Optimal Banking Contracts and Financial Fragility

Huberto M. Ennis
Research Department
Federal Reserve Bank of Richmond
Huberto.Ennis@rich.frb.org

Todd Keister
Research and Statistics Group
Federal Reserve Bank of New York
Todd.Keister@ny.frb.org

October 17, 2011

Abstract

We study a finite-depositor version of the Diamond-Dybvig model of financial intermediation in which the bank and all depositors observe withdrawals as they occur. We derive the (constrained) efficient allocation of resources in closed-form and show that this allocation provides liquidity insurance to depositors. The contractual arrangement that decentralizes this allocation has debt-like features and resembles the type of demand deposits commonly offered by banking institutions. We provide examples where this arrangement admits another equilibrium in which some depositors run on the bank, withdrawing funds regardless of their liquidity needs. A bank run in our setting is always partial, with only those depositors who can withdraw sufficiently early participating. Depositors who are late to withdraw during a run suffer significant discounts from the face value of their deposits. The run, while partial, may involve a large number of depositors and result in significant inefficiencies.
JEL Classification Numbers: G21,G01, D82

We thank seminar participants at the University of Iowa, the Federal Reserve Bank of Richmond, and 2010 Winter Meetings of the Econometric Society for useful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve Bank of Richmond, or the Federal Reserve System.

1 Introduction

Banks and other financial intermediaries engage in maturity transformation by issuing short-term liabilities while investing in assets whose full return is realized over a longer time horizon. In fact, much of an intermediary's liabilities are usually payable on demand, meaning that a holder is entitled to request immediate repayment. These liabilities are also debt-like in the sense that the value of a claim is typically not contingent on the precise timing of withdrawal; during normal times, all holders receive the "face value" of their claim. Such intermediation arrangements appear to be fragile in the sense of being susceptible to runs – events in which liability holders rush *en masse* to redeem their claims. During these episodes, resources are allocated inefficiently and claims that are repaid late in the process are often subjected to considerable discounts. In this paper, we provide a model of financial intermediation that accounts for all of these features simultaneously. We follow the tradition of Wallace (1988) and Green and Lin (2003) in that we place no restrictions on financial arrangements or the allocation of resources other than those explicitly specified as part of the environment.

As events during the recent financial crisis have demonstrated, arrangements based on maturity transformation and debt-like liabilities are not limited to retail banking; they are widespread and integral parts of the modern financial system. A number of such arrangements experienced events resembling a run during the crisis, including wholesale deposits, brokerage accounts, repurchase agreements and money market mutual funds, to name a few.[1] Regulatory reform efforts in the wake of the crisis have focused on limiting maturity transformation by requiring financial institutions to more closely align the maturity structure of their assets with that of their liabilities and to limit the debt-like features of some liabilities. Evaluating the desirability of such reforms requires having a theory of financial arrangements based on maturity transformation that explains why these arrangements arise and what social benefits they offer. Only with such a theory in hand can policy makers properly assess the costs and benefits of different proposals for reform.

Diamond and Dybvig (1983) uncovered some important elements of an economic environment in which financial intermediaries could serve a useful role as providers of maturity transformation while at the same time being subject to financial fragility in the form of runs. In their theory, high-return investment takes time to mature and agents have random needs for immediate liquidity. In this setting, a risk-sharing arrangement resembling a bank can improve the allocation of

[1] For descriptions of some of these events, see Baba *et al*. (2009), Duffie (2010), Gorton and Metrick (2011), and McCabe (2010).

resources. The fact that a depositor's liquidity needs are private information requires the bank to allow depositors to choose when to withdraw their funds. This ability to choose, in turn, creates the potential for inefficient withdrawals and fragility. However, private information alone is not enough to explain the fragility of financial arrangements as an equilibrium outcome. Diamond and Dybvig hint at two other components that would be needed for their theory to successfully explain fragility: (i) some form of first-come, first-served (or *sequential service*) constraint and (ii) a degree of aggregate uncertainty about the total demand for liquidity. While the role of these additional elements was discussed informally in Diamond and Dybvig's seminal work, the precise details of the arguments were left mostly unexplored.

More recently, a literature has developed that attempts to explicitly model these elements of the Diamond-Dybvig theory and investigate the extent to which fragility is an inherent feature of banking and other financial arrangements.[2] The approach in this literature is to fully specify the physical environment and to derive the predicted outcomes by solving an allocation problem with no further restrictions on what can be attained. An important lesson from this literature is that fragility obtains only if the payoffs associated with an institution's short-term liabilities are sufficiently insensitive to the aggregate demand for liquidity. In other words, fragility arises when an intermediary's liabilities are sufficiently debt-like, with many depositors receiving the face value of their claims even when aggregate liquidity demand is high. The insensitivity of payoffs in these models is a direct consequence of sequential service. A bank learns about the level of aggregate liquidity demand by observing individual depositors' actions. In the absence of sequential service, the bank would fully observe the aggregate level of liquidity demand before redeeming any liabilities and would set payments to liability holders accordingly. Under standard assumptions on preferences, the resulting allocation would give depositors no incentive to run. When depositors act sequentially and must be served as they arrive, however, liabilities must be redeemed before the bank is able to fully observe the level of liquidity demand and, hence, payoffs are necessarily coarser with respect to this information.

The degree to which sequential service limits the flow of information about aggregate liquidity demand to the bank depends on the specific details of the environment. Alternative formulations of this process have been proposed in the literature. Green and Lin (2003) present one reasonable formulation and show that the optimal banking arrangement in their setting is *not* fragile. In

[2] See the survey in Ennis and Keister (2010b).

their model, all depositors report to the bank sequentially and either withdraw their deposit or communicate that they do not currently wish to withdraw. The bank conditions the payment it gives to a withdrawing depositor on all previous reports: subsequent withdrawals are adjusted downward if a depositor makes an early withdrawal and upward if she instead reports that she will not withdraw early. In this way, the payments made to depositors may be highly sensitive to the pattern of withdrawal decisions. Green and Lin show that this sensitivity can be sufficient to rule out bank runs as an equilibrium outcome.

Peck and Shell (2003) highlight that the amount of information available to depositors is also relevant for the incidence of fragility. They assume that depositors do not have any information about their place in the order of withdrawals when they decide whether to withdraw early. They also assume that only those depositors intending to make a withdrawal report to the bank. Under this alternative specification, they show that bank runs are possible regardless of whether all depositors report to the bank or only those who wish to withdraw.[3]

In this paper, we propose a new specification of the sequential service constraint that shares some elements with these earlier approaches, but modifies what we regard as extreme, and perhaps unrealistic, aspects of their specifications. In contrast Green and Lin (2003), and in line with Peck and Shell (2003), we assume that depositors only report to the bank when they wish to withdraw. We depart from the previous literature by assuming that each depositor can observe how many withdrawals have already occurred when deciding whether or not to withdraw herself. Observing the number of previous withdrawals provides some information about the depositor's place in the order, but generally does not reveal the exact position. In this sense, our sequential service constraint can be viewed as an intermediate case between the Green-Lin and Peck-Shell formulations. Apart from the specification of the sequential service process, our environment is essentially identical to that in Green and Lin (2003).

We show that this new specification of the sequential service constraint induces some properties in the banking arrangement that resemble those observed in reality. In particular, the bank's payment schedule is initially almost completely insensitive to the order of withdrawals and, in this sense, displays the debt-like property that is common in demand deposit contracts. Furthermore, in situations when withdrawal demand becomes unusually high, depositors who are late to withdraw

[3] Peck and Shell (2003) use a different specification of preferences from Green and Lin (2003), but Ennis and Keister (2009) show that this difference is not essential for their results.

suffer significant discounts relative to the face value of their deposits, a type of "partial suspension of convertibility" that Wallace (1990) argues was a common feature of historical banking panics.[4] We also show that this version of the sequential service constraint restricts the flow of information sufficiently to make a run on the bank consistent with equilibrium. The key ingredient for this result is that when the bank only observes the actions of depositors who withdraw, it is relatively slow to react to a situation in which aggregate withdrawal demand is high, as occurs when depositors run on the bank. This slow reaction is anticipated by depositors and gives an incentive for those who can withdraw relatively early, before the bank reacts much, to do so regardless of their liquidity needs. Based on these findings, we conclude that several common features of banking, such as the face-value property of demand deposits, sharp discounts during crisis and fragility may share the same fundamental cause: the gradual revelation of information inherent in a withdrawal process that takes place sequentially.

The paper is organized as follows. Section 2 describes the environment, including our specification of sequential service. Section 3 derives the efficient allocation of resources in the absence of private information, which serves as a useful benchmark in the rest of the analysis. Section 4 describes the corresponding banking arrangement and Section 5 demonstrates the possibility of runs in our environment. Section 6 provides some discussion and concludes.

