

Bank Portfolio Restrictions and Equilibrium Bank Runs*

by

James Peck

Department of Economics, The Ohio State University, 440 Arps Hall, Columbus, OH 43210-1172. email:
peck.33@osu.edu; telephone: (614) 292-0182; fax: (614) 292-3906.

Karl Shell

Department of Economics, 402 Uris Hall, Cornell University, Ithaca, NY 14853-7601 USA. email:
ks22@cornell.edu; telephone: (607) 255-5277; fax: (607) 255-8838

August 22, 2008

Abstract and Headnote

We extend the usual bank-runs model to capture the role of checking accounts in facilitating transactions. If there are no restrictions on their portfolios, banks are immune to runs, but they ration depositors with positive probability. If, on the other hand, banks are not permitted to hold illiquid assets, then the optimal deposit contract tolerates runs that occur with small probability. For reasonable parameter values, restricted banks overinvest in the liquid asset and they avoid non-run rationing of depositors.

Journal of Economic Literature Classification Numbers: D82, G21, E42.

Keywords: Bank runs, deposit contracts, mechanism design, sunspot equilibrium, narrow banking.

Corresponding Author: James Peck

Proposed Running Head: Narrow Banking

* We dedicate this paper to the memory of Bruce Smith. We are grateful to Neil Wallace for criticism of an earlier draft. We thank Huberto Ennis and Todd Keister for stimulating discussions.

1 Introduction

With the recent collapse of the building society Northern Rock and the investment bank Bear Sterns, the study of bank runs has taken on new urgency. Richmond Fed President Jeffrey Lacker has explicitly connected the Bear Sterns collapse with Diamond-Dybvig “non-fundamental” bank runs.¹ New regulatory oversight of investment banking seems inevitable, with the goal of making the sector less fragile.

In this paper, we revisit the question of whether or not legal restrictions on bank portfolios contribute to the stability of banks. In our model, the equilibrium probability of a bank run is zero if bank portfolios are unrestricted, but these banks do “run out” of cash and ration their depositors with positive probability even though they are immune to runs. However, government policies that restrict banks from holding illiquid assets cause banks to choose deposit contracts that can tolerate runs with small probability.² For reasonable parameters, the restricted banks inefficiently overinvest in the liquid asset, although they avoid “running out” of cash (i.e. non-run rationing of depositors).

Our model is built on the classic model of Diamond and Dybvig (1983) and its successors. Green and Lin (2003) show that the optimal contract can be uniquely implemented, so bank run equilibria do not exist at the optimal deposit contract. Peck and Shell (2003) consider a model in which, among other things, consumers cannot observe their places in line, and show that bank run equilibria can exist at the optimal deposit contract. Ennis and Keister (2007) drop the independence assumption and find equilibria in which some but not all consumers

¹In his June 5, 2008 speech, Laker explicitly links recent events to non-fundamental and fundamental bank runs, even though these events do not directly relate to demand deposit accounts. He concludes that fundamentals are part of the story, but that “determining whether a run is fundamental or self-fulfilling is very difficult.”

²Diamond and Rajan (1998) develop a model in which the possibility of a bank run affects bankers’ bargaining power in renegotiating loan contracts with borrowers. If a run occurs, depositors capture the loans and renegotiate with borrowers directly. It is the *threat* of a run that disciplines bankers. In Diamond and Rajan, a run cannot occur in equilibrium.

run, even if consumers observe their place in line. Ennis and Keister (2006) show that the optimal time-consistent deposit contract allows for partial bank runs. Our model differs substantially from Diamond and Dybvig in the following dimensions:

(1) In Diamond and Dybvig, banks provide insurance against the event that a consumer becomes impatient and must do her consumption “early” (if patient, she consumes “later”). We follow Diamond and Dybvig in building our model upon the stochastic nature of some urgent consumption opportunities³ or urgent needs, but we modify the model to capture the transactions role played by checking accounts. A depositor facing a consumption opportunity is someone who requires immediate liquidity to make an important purchase or someone taking advantage of the convenience of writing a check. The benefits of a demand deposit account would be severely limited if 100% payment were not made by the bank. Either the cash transaction could not be completed or the check would bounce. The consumption opportunity would be lost or deferred.⁴ To capture this need for the bank to make payments at par, we posit that a consumption opportunity takes the form of an indivisible good. Impatient consumers are those who find their best consumption opportunities in the first period, while patient consumers find their best consumption opportunities in the second period. Thus, banks facilitate consumption opportunities in our model, as opposed to merely providing insurance against impatience. In another departure from Diamond and Dybvig and as a proxy for more complete intertemporal analysis, we assume that all consumers value “left-over” cash (beyond the demand for funds to finance these indivisible consumption opportunities) in the final period.⁵ Utility is assumed to be a strictly concave function of

³For a general-equilibrium analysis of this type of intrinsic uncertainty, see Peck (1996). The interest in impulse-demand extends well beyond financial intermediation.

⁴In reality, of course, consumption goods come in various qualities and quantities. However, once the particular good is chosen it is all about whether the debit card will “work” or the check will clear.

⁵Our motivation is somewhat different from theirs, but Jacklin (1987) and Wallace (1996) allow utility to depend in a general way on consumption in each period.

the “left-over” consumption equivalent.

(2) We allow for intrinsic uncertainty: in particular, we allow the proportion of impatient consumers to be random.⁶

(3) Following Cooper and Ross (1998), Peck and Shell (2003), and Ennis and Keister (2006), we allow for a publicly observed extrinsic (sunspots) random variable on which depositors can coordinate their actions. This allows bank runs to be tolerated with positive probability in the optimal contract for the pre-deposit game. It seems to us that it is the combination of intrinsic uncertainty (represented here by the proportion of consumers who are impatient) *and* extrinsic or strategic, market uncertainty (represented here by “sunspots”) that is essential in modeling financial intermediation — especially if one is interested in the potential fragilities due to banking and other intermediation.

(4) We assume that there are two assets: one based on a liquid, lower-return technology and the other based on an illiquid, higher-return technology.⁷ We also allow for the possibility that the government might restrict banks to holding only the liquid asset in their portfolios.

(5) We follow Wallace and others in treating the bank’s deposit contract as a mechanism, but we place some restrictions on this mechanism. While the bank must satisfy sequential service, we do not allow the bank to punish customers who were denied service in period 1 or who refused service in period 1. We also assume the bank is not allowed to offer pure lotteries in its deposit contract. While our class of mechanisms is very broad and includes the possibility of partial suspension of convertibility, the indivisible nature of consumption opportunities⁸ implies that partial suspension will not be employed in the optimal contracts.

⁶See also Diamond and Dybvig (1983), Green and Lin (2003), Wallace (1988, 1990), and Ennis and Keister (2007).

⁷This two-technology approach was used by Wallace (1996). The restrictions on the bank in Wallace’s formal models are different from ours.

