Bank Runs and the Optimality of Limited Banking*

James Peck
The Ohio State University

A. Setayesh
The Ohio State University

October 31, 2021

Abstract

We extend the Diamond-Dybvig model of bank runs to include a specification of *how much* to deposit. When the propensity to run is zero, we prove an equivalence result, that the efficient allocation (satisfying resource, IC, and sequential service constraints) can be achieved in equilibrium as long as the deposit level is above a threshold. Within this range, the lower the deposit level, the more tempted patient depositors are to withdraw early. When the propensity to run is positive and certain conditions are met, the optimal banking system entails less than full deposits and runs on the equilibrium path. We extend the analysis to consider a propensity to run that depends on the risk factor of the run equilibrium.

1. Introduction

In the many years and many published articles following the bank runs paper of Diamond and Dybvig (1983), only a few papers have modeled the decision

*An earlier version of this paper was titled, "A Diamond-Dybvig Model in which the Level of Deposits is Endogenous." We appreciate the helpful comments of the editor, two referees, Yaron Azrieli, Kyle Dempsey, Huberto Ennis, Chao Gu, Todd Keister, Karl Shell, and audiences at the September 2018 Cornell/Penn State Macroeconomics Confererence, at the March 2019 Diamond-Dybvig 36 Conference at Washington University, the Spring 2019 Midwest Economic Theory Conference, the 2019 North American Summer Meeting of the Econometric Society, and at Ohio State. Email addresses are peck.33@osu.edu (Peck) and setayesh.1@osu.edu (Setayesh).

of *whether* to deposit, much less the decision of *how much* to deposit. This is peculiar, considering the vast amount of wealth that is invested by financial firms that do not provide insurance against being an "impatient" consumer with immediate consumption needs. The questions we address here are, how does the opportunity for consumers to invest outside the banking system, in investments that do not provide liquidity insurance, (1) affect the nature of the final allocation, (2) affect the nature of the optimal deposit contract, and (3) affect the fragility of the banking system?

We consider a model with N ex ante identical consumers, who invest some of their endowment directly (i.e., outside of the banking system) and deposit the rest of their endowment with the bank in period 0, before observing whether they will become impatient (requiring consumption in period 1) or patient (able to consume in period 2). In the model, there is a single technology available to the bank and to investments outside the banking system. We use the term "outside" to describe investments outside the banking system. A contract specifies a deposit level, d, and period 1 withdrawals. Due to a sequential service constraint, withdrawals in period 1 depend on the consumer's place in line, but not on the number of consumers yet to withdraw in period 1.

We find that any allocation that the bank can induce (satisfying incentive compatibility and non-negativity constraints) with a given deposit level can also be induced with any higher deposit level. For the typical economy, the incentive compatibility constraint does not bind at the optimal contract when consumers deposit their full endowment. If so, we find that there is an interval of deposit levels yielding the efficient allocation in equilibrium. The intuition for this equivalence result is that the withdrawal amount augments the outside investment of impatient consumers by exactly the amount of consumption needed to achieve the optimal allocation. Since the optimal contract provides insurance against being impatient, by allowing impatient consumers more consumption than their endowment on average, the bank responds to lower deposits by magnifying the ratio of period 1 withdrawals to deposits. As we consider lower and lower deposit levels, at some point the efficient allocation cannot be achieved, due to a failure of incentive compatibility or non-negativity.

We also find that, in equilibria achieving the efficient allocation, lower deposit levels make the banking system more fragile. The lower the deposit level, the more tempted a patient consumer is to withdraw early in the no-run equilibrium, and the more tempted she is to join a run. The main reason is that, when deposits are lower, the bank must offer a larger proportion of its deposits as withdrawals in period 1 (i.e., more maturity transformation), in order to achieve the efficient allocation. A secondary reason is that withdrawals from the bank in period 1 are stored by a patient consumer until consuming in period 2, but she can keep her outside investment until period 2, receiving a higher return.

Our results have implications for the optimal size of the banking system. Under the conditions of our main result (Proposition 3), which are likely to hold especially for large economies, a limited banking system with less than full deposits is optimal. The reason is that the probability that a bank run occurs does not depend on whether deposits are full or partial; however, in the event of a run occurring, the misallocation of resources is smaller with partial deposits. Specifically, patient depositors do not have to liquidate their investments outside the banking system, so welfare is higher with partial deposits than full deposits when a run occurs.[1] When the model is adjusted to allow the propensity to run to vary with the "risk factor" of the run equilibrium, a new tradeoff emerges. Under the conditions of Proposition 3, as we reduce the deposit level starting from full deposits, welfare increases for a given propensity to run. However, the propensity to run can increase, which introduces an effect that lowers welfare.

Finally, we consider the extended model in an online appendix, with the specification that outside investment held until period 2 yields an ε higher return than bank investment held until period 2, where ε is small but can be either positive or negative. We show that our main result is robust to the introduction of nonzero ε. That is, under the conditions of Proposition 3, the optimal contract entails partial deposits and a positive probability of a bank run on the equilibrium path.

Section 2 contains a literature review. Section 3 sets up the model, and Section 4 contains the results for the model with a zero propensity to run. Section 5 contains our most important result, for the model with a positive propensity to run. Section 6 contains an analysis of the model in which the propensity to run is a function of the risk factor of the run equilibrium. Section 7 contains a summary and discussion.