2 The environment

There are two time periods, indexed by $t \in \{0, 1\}$, and a finite number I of depositors. There is a single good that can be consumed in each period. Let $c_i = (c_i^1, c_i^2) \in \mathbb{R}_+^2$ denote the consumption of depositor i in each period and let $c = (c_1, \ldots, c_I)$ denote the complete vector of consumption bundles. A depositor's preferences depend on her type $\omega_i \in \{0, 1\}$. If $\omega_i = 0$, the depositor is *impatient* and only cares about consumption in period 0. If $\omega_i = 1$, the depositor is *patient* and cares about the sum of her consumption in the two periods. Depositor i's utility level is given by

$$u\left(c_i^0 + \omega_i c_i^1\right) = \frac{1}{1-\gamma}\left(c_i^0 + \omega_i c_i^1\right)^{1-\gamma}. \tag{1}$$

As in Diamond and Dybvig (1983), we assume the coefficient of relative risk aversion γ is greater than unity. Each depositor's type ω_i is an independent draw from a Bernoulli distribution, where

[4] For discussions of the specific features of historical suspension schemes, see Friedman and Schwartz (1963, pages 160-8, 328-30), Selgin (1993) and Dwyer and Hasan (2003).

π is the probability of being impatient, and a depositor's type is private information. Let $\omega = (\omega_1, \ldots, \omega_I)$ denote the vector of types for all depositors and let Ω denote the set $\{0, 1\}$, so that we have $\omega_i \in \Omega$ and $\omega \in \Omega^I$.

We use $\theta(\omega)$ to denote the total number of patient depositors in the profile ω, that is

$$\theta(\omega) = \sum_{i=1}^{I} \omega_i.$$

Let $p(\theta)$ denote the probability of the set of profiles ω in which there are exactly θ patient depositors. Since types are independent across depositors, θ is a binomial random variable and we have

$$p(\theta) = C(I, \theta)(1 - \pi)^{\theta} \pi^{I-\theta}, \tag{2}$$

where C is the standard combinatorial function

$$C(I, \theta) = \frac{I!}{\theta!(I - \theta)!}. \tag{3}$$

There is also a bank that has an endowment of I units of the good at date 0 and aims to distribute these resources to maximize depositors' expected utility.[5] A depositor can withdraw from the bank either in period 0 or in period 1, but not both. Depositors are isolated from each other and goods must be consumed immediately after withdrawal in order to give utility. These assumptions imply that no markets exist in which depositors could trade after withdrawing; a depositor simply consumes what she receives from the bank (see Wallace, 1988, on this point). Each unit of the good that is not consumed in period 0 is transformed into $R > 1$ units of the good in period 1.

Depositors' opportunities to withdraw in period 0 arrive sequentially in a random order, with each depositor equally likely to occupy each place in the order. A depositor does not observe her position in this order. When her opportunity to withdraw arrives, she is able to observe how many withdrawals have already been made. She does not, however, observe the decisions of any depositors before her in the order who chose not to withdraw in period 0 and, hence, may still be uncertain about her exact position in the order. Likewise, the bank only observes depositors' actions when they choose to withdraw. When a depositor arrives to withdraw, the bank knows how

[5] It would be straightforward to add an earlier time period to the model in which individuals have private endowments and choose whether to deposit in the bank, as in Peck and Shell (2003) and others. Doing so would not change our results in any way; we bypass this step for simplicity.

many withdrawals it has already processed, but has no information about any depositors who may have already chosen not to withdraw in period 0, nor about the actions that any of the depositors later in the order will take.

When a depositor withdraws, the amount she receives can depend only on the information currently available to the bank. This *sequential service constraint* limits the ability of the bank to make payments contingent on the total demand for early withdrawals and thereby provide risk sharing to depositors. Consider, for example, the first depositor to withdraw in period 0. When she arrives, the bank only knows that at least one depositor is withdrawing; it has no other information that can be used to make inferences about the total number of early withdrawals that will take place. The consumption this depositor receives must, therefore, be the same for all outcomes in which at least one withdrawal occurs; let x_1 denote this amount. Similarly, let x_n denote the amount of consumption received by the n^{th} depositor to withdraw, which must be the same for all outcomes in which at least n withdrawals occur. The restrictions imposed by the sequential service constraint thus imply that the period-0 actions of the bank can be fully described using a *payment schedule* $x = \{x_n\}_{n=1}^{I}$. This schedule must satisfy the feasibility constraints

$$\sum_{n=1}^{I} x_n \leq I \quad \text{and} \quad x_n \geq 0 \text{ for all } n. \tag{4}$$

Notice that the sequence x is a contingent plan; it specifies the period-0 payments the bank will make in all possible scenarios, including the one where all I depositors withdraw early. If fewer than I depositors withdraw in period 0, some of these payments will not be made. After all depositors have made their withdrawal decision in period 0, the bank observes that the period has ended and the economy moves to period 1. At this point, the bank knows how many depositors remain who have not yet withdrawn. Since depositors are risk averse, efficiency requires that the bank divide the matured assets in period 1 evenly among these depositors. As a result, the operation of the bank in our environment is completely summarized by the period-0 payment schedule x.

Our formulation of the sequential service constraint differs from that in Wallace (1988, 1990) and Green and Lin (2003). In those papers, each depositor contacts the bank in period 0 and announces whether or not she wishes to withdraw. In such a setting, the period-0 payment received by a depositor can depend on the entire sequence of actions taken by the depositors before her in the order. We instead follow Peck and Shell (2003) in assuming that the bank only observes the decisions of those depositors who have chosen to withdraw.[6] This formulation slows down the flow of information about the level of total withdrawal demand to the bank and, in so doing, places stronger restrictions on the payment schedule as described above. In the sections that follow, we show how this slower flow of information to the bank has important implications for the form of optimal banking arrangements and for financial fragility.

Our assumption that a depositor is able to observe the number of withdrawals that have already taken place when her opportunity to withdraw arises is new in the literature. In Green and Lin (2003), a depositor observes a signal correlated with her position in the order before making her decision, but does not observe the actions taken by depositors before her in the order. In Peck and Shell (2003), a depositor receives no information about her position in the order or the actions of other depositors. Since the bank necessarily observes the actions of those depositors with whom it interacts, both of these approaches imply that a depositor has less information than the bank about the withdrawal history when making a decision. This asymmetry raises the question of whether the bank might choose to communicate its information to a depositor before she makes her choice. (See Nosal and Wallace, 2009, for an analysis of this issue.) Under our approach, in contrast, depositors and the bank observe exactly the same information about the withdrawal history; there is no scope for the bank to either provide or withhold information about this history to depositors.

3 The efficient allocation

We begin our analysis by deriving the efficient allocation of resources in an environment that is identical to the one described, but where the preference types ω_i can be observed by the bank. This allocation will be a useful benchmark in subsequent sections when we study the equilibria of a game played by depositors with private information.

Given the form of preferences (1) and the return $R > 1$ on investment, efficiency clearly requires that a depositor consume in period 0 if and only if she is impatient. When the bank can observe types ω_i, therefore, it will only permit impatient depositors to withdraw in the early period and its objective function can be written as

$$\max_{x} \sum_{\theta=1}^{I} p(\theta) \left(\sum_{n=1}^{I-\theta} u(x_n) + \theta u \left(\frac{R\phi_x (I-\theta)}{\theta} \right) \right) + p(0) \left(\sum_{n=1}^{I-1} u(x_n) + u(\phi_x (I-1)) \right), \quad (5)$$

[6] Peck and Shell (2003) motivate the approach by saying "[i]t is hard to imagine people visiting their bank for the purpose of telling them that they are not interested in making any transactions at the present time."

where $\phi_x(m)$ is the amount of resources that the bank has remaining after serving m depositors in period 0 under the payment schedule x, that is

$$\phi_x(m) = I - \sum_{n=1}^{m} x_n.$$

The efficient payment schedule is the sequence $\{x_n^*\}$ that maximizes (5) subject to the feasibility constraints in (4). To solve for the efficient schedule, we reformulate (5) as a dynamic programming problem. Define the following conditional probabilities:

$$q_n = \text{Prob}\left[I - \theta(\omega) \geq n \mid I - \theta(\omega) \geq n-1\right].$$

After the bank has encountered $n-1$ impatient depositors in period 0, q_n is the probability that it will meet at least one more. Using the binomial density in (2), this conditional probability can be written as

$$q_n = \frac{\sum_{i=0}^{I-n} p(\theta)}{\sum_{i=0}^{I-n+1} p(\theta)}.$$

We can then derive the efficient payment schedule x_n as a function of these probabilities q_n.

Proposition 1 *The efficient payment schedule sets*

$$x_n^* = \frac{z_{n-1}}{(\phi_n)^{\frac{1}{\gamma}} + 1} \quad \textit{for } n = 1, \ldots, I,$$

where $z_{n-1} = I - \sum_{j<n} x_j^*$ *and the constants* ϕ_n *are defined recursively by* $\phi_I = 0$ *and*

$$\phi_n = q_{n+1}\left({\phi_{n+1}}^{\frac{1}{\gamma}} + 1\right)^{\gamma} + (1 - q_{n+1})(I-n)^{\gamma} R^{1-\gamma}$$

for $n = 1, \ldots, I-1$.

The variable z_{n-1} measures the amount of resources remaining when the bank encounters the n^{th} impatient depositor. The proposition shows how the fraction of the remaining resources this depositor will receive depends on the remaining conditional probabilities q_{n+1}, q_{n+2}, etc., as well as on the parameters R and γ. A proof of the proposition is given in the appendix .