⁸Along with the assumption that there is a continuum of consumers.

We analyze two financial systems: (I) In the unified financial system, the bank provides transactions services and invests in both types of assets. (II) In the separated⁹ financial system, the bank provides transactions services but it invests only in the liquid asset. The separated financial system might emerge because of regulations which restrict bank investment to relatively liquid assets or because the non-regulatory or monitoring costs of financial integration justify the separated financial system.

Our results can be summarized as follows. The unified bank's optimal contract eliminates runs¹⁰, but does allow for a positive probability of non-run rationing¹¹ of impatient customers (that occurs when the realized proportion of impatient customers is relatively large). Intuitively, there is no incentive to run because early withdrawers receive less consumption from illiquid asset returns in period 2. For the bank in the separated financial system, the constrained efficient contract allows for runs.¹² Intuitively, if the fraction of impatient consumers is at its maximum, patient consumers must be receiving less consumption from the liquid asset than impatient consumers, because otherwise there is too much invested in the liquid asset. Thus, the incentive to economize on investments in the liquid asset makes bank

⁹Our separated financial system yields a bank that is close to what advocates of narrow banking recommend, although Friedman would perhaps restrict banks to holding only reserves. See the introduction to Wallace (1996). The differences are largely of interpretation. We are thinking that the separated bank does not hold the most illiquid assets in its portfolio, while the narrow bank is restricted to holding only the most liquid assets. Of course, with only two asset types, these ideas are the same. Nonetheless, advocates of narrow banking might be startled at first by our results.

¹⁰Strictly speaking, there is a run equilibrium at the optimal contract in the unified system, based on the (traditional) planner's problem that selects the best (non-run) equilibrium. However, the run equilibrium can be eliminated at negligible welfare cost. More importantly, in the unified system, a bank run equilibrium cannot occur with positive probability at the fully optimal contract that takes into account depositors' propensity to run.

¹¹This type of rationing is socially desirable since otherwise there would be overinvestment in preparation for the worst case outcome. Champ, Smith, and Williamson (1996) dub as "panics" large shocks that necessitate rationing. Thus, the unified financial system can avoid extrinsic panics of the self-fulfilling Diamond-Dybvig variety, but it opens up the possibility of intrinsic "panics" and non-run rationing.

¹²If a run occurs with small enough probability, then it will be rational for consumers to deposit, ex ante.

runs self fulfilling in the separated system. Compared to the unified system, the separated financial system tends to heavily overinvest in the liquid asset, in order to satisfy incentive compatibility. For this reason, rationing by the restricted bank is impossible or extremely unlikely.

We demonstrate an incentive-based return to scope, in providing both banking and investment services. If the government restricts banks to holding only the liquid asset, then this restriction tends to make the financial system more fragile by allowing for a positive probability of bank runs. The policy implications of this analysis might be surprising — at least to some of the early proponents of narrow banking.

2 The Model

There are three periods and a continuum of consumers (the potential bank depositors) represented by the unit interval. In period 0, each consumer is endowed with y units of the consumption good. A fraction α of the consumers is *impatient*: each of these has a “consumption opportunity” in period 1, yielding incremental utility of \bar{u} for 1 unit of consumption in period 1. If the consumption opportunity goes unfulfilled in period 1, these consumers face a diminished (or discounted) consumption opportunity in period 2, yielding incremental utility of $\beta\bar{u}$ for 1 unit of consumption in period 2, where the scalar β is less than unity. The remaining consumers are *patient*: each of these has a consumption opportunity, yielding incremental utility of \bar{u} for 1 unit of consumption in period 2.¹³ Beyond these urgent “consumption opportunities,” both types of consumers derive utility from additional (“left-over”) consumption in period 2, and can costlessly store consumption across periods. Thus, impatient and patient consumers, respectively, have the reduced-form utility functions:

¹³We could have argued that patient consumers would receive incremental utility of $\beta\bar{u}$ if they took advantage of their consumption opportunity in period 1 instead of period 2. This would not affect our results.

$$U_I(C_I^1, C_I^2) = \begin{cases} \bar{u} + u(C_I^1 + C_I^2 - 1) & \text{if } C_I^1 \geq 1 \\ \beta\bar{u} + u(C_I^1 + C_I^2 - 1) & \text{if } C_I^1 < 1 \end{cases} \quad (1)$$

and

$$U_P(C_P^1, C_P^2) = \bar{u} + u(C_P^1 + C_P^2 - 1),$$

where C_i^t (for “cash”) is the *total withdrawal* of a type i consumer from the bank in period t . I stands for impatient and P stands for patient. The positive scalar \bar{u} is the utility from the (indivisible) consumption opportunity. Specification (1) is based on the assumption that the consumers will always be able to afford their consumption opportunities in period 2, and that \bar{u} is high enough so that it is optimal to undertake available consumption opportunities. We assume that u is an increasing, smooth, and strictly concave function of “terminal” (or “left-over”) consumption, so we have $u' > 0$ and $u'' < 0$.

Let f denote the probability density function for α , the fraction of the consumers who become impatient, which is assumed to be continuous and have support $[0, \bar{\alpha}]$, where $\bar{\alpha} < 1$. In keeping with our assumption that consumers are identical, *ex ante*, we have the following process in mind.¹⁴ First, nature determines α according to f . Then, nature selects each particular consumer to be impatient with probability α and patient with probability $(1 - \alpha)$. A consumer’s type is her private information.

There are two constant-returns-to-scale technologies, an illiquid, higher-yield technology, A , and a liquid, lower-yield technology, B . Investing 1 unit of period-0 consumption in technology A yields R_A units of consumption in period 2. Investing 1 unit of period-0 consumption in technology B yields R_B units of consumption if held until period 2, or 1 unit of consumption if harvested in period 1. We assume that $1 < R_B < R_A$ holds.

¹⁴The continuum model is convenient, but there are technical issues regarding the law of large numbers, which we ignore. Our main results hold with a finite number of consumers, although the expressions and calculations become more complicated.

In period 0, the bank designs the demand-deposit contract, or banking mechanism. We assume that the bank seeks to maximize the *ex-ante* expected utility of consumers. Following most of the literature, we focus for the moment on the post-deposit game, which starts after the mechanism is announced and deposits are in place. The *no-run optimal contract* (NROC) solves the traditional planner’s problem, which imposes the incentive compatibility condition: a patient consumer chooses period 2 *given that all other patient consumers choose period 2*. We can then ask whether the NROC also admits a run equilibrium. Of course, if the run equilibrium is expected, the contract would no longer be optimal, and perhaps no one would be willing to deposit.¹⁵ In Section 5, with the help of extrinsic uncertainty, we analyze equilibrium runs in the full (pre-deposit) game, where the fully optimal contract takes into account depositors’ propensity to run.