2. Literature Review

Most of the papers in the Diamond-Dybvig literature simply assume that consumers have already deposited their entire endowment, and consider the optimal

[1] Andolfatto et al. (2020) also combine a banking system with financial markets, and find that non-bank financial markets improve welfare conditional on a bank run. The mechanism is quite different.

contract that would arise if a run is certain not to occur. There are only a few exceptions. Peck and Shell (2003) consider a model similar to the present model, except that (1) consumers decide whether or not to deposit their entire endowment with the bank, rather than how much to deposit, and (2) the utility functions of impatient and patient consumers can be different.[2] Examples are constructed in which any contract achieving the efficient allocation (highest welfare consistent with resource, incentive compatibility, and sequential service constraints) has a bank run equilibrium to the post-deposit subgame. Bank runs, triggered by sunspots, can occur on the equilibrium path if the probability of a run is small enough. Shell and Zhang (2018) provide a complete characterization of the 2 consumer version of Peck and Shell (2003). The present paper models the decision of *how much* to deposit. Also, we provide a result about when a run occurs on the equilibrium path, but we do not provide a complete characterization of the optimal contract for a positive propensity to run, like Peck and Shell (2003) and Shell and Zhang (2018) do. The connection is discussed further in Section 5.[3]

In Peck and Shell (2010), there are two technologies, one liquid and one illiquid. Both patient and impatient consumers care about "left-over" consumption, with the difference being that an impatient consumer receives an urgent consumption opportunity in period 1 and a patient consumer receives an urgent consumption opportunity in period 2. It is shown that legal restrictions that prevent the bank from accessing the illiquid technology lead to overinvestment in the liquid asset and run equilibria in the post-deposit subgame. In the separated system in Peck and Shell (2010), investment in the illiquid technology is outside the banking system, and the result is greater instability, as in the present model. Shell and Zhang (2019) model the pre-deposit game in a two-consumer version of Peck and Shell (2010), to encorporate runs on the equilibrium path and to characterize when the propensity to run affects the optimal contract (for the separated system). The main difference between these papers and the current paper is that these papers depart from the original Diamond-Dybvig model, in terms of utility functions and timing of impatient consumption. Similarly, Ennis and Keister (2003) consider an endogenous growth model with the Diamond-Dybvig timing within a generation. Banks invest in both storage and capital, and consumers can store income that is not deposited. Contracts are of a simple form that rules out suspension schemes.

[2] In their main model, only consumers wishing to withdraw arrive in period 1, but an example is presented in which all consumers arrive in random order to report their type, without observing their place in line.

[3] Kang (2020) extends this literature by introducing hyperbolic discounting.

The authors show that much of the welfare cost of bank runs fall on future generations, which is not taken into account by banks competing for deposits from the current generation. Thus, although there is a decision of how much to deposit, the model and questions are very different from those of the current paper. In a recent paper by Gu et al. (2019), a dynamic model is studied in which long-lived agents endogenously set up banks and are willing to honor their obligations in order to maintain their reputations. The simple bank contract specifies how much of the endowment is deposited and specifies constant withdrawal amounts in each period. A smaller deposit level reduces the bank's incentive to misbehave, while in our setting, a smaller deposit level increases a patient consumer's incentive to withdraw early. Interestingly, in one of their parameter regions, an equivalence result reminiscent of our Proposition 1 obtains.

Green and Lin (2003) consider a model in which all agents join the line in period 1 to report their type, and all agents know the "clock" time at which they join the line. It is shown that the efficient allocation is uniquely implemented, with no possibility of a bank run. Ennis and Keister (2009a) show that the Green and Lin (2003) implementation result is sensitive to observing the clock and the assumption of independent types. Ennis and Keister (2016) show that the efficient allocation may also have a bank run equilibrium in a model where only those seeking to withdraw in period 1 join the line, but where consumers know their place in line. Andolfatto et al. (2017) extend the space of possible reports, and they assume that all depositors visit the bank in period 1, without knowing their place in line. Essentially, a depositor reports whether she is patient or impatient, and also reports whether a run is in progress. With these more general mechanisms, the efficient allocation is uniquely implemented. A similar result is obtained by Cavalcanti and Monteiro (2016). While these implementation results are important, an essential feature of demand deposit accounts is their convenience, so it is useful to study models in which mechanisms do not involve reports, and an arrival at the bank is a request to withdraw. It would also be interesting to see whether endogenizing the deposit level affects any of these implementation results, or the results regarding commitment (see Ennis and Keister (2009b)).

3. The Model

We consider an otherwise standard Diamond-Dybvig economy, but where consumers decide how much of their endowment to deposit and how much to invest directly. There are three time periods and a single investment technology. Each unit of consumption invested in period 0, by either a consumer or the bank, yields 1 unit if harvested in period 1, and $R > 1$ units if harvested in period 2.

There is a finite number, N, of consumers. In period 0, each consumer receives an endowment of the consumption good, normalized to 1. In period 1, each consumer will privately observe whether she is impatient or patient. An impatient consumer only derives utility from consumption in period 1, and a patient consumer derives utility from consumption in period 2. A patient consumer can costlessly store consumption received in period 1 to period 2. Thus, denoting the consumption received in period 1 by x_1 and the consumption received in period 2 by x_2, an impatient consumer receives utility $u(x_1)$ and a patient consumer receives utility $u(x_1 + x_2)$. As is standard, we assume $u'(x) > 0$, $u''(x) < 0$, and $\lim_{x \to 0} u'(x) = \infty$.

The number of impatient consumers, denoted by α, is a random variable with probability distribution $f(\alpha)$. Impatience can be i.i.d. or correlated across consumers. It is assumed that the distribution of α, *conditional* on being patient, denoted by $f_p(\alpha)$, is the same for all consumers. From Bayes' rule, we have

$$f_p(\alpha) = \frac{(1 - \frac{\alpha}{N})f(\alpha)}{\sum_{a=0}^{N-1}(1 - \frac{a}{N})f(a)}.$$

Here is the timing of the game. At the beginning of period 0, the bank chooses a deposit contract, which specifies a deposit level, d, and period-1 and period-2 withdrawals, C, which we fully describe below. Next, consumers decide whether or not to deposit; any endowment not invested in the bank is invested outside of the bank. At the beginning of period 1, consumers observe their type, impatient or patient, and a public "sunspot" variable, σ. Next, each consumer decides whether or not to withdraw in period 1. Those who withdraw in period 1 arrive in random order, with all orders treated as equally likely at the time of the withdrawal choice. Arriving in period 1 can be interpreted as a statement of impatience, but no formal reports are made. Those who do not withdraw in period 1 do not contact the bank until period 2.[4] We assume that a consumer can, if she wishes, liquidate her outside investment in period 1, to provide consumption over and above whatever she may withdraw from the bank. However, as is standard in the

