\theta, n(\theta))$ were negative, then for some $\theta \in [\theta_1, \theta_2]$ we would have $v(\theta, n(\theta)) < 0$. Since $n(\theta)$ is decreasing in the range $[\theta_1, \theta_2]$, and since $v(\theta, n)$ increases in θ and satisfies single crossing with respect to n (i.e., if $v(\theta, n) < 0$ and $n' > n$, then $v(\theta, n') < 0$), we get that $v(\theta_1, n(\theta_1)) < 0$. Applying Claim 2 and using again the fact that $v(\theta, n)$ increases in θ and satisfies single crossing with respect to n, we get that $\int_{\theta \in c} \mathrm{E}_n v(\theta, \tilde{n}(\theta))$ is also negative. This contradicts the fact that the sum of $\int_{\theta=\theta_1}^{\theta_2} v(\theta, n(\theta))$ and $\int_{\theta \in c} \mathrm{E}_n v(\theta, \tilde{n}(\theta))$, which is simply $\Delta^{r_1}(\theta_B, n(\theta))$, equals 0.

Step 3: We now turn to show that (A17) holds. For any $\theta \in [\theta_1, \theta_2]$, let $\underline{n}^x(\overleftarrow{\theta})$ be a monotone (i.e., weakly decreasing in θ) version of $n^x(\overleftarrow{\theta})$, and let $\theta^x(n)$ be its "inverse" function:

$$\begin{aligned} \underline{n}^x(\overleftarrow{\theta}) &= \min\left(n^x(\overleftarrow{t}) : \theta_1 \le t \le \theta\right) \\ \theta^x(n) &= \min\left(\left(\theta \in [\theta_1, \theta_2] : n^x(\overleftarrow{\theta}) \le n\right) \cup \{\theta_2\}\right). \end{aligned} \tag{A18}$$

(Note that if $n^x(\overleftarrow{\theta}) > n$ for all $\theta \in [\theta_1, \theta_2]$, then $\theta^x(n)$ is defined as θ_2.) Let $A(\theta)$ indicate whether $\underline{n}^x(\overleftarrow{\theta})$ is strictly decreasing at θ (if it is not, then there is a jump in $\theta^x(\underline{n}^x(\overleftarrow{\theta}))$):

$$A(\theta) = \begin{cases} 1 & \text{if } \partial \underline{n}^x(\overleftarrow{\theta})/\partial\theta \text{ exists, and is strictly negative} \\ 0 & \text{otherwise.} \end{cases}$$

Since $n^x(\overleftarrow{\theta}_1) \ge n(\theta_1)$ (Claim 2), we can rewrite the RHS of (A17) as

$$\begin{aligned} &\int_{\theta_1}^{\theta^x(n(\theta_1))} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))\, d\theta + \int_{\theta^x(n(\theta_1))}^{\theta_2} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))(1 - A(\theta))\, d\theta \\ &\quad + \int_{\theta^x(n(\theta_1))}^{\theta_2} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))A(\theta)\, d\theta. \end{aligned} \tag{A19}$$

The third summand in (A19) equals $\int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\overleftarrow{\theta^x(n)}, n)A(\theta^x(n))\, d\theta^x(n)$, and can be written as

$$\int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\overleftarrow{\theta^x(n)}, n) \cdot A(\theta^x(n)) d(\theta^x(n) - \theta(n))$$

$$+ \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\overleftarrow{\theta^x(n)}, n) A(\theta^x(n)) d\theta(n), \tag{A20}$$

where $\theta(n)$ is the inverse function of $n(\theta)$. Note that the second summand simply equals $\int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\overleftarrow{\theta^x(n)}, n) d\theta(n)$. This is because $A(\theta^x(n))$ is 0 no more than countably many times (corresponding to the jumps in $\theta^x(n)$), and because $\theta(n)$ is differentiable.

We now rewrite the LHS of (A17):

$$\int_{\theta_1}^{\theta_2} v(\theta, n(\theta)) d\theta = \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\theta(n), n) d\theta(n) + \int_{\theta(\underline{n}^x(\overleftarrow{\theta}_2))}^{\theta_2} v(\theta, n(\theta)) d\theta. \tag{A21}$$

By Claim 1, we know that

$$\int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\theta(n), n) d\theta(n) > \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\overleftarrow{\theta^x(n)}, n) d\theta(n). \tag{A22}$$