Example. Figure 1 plots the efficient payment schedule when there are 20 depositors and the parameter values are given by $R = 1.1$, $\gamma = 6$, and $\pi = 0.5$. The lower curve in the figure presents, for each value of n, the consumption x_n^* that the n^{th} impatient depositor will receive in period 0. The upper curve in the figure represents the level of consumption that all patient depositors will receive in period 1 if there is a total of $n - 1$ impatient depositors. The fact that this latter curve lies everywhere above the former has the following interpretation. Consider the last depositor in the period-0 decision order, and let the number of impatient depositors before her be given by $n - 1$. If she is impatient she will receive x_n^*, from the lower curve in the figure, while if she is patient she will receive the consumption allocated for patient depositors when there are a total of $n - 1$ impatient depositors, which is the corresponding point on the upper curve. Thus the figure shows that the last depositor in the order always consumes more when she is patient than when she is impatient, regardless of the types of the other depositors. Notice that this feature does not necessarily hold for other depositors. The first depositor in the order, for example, consumes x_1 if she is impatient and can end up receiving any point on the upper curve if she is patient, depending on the total number of impatient depositors. In this example, she will consume more than x_1 when she is patient if the number of impatient depositors turns out to be less than 12, but otherwise will consume less.

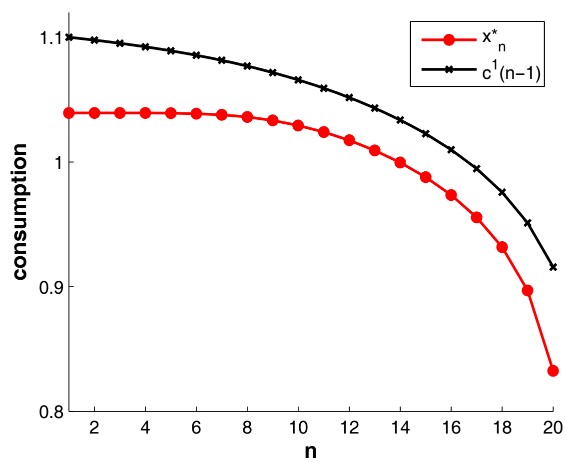

Figure 1: The efficient allocation

Using the solution given in Proposition 1, we can derive some properties of the efficient payment schedule. First, we establish that this schedule offers depositors *liquidity insurance* in the sense that the benefit of the return R on investment is shared by all depositors, even those who consume before this return is realized. In a model with no aggregate uncertainty, Diamond and Dybvig (1983) showed how the efficient level of consumption for all impatient depositors is greater than the per-capita value of the bank's assets in period 0. When there is aggregate uncertainty, as in our model here, the definition of liquidity insurance must be adjusted because impatient depositors consume under different circumstances. In this case, we say that a payment schedule x offers liquidity insurance to the n^{th} impatient depositor if

$$x_n > \frac{z_{n-1}}{I - n + 1}, \tag{6}$$

that is, if the amount given to this depositor is larger than the value of the bank's remaining resources divided by the number of depositors who have not yet withdrawn. The following result shows that the efficient allocation satisfies this expression for all possible withdrawals in period 0 up to $I - 1$.[7] A proof is given in the appendix.

Proposition 2 *The efficient payment schedule x^* offers liquidity insurance in the sense of* (6) *for* $n = 1, \ldots, I - 1$.

Expression (6) has important implications for the shape of the efficient payment schedule depicted in Figure 1. When more depositors consume in the early period, before the return R is realized, providing liquidity insurance for impatient depositors becomes more costly in terms of lowering the consumption of patient depositors. The bank's efficient response to an increase in the likely number of impatient depositors is, therefore, to decrease the consumption of all remaining depositors. As each withdrawal takes place in period 0, the bank becomes more pessimistic about the total number of impatient depositors. As shown in the figure, the bank adjusts to this new information by setting the payment x^*_{n+1} lower than x^*_n. The following proposition shows that this monotonicity is a general feature of the efficient payment schedule.

Proposition 3 *The efficient payment schedule is strictly decreasing, with $x^*_{n+1} < x^*_n$ for* $n = 1, \ldots, I - 1$.

[7] If *all* depositors are impatient, the feasbility constraint (4) implies that the payment made to the last depositor, x_I, cannot be greater than the bank's remaining resources. In other words, it is infeasible for condition (6) to hold for the "final" payment x_I.

Proof: From Proposition 1, we know that the efficient payment schedule satisfies

$$x_n^* = \frac{z_{n-1}}{{\phi_n}^{\frac{1}{\gamma}} + 1} \qquad \text{and} \qquad x_{n+1}^* = \frac{z_n}{{\phi_{n+1}}^{\frac{1}{\gamma}} + 1},$$

where

$$z_n = z_{n-1} - x_n^*.$$

Combining these expressions yields

$$x_{n+1}^* = x_n^* \frac{{\phi_n}^{\frac{1}{\gamma}}}{{\phi_{n+1}}^{\frac{1}{\gamma}} + 1}. \tag{7}$$

Using the second inequality in Lemma 1 (in the appendix), it follows immediately that $x_{n+1}^* < x_n^*$ for $n = 1, \ldots, I-1$. ■

Figure 1 also indicates that the efficient payment schedule x^* is initially quite flat, with x_1^* very close in value to x_2^*, x_3^*, and several more payments. Only as the number of early withdrawals becomes larger do the downward adjustments in x_n^* become visible. In other words, depositors withdrawing in period 0 are initially treated similarly, each receiving close to what might be considered the "face value" of their deposit. This property emerges from the fact that the first few withdrawals provide relatively little information to the bank about the total number of impatient depositors and, hence, have little impact on the efficient payment.

To understand this property better, note that the total number of impatient depositors for the parameter values used in Figure 1 is likely to be close to 10, since each of the 20 depositors has a $1/2$ probability of being impatient. When the first early withdrawal takes place, the bank learns that at least one depositor is impatient, which rules out the (extremely unlikely) event that all 20 depositors are patient. The bank's belief about the distribution of θ is only slightly changed by this information, which leads it to set the payment for the next impatient depositor, x_2^*, only slightly different from x_1^*. In other words, the continuation probability q_{n+1} is very close to 1 for small values of n, which implies that the bank is almost certain that additional withdrawals will be made. Since depositors are risk averse, the efficient plan will approximately equalize the values of all payments that are nearly-certain to be made.

The fact that the flatness of the payment scheduled is directly linked to the behavior of the probabilities q_{n+1} can be readily verified by inspecting expression (7). From the definition of ϕ_n in Proposition 1, we see that ϕ_n converges to $({\phi_{n+1}}^{\frac{1}{\gamma}} + 1)^{\gamma}$ as q_{n+1} converges to unity. It is then clear from expression (7) that x^*_{n+1} converges to x^*_n when q_{n+1} converges to unity.

As the number of early withdrawals increases, each additional withdrawal becomes more informative about the value of θ. After 10 withdrawals have taken place, for example, the continuation probability q_{11} has fallen to 0.7. At this point, there is a significant amount of uncertainty about whether or not an additional early withdrawal will be made. If another depositor arrives to withdraw, the bank's belief about θ changes significantly and, as a result, the payment x^*_{11} is noticeably lower than x^*_{10}. As illustrated in Figure 1, these changes become larger and larger as n increases toward 20. This part of the curve resembles what Wallace (1990) calls a "partial suspension of convertibility," in which some depositors receive significantly less than the face value of their deposits when the realized demand for early withdrawal is high.

4 Banking

We now return to the environment in which preference types ω_i are private information. In this setting, the only way for the bank to make a depositor's consumption contingent on ω_i is to allow her to choose the period in which she withdraws. We consider the following banking arrangement: each depositor chooses whether to withdraw early or late based on her own preference type and on the number of withdrawals made by depositors before her in the order, and the bank makes payments on early withdrawals according the the efficient payment schedule x^* derived in Proposition 1. This arrangement creates a *withdrawal game* for depositors; the remainder of the paper is devoted to studying the equilibria of this game.[8]

4.1 The withdrawal game

A depositor chooses a withdrawal strategy, which assigns a withdrawal period (either 0 or 1) to each combination of her preference type ω_i and the number n representing the opportunity to make the n^{th} withdrawal in period 0,

$$y_i = \Omega \times \{1, \ldots, I\} \to \{0, 1\}.$$

We use y to denote a profile of strategies, one for each depositor, and y_{-i} to denote the strategies of all depositors except i.

[8] In general, a withdrawal game can be defined based on any payment schedule x. Our focus here, however, is exclusively on the properties of the withdrawal game generated by the efficient payment schedule x^*.

Our interest is in the Bayesian Nash equilibria of the withdrawal game, where each depositor chooses the strategy y_i that maximizes her expected utility while correctly anticipating the strategies y_{-i} of other depositors.[9] We study symmetric equilibria, in which all depositors follow the same strategy. Different depositors may still take different actions, of course, as their preference types and withdrawal opportunities will differ.

As described above, impatient depositors only care about consumption in period 0 and, hence, the individual best response to any strategy profile y_{-i} will satisfy $y_i(0,n) = 0$ for all n. In other words, we only need to check what the depositor will choose to do when she is patient.