[4] For simplicity and realism, we do not allow mechanisms where the bank requires all consumers to contact the bank in period 1 to report their types, even patient consumers who do not want to withdraw. See Peck and Shell (2003) for a discussion. Also see Green and Lin (2003), Cavalcanti and Monteiro (2016), Andolfatto et al. (2017), and the appendix of Peck and Shell (2003) for analysis of more complicated mechanisms.

literature, we rule out additional markets in which consumption and unharvested investments are traded in period 1.[5]

Based on the idea that the banking system is competitive, we model the bank as choosing a deposit contract to maximize the expected utility of their depositors. Due to a sequential service constraint, withdrawals in period 1 can only be a function of the history of previous withdrawals. Withdrawals can be characterized by a pair of non-negative functions, $c_1(z, d)$ and $c_2(\alpha_1, d)$. For $z = 1, ..., N$, $c_1(z, d)$ is the withdrawal amount for the z^{th} consumer to arrive in period 1, when all consumers deposit d units.[6] For $\alpha = 0, ..., N-1$, $c_2(\alpha_1, d)$ is the withdrawal amount in period 2, when all consumers deposit d units and the number of consumers withdrawing in period 1 is α_1. The specification, that each consumer withdrawing in period 2 receives the same consumption, follows from expected utility maximization and consumption smoothing. Also due to expected utility maximization, the optimal contract will fully allocate the bank's resources:

$$c_1(N, d) = dN - \sum_{z=1}^{N-1} c_1(z, d), \tag{3.1}$$

$$c_2(\alpha_1, d) = \frac{[dN - \sum_{z=1}^{\alpha_1} c_1(z, d)]R}{N - \alpha_1}. \tag{3.2}$$

We now characterize the bank's optimal contract by solving the associated planner's problem, of maximizing welfare subject to the resource constraints (3.1) and (3.2), and the incentive compatibility constraint, that a patient consumer weakly prefers to wait until period 2, given that other patient consumers also wait. This is the optimal contract if the no-run equilibrium, in which the patient consumers withdraw in period 2, is selected in the post-deposit subgame, for all realizations of σ.

Notice that an impatient consumer withdrawing c_1 from the bank in period 1 receives utility, $u(1 - d + c_1)$. Also notice that a patient consumer withdrawing c_1 from the bank in period 1 will store this consumption until period 2 and liquidate her outside investment in period 2, thereby receiving utility, $u((1 - d)R + c_1)$. A patient consumer withdrawing c_2 from the bank in period 2 will liquidate her

[5] Jacklin (1987) has shown that allowing such markets undermines the ability of the bank to provide insurance against being impatient.

[6] Our notation presumes that all N consumers will choose to deposit, which must hold in equilibrium. Also, it will be convenient to keep track of the deposit level d in notation for withdrawals.

outside investment in period 2, thereby receiving utility, $u((1-d)R+c_2)$. Given the contract, it will be convenient to use the notation, $x_1(z,d) \equiv 1-d+c_1(z,d)$ to denote the overall consumption of an impatient consumer who withdraws in position z in period 1 and $x_2(\alpha_1,d) \equiv (1-d)R+c_2(\alpha_1,d)$ to denote the overall consumption of a patient consumer who withdraws in period 2 when the number of consumers withdrawing in period 1 is α_1. Then we can write the welfare associated with a contract, (C,d), when the patient consumers withdraw in period 2, in which case $\alpha_1=\alpha$ holds, as

$$\begin{aligned} \widehat{W}(C,d) &= \sum_{\alpha=0}^{N-1} f(\alpha)\left[\sum_{z=1}^{\alpha} u(x_1(z,d)) + (N-\alpha)u(x_2(\alpha,d))\right] \\ &\quad + f(N)\left[\sum_{z=1}^{N} u(x_1(z,d))\right]. \end{aligned}$$

Given the contract, the incentive compatibility constraint for a patient consumer is given by

$$\sum_{\alpha=0}^{N-1} f_p(\alpha)\left[\frac{1}{1+\alpha}\sum_{z=1}^{\alpha+1} u((1-d)(R-1)+x_1(z,d))\right] \leq \sum_{\alpha=0}^{N-1} f_p(\alpha)u(x_2(\alpha,d)). \tag{3.3}$$

The reason for the term, $(1-d)(R-1)$, in (3.3) is that a patient consumer who withdraws in period 1 receives this additional consumption because her outside investment is harvested in period 2 rather than period 1.

4. Optimal Contracts with a Zero Propensity to Run

The bank's optimal contract, when the propensity to run is zero, specifies a deposit level and withdrawals that solve the following problem.

$$\begin{aligned} & \max \widehat{W}(C,d) \\ & \text{subject to} \\ & (3.1), (3.2), (3.3) \\ c_1(z,d) \geq\ & 0 \quad \text{for all } z \end{aligned} \tag{4.1}$$

Sometimes it is useful to consider the withdrawals that solve (4.1) for fixed d. The solution to (4.1) for $d=1$ is the solution to the planner's problem studied in the previous literature in which consumers deposit their entire endowment. See, for example, Peck and Shell (2003) and Shell and Zhang (2018). Denote the correponding efficient allocation by $x^* = \{x_1^*(z)|_{z=1}^N, x_2^*(\alpha)|_{\alpha=0}^{N-1}\}$.[7]

When impatient and patient consumers have the same utility function, given by $u(\cdot)$ here, incentive compatibility often does not bind when consumers deposit their entire endowment. Lemma 1, below, shows that any allocation satisfying the constraints in (4.1) with deposit level d' must satisfy the constraints with deposit level $d'' > d'$. It follows that the allocation yielding the highest welfare across all PBE of the game is x^*. Proposition 1 shows that, if incentive compatibility does not bind when consumers deposit their entire endowment, then for any d sufficiently close to 1, there is a PBE with deposit level d yielding the efficient allocation, x^*. The intuition for this equivalence result is that the bank can negate the effect of outside investment, by subtracting it from x^* and allowing consumers to withdraw the difference.