Thus, to conclude the proof, we need to show that

$$\int_{\theta(\underline{n}^x(\overleftarrow{\theta}_2))}^{\theta_2} v(\theta, n(\theta)) d\theta > \int_{\theta_1}^{\theta^x(n(\theta_1))} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta})) d\theta + \int_{\theta^x(n(\theta_1))}^{\theta_2} v(\overleftarrow{\theta}, n^x(\overleftarrow{\theta}))(1 - A(\theta)) d\theta$$

$$+ \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} v(\overleftarrow{\theta^x(n)}, n) A(\theta^x(n)) d(\theta^x(n) - \theta(n)). \tag{A23}$$

First, note that, taken together, the integrals in the RHS of (A23) have the same length (measured with respect to θ) as the integral in the LHS of (A23). (That is, if we replace $v(\cdot, \cdot)$ with 1 in all the four integrals, (A23) holds with equality.) This is because the length of the LHS of (A23) is the length of the LHS of (A17), which is $(\theta_2 - \theta_1) - \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} 1 \, d\theta(n)$ (by (A21)); similarly, the length of the RHS of (A23) is the length of the RHS of (A17), which is $(\theta_2 - \theta_1) - \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} 1 \, A(\theta^x(n) d\theta(n) = \int_{n(\theta_1)}^{\underline{n}^x(\overleftarrow{\theta}_2)} 1 \, d\theta(n)$ (by (A19) and (A20)). Second, note that the weights on values of $v(\theta, n)$ in the RHS of (A23) are always positive since, by Lemma A2, $\theta^x(n) - \theta(n)$ is a weakly decreasing function (note that whenever $A(\theta^x(n)) = 1$, $\theta^x(n)$ is simply the inverse function of $n^x(\overleftarrow{\theta})$).

By Claim 1, and since $v(\theta, n)$ is increasing in θ, the differences in θ push the RHS of (A23) to be lower than the LHS. As for the differences in n, the RHS of (A23) sums values of $v(\theta, n)$ with θ in d^A and n above $\underline{n}^x(\overleftarrow{\theta}_2)$, while the LHS of (A23) sums values of $v(\theta, n)$ with θ in d^B and n below $\underline{n}^x(\overleftarrow{\theta}_2)$. Consider two cases. (1) If $v(\theta(\underline{n}^x(\overleftarrow{\theta}_2)), \underline{n}^x(\overleftarrow{\theta}_2)) < 0$, then, since v satisfies single crossing, all the

values of v in the RHS of (A17) are negative. But then, since the LHS of (A17) is nonnegative (Claim 3), (A17) must hold. (2) If $v(\theta(\underline{n}^x(\tilde{\theta}_2)), \underline{n}^x(\tilde{\theta}_2)) \geq 0$, then, since v decreases in n whenever it is nonnegative (see Figure 2), all values of v in the LHS of (A23) are greater than those in the RHS. Then, since the integrals on both sides have the same length, (A23) must hold, and thus (A17) holds. Q.E.D.

Proof of Lemma A1: (i) Continuity in θ_i holds because a change in θ_i only changes the limits of integration $[\theta_i - \varepsilon, \theta_i + \varepsilon]$ in the computation of Δ^{r_1} and because the integrand is bounded. (ii) The continuity with respect to a holds because v is bounded, and Δ^{r_1} is an integral over a segment of θ's. The function $\Delta^{r_1}(\theta_i + a, (\tilde{n} + a)(\cdot))$ is nondecreasing in a because as a increases, the agent sees the same distribution of n, but expects θ to be higher. More precisely, the only difference between the integrals that are used to compute $\Delta^{r_1}(\theta_i + a, (\tilde{n} + a)(\cdot))$ and $\Delta^{r_1}(\theta_i + a', (\tilde{n} + a')(\cdot))$ is that in the first we use $v(\theta + a, n)$, while in the second we use $v(\theta + a', n)$. Since $v(\theta, n)$ is nondecreasing in θ, $\Delta^{r_1}(\theta_i + a, (\tilde{n} + a)(\cdot))$ is nondecreasing in a. (iii) If over the limits of integration there is a positive probability that $n < 1/r_1$ and $\theta < \bar{\theta}$, so that $v(\theta, n)$ strictly increases in θ, $\Delta^{r_1}(\theta_i + a, (\tilde{n} + a)(\cdot))$ strictly increases in a. Q.E.D.