The banking mechanism must respect the restrictions required by the timing of the post-deposit game, described as follows. At the beginning of period 1, each consumer (now a depositor) learns her type and decides whether to arrive at the bank in period 1 or period 2. Consumers who choose period 1 are assumed to arrive in random order. Let z_j denote the position of consumer j in the queue. Because of the sequential service constraint, consumption must be allocated to consumers as they arrive to the head of the queue, as a function of the history of transactions up until that point. We further assume that consumer j ’s withdrawal can only be a function of her position, z_j ¹⁶, and that she has an opportunity to

¹⁵A consumer could invest her endowment herself, instead of dealing with the bank. It does not matter whether or not we allow a consumer to access technology B privately, but we do require that unharvested “trees” cannot be traded. This is to rule out the case in which a patient depositor (claiming to be impatient) trades period-1 consumption withdrawn from the bank for unharvested trees. Jacklin (1987) has shown that such a market undermines the optimal contract, and his argument applies to our setting as well. See Haubrich (1988) for a more general analysis. Ruling out this asset market is merely to posit that only banks can provide the liquidity necessary to pay for urgent consumption opportunities.

¹⁶From a mechanism design standpoint, it might seem strange not to allow consumers to send messages to the bank. This is almost without loss of generality, since consumers will send whatever message gives them the most consumption. However, we are ruling out the bank offering a lottery to learn a consumer’s type, and punishing consumers in period 1 after a patient consumer has arrived. Besides being costly to

refuse to withdraw and return without prejudice in period 2. The bank cannot keep track of how many consumers have refused.¹⁷ Let α_1 denote the measure of consumers who have made a withdrawal in period 1. In period 2, the bank chooses how to divide its remaining resources between those who have withdrawn in period 1 and those who have not.¹⁸

A contract specifies the fraction of a consumer's endowment invested in technology B , denoted by γ ; her withdrawal in period 1 as a function of her arrival position, denoted by $c^1(z)$; and her withdrawal in period 2 from technology B investments as a function of α_1 and whether the consumer made a withdrawal in period 1 or not, denoted respectively by $c_I^2(\alpha_1)$ and $c_P^2(\alpha_1)$.¹⁹ That is, a consumer who receives $c_I^2(\alpha_1)$ from technology B investments receives a total withdrawal in period 2 of $C_I^2(\alpha_1) = c_I^2(\alpha_1) + (1 - \gamma)R_A y$. Similarly, a consumer who receives $c_P^2(\alpha_1)$ from technology B investments receives a total withdrawal in period 2 of $C_P^2(\alpha_1) = c_P^2(\alpha_1) + (1 - \gamma)R_A y$. We assume that parameters are such that nonnegativity constraints $C_I^2(\alpha_1) \geq 0$ and $C_P^2(\alpha_1) \geq 0$ never bind.

For the mechanism to be feasible, all remaining resources must be distributed in period 2. The resource constraint is given by

$$\alpha_1 c_I^2(\alpha_1) + (1 - \alpha_1) c_P^2(\alpha_1) = [\gamma y - \int_0^{\alpha_1} c^1(z) dz] R_B. \quad (2)$$

Thus, the space of deposit contracts or mechanisms, M , is given by

$$M = \{ \gamma, c^1(z), c_I^2(\alpha_1), c_P^2(\alpha_1) \mid \text{Equation (2) holds for all } \alpha_1 \}.$$

implement, these lotteries and punishments hardly correspond to accepted conservative banking practices.

¹⁷Thus, z_j should really be interpreted as the measure of consumers who have already withdrawn from the bank in period 1 before consumer j has an opportunity to withdraw. The purpose of this restriction is to disallow the bank from telling a customer during a run that they may not withdraw in period 1 and that they also forfeit their claims to consumption in period 2.

¹⁸In principle, what a consumer receives in period 2 could depend on how much she received in period 1, and not just on whether she made a withdrawal. This distinction is not important here, since the optimal contract always provides 1 unit of consumption in period 1 until the bank runs out.

¹⁹Here, in an abuse of notation, the subscripts I and P refer to consumers *claiming* to be impatient and patient respectively.

We analyze bank behavior in each of the two financial systems: (I) In the *separated financial system*, consumers place a fraction $(1 - \gamma)$ of their wealth in technology A , whose return cannot be touched by the bank. In terms of resource constraint (2), this is equivalent to imposing the additional constraints: $c_P^2(\alpha_1) \geq 0$ and, more importantly, $c_I^2(\alpha_1) \geq 0$. Combined with incentive compatibility, these additional constraints give rise to overinvestment in technology B and the possibility of bank runs. (II) In the *unified financial system*, the bank is able to invest in both technologies. This allows the bank more flexibility in smoothing consumption and preventing runs. For example, when $\bar{\alpha}$ consumers arrive in period 1 (the worst case scenario), the bank can liquidate all of its technology B holdings, but differentially reward consumers from technology A in period 2. Consumers who arrive in period 1 might receive less than $(1 - \gamma)R_{Ay}$ in period 2, while consumers who wait might receive more than $(1 - \gamma)R_{Ay}$. In terms of resource constraint (2) this is equivalent to allowing $c_P^2(\alpha_1)$, or, more importantly, $c_I^2(\alpha_1)$, to be negative.

The unified system can be interpreted in several ways. The most straightforward interpretation is that the bank is allowed to hold technology A assets (stocks or mutual funds) as part of its portfolio. Another interpretation is that the bank can write subordinated debt contracts with firms investing in technology A , whereby the bank receives period-2 consumption in the event that sufficiently many consumers arrive in period 1.²⁰

We next define “run equilibrium” in the post-deposit game.

Definition 2.1: Consider either a unified financial system or a separated financial system, and a contract $m \in M$. Then the *post-deposit game* is said to have a *run equilibrium* if there is a Bayes-Nash equilibrium in which all consumers arrive in period 1 and a positive measure of patient consumers withdraw in period 1.

²⁰We ignore potential moral hazard problems on the part of the bankers. One can imagine, though, that bankers have a stronger incentive to cheat or undertake improper investments in the unified system. See Calomiris and Kahn (1991) for an explicit analysis of moral hazard issues and embezzlement in banking.

Given a contract, the solution concept in Definition (2.1) is Bayes-Nash equilibrium. However, it should be clear from the constructions given below that the equilibrium strategies are consistent with sequential equilibrium as well. Our definition of run equilibrium requires all patient consumers to choose period 1. It is not required that all withdraw. The bank might very well offer zero consumption after $\bar{\alpha}$ of the consumers have made withdrawals (and hence run is known to be in progress). We require a positive measure of patient consumers to withdraw, to rule out the degenerate case in which patient consumers arrive in period 1 with the intention of refusing all offers, since this is equivalent to waiting until period 2.