Lemma 1: *Suppose that, for deposit level d', there is a contract satisfying the constraints in (4.1) yielding the allocation $x = \{x_1(z)|_{z=1}^N, x_2(\alpha)|_{\alpha=0}^{N-1}\}$ if patient consumers withdraw in period 2. Then, for all deposit levels $d'' > d'$, there is a contract satisfying the constraints in (4.1) yielding the same allocation x.*

Proof. Since the contract for deposit level d' yields the allocation x, it must be given by

$$\begin{aligned} c_1(z, d') &= x_1(z) - (1 - d'), && (4.2) \\ c_2(\alpha_1, d') &= x_2(\alpha_1) - (1 - d')R. && (4.3) \end{aligned}$$

The contract for the deposit level d'' yields the allocation x if and only if it is given by

$$\begin{aligned} c_1(z, d'') &= x_1(z) - (1 - d''), && (4.4) \\ c_2(\alpha_1, d'') &= x_2(\alpha_1) - (1 - d'')R. && (4.5) \end{aligned}$$

We first show that if (4.2) and (4.3) satisfy the resource constraints (3.1) and (3.2), then (4.4) and (4.5) satisfy the resource constraints. From (3.1) and (4.2),

[7] We refer to x^* as the efficient allocation, and uniqueness is guaranteed under CRRA utility and independent types. However, our results go through for any efficient allocation if x^* is not unique.

we have

$$c_1(N,d') = d'N - \sum_{z=1}^{N-1}[x_1(z) - (1-d')] = x_1(N) - (1-d'),$$

which can be simplified to

$$\sum_{z=1}^{N} x_1(z) = N. \tag{4.6}$$

For deposit level d'', (3.1) is satisfied if and only if we have

$$x_1(N) - (1-d'') = d''N - \sum_{z=1}^{N-1}[x_1(z) - (1-d'')],$$

which is equivalent to (4.6).

From (3.2) and (4.3), we have

$$c_2(\alpha_1, d') = \frac{d'NR - \sum_{z=1}^{\alpha_1}[x_1(z) - (1-d')]R}{N-\alpha_1} = x_2(\alpha_1) - (1-d')R,$$

which can be simplified to

$$NR = R\sum_{z=1}^{\alpha_1} x_1(z) + (N-\alpha_1)x_2(\alpha_1). \tag{4.7}$$

For deposit level d'', (3.2) is satisfied if and only if we have

$$x_2(\alpha_1) - (1-d'')R = \frac{d''NR - \sum_{z=1}^{\alpha_1}[x_1(z) - (1-d'')]R}{N-\alpha_1},$$

which is equivalent to (4.7). Thus, (4.4) and (4.5) satisfy the resource constraints.

Since (4.2) and (4.3) satisfy the incentive compatibility constraint, we have

$$\sum_{\alpha=0}^{N-1} f_p(\alpha)\left[\frac{1}{1+\alpha}\sum_{z=1}^{\alpha+1} u((1-d')(R-1) + x_1(z))\right] \leq \sum_{\alpha=0}^{N-1} f_p(\alpha)u(x_2(\alpha)).$$

Because utility is strictly increasing and $d'' > d'$ holds, it follows that

$$u((1-d')(R-1) + x_1(z)) > u((1-d'')(R-1) + x_1(z))$$

holds, which implies

$$\sum_{\alpha=0}^{N-1} f_p(\alpha)\left[\frac{1}{1+\alpha}\sum_{z=1}^{\alpha+1} u((1-d'')(R-1)+x_1(z))\right] < \sum_{\alpha=0}^{N-1} f_p(\alpha)u(x_2(\alpha)).$$

Thus, the incentive compatibility constraint with deposit level d'' holds (and is not binding).

From (4.2), (4.3), (4.4), and (4.5), and from the fact that $d'' > d'$ holds, it follows immediately that the non-negativity constraints being satisfied with deposit level d' implies that the non-negativity constraints must be satisfied with deposit level d''. Thus, we have constructed a contract with deposit level d'', satisfying non-negativity and incentive compatibility, yielding the allocation x if the patient withdraw in period 2. ■

Proposition 1 (Equivalence): *If IC is not binding when consumers deposit their entire endowment, so*

$$\sum_{\alpha=0}^{N-1} f_p(\alpha)\left[\frac{1}{1+\alpha}\sum_{z=1}^{\alpha+1} u(x_1^*(z))\right] < \sum_{\alpha=0}^{N-1} f_p(\alpha)u(x_2^*(\alpha)) \tag{4.8}$$

holds, then there exists $d^* < 1$ such that, for all $d > d^*$, there is a contract satisfying the constraints in (4.1) yielding the allocation x^*.

Proof. Consider the allocation x^* and suppose that the incentive compatibility constraint for $d = 1$ holds strictly. By continuity of the utility function,

$$\sum_{\alpha=0}^{N-1} f_p(\alpha)\left[\frac{1}{1+\alpha}\sum_{z=1}^{\alpha+1} u((1-d)(R-1)+x_1^*(z))\right] < \sum_{\alpha=0}^{N-1} f_p(\alpha)u(x_2^*(\alpha)) \tag{4.9}$$

holds for all d close enough to 1. Also, since utility is strictly increasing in consumption, the left side of (4.9) is strictly decreasing in d. Let $d_{IC}^* < 1$ denote the value of d such that (4.9) holds as an equality, if such a solution exists, and let $d_{IC}^* = 0$ if the right side always exceeds the left side.

Since $\lim_{x\to 0} u(x) = \infty$ holds, we have $x_1^*(z) > 0$ for all z and $x_2^*(\alpha) > 0$ for all α. The non-negativity constraints are given by

$$\begin{aligned} c_1(z,d) &= x_1^*(z) - (1-d) \geq 0 \text{ for all } z \text{ and} \\ c_2(\alpha_1,d) &= x_2^*(\alpha_1) - (1-d)R \geq 0 \text{ for all } \alpha_1. \end{aligned}$$

The left side of each inequality is increasing in d and strictly positive for $d = 1$. It follows that there is a deposit level, $d^*_{NN} < 1$, above which all non-negativity constraints are satisfied. Define $d^* = \max[d^*_{IC}, d^*_{NN}]$. ■

In preparation for the analysis of bank runs that occur with positive probability in equilibrium, in the next section, we define here the notion of the temptation to join a run, and the notion of a run equilibrium to the post-deposit subgame.