Proof of Lemma A2: Consider some $\mu > 0$, and consider the transformation

$$\hat{\varepsilon}_i = \begin{cases} \varepsilon_i + \mu & \text{if } \varepsilon_i \leq \varepsilon - \mu \\ \varepsilon_i + \mu - 2\varepsilon & \text{if } \varepsilon_i > \varepsilon - \mu \end{cases}. \tag{A24}$$

In defining $\hat{\varepsilon}_i$, we basically add μ to ε_i modulo the segment $[-\varepsilon, \varepsilon]$. To write the CDF of $\tilde{n}(\cdot)$ at any level of the fundamentals (see (A1)), we can use either the variable ε_i or the variable $\hat{\varepsilon}_i$. In particular, we can write the CDF at θ as follows:

$$F_\theta(n) = \text{prob}\left[\lambda + (1-\lambda)\cdot \int_{i=0}^{1} s_i(\theta + \hat{\varepsilon}_i)\, di \leq n\right], \tag{A25}$$

and the CDF at $\theta + \mu$ as

$$F_{\theta+\mu}(n) = \text{prob}\left[\lambda + (1-\lambda)\cdot \int_{i=0}^{1} s_i(\theta + \mu + \varepsilon_i)\, di \leq n\right]. \tag{A26}$$

Now, following (A24), we know that with probability $1 - \mu/2\varepsilon$, $\mu + \varepsilon_i = \hat{\varepsilon}_i$. Thus, since we have a continuum of patient players, we know that for exactly $1 - \mu/2\varepsilon$ of them, $s_i(\theta + \mu + \varepsilon_i) = s_i(\theta + \hat{\varepsilon}_i)$. For the other $\mu/2\varepsilon$ patient players, $s_i(\theta + \mu + \varepsilon_i)$ can be anything. Thus, with probability 1,

$$\left(1-\frac{\mu}{2\varepsilon}\right)\cdot\int_{i=0}^{1}s_i(\theta+\hat{\varepsilon}_i)\,di+\frac{\mu}{2\varepsilon}\cdot 0\le\int_{i=0}^{1}s_i(\theta+\mu+\varepsilon_i)\,di. \tag{A27}$$

This implies that

$$\begin{aligned}&\text{prob}\left[\lambda+(1-\lambda)\cdot\left(1-\frac{\mu}{2\varepsilon}\right)\cdot\int_{i=0}^{1}s_i(\theta+\hat{\varepsilon}_i)\,di\le n\right]\\&\quad\ge\text{prob}\left[\lambda+(1-\lambda)\cdot\int_{i=0}^{1}s_i(\theta+\mu+\varepsilon_i)\,di\le n\right],\end{aligned} \tag{A28}$$

which means that

$$\begin{aligned}&\text{prob}\left[\lambda+(1-\lambda)\cdot\int_{i=0}^{1}s_i(\theta+\hat{\varepsilon}_i)\,di\le n+(1-\lambda)\cdot\frac{\mu}{2\varepsilon}\right]\\&\quad\ge\text{prob}\left[\lambda+(1-\lambda)\cdot\int_{i=0}^{1}s_i(\theta+\mu+\varepsilon_i)\,di\le n\right].\end{aligned} \tag{A29}$$

Using (A25) and (A26), we get

$$F_\theta\left(n+(1-\lambda)\frac{\mu}{2\varepsilon}\right)\ge F_{\theta+\mu}(n). \tag{A30}$$

Let $x\in[0,1]$. Then (A30) must hold for $n=n^x(\theta+\mu)$:

$$F_\theta\left(n^x(\theta+\mu)+(1-\lambda)\frac{\mu}{2\varepsilon}\right)\ge F_{\theta+\mu}(n^x(\theta+\mu))\ge x. \tag{A31}$$

This implies that

$$n^x(\theta+\mu)+(1-\lambda)\frac{\mu}{2\varepsilon}\ge n^x(\theta). \tag{A32}$$

This yields

$$\frac{n^x(\theta+\mu)-n^x(\theta)}{\mu}\ge-\frac{1-\lambda}{2\varepsilon},\quad\text{or}\quad\frac{\partial n^x(\theta)}{\partial\theta}\ge-\frac{1-\lambda}{2\varepsilon}. \tag{A33}$$

Repeating the same exercise after defining a variable $\hat{\varepsilon}_i$ by subtracting μ from ε_i modulo the segment $[-\varepsilon,\varepsilon]$, we obtain

$$\frac{n^x(\theta-\mu)-n^x(\theta)}{\mu}\ge-\frac{1-\lambda}{2\varepsilon},\quad\text{or}\quad\frac{\partial n^x(\theta)}{\partial\theta}\le\frac{1-\lambda}{2\varepsilon}. \tag{A34}$$

Since $\frac{\partial\overleftarrow{\theta}}{\partial\theta}=-1$, (A33) and (A34) imply that $\left|\frac{\partial n^x(\overleftarrow{\theta})}{\partial\theta}\right|\le\frac{1-\lambda}{2\varepsilon}$. Q.E.D.