3 The Unified System

In this section, we describe the planner’s problem, the solution to which yields the NROC for the unified system. The NROC attains the full-information optimal outcome in equilibrium. Although, strictly speaking, the post-deposit game has a run equilibrium at the NROC, the run equilibrium can be eliminated at negligible welfare cost.²¹ On the other hand, we demonstrate below that the NROC entails a positive probability of (non-run) rationing of consumers in period 1.

We restrict attention to environments in which it is beneficial to provide for consumption opportunities whenever the resources are available. It is then desirable that the impatient consumers choose period 1 and the patient consumers choose period 2 for making their urgent withdrawals. Also, impatient consumers unable to withdraw in period 1 and patient consumers should take advantage of their consumption opportunities in period 2.²² Since the amount of a withdrawal in period 1 greater than one unit would be stored for period 2,

²¹See Theorem (3.2). The run equilibrium is also eliminated by assuming that a patient depositor will choose not to run when indifferent between running and not running. When we introduce sunspots and the propensity to run in section 5, it will follow that the probability of a bank run, at the optimal contract to the full pre-deposit game, is zero.

²²These conditions will be met if $\beta\bar{\alpha}$ is large, relative to the marginal utility u' of “left-over” consumption.

there is no reason for the bank to provide more than one unit in period 1; hence we have $y\gamma \leq \bar{\alpha}$. Thus we can restrict our search to contracts in which we have

$$c^1(z) = 1 \quad \text{for } z \leq \gamma y. \quad (3)$$

Given (3) and the fact that patient consumers wait until period 2, the *ex ante* welfare W is given by

$$\begin{aligned} W = & \int_0^{\gamma y} [\bar{u} + (1 - \alpha)u((1 - \gamma)yR_A + c_P^2(\alpha) - 1) + \alpha u((1 - \gamma)yR_A + c_I^2(\alpha))] f(\alpha) d\alpha \\ & + \int_{\gamma y}^{\bar{\alpha}} [(1 - \alpha + \gamma y)\bar{u} + (\alpha - \gamma y)\beta\bar{u} + (1 - \alpha)u((1 - \gamma)yR_A + c_P^2(\alpha) - 1) \\ & + (\alpha - \gamma y)u((1 - \gamma)yR_A + c_P^2(\alpha) - 1) + \gamma y u((1 - \gamma)yR_A + c_I^2(\alpha))] f(\alpha) d\alpha. \end{aligned} \quad (4)$$

Maximand (4) captures the fact that impatient consumers who are rationed in period 1 cannot be prevented from receiving the (higher) consumption that the patient consumers receive in period 2. The only incentive compatibility constraint to worry about is that a patient consumer must be better off waiting until period 2 than accepting one unit in period 1, given that the other patient consumers wait. Conditional on being patient and being offered c^1 , the conditional density for α is denoted by f_P .²³ Thus, we have

$$\begin{aligned} & \int_0^{\bar{\alpha}} u(c_P^2(\alpha) + (1 - \gamma)yR_A - 1) f_P(\alpha) d\alpha \geq \int_0^{\gamma y} u(c_I^2(\alpha) + (1 - \gamma)yR_A) f_P(\alpha) d\alpha \\ & + \int_{\gamma y}^{\bar{\alpha}} (\gamma y/\alpha) u(c_I^2(\alpha) + (1 - \gamma)yR_A) + (1 - \gamma y/\alpha) u(c_P^2(\alpha) + (1 - \gamma)yR_A - 1) f_P(\alpha) d\alpha. \end{aligned} \quad (5)$$

²³In general, $f_P(\alpha)$ is different from $f(\alpha)$. For example, if nothing is learned from observing c^1 , then $f_P(\alpha)$ can be calculated as

$$f_P(\alpha) = \frac{(1 - \alpha)f(\alpha)}{\int_0^{\bar{\alpha}} (1 - a)f(a) da}.$$

Resource constraint (2) can be simplified to yield

$$\begin{aligned} \alpha_1 c_I^2(\alpha_1) + (1 - \alpha_1) c_P^2(\alpha_1) &= (\gamma y - \alpha_1) R_B & \text{if } \alpha_1 \leq \gamma y \\ \gamma y c_I^2(\alpha_1) + (1 - \gamma y) c_P^2(\alpha_1) &= 0 & \text{if } \alpha_1 > \gamma y. \end{aligned} \quad (6)$$

An NROC under the unified system is the solution to the following problem:

$$\begin{aligned} &\max_{\gamma, c_I^2(\alpha_1), c_P^2(\alpha_1)} W \\ &\text{subject to 5 and 6.} \end{aligned} \quad (7)$$

The next theorem establishes that an NROC necessarily rations consumers in period 1 when the fraction of impatient consumers arriving in period 1 is sufficiently large, i.e. when α_1 is close to $\bar{\alpha}$, equal to $\bar{\alpha}$, or greater than $\bar{\alpha}$. The intuition for this result is that, if consumers were never rationed (no matter the realization of α) in period 1, then society through over-caution would be over-investing in the liquid technology, B .²⁴

Theorem 3.1: An NROC in the unified system satisfies $\gamma y < \bar{\alpha}$. The “first” γy impatient consumers to arrive are fully served by the bank in period 1. There is a positive probability that $\alpha > \gamma y$ holds, in which case $(\alpha - \gamma y)$ impatient consumers are rationed. Patient consumers do not withdraw in period 1, and we have full consumption smoothing, i.e.,

$$c_I^2(\alpha_1) = c_P^2(\alpha_1) - 1 \quad \text{for all } \alpha_1 \leq \gamma y. \quad (8)$$

Proof: Given γ , the functions $\{c_I^2(\alpha_1), c_P^2(\alpha_1)\}$ that maximize W subject only to resource constraint (6) entail full consumption smoothing (8). However, the allocation defined by (6) and (8) also satisfies the incentive compatibility constraint (5), and therefore solves the

²⁴This is similar to design of simpler non-financial economic institutions. If, for example, the price of electricity is fixed, it would generally be foolish to build generating capacity sufficient to meet every *possible* level of demand.

more tightly constrained problem (7). Plugging (6) and (8) into the expression for W , and differentiating with respect to γ , we have

$$\left(\frac{\partial W}{\partial \gamma}\right)_{\gamma=\bar{\alpha}/y} = \int_0^{\bar{\alpha}} y(R_B - R_A)u'[c_P^2(\alpha) - 1 + (y - \bar{\alpha})R_A]f(\alpha)d\alpha < 0.$$

Clearly, any contract for which we have $\gamma y > \bar{\alpha}$ is inferior to the one characterized by (6) and (8), since the former provides fewer resources available in period 2, with no compensating advantage in terms of consumption smoothing in period 2 or provision of consumption in period 1. \square

The proof of Theorem (3.1) shows that an NROC offers complete consumption smoothing, according to (6) and (8). Since the patient consumers receive the same consumption, whether they arrive in period 1 or period 2, it is obviously incentive compatible. One advantage of the indivisibility of consumption opportunities is that the NROC has a simple, realistic solution in which there is full, but not partial, suspension of convertibility.