Definition 1: *Consider an economy in which incentive compatibility is not binding at x^* when consumers deposit their entire endowment, so (4.8) holds. For $d \geq d^*$, consider the contract achieving the optimal allocation x^* (when the patient consumers wait). The* ***temptation to join a run*** *is defined to be the utility advantage of withdrawing in period 1, relative to withdrawing in period 2, when all other consumers withdraw in period 1, given by*

$$\frac{1}{N}\sum_{z=1}^{N} u((1-d)(R-1) + x_1^*(z)) - u(x_2^*(N-1)). \tag{4.10}$$

An equilibrium of the post-deposit subgame in which all consumers withdraw in period 1 is called a ***run equilibrium****.*

The post-deposit game has a run equilibrium if and only if the temptation to join a run is non-negative.[8] Consider contracts yielding the optimal allocation x^*, for various deposit levels, $d \geq d^*$. Since utility is increasing, expression (4.10) is strictly decreasing in d, so the lower the deposit level the higher the temptation to join a run. Define d^*_{NR} to be the minimum deposit level above which the optimal allocation, x^* can be uniquely implemented. There are three possible mutually exclusive cases. First, if there exists a threshold, $d^*_{NR} \in [d^*, 1)$, for which expression (4.10) equals zero, then the post-deposit subgame has a run equilibrium for $d \leq d^*_{NR}$, and the optimal allocation x^* can be uniquely implemented for any $d > d^*_{NR}$. Second, if the temptation to join a run is non-negative for $d = 1$, then the post-deposit subgame has a run equilibrium for all $d \geq d^*$; in this case, define $d^*_{NR} = 1$. Third, if the temptation to join a run is negative for $d = d^*$, then the optimal allocation x^* can be uniquely implemented for any $d \in [d^*, 1]$; in this case, define $d^*_{NR} = d^*$. The argument in this paragraph establishes the following.

[8] The magnitude of the temptation to join a run, and not simply whether it is positive or negative, is related to the notion of the risk factor of the run equilibrium, which is considered in Section 6.

Proposition 2: *Consider an economy in which incentive compatibility is not binding at x^* when consumers deposit their entire endowment, so (4.8) holds. Then for all d', d'' such that $d^* < d' < d''$ holds, the temptation to join a run is higher with deposit level d' than with deposit level d''. The post-deposit subgame has a run equilibrium for $d \leq d^*_{NR}$, and the optimal allocation x^* can be uniquely implemented for any $d > d^*_{NR}$.*

Consider the typical economy, where incentive compatibility is not binding at x^* when consumers deposit their entire endowment. Ennis and Keister (2016) have shown that, for CRRA utility, $x_1^*(z)$ is strictly decreasing in z. This captures the idea that the optimal contract offers liquidity insurance to the impatient, so it would be surprising not to see this pattern more generally. The following Corollary to Proposition 2 establishes that, if this monotonicity holds and non-negativity binds before incentive compatibility, then there is a range of deposit levels for which the post-deposit game has a run equilibrium (the optimal contract is "fragile" according to the terminology of Ennis and Keister).

Corollary to Proposition 2: *Consider an economy in which (4.8) holds. Also assume that $x_1^*(z)$ is strictly decreasing in z, and that $d^*_{IC} \leq d^*_{NN}$ holds. Then we have $d^* < d^*_{NR}$, and the post-deposit subgame has a run equilibrium for deposit levels between d^* and d^*_{NR}.*

Proof. At deposit level d^*, non-negativity binds. Since $x_1^*(z)$ is strictly decreasing in z, it follows that $c_1(z, d^*)$ is strictly decreasing in z, so we must have $c_1(N, d^*) = 0$. Expression (4.10) must be positive, because if all other consumers withdraw in period 1, we have $c_1(N, d^*) = 0$, and therefore, $c_2(N-1, d^*) = 0$. Waiting until period 2 ensures a withdrawal of zero, so joining a run must yield higher utility. By continuity, the post-deposit subgame has a run equilibrium for deposit levels greater than, but sufficiently close to, d^*. ■

5. Main Result: Optimal Contracts with a Positive Propensity to Run

In this section, we consider the model with a positive propensity to run. It is without loss of generality to assume that the sunspot variable is uniformly distributed over the unit interval. Then, given a contract that satisfies (3.1), (3.2), and (3.3), a propensity to run of s means that, if the post-deposit subgame has a run equilibrium, all patient consumers withdraw in period 1 whenever $\sigma \leq s$ holds, and all patient consumers withdraw in period 2 whenever $\sigma > s$ holds. If the post-deposit subgame does not have a run equilibrium, all patient consumers withdraw in period 2 for all realizations of σ.

Welfare, conditional on a run not taking place, is $\widehat{W}(C,d)$. Welfare, conditional on a run taking place, is denoted by $W^R(C,d)$, given by

$$W^R(C,d) = \sum_{\alpha=0}^{N} f(\alpha) \left[\frac{\alpha}{N} \sum_{z=1}^{N} u(x_1(z,d)) + \frac{N-\alpha}{N} \sum_{z=1}^{N} u((1-d)(R-1) + x_1(z,d)) \right].$$

The last term in brackets is due to the fact that a patient consumer who runs does not liquidate her outside investment until period 2, and therefore receives the additional consumption, $(1-d)(R-1)$. Overall welfare is given by $W(C,d,s)$, where we have

$$\begin{aligned} W(C,d,s) &= (1-s)\widehat{W}(C,d) + sW^R(C,d) \qquad \text{if } (C,d) \text{ allows a run equilibrium} \\ &= \widehat{W}(C,d) \qquad\qquad\qquad\qquad \text{if } (C,d) \text{ does not allow a run equilibrium.} \end{aligned}$$

The optimal contract, taking into account the propensity to run, is the solution to the problem of choosing (C,d) to maximize $W(C,d,s)$, subject to (3.1), (3.2), and (3.3). We refer to this contract as the s-optimal contract.