Proof of Theorem 2: The equation that determines $\theta^*(r_1)$ (away from the limit) is

$$f(\theta^*, r_1) = \int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))u\left(\frac{1-r_1 n}{1-n}R\right) - u(r_1)\right] - \int_{n=1/r_1}^{1} \frac{1}{nr_1}u(r_1) = 0, \tag{A35}$$

where $\theta(\theta^*, n) = \theta^* + \varepsilon(1 - 2\frac{n-\lambda}{1-\lambda})$ is the inverse of $n(\theta, \theta^*)$. We can rewrite (A35) as

$$\hat{f}(\theta^*, r_1) = r_1 \int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))u\left(\frac{1-r_1 n}{1-n}R\right)\right] - u(r_1)\cdot\left[1 - \lambda r_1 + \ln(r_1)\right] = 0. \tag{A36}$$

We can see immediately that $\frac{\partial \hat{f}(\theta^*, r_1)}{\partial \theta^*} > 0$. Thus, by the implicit function theorem, in order to prove that $\frac{\partial \theta^*}{\partial r_1} > 0$, it suffices to show that $\frac{\partial \hat{f}(\theta^*, r_1)}{\partial r_1} < 0$. This derivative is given by

$$\begin{aligned}\frac{\partial \hat{f}(\theta^*, r_1)}{\partial r_1} = &\int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))u\left(\frac{1-r_1 n}{1-n}R\right)\right] \\ &+ r_1 \int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))\left(\partial u\left(\frac{1-r_1 n}{1-n}R\right)\Big/\partial r_1\right)\right] \\ &- u'(r_1)\cdot\left[1 - \lambda r_1 + \ln(r_1)\right] - (u(r_1)/r_1)\cdot\left[1 - \lambda r_1\right].\end{aligned} \tag{A37}$$

The last two terms are negative. Thus, it suffices to show that the sum of the first two is negative. One can verify that $\partial u(\frac{1-r_1 n}{1-n}R)/\partial r_1 = \frac{n(1-n)}{r_1 - 1}(\partial u(\frac{1-r_1 n}{1-n}R)/\partial n)$. Thus, the sum equals

$$\begin{aligned}&\int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))u\left(\frac{1-r_1 n}{1-n}R\right)\right] \\ &+ r_1 \int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))\cdot\frac{n(1-n)}{r_1 - 1}\left(\partial u\left(\frac{1-r_1 n}{1-n}R\right)\Big/\partial n\right)\right].\end{aligned} \tag{A38}$$

Using integration by parts, this equals

$$\begin{aligned}&\int_{n=\lambda}^{1/r_1} \left[p(\theta(\theta^*, n))u\left(\frac{1-r_1 n}{1-n}R\right)\right] - r_1 p(\theta(\theta^*, \lambda))\frac{\lambda(1-\lambda)}{r_1 - 1}u\left(\frac{1-r_1\lambda}{1-\lambda}R\right) \\ &- \frac{r_1}{r_1 - 1}\int_{n=\lambda}^{1/r_1} \left[p'(\theta(\theta^*, n))\cdot\frac{-2\varepsilon}{1-\lambda}n(1-n) + p(\theta(\theta^*, n))(1-2n)\right]u\left(\frac{1-r_1 n}{1-n}R\right)\end{aligned} \tag{A39}$$

$$= -r_1 p(\theta(\theta^*, \lambda)) \frac{\lambda(1-\lambda)}{r_1 - 1} u\left(\frac{1-r_1\lambda}{1-\lambda} R\right) - \int_{n=\lambda}^{1/r_1} \frac{1-2nr_1}{r_1-1} p(\theta(\theta^*, n)) u\left(\frac{1-r_1 n}{1-n} R\right)$$
$$+ \frac{r_1}{r_1 - 1} \int_{n=\lambda}^{1/r_1} p'(\theta(\theta^*, n)) \cdot \frac{2\varepsilon}{1-\lambda} n(1-n) u\left(\frac{1-r_1 n}{1-n} R\right) \tag{A40}$$

$$= -r_1 p(\theta(\theta^*, \lambda)) \frac{\lambda(1-\lambda)}{r_1 - 1} u\left(\frac{1-r_1\lambda}{1-\lambda} R\right)$$
$$- p(\theta(\theta^*, \lambda)) \int_{n=\lambda}^{1/r_1} \frac{(1-2nr_1)}{r_1-1} u\left(\frac{1-r_1 n}{1-n} R\right)$$
$$+ 2\varepsilon \int_{n=\lambda}^{1/r_1} \frac{r_1 p'(\theta(\theta^*, n)) n(1-n) + (1-2nr_1) \int_{x=\lambda}^{n} (p'(\theta(\theta^*, x)))}{(r_1 - 1)(1-\lambda)} u\left(\frac{1-r_1 n}{1-n} R\right). \tag{A41}$$