Theorem 3.2: The unified system is immune to bank runs, in the following sense: for all $\varepsilon > 0$, there is a contract for which the post-deposit game has a dominant-strategy equilibrium yielding welfare within ε of the solution to (7), and for which there is no run equilibrium.

Proof: First, note that a solution to (7) must exist. Given γ , a sufficient condition for $\{c_I^2(\alpha_1), c_P^2(\alpha_1)\}$ to solve problem (7) is for (6) and (8) to hold. The specification of consumption for $\alpha_1 > \gamma y$ does not affect the objective or the incentive compatibility constraint. Plugging (6) and (8) into the expression for W , problem (7) can be transformed into an equivalent unconstrained problem of choosing γ to maximize W , which must have a solution satisfying $\gamma y < \bar{\alpha}$.

Construct the contract, $(\gamma', c_I^2(\alpha_1)', c_P^2(\alpha_1)')$, as follows. Let γ' be as in the solution to (7). Consumptions are determined by the resource equation, (5), and

$$c_I^2(\alpha_1)' = c_P^2(\alpha_1)' - 1 - \varepsilon' \quad \text{for all } \alpha_1,$$

where ε' is a positive constant. Because $c_P^2(\alpha_1)' - c_I^2(\alpha_1)' > 1$ holds for all α_1 , patient consumers have a dominant strategy not to withdraw in period 1. If ε' is small enough, impatient consumers have a dominant strategy to withdraw when offered one unit of consumption in period 1 (and not “withdraw” when zero is offered). Thus, there is no run equilibrium.

It is easy to see that, as $\varepsilon' \rightarrow 0$, consumptions are converging to the solution of (7), so we can always find ε such that equilibrium yields welfare within ε of the solution to (7). \square

In the next example, we specify the following basic parameters: the utility function $u(\cdot)$ for “left-over” consumption, the utility \bar{u} from satisfying the urgent opportunity, the interest factor on the illiquid asset R_A , the discount factor β for delayed consumption opportunities, and the density function $f(\alpha)$ for the proportion who are impatient.

Example 3.3:

$$y = 10, \quad u(c) = 100 \log(c) - 249, \quad \bar{u} = 20, \quad R_A = 1.1, \quad \beta = 0.7,$$

$$\text{uniform distribution with } \bar{\alpha} = 0.5: \quad f(\alpha) = \begin{cases} 2 & \text{for } \alpha \in [0, 0.5] \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

For the unified financial system, we compute γ , the proportion of wealth invested in technology B and the *ex-ante* welfare of consumers W for different values of the interest factor

R_B on the liquid technology. Our results are summarized in Table 1.²⁵ An increase in R_B increases welfare, due to the higher yield on technology B investments, and increases γ , because reducing the probability of rationing consumers in period 1 is now less costly, because the gap between the yields in technologies A and B has been reduced. The optimal γ is less than 0.05, which follows from Theorem (3.1) and the parameter specification, $\bar{\alpha}/y = 0.5/10 = 0.05$. The fact that γ is close to 0.05 indicates that the probability of rationing is small at the optimal contract. The intuition is that, at the margin, the loss of $(1 - \beta)\bar{u}$ multiplied by the probability of rationing is traded off against the difference in yields, $R_A - R_B$. Whenever the consumption opportunity is important relative to $R_A - R_B$ rationing will be unlikely.

	γ	W
$R_B = 1.05$	0.04544	0.8942
$R_B = 1.08$	0.04807	0.9599

Table 1: **The Liquidity Proportion γ and Welfare W in the Unified System**

4 The Separated System

In order to evaluate the impact of portfolio restrictions on banks, we assume that the bank cannot gain access to the funds invested in technology A . A comparison of the separated system with the unified system might indicate how costly these restrictions would be. As in Section 3, we restrict attention to environments in which it is optimal to provide one unit of consumption to consumers arriving in period 1, whenever technology B assets are

²⁵Computations were made using Maple version 5. The code is available from the authors for the purposes of replicating the results.

available.²⁶ Since second-period withdrawals from the bank must come from technology B investments that were not harvested in period 1, the separated system is quite different from the unified system. When α is high enough, some impatient consumers are rationed in the unified system, yet full consumption smoothing is optimal. When α is high in the separated system, it may be impossible to provide those arriving in period 2 with one unit of consumption from technology B investments at the optimal γ , so full consumption smoothing might be impossible. Excessive investment in technology B might be necessary in order to satisfy incentive compatibility, so the NROC may require $\gamma y > \bar{\alpha}$. Finally, the NROC might be subject to bank runs, which can only be avoided at significant welfare cost.

In the separated system, ex-ante welfare, the incentive compatibility constraint, and the resource constraint are as given in expressions (4), (5), and (6). The restriction that the bank cannot gain access to investments of technology A is expressed simply as follows:

$$c_I^2(\alpha_1) \geq 0 \text{ and } c_P^2(\alpha_1) \geq 0 \text{ for all } \alpha_1. \quad (10)$$

Notice that constraints (6) and (10) imply $c_I^2(\alpha_1) = c_P^2(\alpha_1) = 0$ for $\alpha_1 > \gamma y$. If all of the technology B investments are liquidated in period 1, then withdrawals from the bank must be zero in period 2. An NROC under the separated system is a solution to the following planner's problem:

$$\begin{aligned} & \max_{\gamma, c_I^2(\alpha_1), c_P^2(\alpha_1)} W & (11) \\ & \text{subject to (5), (6), and (10).} \end{aligned}$$

In the unified system, all of the technology B investments are harvested and some consumers are rationed in period 1 when α is sufficiently high. In the separated system when

²⁶That is, $c^1(z) = 1$ holds for $z \leq \bar{\alpha}$. Not wanting to hoard technology B assets is a stronger assumption in the separated system than in the unified system. We will see that incentive compatibility binds in the separated system, and refusing to liquidate technology B assets allows the bank to reduce its technology B overinvestment while still satisfying incentive compatibility. However, if we were to rewrite the expressions in problem (11) to allow for the possibility of hoarding technology B assets, Theorem (4.2) and the overinvestment component of Theorem (4.3) continue to hold. Details are available from the authors.

α is close to $\bar{\alpha}$, it is typically the case that not all technology B investments are harvested and no one is rationed. The intuition is that more investment in technology B is needed in order to satisfy incentive compatibility without using technology A resources. While some technology B investments may remain for patient consumers in state $\bar{\alpha}$, the next lemma establishes that in state $\bar{\alpha}$ the consumers who choose period 1 withdraw more than those who choose period 2. Denote the solution to (11) by $m^* = \{\gamma^*, (c_I^2(\alpha_1))^*, (c_P^2(\alpha_1))^*\}$.