If $d^*_{NR} < 1$ holds, the situation is quite simple. Any deposit level, $d > d^*_{NR}$, along with the withdrawal schedule yielding the allocation x^*, is an s-optimal contract for $s > 0$. That is, the optimal allocation x^* can be implemented, without risk of a run equilibrium for the post-deposit subgame. Deposit levels between d^* and d^*_{NR}, where the post-deposit subgame also has a run equilibrium, are optimal if the propensity to run is zero, but they cannot be s-optimal for $s > 0$.

Our main result considers the more interesting case, in which $d^*_{NR} = 1$ holds. In other words, if the bank is constrained to require full deposits, then the optimal contract, which yields the allocation x^* in the no-run equilibrium of the post-deposit subgame, also has a run equilibrium. From the work of Ennis and Keister (2009a) and others, we now know that this case is very likely to hold when the number of consumers is fairly large.

If $d^*_{NR} = 1$ holds, the s-optimal contract may depend on s, and welfare will be strictly less than the welfare associated with the allocation, x^*.[9] If s is large,

[9] Welfare will be strictly less, because either (i) a run will occur on the equilibrium path, and

then the s-optimal contract will eliminate the possibility of a run, by reducing the amount of insurance offered against being impatient. For smaller s, then the s-optimal contract will tolerate a positive probability of runs on the equilibrium path. Shell and Zhang (2018) perform an elegant and thorough analysis of economies with two consumers and full deposits, characterizing the optimal contract that takes into account a positive propensity to run. An important takeaway from Shell and Zhang (2018) is that, when the incentive compatibility constraint is not binding at the optimal contract with a zero propensity to run, then for a small positive propensity to run, the optimal overall contract balances welfare when a run occurs and welfare when a run does not occur. Thus, the probability of a run affects the contract. Here, for $d > d^*$, incentive compatibility does not bind, and we would expect to see the Shell and Zhang (2018) finding that the run probability affects the optimal contract.

It will be useful to adopt the notation, (s, d)-optimal contract to denote the optimal contract when the propensity to run is s and the deposit level is constrained to be d. Thus, an $(s, 1)$-optimal contract is the optimal withdrawal scheme when the propensity to run is s and the bank is constrained to require full deposits, as in the previous literature. Our main result, Proposition 3 below, provides conditions under which the optimal banking system entails partial deposits and runs on the equilibrium path. These conditions are likely to hold when the number of consumers is fairly large and s is fairly small.[10] Note that, when s is very small, the condition, that an $(s, 1)$-optimal contract is such that the post-deposit subgame has a run equilibrium, can be expressed simply as $d^*_{NR} = 1$.[11]

Proposition 3: *Suppose that any $(s, 1)$-optimal contract is such that (i) incentive compatibility is not binding and (ii) the post-deposit subgame has a run equilibrium. Then, when the deposit level is not constrained, any s-optimal contract*

when a run occurs, welfare is even less than welfare under autarky, or (ii) the contract will be altered so that the post-deposit subgame does not have a run equilibrium, in which case welfare is below the level associated with x^*.

[10] Here is the intuition. Suppose n is very large, s is very small, and impatience is i.i.d. Then the s-optimal contract will anticipate the no-run equilibrium and a known fraction of impatient. Under the standard assumption that the coefficient of relative risk aversion is greater than one, period 1 withdrawals will be greater than deposits until the "known" fraction of withdrawals occurs. If more withdrawals occur, the bank is nearly certain that period 1 withdrawals will soon cease, so the contract continues to offer generous withdrawals. Thus, the contract admits a run equilibrium, since very little will be left for period 2.

[11] In that case, it can be shown that, as s approaches zero, the s-optimal deposit level approaches d^* and the allocation in the no-run equilibrium approaches x^*.

entails less than full deposits, and a run occurs with positive probability on the equilibrium path.

Proof of Proposition 3. First, we show that any s-optimal contract must exhibit partial deposits, $d < 1$. Let $C^{s,1}$ be the withdrawal schedule in an $(s, 1)$-optimal contract, and denote by $x^{(s,1)*}$ the allocation that arises in the corresponding no-run equilibrium.[12] Since incentive compatibility is not binding, an argument identical to that in the proof of Proposition 1 establishes that there is a contract (C', d'), where we have $d' < 1$, satisfying incentive compatibility and nonnegativity, and yielding the allocation $x^{(s,1)*}$ in the corresponding no-run equilibrium. The withdrawal schedule in this contract is given by

$$\begin{aligned} c_1(z, d') &= x_1^{(s,1)*}(z) - (1 - d') \\ c_2(\alpha_1, d') &= x_2^{(s,1)*} - (1 - d')R. \end{aligned}$$

In the no-run equilibrium, which occurs with probability $(1 - s)$, consumption is the same across both contracts, $(C^{s,1}, 1)$ and (C', d'). In the run equilibrium, which occurs with probability s under both contracts, the patient consumers consume more under (C', d') than under $(C^{s,1}, 1)$, by exactly $(1 - d')(R - 1)$. Thus, the contract (C', d') provides higher welfare than $(C^{s,1}, 1)$, so the full deposit contract $(C^{s,1}, 1)$ cannot be s-optimal.

Next, we show that any s-optimal contract tolerates runs with positive probability. Suppose not, so there is an s-optimal contract, (C'', d''), that does not have an associated run equilibrium. We demonstrated above that $d'' < 1$ must hold. From Lemma 1, there is another contract, $(C, 1)$, with full deposits, satisfying incentive compatibility and non-negativity constraints, and yielding the same allocation in the no-run equilibrium as (C'', d''). If we were to use the allocation in the no-run equilibrium (instead of x^*) in the expression for the temptation to join a run in Definition 1, it is clear that the temptation to join a run is smaller with $(C, 1)$ than with (C'', d''). Thus, since (C'', d'') does not have a run equilibrium, neither does $(C, 1)$. It follows that $(C, 1)$ yields the same welfare as (C'', d''), and is therefore an s-optimal contract, contradicting the fact that an s-optimal contract must have a deposit level less than one. ■

[12]Since the propensity to run is positive, in general, $x^{(s,1)*}$ will be a different allocation than x^*.