Consider the decision faced by a patient depositor who has an opportunity to make the n^{th} withdrawal in period 0. If she chooses to withdraw, she will receive x_n^*. If she waits, the amount she receives at $t = 1$ will depend on the total number of withdrawals that take place in period 0, which is not yet known. Let $\widehat{\theta}$ denote the number of depositors who wait until $t = 1$ to withdraw in this case, including herself. Then $I - \widehat{\theta}$ is the total number of early withdrawals. Note that $\widehat{\theta}$ is a random variable that depends on both the number of patient depositors θ and the profile of withdrawal strategies followed by the other depositors, y_{-i}. Let $p_n\left(\widehat{\theta}; y_{-i}\right)$ denote the probability this depositor assigns to the event that exactly $\widehat{\theta}$ depositors (including herself) wait until $t = 1$ to withdraw. Given this belief, she computes the expected utility of waiting to withdraw, denoted $z(n; y_{-i})$, as

$$z(n; y_{-i}) = \sum_{\widehat{\theta}=1}^{I} p_n\left(\widehat{\theta}; y_{-i}\right) u\left(\frac{R\phi_{x^*}\left(I - \widehat{\theta}\right)}{\widehat{\theta}}\right). \tag{8}$$

An equilibrium of the withdrawal game is a strategy profile y^* satisfying

$$y_i^*(\omega_i, n) = \left\{ \begin{array}{ll} 0 & \text{if } \omega_i = 0 \text{ or } u(x_n^*) > z\left(n; y_{-i}^*\right) \\ 1 & \text{if } \omega_i = 1 \text{ and } u(x_n^*) < z\left(n; y_{-i}^*\right) \end{array} \right\} \text{ for all } n, \text{ for all } i.$$

In other words, each depositor is optimally choosing when to withdraw given her beliefs about the total number of early withdrawals, and those beliefs are generated by the actions prescribed in the equilibrium strategy profile y^*. Impatient depositors always withdraw early, and patient depositors withdraw early when the value of the available payment, $u(x_n^*)$, exceeds the expected value of waiting, $z\left(n; y_{-i}^*\right)$. Note that when a patient agent is indifferent, because $u(x_n^*) = z\left(n; y_{-i}^*\right)$, either action ($y_i = 0$ or $y_i = 1$) is consistent with equilibrium.

[9] While the structure of the game implies that there are sequential moves by the players, it is not hard to see that the extensive form representation of the game has no proper subgames.

4.2 Incentive compatibility

One strategy of particular interest is the *no-run strategy*, in which a depositor withdraws early if and only if she is impatient

$$y_i^0(\omega_i, n) = \omega_i \quad \text{for all } n. \tag{9}$$

Let y^0 denote the no-run strategy profile. If y^0 is an equilibrium of the withdrawal game, this equilibrium achieves the efficient allocation of resources derived in Section 3.

In this case, the number of depositors who wait to withdraw, $\widehat{\theta}$, is the same as the number of patient depositors θ. A depositor's initial belief about θ is given by (2). To illustrate how the posterior beliefs $p_n\left(\theta; y_{-i}^0\right)$ are formed, suppose that $n = 1$, meaning that no withdrawals have occurred yet. This situation could arise because the depositor is first in the order, in which case it would convey no information about θ. However, it also could arise because the depositor is later in the order but all of the depositors before her were patient, in which case θ is likely to be high. The depositor weighs the relative likelihood of these different situations in updating her beliefs about both θ and her position in the order according to Bayes' rule.

Define Z_k^{n-1} to be the set of all type profiles ω in which there is a patient depositor in the k^{th} position with exactly $n-1$ impatient depositors ahead of her in the order, that is

$$Z_k^{n-1} = \left\{\omega : \omega_k = 1 \ \wedge \ \sum_{j=1}^{k-1}(1-\omega_j) = n-1\right\}.$$

The depositor knows that the realized profile must lie in this set for some value of k. The following proposition derives her posterior belief about the number of patient depositors θ; a proof of the result is given in the appendix.

Proposition 4 *A patient depositor who anticipates that all other depositors are following* (9) *and who has the opportunity to make the* n^{th} *withdrawal in period* 0 *will assign probability*

$$p_n\left(\theta; y_{-i}^0\right) = \rho_n(\theta; I) \equiv \frac{\sum_{\omega\in\{\omega:\theta(\omega)=\theta\}}\left(p(\omega)\sum_{k=n}^{I}\mathcal{I}_{\left\{\omega\in Z_k^{n-1}\right\}}\right)}{\sum_{\omega\in\Omega^I}\left(p(\omega)\sum_{k=n}^{I}\mathcal{I}_{\left\{\omega\in Z_k^{n-1}\right\}}\right)}$$

to the event that exactly θ *of the* I *depositors are patient, where* $\mathcal{I}_{\{A\}}$ *is the indicator function for the set* A.

Using the distribution $p_n\left(\theta; y^0_{-i}\right)$, we can obtain the value of $z\left(n; y^0_{-i}\right)$ from equation (8). The no-run strategy profile is an equilibrium of the withdrawal game if and only if:

$$u\left(x^*_n\right) \leq z\left(n; y^0_{-i}\right) \qquad \text{for } n = 1, \ldots, I. \tag{10}$$

The left-hand side of these inequalities represents the value of withdrawing early and receiving the payment x^*_n for sure. The right-hand side is the expected utility of waiting to withdraw when all other depositors are following (9). When these inequalities are satisfied for all values of $n,$ a patient depositor will find it optimal to withdraw in period 1 regardless of the number of withdrawals that have taken place when her opportunity to withdraw arrives. In this case, there is an equilibrium under the efficient payment schedule in which all depositors follow the strategy (9) and the efficient allocation of resources obtains. It is straightforward to see that the equilibrium inequalities (10) hold whenever the efficient allocation is incentive compatible.

Our focus in the rest of the paper is on situations in which the efficient allocation x^* is incentive compatible. We follow an approach that is common in the literature by first solving for the efficient payment schedule without imposing the incentive compatibility conditions, using Proposition 1, and then verifying that the solution satisfies (10). In the examples we present in the next section, the efficient payment schedule x^* always satisfies (10). In other words, the constraints imposed by private information are not binding on the efficient allocation of resources in these cases. This fact illustrates the benefits of banking arrangements that resemble demand deposit contracts, in which depositors are allowed to choose when to withdraw their funds from the bank. As in Diamond and Dybvig (1983) and others, such an arrangement allows the bank to potentially achieve the same allocation of resources it would choose if it could observe individual depositor's consumption preferences, even though these preferences are private information. However, such arrangements may also open the door to financial fragility in the sense of admitting other equilibria in which some patient depositors "run" on the bank and withdraw in period $0,$ leading to an inferior allocation. In the next section, we investigate whether such fragility arises in our model.

5 Fragility

In this section, we investigate whether the withdrawal game defined above can have other Bayesian Nash equilibria, in which some depositors choose to withdraw in period 0 even when they are patient. If such equilibria exist, the bank is fragile in the sense that attempting to implement the efficient allocation of resources using the arrangement described above could lead to an inefficient outcome that resembles a run on the bank. We first show that there cannot be an equilibrium in which all depositors attempt to withdraw early. We then construct an example of a *partial run* equilibrium, in which some patient depositors withdraw early but others do not.

5.1 No full-run equilibrium

The type of bank run studied in most of the literature has all depositors attempting to withdraw their funds in the early period.[10] We refer to this strategy profile,

$$y_i(\omega_i, n) = 0 \quad \text{for all } \omega_i, n, \text{ for all } i, \tag{11}$$

as a *full run*. It is fairly easy to see that a full run cannot be part of an equilibrium in our model. Consider a depositor with $n = I$, meaning that all of the other $I - 1$ depositors have already withdrawn when her opportunity to withdraw in period 0 arises. If she chooses to withdraw in period 0, she will receive all of the bank's remaining resources, $\phi_{x^*}(I-1)$. If she waits until period 1 to withdraw, however, she will receive the matured value of these resources $R\phi_{x^*}(I-1)$. Because $R > 1$, a patient depositor in this situation would always strictly prefer to wait to withdraw. In other words, the full-run strategy profile (11) is strictly dominated by a strategy profile with $y_i(1, I-1) = 1$ and, hence, cannot be part of an equilibrium. The following proposition records this result.

Proposition 5 *There is no equilibrium in which depositors play the strategy profile* (11).

This aspect of our model is similar in spirit to one of the key features of Green and Lin (2003). In their setting, depositors receive a signal about their position in the decision order. If a depositor's signal indicates that she is very likely to be last in the order, she faces a decision that is similar to the one described above and will strictly prefer to wait until period 1 if she is patient. Depositors in our model do not observe their place in the order; however, a depositor can infer this information from n when all other depositors are playing the full-run strategy. For this reason, the Green-Lin logic for the last depositor applies in our setting under the strategy profile (11) and shows that a

[10] See, for example, Diamond and Dybvig (1983), Cooper and Ross (1998), and Peck and Shell (2003). Some recent exceptions are Ennis and Keister (2009, 2010a), Gu (2011) and Azreili and Peck (2011).

full-run equilibrium cannot exist. Green and Lin (2003) then use a backward-induction argument to show that the no-run strategy profile is the *unique* equilibrium in their setting; no type of run equilibrium can exist. In the next subsection, we show this stronger result does not hold in our environment.

5.2 Partial-run equilibria

Proposition 5 shows that if a run equilibrium exists in our environment, it must be partial, with only some depositors participating. One possibility is for depositors to run until the total number of withdrawals reaches some critical level and then for the run to stop. Consider the following class of strategy profiles

$$y_i^{n^*}(\omega_i, n) = \begin{cases} 0 & \text{for} \quad n \leq n^* \\ \omega_i & \phantom{\text{for}} \quad n > n^* \end{cases} \quad \text{for some } 1 \leq n^* \leq I-1, \quad \text{for all } i. \tag{12}$$

Note that setting $n^* = 0$ would correspond to a no-run strategy profile, which is consistent with equilibrium as long as the incentive compatibility conditions (10) hold. Setting $n^* = I$ would correspond to a full-run strategy profile, which Proposition 5 shows is inconsistent with equilibrium. For values of n^* between 1 and $I-1,$ this profile represents a *partial run*.