We now derive an upper bound for the sum of the first two terms in (A41). Since $u(\frac{1-r_1 n}{1-n} R)$ is decreasing in n, this sum is less than

$$- r_1 p(\theta(\theta^*, \lambda)) \frac{\lambda(1-\lambda)}{r_1 - 1} u\left(\frac{1-r_1\lambda}{1-\lambda} R\right)$$
$$- p(\theta(\theta^*, \lambda)) \int_{n=\lambda}^{1/r_1} \frac{\left(1 - 2\frac{\lambda + 1/r_1}{2} r_1\right)}{r_1 - 1} u\left(\frac{1-r_1 n}{1-n} R\right) \tag{A42}$$

$$= -r_1 p(\theta(\theta^*, \lambda)) \frac{\lambda(1-\lambda)}{r_1 - 1} u\left(\frac{1-r_1\lambda}{1-\lambda} R\right) - p(\theta(\theta^*, \lambda)) \int_{n=\lambda}^{1/r_1} \frac{-\lambda r_1}{r_1 - 1} u\left(\frac{1-r_1 n}{1-n} R\right) \tag{A43}$$

$$< -r_1 p(\theta(\theta^*, \lambda)) \frac{\lambda(1-\lambda)}{r_1 - 1} u\left(\frac{1-r_1\lambda}{1-\lambda} R\right)$$
$$- p(\theta(\theta^*, \lambda)) \frac{-\lambda r_1}{r_1 - 1} \left(\frac{1}{r_1} - \lambda\right) u\left(\frac{1-r_1\lambda}{1-\lambda} R\right) \tag{A44}$$

$$= -\lambda p(\theta(\theta^*, \lambda)) u\left(\frac{1-r_1\lambda}{1-\lambda} R\right). \tag{A45}$$

Thus, (A41) is smaller than

$$-\lambda p(\theta(\theta^*, \lambda)) u\left(\frac{1-r_1\lambda}{1-\lambda} R\right)$$
$$+ 2\varepsilon \int_{n=\lambda}^{1/r_1} \frac{r_1 p'(\theta(\theta^*, n)) n(1-n) + (1-2nr_1) \int_{x=\lambda}^{n} (p'(\theta(\theta^*, x)))}{(r_1 - 1)(1-\lambda)} u\left(\frac{1-r_1 n}{1-n} R\right). \tag{A46}$$

The first term is negative. The second can be positive, but if it is, it is small when ε is small. Thus, there exists $\underline{\varepsilon} > 0$, such that for each $\varepsilon < \underline{\varepsilon}$, the claim of the theorem holds. Q.E.D.

Proof of Theorem 3: The expected utility of a representative agent, $\mathrm{EU}(r_1)$, (away from the limit) is

$$\begin{aligned}\mathrm{EU}(r_1) = &\int_0^{\tilde{\theta}(r_1,\theta^*(r_1))} \frac{1}{r_1} u(r_1)\, d\theta + \int_{\tilde{\theta}(r_1,\theta^*(r_1))}^{\theta^*(r_1)+\varepsilon} n\big(\theta,\theta^*(r_1)\big)\cdot u(r_1)\\ &+ \big(1 - n\big(\theta,\theta^*(r_1)\big)\big)\cdot p(\theta)\cdot u\left(\frac{1 - n\big(\theta,\theta^*(r_1)\big)r_1}{1 - n\big(\theta,\theta^*(r_1)\big)}R\right) d\theta\\ &+ \int_{\theta^*(r_1)+\varepsilon}^{\bar{\theta}} \lambda\cdot u(r_1) + (1-\lambda)\cdot p(\theta)\cdot u\left(\frac{1-\lambda r_1}{1-\lambda}R\right) d\theta\\ &+ \int_{\bar{\theta}}^{1} \lambda\cdot u(r_1) + (1-\lambda)\cdot u\left(\frac{R-\lambda r_1}{1-\lambda}\right) d\theta, \end{aligned} \tag{A47}$$