6. Allowing the Run Probability to Depend on the Risk Factor

The previous analysis assumes that the propensity to run is constant. However, it stands to reason that the probability of selecting an equilibrium of the post-deposit subgame could depend on the payoffs associated with the two equilibria, as well as the payoffs associated with being "out of step" and anticipating the wrong equilibrium. Ennis and Keister (2005a, 2005b) consider government policies in games with multiple equilibia. They define the *risk factor* of an equilibrium to be the smallest probability p such that, if a player believes that the equilibrium is selected with at least probabilty p, then the player prefers to play according to the equilibrium. The idea is that the higher the risk factor, the lower the probability of the equilibrium being selected. We are interested in the extent to which the conclusions of Proposition 3 are robust to a propensity to run that is not constant.

In our setting, let U_{rr} denote the expected utility of a patient consumer who withdraws in period 1 when all other patient consumers run; let U_{rn} denote the expected utility of a patient consumer who withdraws in period 1 when all other patient consumers wait; let U_{nr} denote the expected utility of a patient consumer who waits when all other patient consumers run; and let U_{nn} denote the expected utility of a patient consumer who waits when all other patient consumers wait. Then the risk factor of the run equilibrium is given by

$$p_r = \frac{U_{nn} - U_{rn}}{(U_{rr} - U_{nr}) + (U_{nn} - U_{rn})}$$

and the risk factor of the no-run equilibrium is given by $p_n = 1 - p_r$. The expressions for these utilities are given by

$$\begin{aligned}
U_{rr} &= \frac{1}{N}\sum_{z=1}^{N} u((1-d)(R-1) + x_1(z,d)), \\
U_{rn} &= \sum_{\alpha=0}^{N-1} f_p(\alpha)\left[\frac{1}{1+\alpha}\sum_{z=1}^{\alpha+1} u((1-d)(R-1) + x_1(z,d))\right], \\
U_{nr} &= u(x_2(N-1,d)), \\
U_{nn} &= \sum_{\alpha=0}^{N-1} f_p(\alpha) u(x_2(\alpha,d)).
\end{aligned}$$

We assume that the propensity to run is a weakly decreasing and differentiable function of the risk factor, denoted by $s(p_r)$.[13] Although not required for the analysis, it is natural to impose $s(0) = 1$ and $s(1) = 0$. An important new element is that, since the contract affects the risk factor, it will affect the probability that each equilibrium is selected.

Denote by C^d an optimal withdrawal schedule when the deposit level is constrained to be d. That is, C^d maximizes welfare, holding d fixed but taking into account the effect on the risk factor and therefore the probability of the run equilibrium. Since we are assuming that the post-deposit subgame has a run equilibrium, the fully optimal contract is of the form (C^d, d), where d maximizes

$$(1 - s(p_r))\widehat{W}(C^d, d) + s(p_r)W^R(C^d, d). \tag{6.1}$$

Differentiating (6.1) with respect to d yields

$$\begin{aligned} &\frac{\partial s(p_r)}{\partial p_r}\frac{\partial p_r}{\partial d}\left[W^R(C^d, d) - \widehat{W}(C^d, d)\right] \\ &+(1 - s(p_r))\frac{\partial \widehat{W}(C^d, d)}{\partial d} + s(p_r)\frac{\partial W^R(C^d, d)}{\partial d}. \end{aligned} \tag{6.2}$$

If expression (6.2) is negative, evaluated at $d = 1$, then our main result that the optimal banking system entails partial deposits continues to hold when the propensity to run depends on the risk factor.

When is expression (6.2) negative, evaluated at $d = 1$? If the assumptions of Proposition 3 hold and $s(p_r)$ is positive, it follows from Proposition 3 that the expressions on the bottom row of (6.2) sum to a negative number. The first expression in (6.2) is likely to be positive, since $\frac{\partial s(p_r)}{\partial p_r}$ is non-negative, $\frac{\partial p_r}{\partial d}$ is likely to be negative,[14] and welfare in the no-run equilibrium exceeds welfare in the run equilibrium. The function, $s(p_r)$, and its derivative are determined by factors outside the model itself. They reflect how the behavior of consumers in choosing between equilibria responds to the risk factor, and how quickly behavior changes as the risk factor increases. Expression (6.2), evaluated at $d = 1$, will be

[13] Ennis and Keister (2005b) show that a learning process leads to selection probabilities of this sort. That is, the equilibrium with the lower risk factor (the risk-dominant equilibrium) is not selected with probability one.

[14] As d increases, p_r increases, holding the allocation fixed. We believe that it is unlikely that the indirect effect of d on the run and no-run allocations, which in turn affects p_r, could dominate the direct effect. However, in the unlikely event that $\frac{\partial p_r}{\partial d}$ is positive, then it would follow that expression (6.2) is negative and the optimal banking system entails partial deposits.

negative when $\frac{\partial s(p_r)}{\partial p_r}$ is sufficiently small, given that $s(p_r)$ is small enough that the assumptions of Proposition 3 hold. In the extreme case, when $s(p_r)$ is constant and small enough for C^1 to admit a run equilibrium to the post-deposit subgame, the first term in (6.2) is zero, and we are back to the setting of Proposition 3.

Notice that a new tradeoff emerges when the propensity to run is not constant. Let us impose the assumptions of Proposition 3, that C^1 admits a run equilibrium to the post-deposit subgame and that incentive compatibility does not bind. Then reducing d below unity raises welfare conditional on a run taking place (because the patient are able to receive the return of R on their outside investment). However, reducing d alters the risk factor and makes the run equilibrium more likely to occur.