Lemma 4.1: Any NROC in the separated system, which solves problem (11), satisfies $(c_P^2(\bar{\alpha}))^* < 1$.

Proof: Suppose instead that $(c_P^2(\bar{\alpha}))^* > 1$ holds. Since resources remain in period 2, it follows that $\bar{\alpha} \leq \gamma^*y$ holds. Therefore, it is possible to increase welfare by reducing γ and achieving full consumption-smoothing, $(c_P^2(\alpha_1))^* = 1 + (c_I^2(\alpha_1))^*$ for all α_1 . It is easy to see that incentive compatibility and nonnegativity are satisfied, contradicting the fact that m^* solves (11).

Now suppose that $(c_P^2(\bar{\alpha}))^* = 1$ holds. We must have $(c_I^2(\bar{\alpha}))^* = 0$, or else (just as in the previous case) we can increase welfare by reducing γ , while maintaining full consumption-smoothing. It follows that we must have

$$(c_P^2(\alpha_1))^* - (c_I^2(\alpha_1))^* \geq 1 \tag{12}$$

for almost all α_1 . Otherwise, for a positive-measure set of realizations of α_1 , $c_P^2(\alpha_1)$ can be increased and $c_I^2(\alpha_1)$ can be reduced to satisfy the resource and nonnegativity constraints, which increases welfare and relaxes the incentive compatibility constraint. If inequality (12) is strict for a positive-measure set of realizations of α_1 , then welfare can be feasibly increased by choosing $(c_P^2(\alpha_1))^*$ and $(c_I^2(\alpha_1))^*$ to satisfy full consumption-smoothing (where (12) holds as an equality) and the resource constraint, (6). Therefore, we have for all α_1 ,

$$(c_P^2(\alpha_1))^* - (c_I^2(\alpha_1))^* = 1. \tag{13}$$

Treating γ as a parameter, and solving the equations (6) and (13) for consumptions, we can define welfare as a function of γ , $W(\gamma)$. Because m^* solves (11), $W(\gamma)$ must be maximized at $\gamma = \gamma^*$. Applying the envelope theorem, one can show:

$$W'(\gamma^*) = y(R_B - R_A) \int_0^{\bar{\alpha}} u'[(1 - \gamma^*)yR_A + (\gamma^*y - \alpha)R_B + \alpha - 1]f(\alpha)d\alpha < 0.$$

It follows that reducing γ improves welfare, contradicting the fact that γ^* is part of the NROC, m^* . \square

The intuition behind Lemma (4.1) is that too much is invested in technology B if patient consumers arriving in period 2 receive at least 1 unit of consumption in state $\bar{\alpha}$. Reducing γ does not lead to rationing, and yields a higher return on investment. The only reason to save any consumption at all for period 2 in state $\bar{\alpha}$ is to satisfy non-negativity and incentive compatibility constraints. Since these constraints are not binding if $(c_P^2(\bar{\alpha}))^* \geq 1$ holds, too much has been invested in technology B . Applying Lemma (4.1), we next show that the NROC in the separated system always admits a run equilibrium.

Theorem 4.2: Any NROC in the separated financial system has a run equilibrium.

Proof: We know that the NROC satisfies $c^1(z) = 1$ for all $z \leq \bar{\alpha}$. Since we are considering the possibility of bank runs here, a patient consumer's decision to arrive in period 1 must also take into account $c^1(z)$ for all $z > \bar{\alpha}$. Let z^* be the smallest value of z , greater than or equal to $\bar{\alpha}$, such that the following inequality holds

$$c^1(z) \leq \frac{[\gamma y - \bar{\alpha} - \int_{\bar{\alpha}}^z c^1(a)da]R_B}{1 - z}. \quad (14)$$

If there is no value of z satisfying (14), define z^* to equal 1. From Lemma (4.1), we know that $(c_P^2(\bar{\alpha}))^* < 1$ holds at the NROC. We must also have $(c_I^2(\bar{\alpha}))^* = 0$, or else higher welfare can be achieved by transferring consumption from those who arrived in period 1 to those

who did not, while continuing to satisfy the constraints.²⁷ It follows that, setting $z = \bar{\alpha}$, the left side of (14) is equal to unity, and the right side of (14) is equal to $(c_P^2(\bar{\alpha}))^*$. Since inequality (14) is not satisfied, we have $z^* > \bar{\alpha}$.

We claim that there is a run equilibrium, in which all consumers arrive in period 1. Those for whom $z_j < z^*$ holds accept $c^1(z_j)$, and those for whom $z_j \geq z^*$ holds refuse $c^1(z_j)$ and do not withdraw in period 1.

Without loss of generality, we can assume $(c_P^2(\alpha_1))^* = 0$ for $\alpha_1 > \bar{\alpha}$, because giving period-2 consumption to those who withdraw in period 1 only increases the incentive to run. Thus, for $\alpha_1 > \bar{\alpha}$, second period consumption is given by

$$(c_P^2(\alpha_1))^* = \frac{[\gamma y - \bar{\alpha} - \int_{\bar{\alpha}}^{\alpha_1} c^1(a) da] R_B}{1 - \alpha_1}. \quad (15)$$

Differentiating with respect to α_1 in (15) yields

$$\frac{\partial (c_P^2(\alpha_1))^*}{\partial \alpha_1} = \frac{R_B((c_P^2(\alpha_1))^* - c^1(\alpha_1))}{1 - \alpha_1},$$

which is negative for $\alpha_1 < z^*$, since inequality (14) does not hold. Thus $(c_P^2(\alpha_1))^*$ is decreasing in α_1 for $\alpha_1 < z^*$.

Given the acceptance/refusal behavior specified above, everyone will refuse $c^1(z^*)$, so we have $\alpha_1 = z^*$ with probability 1. By (14) and (15), it is a best response for consumer j to refuse if $z_j = z^*$, and $z_j > z^*$ is irrelevant. If $z_j < z^*$ holds, (14) and (15) imply $c^1(z_j) > (c_P^2(z_j))^* > (c_P^2(z^*))^*$, where the last inequality follows from the fact that $(c_P^2(\alpha_1))^*$ is continuous and decreasing in α_1 for $\alpha_1 < z^*$. Therefore, it is a best response for consumer j to accept $c^1(z_j)$. \square

²⁷Remember, the NROC is the one that provides the highest welfare, based on the equilibrium in which the patient consumers wait until period 2.