where $\tilde{\theta}(r_1,\theta^*(r_1)) \equiv \theta^*(r_1) + \varepsilon\left(1 - 2\frac{1-\lambda r_1}{(1-\lambda)r_1}\right)$ is the level of fundamentals, at which $n(\theta,\theta^*) = 1/r_1$, that is, below which the bank cannot pay all agents who run. Thus

$$\begin{aligned}\frac{\partial \mathrm{EU}(r_1)}{\partial r_1} = &\int_0^{\tilde{\theta}(r_1,\theta^*(r_1))} \frac{r_1 u'(r_1) - u(r_1)}{(r_1)^2} d\theta\\ &+ \int_{\tilde{\theta}(r_1,\theta^*(r_1))}^{\theta^*(r_1)+\varepsilon} n\big(\theta,\theta^*(r_1)\big)\\ &\cdot \left(u'(r_1) - R\cdot p(\theta)\cdot u'\left(\frac{1-n\big(\theta,\theta^*(r_1)\big)r_1}{1-n\big(\theta,\theta^*(r_1)\big)}R\right)\right) d\theta\\ &+ \frac{1-\lambda}{2\varepsilon}\cdot\frac{\partial\theta^*(r_1)}{\partial r_1}\int_{\tilde{\theta}(r_1,\theta^*(r_1))}^{\theta^*(r_1)+\varepsilon}\left(u(r_1) - p(\theta)\cdot u\left(\frac{1-n\big(\theta,\theta^*(r_1)\big)r_1}{1-n\big(\theta,\theta^*(r_1)\big)}R\right)\right.\\ &\left. - \frac{(r_1-1)\cdot R\cdot p(\theta)}{1-n\big(\theta,\theta^*(r_1)\big)}u'\left(\frac{1-n\big(\theta,\theta^*(r_1)\big)r_1}{1-n\big(\theta,\theta^*(r_1)\big)}R\right)\right) d\theta\\ &+ \lambda\cdot\int_{\theta^*(r_1)+\varepsilon}^{\bar{\theta}}\left(u'(r_1) - R\cdot p(\theta)\cdot u'\left(\frac{1-\lambda r_1}{1-\lambda}R\right)\right) d\theta\\ &+ \lambda\cdot\int_{\bar{\theta}}^{1}\left(u'(r_1) - u'\left(\frac{R-\lambda r_1}{1-\lambda}\right)\right) d\theta. \end{aligned} \tag{A48}$$

When $r_1 = 1$, (A48) becomes

$$\int_0^{\theta^*(1)-\varepsilon} (u'(1) - u(1))d\theta + \int_{\theta^*(1)-\varepsilon}^{\theta^*(1)+\varepsilon} n(\theta, \theta^*(1)) \cdot (u'(1) - R \cdot p(\theta) \cdot u'(R))d\theta$$
$$+ \frac{1-\lambda}{2\varepsilon} \cdot \frac{\partial\theta^*(1)}{\partial r_1} \int_{\theta^*(1)-\varepsilon}^{\theta^*(1)+\varepsilon} (u(1) - p(\theta) \cdot u(R))d\theta + \lambda \cdot$$
$$\int_{\theta^*(1)+\varepsilon}^{\bar{\theta}} (u'(1) - R \cdot p(\theta) \cdot u'(R))d\theta + \lambda \cdot \int_{\bar{\theta}}^{1} \left(u'(1) - u'\left(\frac{R-\lambda}{1-\lambda}\right)\right) d\theta. \quad \text{(A49)}$$

Here, the second term and the fourth term are positive because $u'(1) > Ru'(R)$ for all $R > 1$, and because $p(\theta) < 1$. The fifth term is positive because $\frac{R-\lambda}{1-\lambda} > 1$, and because $u''(c)$ is negative. The third term is zero by the definition of $\theta^*(1)$. The first term is negative. This term, however, is small for a sufficiently small $\theta^*(1)$. As ε goes to 0, $\theta^*(1)$ converges to $\underline{\theta}(1)$. Thus, for a sufficiently small $\underline{\theta}(1)$, the claim of the theorem holds.

As ε goes to 0, the condition becomes

$$-\underline{\theta}(1)(u(1) - u'(1)) + \lambda \cdot \int_{\underline{\theta}(1)}^{1} (u'(1) - R \cdot p(\theta) \cdot u'(R))d\theta > 0. \quad \text{(A50)}$$

This is equivalent to an upper bound on $\underline{\theta}(1)$. Q.E.D.