Recall that $\widehat{\theta}$ denotes the total number of depositors who wait until period 1 to withdraw. Under a partial-run strategy profile, $\widehat{\theta}$ is bounded above by $\min[\theta, I-n^*]$. The realized value of $\widehat{\theta}$ will be less than the number of patient depositors θ if one or more of the first n^* depositors in the decision order is patient and, following (12), withdraws in period 0.

To determine if (12) is consistent with equilibrium for some value of n^*, we need to compare the expected utility a depositor receives from following this strategy to the expected utility associated with deviating to the best alternative. In particular, for $n \leq n^*$, we need to consider whether a patient depositor would be better off waiting until period 1 to withdraw, while for $n > n^*$ we need to consider whether a patient depositor would be better off withdrawing in period 0. We address these two cases in turn.

(i) $n \leq n^*$: If the depositor follows the strategy in (12), she will withdraw early and receive x_n^*. If she deviates, she will receive the period-2 payment associated with $\widehat{\theta}$ depositors waiting to withdraw. Comparing the expected utility of these two outcomes requires deriving the depositor's belief about $\widehat{\theta}$ under the assumption that (a) she is patient, (b) all other depositors follow the strategy profile in (12), and (c) she deviates from (12) and waits until period 1. In this case, the value of n fully reveals the depositor's position in the decision order: she must be the n^{th} depositor. It does not, however, give her any information about the profile of types ω. The depositor knows that at least n^* depositors will withdraw early, and that the remaining $I - n^* - 1$ depositors will each withdraw early if and only if they are impatient.[11] Her posterior belief about $\widehat{\theta}$ is then given by

$$p_n\left(\widehat{\theta}; y_{-i}^{n^*}\right) = 0 \qquad \text{for } \widehat{\theta} > I - n^* \text{ and } n \leq n^* \tag{13}$$

and

$$p_n\left(\widehat{\theta}; y_{-i}^{n^*}\right) = C\left(I - n^* - 1, \widehat{\theta} - 1\right) \pi^{I-n^*-\widehat{\theta}} \left(1-\pi\right)^{\widehat{\theta}-1} \tag{14}$$
$$\text{for } \widehat{\theta} = \{1, \ldots, I - n^*\} \text{ and } n \leq n^*,$$

where C is the combinatorial function (3). In other words, the depositor knows that $\widehat{\theta}$ will be at least 1 (herself) and can be up to $I - n^*$ if all of the other depositors who do not participate in the run are patient. The types of these other depositors are i.i.d. Bernoulli trials. Notice that $p_n\left(\widehat{\theta}; y_{-i}^{n^*}\right)$ is independent of n for all $n < n^*$; the depositor's precise position in the order does not affect her posterior belief about $\widehat{\theta}$ in this case.

The expected utility of deviating from (12) and withdrawing in period 1 is given by $z\left(n; y_{-i}^{n^*}\right)$ as defined in (8), using the distribution $p_n\left(\widehat{\theta}; y_{-i}^{n^*}\right)$ presented in (13) – (14). A necessary condition for (12) to be consistent with equilibrium is that

$$u\left(x_n^*\right) \geq z\left(n; y_{-i}^{n^*}\right) \qquad \text{for } n \leq n^*.$$

Notice that $z\left(n; y_{-i}^{n^*}\right)$ is independent of n for $n \leq n^*$, while Proposition 3 establishes that $u\left(x_n^*\right)$ is strictly decreasing in n. Therefore, the condition above will be satisfied for all $n \leq n^*$ if and only if it is satisfied for n^*, that is,

$$u\left(x_{n^*}^*\right) \geq z\left(n^*; y_{-i}^{n^*}\right). \tag{15}$$

When this condition holds, a depositor who expects all other depositors to follow the strategy in (12) and observes a value of n less than or equal to n^* will also be willing to follow (12). In other words, condition (15) is a necessary condition for the strategy profile y^{n^*} to be an equilibrium. We

[11] Notice that the n^* withdrawals that constitute the partial run will take place regadless of whether or not this individual depositor chooses to participate in the run. If she deviates from (12), other depositors will continue the run until n^* is reached.

record this result, which is useful for constructing the examples below, as a proposition.

Proposition 6 *If the strategy profile y^{n^*} is an equilibrium, then $u\left(x^*_{n^*}\right) \geq z\left(n^*; y^{n^*}_{-i}\right)$.*

(ii) $n > n^*$: When the depositor's opportunity to withdraw arrives later, after the run has ended, she can no longer be certain about her position in the order. Instead, she must form beliefs about how many of the $I - n^*$ no-run depositors are impatient and what her position within the order of these depositors might be. This inference problem is similar in structure to the one discussed in the context of incentive compatibility in Section 3. In this case, the depositor's posterior belief about $\widehat{\theta}$ is given by

$$p_n\left(\widehat{\theta}; y^{n^*}_{-i}\right) = \rho_{n-n^*}\left(\widehat{\theta}; I - n^*\right) \quad \text{for } \widehat{\theta} = \{1, \ldots, I - n^*\} \text{ and } n > n^*. \tag{16}$$

The expression $\rho_{n-n^*}\left(\widehat{\theta};, I - n^*\right)$ is the probability distribution presented in Proposition 4 applied to the *subset* of $I - n^*$ depositors who do not participate in the run. Note that the index on this distribution is adjusted from n to $n - n^*$, since the depositor in question observes $n - 1 - n^*$ withdrawals from the $I - n^*$ no-run depositors.

A patient depositor following the strategy in (12) would wait until period 1 to withdraw and would have expected utility given by $z\left(n; y^{n^*}_{-i}\right)$ in (8), with the distribution $p_n\left(\widehat{\theta}; y^{n^*}_{-i}\right)$ now given by (16). If she instead deviates and withdraws early, she will receive x^*_n for sure. The strategy profile y^{n^*} is consistent with a best response for these depositors if

$$u\left(x^*_n\right) \leq z\left(n; y^{n^*}_{-i}\right) \quad \text{for } n = n^* + 1, \ldots, I. \tag{17}$$

If conditions (15) and (17) both hold for some $1 \leq n^* \leq I - 1$, the partial-run strategy profile (12) is a Bayesian Nash equilibrium of the withdrawal game. Our main result in this section is that, for some parameter values, such a partial run equilibrium exists.

Proposition 7 *For some parameter values it is the case that the efficient payment schedule (i) generates an incentive compatible allocation and (ii) admits a partial run equilibrium with $1 \leq n^* \leq I - 1$.*

The proof is by example. We present the example in the next subsection, followed by a detailed discussion of the intuition behind it.

5.3 An Example

Our example is based on the same parameter values that were used to illustrate the efficient allocation in Figure 1: $I = 20$, $R = 1.1$, $\gamma = 6$, and $\pi = 0.5$. Figure 2 plots three curves, two of which are simply the expected utility associated with the consumption levels from the earlier figure. The third curve is the auxiliary function $z_n(n) \equiv z\left(n; y_{-i}^n\right)$ that gives the expected utility from waiting to withdraw for a patient depositor who has the opportunity to make the n^{th} withdrawal if she anticipates all other depositors are playing the partial run strategy with the critical withdrawal (after which the run stops) given by n.

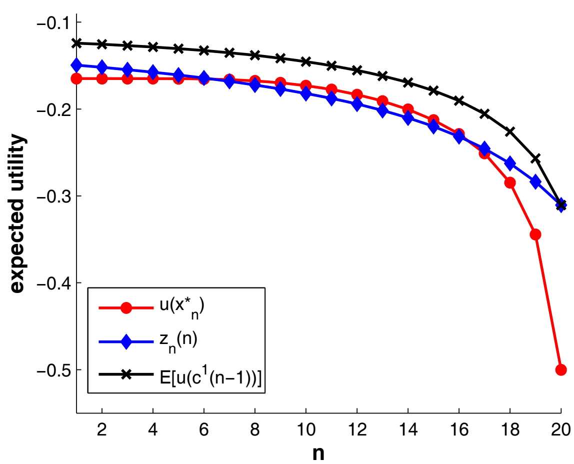

Figure 2: A partial run equilibrium

Note that the necessary condition identified in Proposition 6 – that $u(x_n^*) \geq z_n(n)$ hold – is satisfied for values of n between 7 and 16. Let us concentrate attention on the partial run strategy with the largest possible threshold n^*; in this case, $n^* = 16$. The fact that the necessary condition holds implies that depositors with $n \leq 16$ will all chose to play according to the partial run strategy profile y^{16}. Hence, to show that y^{16} is an equilibrium strategy profile, we only need to show that depositors with $n > 16$ will also prefer to follow this strategy, which specifies that they wait until period 1 if patient. This is done in Figure 3, which plots the incentive to run

$$u\left(x_n^*\right) - z\left(n; y_{-i}^{16}\right)$$

for all n between 1 and 20 (see the solid red curve). The discussion above demonstrated that this value is positive for n between 1 and 16; this fact is also reflected in Figure 3. The new information in the figure is that the expression is negative for $n = 17$ through 20, which verifies that a depositor who has any one of these opportunities to withdraw in period 0 would choose to wait, in accordance with the candidate equilibrium strategy profile. Hence, the figure establishes that strategy profile (12) with $n^* = 16$ comprises an equilibrium of the withdrawal game.