There is an intuitive explanation for why run equilibria always exist in the separated system. In an NROC for the separated financial system, $(c_P^2(\bar{\alpha}))^*$ must be less than 1, or else too much is invested in technology B . Therefore, in the event of a run, consumers arriving in period 2 receive less than one unit. Consumers arriving in period 1 are better off, since they receive 1 unit of consumption if $z_j \leq \bar{\alpha}$, and they can refuse to withdraw and delay their arrival until period 2 otherwise. The proof is a bit more intricate, since we must rule out the possibility that a fraction greater than $\bar{\alpha}$ of the consumers withdraw in period 1, possibly leaving more than one unit of consumption per capita in period 2. Another perspective is that the desire to economize on technology B investments causes the incentive compatibility constraint to bind, so that a patient consumer is indifferent between period 1 arrival and period 2 arrival, assuming other patient consumers wait. If instead the other patient consumers arrive in period 1, those who wait are worse off, so that incentive compatibility is no longer satisfied.

The following theorem provides a (seemingly weak) sufficient condition for the optimal liquid asset investment under the separated financial system to be greater than it is under the unified financial system and for the absence of non-run rationing of impatient consumers of the restricted bank.

Theorem 4.3 (Overinvestment in the Liquid Asset): If $\bar{\alpha} < 1/R_B$ holds, then any NROC for the separated financial system does not ration consumers in period 1 in the no-run equilibrium, and invests more in technology B than any NROC for the unified financial system.

Proof: From Theorem (3.1), an NROC in the unified system satisfies $\gamma < \bar{\alpha}/y$. In the separated system, an NROC must invest at least enough in technology B to provide 1 unit of consumption in period 2 when everyone is patient. That is, we must have $(c_P^2(0))^* \geq 1$, or else all patient consumers will choose period 1. Thus, the optimal fraction invested in

technology B for the separated system, γ^* , satisfies $\gamma^* \geq 1/(R_B y)$. Since $R_B \bar{\alpha} < 1$ holds, we draw the following conclusions. First, $\gamma^* > \bar{\alpha}/y$ holds, so consumers are not rationed in period 1 unless there is a run. Second, γ^* exceeds the optimal fraction invested in technology B for the unified system. \square

This overinvestment in the liquid asset, B , is likely to be substantial, as long as the maximum fraction of impatient consumers is relatively small. For example, if we have $\bar{\alpha} = 0.5$, $y = 10$, and $R_B = 1.08$, then in the unified system, the fraction of resources invested in technology B is less than 0.05, as a consequence of Theorem (3.1). In the separated system, the fraction of resources invested in technology B is more than 0.09, as consequence of the fact that $(c_p^2(0))^* \geq 1$ must hold, which implies $\gamma^* \geq 1/(R_B y)$. Notice that this lower bound on overinvestment (at least 80% more than under the unified system) applies to all utility functions consistent with our maintained assumptions.

Next we compute the NROC in the separated system, for the parameters of Example (3.3).

Example 4.4: Consider the parameter values specified in Example (3.3). We also specify $R_B = 1.08$.

It can be shown that the NROC satisfies $(c_I^2(\alpha_1))^* = 0$ for all α_1 . The intuition is that giving consumption to impatient consumers in period 2 hurts incentive compatibility and forces more overinvestment in technology B , outweighing any consumption-smoothing

advantages.²⁸ Since Theorem (4.3) applies, it follows from (6) that we have

$$(c_P^2(\alpha_1))^* = \frac{(10\gamma - \alpha_1)1.08}{1 - \alpha_1}. \quad (16)$$

Finding the NROC now reduces to finding the value of γ that maximizes welfare subject to the incentive compatibility constraint. The optimal γ will cause the incentive compatibility constraint to hold with equality, yielding:

$$\gamma^* = 0.09445 \quad \text{and} \quad W^* = 0.8688. \quad (17)$$

Now compare the NROC in the unified system to the NROC in the separated system, for parameters given by (9) and $R_B = 1.08$. From Table 1 and (17), we see that welfare is clearly higher in the unified system. Problem (11) involves maximizing over a strictly smaller set than problem (7), due to the additional non-negativity constraints. Technology B investment in the separated system is nearly double that in the unified system.

Several conclusions can be drawn from our example. Efficiency can be enhanced by eliminating legal restrictions on bank portfolios, such as the restrictions imposed by the weak form of narrow banking. Moreover, our analysis suggests that in this case eliminating regulatory and other government restrictions would cause a major shift from liquid, lower-yield investments to illiquid, higher-yield investments. This shift would necessitate occasional scarcities of liquidity (in the model, non-run rationing of period 1 consumption). Allowing banks to freely invest in all technologies eliminates bank runs triggered by extrinsic shocks, but it introduces the possibility of planned “running out” of bank cash due to intrinsic shocks.²⁹

²⁸This result is derived by first obtaining a lower bound for the multiplier on the incentive compatibility constraint, which allows us to obtain a lower bound for the multiplier on each state- α non-negativity constraint. Since these multipliers are positive, the constraints must bind, so the impatient receive no consumption from the bank in period 2. If the parameters of the example are changed, it is possible that some of the non-negativity constraints do not bind. Also, it can be shown that the bank cannot do better by refusing to liquidate some technology B assets. (Details are available from the authors.)

²⁹Our analysis does not include deposit insurance or model moral hazard issues for bank management, nor

5 Sunspots and the Propensity to Run

Run equilibria in Diamond and Dybvig (1983) are not really equilibria, because consumers would not deposit their funds if they knew that a run would take place.³⁰ Diamond and Dybvig suggest that a run could take place in equilibrium with positive probability, triggered by some extrinsic random variable “sunspots,” as long as the probability of the run is sufficiently small. We calculate here what the “sufficiently small” probability is for an example of a bank in the separated financial system. The fully optimal contract (as opposed to the NROC) must consider the probability that a run will take place. Cooper and Ross (1998) formalize this notion, in a model that restricts contracts to a particularly simple class. Peck and Shell (2003) perform a similar analysis to that presented here. Since the NROC in the separated financial system always has a run equilibrium (Theorem 4.2), and since equilibrium bank runs are considerably rarer in Peck and Shell (2003), it is worthwhile to perform the corresponding calculations here.

The complete (pre-deposit) game is characterized as follows. At the beginning of period 0, the bank announces a deposit contract, $m \in M$.³¹ Then consumers, before observing their types, decide whether to deposit. At the beginning of period 1, each consumer learns her type and observes a sunspot variable s distributed uniformly on $[0,1]$.³² Sunspots do not affect preferences (including the likelihood of being impatient), endowments, or technology. Next, consumers decide whether to arrive in period 1 or period 2, and the game proceeds as before. Note that the period in which a consumer arrives can depend on the realization

does it include credit chains (or other systemic problems) as analyzed by Kiyotaki and Moore (1997).

³⁰See Postlewaite and Vives (1987).

³¹We avoid the complication of having the mechanism depend on the measure of consumers that deposit. Instead, assume that the autarky outcome occurs if the measure of depositors is less than one.