7. Summary and Discussion

We have extended the Diamond-Dybvig model to include a choice of *how much* to deposit. For typical economies studied in the literature in which impatient and patient consumers have the same utility function and consumers deposit their entire endowments, incentive compatibility does not bind in the no-run equilibrium under the optimal contract. If this is the case, then the efficient allocation, x^*, can also be achieved in an equilibrium of the present model, where consumers deposit only a fraction of their endowments with the bank, according to the equivalence result given in Proposition 1. The bank offers a contract that magnifies the extent to which impatient consumers can withdraw more than their deposit (on average), in order to provide the allocation x^* once the non-deposited investments are considered. However, the less consumers deposit, the more tempted patient consumers are to withdraw early.

Our main result, Proposition 3, considers a positive propensity to run, s. Suppose that, taking into account the propensity to run, the $(s, 1)$-optimal contract is such that incentive compatibility is not binding and there exists a run equilibrium in the post-deposit subgame. In that case, we show that any s-optimal contract specifies partial deposits and tolerates a positive probability of runs on the equilibrium path. These conditions are likely to be met when the number of consumers is not too small and when s is not too large.

In this paper, our equilibrium selection is based on a single parameter, the propensity to run, s. More generally, the propensity to run could depend on the deposit level. We follow Ennis and Keister (2005a, 2005b) by allowing the propensity to run to be a decreasing function of the risk factor of the run equilibrium.[15] We show that limited banking is optimal under certain conditions. If this is the case, then a system that allows outside investments can actually be more efficient, but more vulnerable to runs, than a system that does not allow outside investments! One might think that when depositors are allowed to invest a fraction of their endowments outside the banking system, they would be hedging against the risk of a run occurring, but losing out on some of the services provided by banks. Thus, one might think that this would improve the stability of the financial system at the expense of lost efficiency. However, the opposite could be true, with reduced stability (runs more likely) but higher efficiency!

In the online appendix, we present the extended model in which the difference between the return on outside investment and the return on bank investment is ε. We show that the conclusion of Proposition 3 is robust to differences between the returns on bank investments and outside investments. Whether ε is positive or negative, the conclusions extend as long as ε is small in magnitude. Even when ε is negative, so that bank investments have a higher return, then any s-optimal contract specifies partial deposits. The intuition is that, when ε is negative, reducing the deposit level below one increases welfare when a run takes place (since patient consumers keep more investments until period 2), but reduces welfare when a run does not take place (since bank investments yield a higher return than outside investments). When ε is small enough in magnitude, the overall effect is to increase welfare.

References

[1] Andolfatto, David, Aleksander Berentsen, and Fernando M. Martin. 2020. "Money, Banking, and Financial Markets." *Review of Economic Studies* 87(5): 2049-2086.

[2] Andolfatto, David, Ed Nosal, and Bruno Sultanum. 2017. "Preventing Bank Runs." *Theoretical Economics* 12: 1003-1028.

[15] In the same spirit, Ennis and Keister (2003) extend their analysis to allow the probability of a run to depend on the payoff of joining a run. The bank responds by holding a more illiquid portfolio than what it would hold otherwise.

[3] Cavalcanti, Ricardo and Paulo Klinger Monteiro. 2016. "Enriching Information to Prevent Bank Runs." *Economic Theory* 62: 477-494.

[4] Diamond, Douglas W. and Phillip H. Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy*, 91(3): 401-19.

[5] Ennis, Huberto M. and Todd Keister. 2003. "Economic Growth, Liquidity, and Bank Runs." *Journal of Economic Theory* 109: 220-245.

[6] Ennis, Huberto M. and Todd Keister. 2005a. "Optimal Fiscal Policy Under Multiple Equilibria." *Journal of Monetary Economics* 52: 1359-1377.

[7] Ennis, Huberto M. and Todd Keister. 2005b. "Government Policy and the Probability of Coordination Failures." *European Economic Review* 49: 939-973.

[8] Ennis, Huberto M. and Todd Keister. 2009a. "Run Equilibria in the Green-Lin Model of Financial Intermediation." *Journal of Economic Theory* 144: 1996-2020.

[9] Ennis, Huberto M. and Todd Keister. 2009b. "Bank Runs and Institutions: The Perils of Intervention." *American Economic Review* 99(4): 1588-1607.

[10] Ennis, Huberto M. and Todd Keister. 2016. "Optimal Banking Contracts and Financial Fragility." *Economic Theory* 61: 335-363.

[11] Gertler, Mark and Nobuhiro Kiyotaki. 2015. "Banking, Liquidity, and Bank Runs in an Infinite Horizon Economy." *American Economic Review* 105(7): 2011-2043.

[12] Green, Edward J. and Ping Lin. 2003. "Implementing Efficient Allocations in a Model of Financial Intermediation." *Journal of Economic Theory* 109: 1-23.

[13] Gu, Chao, Cyril Monnet, Ed Nosal, and Randall Wright. 2019. "On the Instability of Banking and Financial Intermediation." Working paper.

[14] Jacklin, Charles J. "Demand Deposits, Trading Restrictions, and Risk Sharing." In *Contractural Arrangements for Intertemporal Trade*, edited by Edward J. Prescott and Neil Wallace. Minneapolis: University of Minnesota Press, 1987.

[15] Kang, Minwook. 2020. "Demand deposit contracts and bank runs with present biased preferences." *Journal of Banking and Finance* 119: Article 105901.

[16] Peck, James and Karl Shell. 2003. "Equilibrium Bank Runs." *Journal of Political Economy*, 111(1): 103-23.

[17] Peck, James and Karl Shell. 2010. "Could Making Banks Hold Only Liquid Assets Induce Bank Runs?" *Journal of Monetary Economics*, 57: 420-427.

[18] Shell, Karl and Yu Zhang. 2018. "Bank Runs: The Pre-Deposit Game." *Macroeconomic Dynamics*, 22(4): 1-18.

[19] Shell, Karl and Yu Zhang. 2019. "On Sunspots, Bank Runs, and Glass-Steagall." *International Journal of Economic Theory*, 15(1): 13-25.