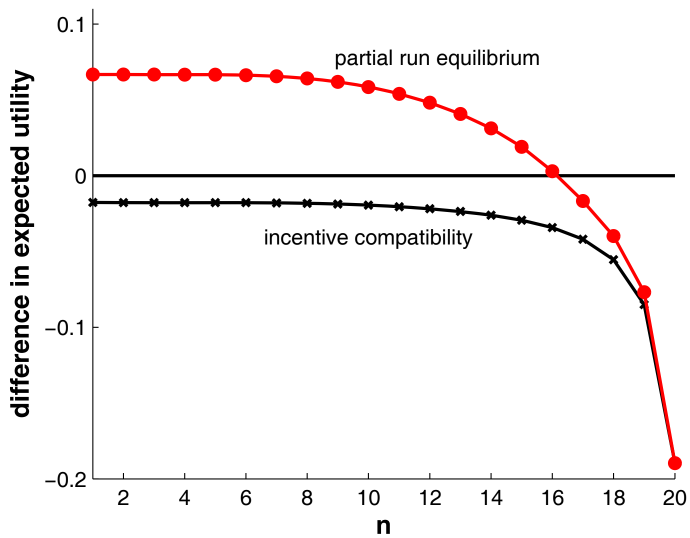

Figure 3: Individual incentive to run

Figure 3 also verifies that the efficient allocation is incentive compatible in this example. The dashed black curve plots

$$u\left(x_n^*\right) - z\left(n; y_{-i}^{0}\right);$$

that is, the gain in expected utility from following the run strategy for each depositor under the assumption that all other depositors follow the no-run strategy. The fact that this line is negative everywhere demonstrates that the no-run strategy profile is also an equilibrium of the withdrawal game. Together, the two curves in Figure 3 thus establish the result in Proposition 7. It is worth emphasizing that there is nothing special about the parameter values used in this example; it is easy to construct similar examples using a wide range of parameter values.

5.4 Intuition

To gain intuition for why the partial-run strategy profile in (12) is an equilibrium, it is useful to examine the behavior of the "critical" depositor whose opportunity to withdraw in period 0 is n^*. This depositor follows the run strategy even though she believes that no one after her will run. This type of behavior is inconsistent with equilibrium under Green and Lin (2003)'s formulation of the sequential service constraint, a fact that is crucial for their backward-induction argument. Understanding why it can arise here thus illustrates a key difference between our environment and theirs.

Suppose we compare this depositor's equilibrium belief about the number of early withdrawals with the beliefs used by the bank in designing the payment schedule x^*. The efficient payment x^*_{16} depends on the probability that there will be a 17^{th} early withdrawal (and an 18^{th}, and so on). Given that the bank expects only impatient agents to withdraw early, and each depositor has an independent, one-half probability of being impatient, the bank considers a 17^{th} early withdrawal fairly unlikely. In other words, when faced with a 16^{th} early withdrawal, the bank believes that the four depositors it has not yet seen are likely to have already had their opportunity to withdraw and decided to wait because they are patient. In this sense, the bank is "optimistic" that the 16^{th} early withdrawal will be the last one and the payment x^*_{16} is chosen based on this optimism.

In the partial run equilibrium, however, the depositor making the 16^{th} withdrawal recognizes that a run is underway and the number of early withdrawals is thus likely to be much larger than the number of impatient depositors. Importantly, she knows that four depositors have not yet had the opportunity to contact the bank in period 0. She expects that, on average, two of these depositors will be impatient and, hence, she realizes that the early withdrawals are unlikely to end with her. Any additional withdrawals will further deplete the bank's resources, lowering the consumption she would receive if she were to wait and withdraw in period 1. Her more "pessimistic" belief about the number of early withdrawals thus makes running – and accepting the payment based on the bank's optimistic belief – an attractive strategy.[12]

[12] It is easy to see that the incentive to run for depositors making withdrawals before the 16^{th} are even stronger. The payment that the depositor making the first withdrawal receives, for example, is based on the expectation that the number of early withdrawals will be, on average, around 10. In equilibrium, however, this depositor anticipates that there will be at least 16 and on average 18 early withdrawals. This large gap in beliefs makes running very attractive. The incentives to participate in the run for depositors making the second through the 15^{th} withdrawal

Notice how the limited flow of information implied by our formulation of the sequential service constraint makes this divergence in beliefs possible. When a depositor makes the 16^{th} withdrawal, the bank does not know that she is 16^{th} in the order. In fact, the efficient payments are based, implicitly, on the belief that if a 16^{th} impatient depositor demands an early payment, she is likely to be the *last* depositor in the order. This sharp divergence of beliefs arises because the bank does not observe the actions of depositors who choose not to withdraw in period 0. In Green and Lin (2003), patient depositors are expected to announce to the bank that they will not withdraw when their turn in the order comes and, hence, the bank knows that the depositor making the 16^{th} withdrawal is the 16^{th} depositor in the order. Both the bank and that depositor thus know that there are four depositors who still have an opportunity to contact the bank in period 0. Since types are independent, if these depositors will report truthfully then the bank and the depositor must have the same belief about the number of additional early withdrawals, regardless of what strategies the earlier depositors have followed.[13] Because of this agreement in beliefs, the payment offered to each depositor in the Green-Lin setup appears "appropriate" given her beliefs and, as a result, she will choose to report truthfully. This reasoning is central to the unique implementation result in Green and Lin (2003). In contrast, the divergence in beliefs in our model arises naturally whenever one or more depositors follows a non-truthful strategy. The example presented here shows how this divergence in beliefs can be strong enough to generate a run equilibrium in the withdrawal game.

5.5 Discussion

In earlier work (Ennis and Keister, 2009), we showed that partial run equilibria can arise in the model with Green and Lin's (2003) formulation of the sequential service constraint when depositors' preference types are correlated. The intuition behind this result is similar to that described above. In particular, there is a critical depositor who chooses to run even though she expects everyone after her to report truthfully. The payment to this depositor is again designed based on the expectation that her withdrawal will likely be the last one in period 0, while the critical depositor believes that additional early withdrawals are very likely. As described above, this divergence in beliefs makes withdrawing early, and receiving the payment based on the more optimistic belief,

lie somewhere in between those of the one making the first and the 16^{th} withdrawal, as illustrated in Figure 3.

13 See also Andolfatto *et al.* (2007) and Nosal and Wallace (2009) on this point.

an attractive choice. In our earlier work, the difference in beliefs was generated by the correlation structure in types, while in the present paper it is generated by the fact that the bank does not observe the actions of depositors who choose not to withdraw.

Ennis and Keister (2009) followed Green and Lin (2000) in assuming that a depositor knows her position in the order and do not observe any information about the actions of other depositors. Here, in contrast, we assume a depositor observes the number of withdrawals that have already been made when her opportunity to withdraw arrives and must use this information to make inferences about her position in the order. This approach, which is new in the literature, gives the withdrawal game a more dynamic flavor. It also potentially opens the door to issues of signaling, in which a depositor takes into account the effect her action will have on the actions of subsequent depositors (see Andolfatto *et al*., 2007). Such effects play a minimal role in our setting, however, because depositors only observe withdrawals and receive no information when another depositor chooses not to withdraw. Under the strategy profile y^{n^*}, for example, there is no way for the first depositor in the order to "signal" that she is deviating from this profile. If she chooses not to participate in the run, this action is unobserved and the next depositor, observing that no withdrawals have taken place yet, will incorrectly infer that she is first in the order. Similar, if a patient depositor deviates from the no-run strategy profile by withdrawing early, other depositors will infer that she was impatient rather than suspecting a deviation. For these reasons, many of the complications typically associated with dynamic games do not arise in our setting.

We have followed a common approach in the literature by assuming the bank makes payments according to the efficient payment schedule x^* and asking whether the resulting withdrawal game admits a run equilibrium. If the bank placed positive probability on the event of a run, it would likely choose to follow a different payment schedule. In particular, the bank would select a payment schedule x that balances the trade-off between having a less efficient allocation than the one generated by x^* when a run does not occur and having an improved outcome if a run indeed occurs. If the probability of a run is large enough, the bank may choose a payment schedule that makes the partial-run strategy inconsistent with equilibrium (see Cooper and Ross, 1998, and Ennis and Keister, 2006). In other words, the possibility of having a banking system that is immune to runs is available in the economy. However, this immunity can only be achieved by distorting the payment schedule away from the efficient no-run allocation generated by x^*. If the probability of a run is small, the payment schedule chosen by the bank will be close to the efficient payment schedule

x^*. In this case, a continuity argument can be used to show that the type of run equilibrium we construct in this section will survive. In summary, our results show that when the (perceived) probability of a run is sufficiently small, the bank will choose a deposit contract that exposes the economy to financial fragility.