Proof of Theorem 4: At the limit, as ε goes to 0 and $\bar{\theta}$ goes to 1 $\partial \mathrm{EU}(r_1)/\partial r_1$ becomes (see equation (6))

$$\int_0^{\theta^*(r_1)} \frac{r_1 u'(r_1) - u(r_1)}{(r_1)^2} d\theta - \frac{\partial\theta^*(r_1)}{\partial r_1}$$
$$\times \left(\lambda u(r_1) + (1-\lambda)p(\theta^*(r_1)) \cdot u\left(\frac{1-\lambda r_1}{1-\lambda}R\right) - \frac{1}{r_1}u(r_1)\right)$$
$$+ \lambda \cdot \int_{\theta^*(r_1)}^{1} \left(u'(r_1) - R \cdot p(\theta) \cdot u'\left(\frac{1-\lambda r_1}{1-\lambda}R\right)\right) d\theta. \quad \text{(A51)}$$

Thus, the first-order condition that determines the optimal r_1 converges to[18]

$$u'(r_1) - R \cdot u'\left(\frac{1-\lambda r_1}{1-\lambda}R\right) E\left[p(\theta) \mid \theta > \theta^*(r_1)\right]$$
$$= \frac{\frac{\partial\theta^*(r_1)}{\partial r_1}\left(\lambda u(r_1) + (1-\lambda)p\left(\theta^*(r_1)\right) \cdot u\left(\frac{1-\lambda r_1}{1-\lambda}R\right) - \frac{1}{r_1}u(r_1)\right) + \int_0^{\theta^*(r_1)} \frac{u(r_1) - r_1 u'(r_1)}{(r_1)^2} d\theta}{\lambda\left(1 - \theta^*(r_1)\right)}. \quad \text{(A52)}$$

Since the RHS of the equation is positive, and since $E[p(\theta) \mid \theta > \theta^*(r_1)] > E_\theta[p(\theta)]$, the optimal r_1 is lower than c_1^{FB}, which is determined by equation (1). Q.E.D.

[18] We can employ the limit first order condition since $\partial \mathrm{EU}(r_1)/\partial r_1$ is continuous in ε at 0 and in $\bar{\theta}$ at 1, and since the limit of solutions of first order conditions near the limit must be a solution of the first order condition at the limit.

Appendix B: Discussion on the Assumption of an Upper Dominance Region

A crucial condition that leads to a unique equilibrium in our model is the existence of an upper dominance region (implied by the assumption on the technology or by the alternative of an external lender). Indeed, the range of fundamentals in which waiting is a dominant action can be taken to be arbitrarily small, thereby representing extreme situations where fundamentals are extraordinarily good. Thus, this assumption is rather weak. Nevertheless, we now explore the model's predictions when this assumption is not made.

First, we note that without an upper dominance region, our model has multiple equilibria. Two are easy to point out: One is the trivial, bad equilibrium, in which agents run for any signal. The other is the threshold equilibrium that is the unique equilibrium of our model. In addition, we cannot preclude the existence of other equilibria in which agents run at all signals below θ^*, and above they sometimes run and sometimes wait.

Importantly, the unique equilibrium characterized in Theorem 1 is the only one that survives three different equilibrium selection criteria (refinements). Thus, if we adopt any of these selection criteria, we can still analyze the model without the assumption of an upper dominance region, and obtain the same results. We now list these refinements:

EQUILIBRIUM SELECTION CRITERION A: *The agents coordinate on the best equilibrium.*

By "best" equilibrium we mean the equilibrium that Pareto-dominates all others. In our model, this is also the one at which the set of signals at which agents run is the smallest.[19]

EQUILIBRIUM SELECTION CRITERION B: *The agents coordinate on an equilibrium in which when they observe signals that are extremely high* ($\theta_i \in [1-\varepsilon, 1]$), *they do not run.*

The idea is that a panic-based run does not happen when agents know that the fundamentals are excessively good. As with the assumption of the upper dominance region, this is sufficient in order to rule out runs in a much larger range of parameters.

EQUILIBRIUM SELECTION CRITERION C: *The equilibrium on which patient agents coordinate has monotonic strategies and is nontrivial (i.e., agents' actions depend on their signals).*

[19] An equilibrium with this property exists (and "smallest" is well defined) if ε is small enough, and is the same as the equilibrium characterized in Theorem 1. The reason is that we can show by iterative dominance that patient agents must run below θ^*. Thus, an equilibrium, which has the property that patient agents run below θ^* and never run above, is the one with the smallest set of signals at which agents run. (Note also that since patient agents are small and identical, they all behave in the same manner.)