Whatever probability the bank initially assigns to the event of a run, it can never actually be sure that a run is taking place, even *ex post.* This is another feature our model shares with much of the existing literature: a run is observationally equivalent to an event that occurs with positive probability under the no-run strategy profile (in which many depositors are impatient).[14] There is, however, a difference in how quickly the bank is able to infer that something unusual is happening. In our setting, a run is initially equivalent to almost *all* events under the no-run strategy profile. Suppose depositors follow the partial-run strategy profile in (12). As the first few depositors arrive to withdraw, the bank is only able to infer that at least a few depositors are withdrawing, an event which is very likely to happen under the no-run strategy profile as well. In the Green and Lin (2003) specification of the sequential service constraint, in contrast, the bank will observe the fact that no depositors are choosing to wait to withdraw, which is unlikely to occur if depositors are following the no-run strategy profile. Relative to the Green-Lin approach, our specification slows down the flow of information about depositors' strategies to the bank. This slower flow of information makes the bank slower to react to a run, which in turn tends to increase the incentive for individual depositors to participate in the run.

6 Concluding Remarks

We study a model of financial intermediation in the tradition of Diamond and Dybvig (1983), introducing an explicit sequential service constraint with novel features. Some aspects of our approach could be considered more realistic than those in the existing literature, such as the fact that a depositor's actions are observed by both the bank and other depositors, but only when she chooses to withdraw. We show that our specification is tractable and generates appealing results. In particular, we do not impose any constraints on contractual arrangements other than those created by the information frictions in the underlying environment. Even though this approach is consistent

[14] Ennis and Keister (2010a) present a model in which a partial run may occur and the bank is able to infer in equilibrium that a run is underway. The key features of the model are that there is no uncertainty about the fraction of agents who will be impatient and the bank is unable to commit to a pre-specified payment schedule.

with contracts that in principle could present very complex patterns, the optimal arrangement in our model resembles in many ways a traditional banking contract. In particular, depositors who withdraw early in the course of events all receive approximately the face value of their deposits. Only if the number of withdrawals becomes unexpectedly large do depositors begin experiencing discounts in what they receive from the bank.

In addition to highlighting the debt-like features of the optimal contract, we show that this arrangement is fragile in the sense of being susceptible to a self-fulfilling run by depositors. The previous literature has shown that in environments similar to ours, fragility may not arise if contractual arrangements are sufficiently flexible. Our specification of the sequential service constraint suggests that the debt-like feature of banking contracts and financial fragility may both have a common origin in the gradual revelation of information that is inherent is real world banking arrangements.

Appendix A. Proofs

Proposition 1: *The efficient payment schedule sets*

$$x_n = \frac{z_{n-1}}{(\phi_n)^{\frac{1}{\gamma}} + 1} \text{ for } n = 1, \ldots, I,$$

where $z_{n-1} = I - \sum_{j<n} x_j$ *and the constants* ϕ_n *are defined recursively by* $\phi_I = 0$ *and*

$$\phi_n = q_{n+1} \left(\phi_{n+1}{}^{\frac{1}{\gamma}} + 1 \right)^{\gamma} + (1 - q_{n+1}) (I - n)^{\gamma} R^{1-\gamma}$$

for $n = 1, \ldots, I - 1$.

Proof: Let V_n denote the sum of the expected utilities of all depositors who have not yet consumed when the bank encounters the n^{th} impatient depositor, conditional on the bank dividing the available resources z_{n-1} efficiently among these depositors. These values must satisfy the following recursive equation:

$$V_n (z_{n-1}) = \max_{\{x_n\}} \left\{ \begin{array}{c} \frac{(x_n)^{1-\gamma}}{1-\gamma} + q_{n+1} V_{n+1} (z_{n-1} - x_n) + \\ (1 - q_{n+1}) (I - n) \frac{1}{1-\gamma} \left(\frac{R(z_{n-1} - x_n)}{I-n} \right)^{1-\gamma} \end{array} \right\}, \tag{18}$$

for $n = 1, \ldots I$.

If all I depositors are impatient, the bank will give all of the remaining resources to the last depositor when she reports. We therefore have the following terminal condition

$$V_I (z_{I-1}) = \frac{1}{1-\gamma} (z_{I-1})^{1-\gamma}.$$

The combination of this equation, the initial condition $z_0 = I$, and equation (18) constitutes the dynamic programming problem whose solution gives the efficient payment schedule.

Consider the decision problem faced by the bank if it faces an $(I-1)^{th}$ impatient depositor. Given z_{I-2}, the maximization problem in (18) reduces to

$$\max_{\{x_{I-1}\}} \frac{(x_{I-1})^{1-\gamma}}{1-\gamma} + q_I \frac{(z_{I-2} - x_{I-1})^{1-\gamma}}{1-\gamma} + (1 - q_I) \frac{(R (z_{I-2} - x_{I-1}))^{1-\gamma}}{1-\gamma}.$$

The solution to this problem sets

$$x_{I-1} = \frac{z_{I-2}}{(\phi_{I-1})^{\frac{1}{\gamma}} + 1},$$

where

$$\phi_{I-1} \equiv q_I + (1 - q_I) R^{1-\gamma}. \tag{19}$$

Substituting the solution back into the objective function and doing some straightforward algebra yields the value function

$$V_{I-1}(z_{I-2}) = \frac{(z_{I-2})^{1-\gamma}}{1-\gamma} \left((\phi_{I-1})^{\frac{1}{\gamma}} + 1 \right)^{\gamma}.$$

The function V_{I-1} captures the utility of the last two depositors to report to the bank in the event that at least $I-1$ depositors are impatient. In this case, the $(I-1)^{th}$ depositor to report is necessarily impatient. The I^{th} may also be impatient, reporting in period 0, or patient, in which case she will report in period 1. The probabilities of these events (given by q_I) are contained in the constant ϕ_{I-1}.

It is straightforward to use this same procedure to show that, for any $n < I$, the solution to the maximization problem in (18) sets

$$x_n = \frac{z_{n-1}}{\phi_n^{\frac{1}{\gamma}} + 1},$$

where

$$\phi_n = q_{n+1} \left(\phi_{n+1}^{\frac{1}{\gamma}} + 1 \right)^{\gamma} + (1 - q_{n+1}) (I-n)^{\gamma} R^{1-\gamma}. \tag{20}$$

Note that condition (19) emerges naturally from (20) using the “terminal” value $\phi_I = 0$. ■

Proposition 2: *The efficient payment schedule* x^* *offers liquidity insurance in the sense of* (6) *for* $n = 1, \ldots, I-1$.

Proving this proposition requires establishing $I - n > \phi_n^{\frac{1}{\gamma}}$ for $n = 1, \ldots, I-1$. Using the fact that $\gamma > 1$ implies $R^{\frac{1-\gamma}{\gamma}} < 1$, these inequalities are an immediate consequence of the following lemma:

Lemma 1: The inequalities

$$(I-n) R^{\frac{1-\gamma}{\gamma}} < \phi_n^{\frac{1}{\gamma}} < \phi_{n+1}^{\frac{1}{\gamma}} + 1 \tag{21}$$

hold for $n = 1, \ldots I-1$.

Proof of the lemma: The proof is by backward induction. First, consider the case of $n = I-1$.

Proposition 1 defines $\phi_I = 0$ and

$$\phi_{I-1} = q_I + (1 - q_I)\, R^{1-\gamma}.$$

Note that $\gamma > 1$ implies $R^{1-\gamma} < 1$. Together with the fact that $0 < q_I < 1$, the definition of ϕ_{I-1} above thus implies

$$R^{\frac{1-\gamma}{\gamma}} < \phi_{I-1}{}^{\frac{1}{\gamma}} < 1,$$

which shows that (21) holds for $n = I - 1$.

Next, suppose (21) holds for some $n \leq I - 1$; we need to show that the inequalities then must also hold for $n - 1$. Together with $R^{\frac{1-\gamma}{\gamma}} < 1$, the first inequality in (21) implies

$$(I - n + 1)\, R^{\frac{1-\gamma}{\gamma}} < \phi_n{}^{\frac{1}{\gamma}} + 1.$$

From the definition of ϕ in Proposition 1, we have

$$\phi_{n-1} = q_n \left(\phi_n{}^{\frac{1}{\gamma}} + 1\right)^{\gamma} + (1 - q_n)\,(I - n + 1)^{\gamma}\, R^{1-\gamma}.$$

Since $0 < q_n < 1$, these last two expressions imply

$$(I - n + 1)^{\gamma}\, R^{1-\gamma} < \phi_{n-1} < \left(\phi_n{}^{\frac{1}{\gamma}} + 1\right)^{\gamma}$$

or

$$(I - n + 1)\, R^{\frac{1-\gamma}{\gamma}} < \phi_{n-1}{}^{\frac{1}{\gamma}} < \phi_n{}^{\frac{1}{\gamma}} + 1,$$

which establishes that (21) holds for $n - 1$ as well. ■

Proposition 4: *A patient depositor who anticipates that all other depositors are following* (9) *and who has the opportunity to make the n^{th} withdrawal in period 0 will assign probability*

$$p_n\left(\theta; y^0\right) = \rho_n\left(\theta; I\right) \equiv \frac{\sum\limits_{\omega \in \{\omega : \theta(\omega) = \theta\}} \left(p\left(\omega\right) \sum\limits_{k=n}^{I} \mathcal{I}_{\left\{\omega \in Z_k^{n-1}\right\}}\right)}{\sum\limits_{\omega \in \Omega^I} \left(p\left(\omega\right) \sum\limits_{k=n}^{I} \mathcal{I}_{\left\{\omega \in Z_k^{n-1}\right\}}\right)}$$

to the event that exactly θ of the I depositors are patient, where $I_{\{A\}}$ is the indicator function for the set A.

Proof: Let σ denote the depositor's position in the order of potential withdrawals in period 0; recall that this position is not observed by the depositor. The depositor is concerned with two independent random events: the ordered sequence of preference types ω and her position σ in the order, which is drawn from the uniform distribution on $\{1, \ldots I\}$. The probability of a pair (ω, σ) is given by

$$P(\omega, \sigma) = \frac{p(\omega)}{I},$$

where

$$p(\omega) = (1-\pi)^{\theta(\omega)} \pi^{1-\theta(\omega)}.$$

Recall that we have defined Z_k^{n-1} as the set of all type profiles ω in which there is a patient depositor in the k^{th} position with exactly $n-1$ impatient depositors ahead of her in the order. Now define the events

$$\begin{aligned} B_k^{n-1} &= \left\{(\omega, \sigma) : \omega \in Z_k^{n-1}\right\}, \\ C_k &= \{(\omega, \sigma) : \sigma = k\}, \end{aligned}$$

and

$$A_k^{n-1} = B_k^{n-1} \bigcap C_k.$$

The event A_k^{n-1} is the set of pairs (ω, σ) in which the depositor in question is in the k^{th} position with exactly $n-1$ impatient depositors ahead of her in the order. Note that since the events B_k^{n-1} and C_k are independent and $P(C_k) = 1/I$ for all values of k, we have

$$P\left(A_k^{n-1}\right) = \frac{P\left(B_k^{n-1}\right)}{I} = \sum_{\omega \in \Omega^I} \frac{p(\omega)\, \mathcal{I}_{\left\{\omega \in Z_k^{n-1}\right\}}}{I}.$$

The union

$$A^{n-1} = \bigcup_{k=n}^{I} A_k^{n-1}$$

contains all of the profiles ω that the depositor in question considers to be possible together with the positions σ that she could conceivably occupy in each profile. Because the sets A_k^{n-1} are disjoint for different values of k, we have

$$\begin{aligned} P\left(A^{n-1}\right) &= \sum_{k=n}^{I} P\left(A_k^{n-1}\right) \\ &= \sum_{k=n}^{I} \sum_{\omega\in\Omega^I} \frac{p(\omega)\,\mathcal{I}_{\left\{\omega\in Z_k^{n-1}\right\}}}{I} \\ &= \sum_{\omega\in\Omega^I} p(\omega) \sum_{k=n}^{I} \frac{\mathcal{I}_{\left\{\omega\in Z_k^{n-1}\right\}}}{I}. \end{aligned} \tag{22}$$

In other words, the probability of the set A^{n-1} can be obtained by summing the probabilities of all profiles ω, each weighted by the fraction of the I positions in the order that match the observed criteria (that is, have a patient depositor with exactly $n-1$ impatient depositors ahead of her in the order). Notice that any profile ω that is not in the set Z_k^{n-1} for some value of k receives zero weight in this sum.

Next, define

$$D^{\omega} = \{(\widehat{\omega},\sigma) : \widehat{\omega} = \omega\}.$$

The set D^{ω} contains all of the pairs (ω,σ) that correspond to a particular profile of preference types ω. It is straightforward to see that the prior probability of this set is given by $p(\omega)$ .The posterior probability we want to calculate is

$$P\left(D^{\omega} \mid A^{n-1}\right),$$

that is, the probability of a particular profile ω conditional on the depositor observing that there are $n-1$ impatient depositors ahead of her in the order. Applying Bayes' rule yields

$$P\left(D^{\omega} \mid A^{n-1}\right) = \frac{P\left(A^{n-1} \mid D^{\omega}\right) P\left(D^{\omega}\right)}{P\left(A^{n-1}\right)}. \tag{23}$$

By definition

$$P\left(A^{n-1} \mid D^{\omega}\right) P\left(D^{\omega}\right) = P\left(A^{n-1} \bigcap D^{\omega}\right) = P\left(\bigcup_{k=n}^{I} \left(\left(B_k^{n-1} \bigcap D^{\omega}\right) \bigcap C_k\right)\right). \tag{24}$$

For a given profile ω, we have

$$B_k^{n-1} \bigcap D^{\omega} = \left\{ \begin{array}{ccc} \varnothing & & B_k^{n-1} = \varnothing \\ \varnothing & \text{if} & \omega \notin Z_k^{n-1} \\ D^{\omega} & & \omega \in Z_k^{n-1} \end{array} \right\},$$

Again using the fact that σ is independent of ω and that $P(C_k) = 1/I$ for all k, this implies

$$P\left(A^{n-1} \bigcap D^{\omega}\right) = P\left(D^{\omega}\right) \frac{\sum_{k=n}^{I} \mathcal{I}_{\left\{\omega \in Z_k^{n-1}\right\}}}{I}. \tag{25}$$

In other words, the probability of the intersection of the sets A^{n-1} and D^{ω} is equal to the probability of D^{ω} multiplied by the fraction of the I positions in the order that the depositor could conceivably occupy when the realized profile is ω.

Substituting (22), (24) and (25) into expression (23) yields the posterior belief over profiles ω,

$$P\left(D^{\omega} \mid A^{n-1}\right) = \frac{p(\omega) \sum_{k=n}^{I} \mathcal{I}_{\left\{\omega \in Z_k^{n-1}\right\}}}{\sum_{\omega \in \Omega^I} \left(p(\omega) \sum_{k=n}^{I} \mathcal{I}_{\left\{\omega \in Z_k^{n-1}\right\}}\right)}.$$

This expression shows that the posterior probability of a profile ω is proportional to its prior probability multiplied by the fraction of positions in the order the depositor could conceivably occupy if ω were the true profile.

Finally, the posterior probability distribution over θ, the number of patient depositors, is given by

$$p_n(\theta; I) = \sum_{\left\{\omega : \theta(\omega) = \widehat{\theta}\right\}} P\left(D^{\omega} \mid A^{n-1}\right).$$

■

References

Andolfatto, David and Ed Nosal (2008) "Bank incentives, contract design, and bank runs," *Journal of Economic Theory* 142, 28-47.

Andolfatto, David, Ed Nosal, and Neil Wallace (2007) "The role of independence in the Green-Lin Diamond-Dybvig Model," *Journal of Economic Theory* 137, 709-715.

Azrieli, Yaron and James Peck (2011) "A bank runs model with a continuum of types," mimeo.

Baba, Naohiko, Robert N, McCauley, and Srichander Ramaswamy (2009) "US dollar money market funds and non-US banks," BIS *Quarterly Review*, March, 64-81.

Cooper, Russell and Thomas W. Ross (1998) "Bank runs: liquidity costs and investment distortions," *Journal of Monetary Economics* 41, 27-38.

Diamond, Douglas W. and Phillip H. Dybvig (1983) "Bank runs, deposit insurance, and liquidity," *Journal of Political Economy* 91, 401-419.

Duffie, Darrell (2010) "The failure mechanics of dealer banks," *Journal of Economic Perspectives* 24, 51-72.

Dwyer, Gerald P., Jr., and Iftekhar Hasan (2007) "Suspension of payments, bank failures, and the nonbank public's losses," *Journal of Monetary Economics* 54, 565–580

Ennis, Huberto M. and Todd Keister (2010a) "Banking panics and policy responses," *Journal of Monetary Economics* 57, 404-419.

Ennis, Huberto M. and Todd Keister (2010b) "On the fundamental reasons for bank fragility," Federal Reserve Bank of Richmond *Economic Quarterly* 96, 33-58.

Ennis, Huberto M. and Todd Keister (2009) "Run equilibria in the Green-Lin model of financial intermediation," *Journal of Economic Theory* 144, 1996-2020.

Ennis, Huberto M. and Todd Keister (2006) "Bank runs and investment decisions revisited," *Journal of Monetary Economics* 53, 217-232.

Friedman, Milton and Anna J. Schwartz (1963) *A Monetary History of the United States*, Princeton University Press, Princeton, NJ.

Gorton, Gary and Andrew Metrick (2011) " Securitized banking and the run on repo," *Journal of Financial Economics*, forthcoming.

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

Green, Edward J. and Ping Lin (2000) “Diamond and Dybvig’s classic theory of financial intermediation: What’s missing?” Federal Reserve Bank of Minneapolis *Quarterly Review* 24, 3-13.

Gu, Chao (2011) “Noisy sunspots and bank runs,” *Macroeconomic Dynamics* 15, 398-418.

Jacklin, Charles (1987) “Demand deposits, trading restrictions, and risk sharing,” in: E. Prescott and N. Wallace, eds., *Contractual arrangements for intertemporal trade*, University of Minnesota Press, Minneapolis, MN, 26-47.

McCabe, Patrick E. (2010) “The cross section of money market fund risks and financial crises,” FEDS Working Paper No. 2010-51.

Nosal, Ed and Neil Wallace (2009) “Information revelation in the Diamond-Dybvig banking model,” Federal Reserve Bank of Chicago Policy Discussion Paper Series, December.

Peck, James and Karl Shell (2003) “Equilibrium bank runs,” *Journal of Political Economy* 111, 103-123.

Selgin, George (1993) “In defense of bank suspension,” *Journal of Financial Services Research* 7, 347-364.

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 (Fall), 3-16.

Wallace, Neil (1990) “A banking model in which partial suspension is best,” *Federal Reserve Bank of Minneapolis Quarterly Review* 14 (Fall), 11-23.