Note that while some other papers in the literature use equilibrium selection criteria (the "best equilibrium" criterion), they always select equilibria with no panic-based runs. That is, in their best equilibrium, runs occur only when early withdrawal is the agents' dominant action, that is, when $\theta < \underline{\theta}(r_1)$ (see, e.g., Goldfajn and Valdes (1997), and Allen and Gale (1998)). In our model, by contrast, the equilibrium does have panic-based bank runs (even if it is not a unique equilibrium): Agents run whenever $\theta < \theta^*(r_1)$. Therefore, we can analyze the interdependence between the banking contract and the probability that agents will lose their confidence in the solvency of the bank. This is the main novelty of our paper, and it is thus maintained even if the model had multiple equilibria (which would be the case had the assumption of an upper dominance region been dropped).

REFERENCES

Allen, Franklin, and Douglas Gale, 1998, Optimal financial crises, *Journal of Finance* 53, 1245–1284.

Alonso, Irasema, 1996, On avoiding bank runs, *Journal of Monetary Economics* 37, 73–87.

Bernardo, Antonio E., and Ivo Welch, 2004, Liquidity and financial market runs, *Quarterly Journal of Economics* 119, 135–158.

Bougheas, Spiros, 1999, Contagious bank runs, *International Review of Economics and Finance* 8, 131–146.

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

Carlsson, Hans, and Eric van Damme, 1993, Global games and equilibrium selection, *Econometrica* 61, 989–1018.

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

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

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

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

Demirguc-Kunt, Asli, Enrica Detragiache, and Poonam Gupta, 2000, Inside the crisis: An empirical analysis of banking systems in distress, World Bank Working paper 2431.

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

Goldfajn, Ilan, and Rodrigo O. Valdes, 1997, Capital flows and the twin crises: The role of liquidity, IMF Working paper 97-87,

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

Gorton, Gary, and Andrew Winton, 2003, Financial intermediation, in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds.: *Handbook of the Economics of Finance* (North Holland, Amsterdam).

Green, Edward J., and Ping Lin, 2003, Implementing efficient allocations in a model of financial intermediation, *Journal of Economic Theory* 109, 1–23.

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

Judd, Kenneth L., 1985, The law of large numbers with a continuum of i.i.d. random variables, *Journal of Economic Theory* 35, 19–25.

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

Kindleberger, Charles P., 1978, *Manias, Panics, and Crashes: A History of Financial Crises* (Basic Books, New York).

Krugman, Paul R., 2000, Balance sheets, the transfer problem, and financial crises, in Peter Isard, Assaf Razin, and Andrew K. Rose, eds.: *International Finance and Financial Crises* (Kluwer Academic Publishers, Boston, MA).

Lindgren, Carl-Johan, Gillian Garcia, and Matthew I. Saal, 1996, *Bank Soundness and Macroeconomic Policy* (International Monetary Fund, Washington, DC).

Loewy, Michael B., 1998, Information-based bank runs in a monetary economy, *Journal of Macroeconomics* 20, 681–702.

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

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

Morris, Stephen, and Hyun S. Shin, 2003a, Global games: Theory and applications, in Mathias Dewatripont, Lars P. Hansen, and Stephen J. Turnovsky, eds.: *Advances in Economics and Econometrics* (Cambridge University Press, Cambridge).

Morris, Stephen, and Hyun S. Shin, 2003b, Liquidity black holes, mimeo, Yale University.

Peck, James, and Karl Shell, 2003, Equilibrium bank runs, *Journal of Political Economy* 111, 103–123.

Postlewaite, Andrew, and Xavier Vives, 1987, Bank runs as an equilibrium phenomenon, *Journal of Political Economy* 95, 485–491.

Radelet, Steven, and Jeffrey D. Sachs, 1998, The East Asian financial crisis: Diagnosis, remedies, prospects, *Brookings Papers on Economic Activity* 1, 1–74.

Rochet, Jean-Charles, and Xavier Vives, 2003, Coordination failures and the lender of last resort: Was Bagehot right after all, mimeo, INSEAD.

Temzelides, Theodosios, 1997, Evolution, coordination, and banking panics, *Journal of Monetary Economics* 40, 163–183.

Wallace, Neil, 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* 12, 3–16.

Wallace, Neil, 1990, A banking model in which partial suspension is best, *Federal Reserve Bank of Minneapolis Quarterly Review* 14, 11–23.