³²The uniformity assumption is without loss of generality.

of the sunspot variable s as well as the realization of her type. The space of mechanisms is unchanged; the solution concept is Bayes-Nash equilibrium. We have a run equilibrium if, for some set of realizations of s occurring with positive probability, all consumers arrive at the bank in period 1 and a positive measure of the patient consumers withdraw in period 1.

Suppose that the economy has a “propensity” to run, in the following sense: whenever we have $s < s_0$, then all consumers choose to arrive at the bank in period 1, and accept consumption offers of at least one unit, whenever such a run is consistent with equilibrium, and patient depositors are strictly better off running than not running in the run equilibrium.³³ If a run is not consistent with equilibrium, then all patient consumers wait.³⁴ When we have $s \geq s_0$, the equilibrium is selected in which all patient consumers wait for the second period. Of course, incentive compatibility must be satisfied. We will calculate the cutoff value of s_0 below which the optimal contract in the separated system has a run equilibrium. We work in the context of our example, defined by the parameters in (9), and $R_B = 1.08$.

There are two approaches to eliminating runs within the separated financial system. One approach, which turns out to be the best here, is to continue to provide 1 unit of consumption in period 1,³⁵ and increase γ to the point at which $c_P^2(0.5) = 1$. From (16), we can calculate the optimal γ to avoid runs, $\gamma^{**} = 0.09630$. It is now feasible and welfare-maximizing to offer full consumption smoothing, as in (13). Welfare in this contract that eliminates runs is $W^{norun} = 0.8651$. The other approach to eliminating runs is to ration consumers in period 1, thereby retaining enough consumption to guarantee that $c_P^2(0.5) = 1$. This approach is

³³We impose the condition that the run equilibrium is only selected if patient depositors strictly prefer to run when all other patient depositors run. Otherwise, the bank’s optimization problem would have no solution for some propensities to run.

³⁴We restrict attention to contracts for which there is an equilibrium in which all patient consumers wait. See Ennis and Keister (2003) for an analysis of more general equilibrium-selection mechanisms. For example, as in Ennis and Keister, the propensity to run could depend on the strength of the incentive to run provided by the bank’s contract, which in turn would influence the selection of the optimal contract.

³⁵That is, $c^1(z) = 1$ for $z \leq \bar{\alpha}$, and $c^1(z) = 0$ for $z > \bar{\alpha}$.

too costly here. However, if we change the density function, f , so that high realizations of α are sufficiently unlikely, then the best way to avoid runs might be to make the impatient consumers suffer in the extremely unlikely event of large α .

We now determine the optimal contract that has a run equilibrium, given that a run occurs with probability s_0 and that the equilibrium in which patient consumers wait occurs with probability $(1 - s_0)$. Patient consumers must be induced to wait when $s \geq s_0$, and there is no attempt to provide incentives when $s < s_0$, so the incentive compatibility constraint is unchanged. It can be shown that the incentive compatibility constraint binds, so investment in technology B is unchanged, $\gamma = \gamma^* = 0.09445$. In other words, the probability of a run has no effect whatsoever on the optimal contract, as long as the probability is small enough so that the optimal contract tolerates runs. When we have $s < s_0$, there is no reason to punish consumers for running. It turns out that everyone will receive 1 unit of consumption in period 1, until the bank runs out. The gain from providing for the consumption opportunities of a significant minority of impatient consumers outweighs the foregone yield of R_B . Thus, the probability of receiving 1 unit in period 1 is γy , which equals 0.09445, so the proportion of consumers taking advantage of their undiscounted consumption opportunity in state α is $(1 - \alpha + \alpha \gamma y)$. Conditional on a run taking place, welfare is $W^{run} = 0.1978$. Overall welfare is a weighted average of W^{run} and the welfare when a run does not take place, $W^* = 0.8688$. Thus, the cutoff value of s_0 solves

$$s_0 W^{run} + (1 - s_0) W^* = W^{norun}. \quad (18)$$

From (18), we calculate the cutoff value of s_0 to be 0.005521. If the propensity to run is less than 0.5521%, then it is better to tolerate the unlikely event of a run than to increase technology B investment to prevent runs. In other words, the cost of eliminating this fragility from the financial system is too high, in terms of efficiency when the system is working

smoothly.³⁶ On the other hand, if the propensity to run is greater than 0.5521%, then the bank will not tolerate runs. In this case, the equilibrium allocation differs from any equilibrium allocation of the the NROC in the post-deposit game.

6 Concluding Remarks

What are the differences between our model and the closely related narrow banking model of Neil Wallace (1996)? Wallace’s technologies are slightly different from ours, allowing some liquidation value of technology A in period 1, and having technology B ’s return between period 0 and 1 equal the return between period 1 and 2. These differences are probably not crucial. Wallace’s utility functions do not exhibit any indivisibilities (such as our urgent consumption opportunities), but they are otherwise more general than ours. We argue below that the indivisibility is probably not crucial for most of our basic results. The extra structure we impose on utility functions allows us to maintain tractability while generalizing Wallace (1996) in other directions. In particular, we allow α to be stochastic and only impose the condition that the distribution be continuous, while Wallace requires α to be deterministic. We believe that the introduction of stochastic aggregate fundamentals, in our case the random α , dramatically changes the problem.³⁷

Wallace’s working definition of narrow banking is that all of the bank’s obligations to depositors must be met for *every* possible pattern of withdrawals. This amounts to a restriction that the space of contracts is limited to a menu of consumption bundles that is independent of the history of withdrawals. On the other hand, Wallace’s banks are permitted to invest in both technologies. Wallace shows that, without subordinated debt contracts,

³⁶See the papers by Lagunoff and Schreft (1998) and Allen and Gale (1998) for an analysis of financial crises based on local interactions. Our notion of fragility is distinct from theirs, yet not altogether unrelated.

³⁷The main intrinsic uncertainty in our formal model comes from the urgent transaction opportunities of consumers. We expect quite similar results in the case where there is also uncertainty about the investment returns, R_A and R_B .

any allocation achievable with narrow banking is also achievable with autarky, so that narrow banking eliminates the role for banks. On the other hand, our notion of the separated financial system is closer to the description of narrow banking in the introduction to Wallace (1996), where banks are (only) restricted to holding liquid (short-term) assets. We allow deposit contracts to specify withdrawals that are fully contingent on the history, but restrict the banks' portfolios to technology B assets. Because our portfolio restrictions are different from Wallace's restrictions on the banking contract, banks play a very important role in our separated financial system.

We believe that the indivisibility in our model is well-motivated, meant to capture the idea that checks must clear at par in order to perform their transactions function. The indivisibility assumption also simplifies the analysis. It facilitates calculation of the optimal contract. We conjecture, however, that the indivisibility assumption is not crucial for our basic results. For the discussion below, we drop the indivisibility assumption — supposing instead that the impatient consumers have smooth utility functions over period-1 *and* period-2 consumption, and the patient consumers have smooth utility functions over period-2 consumption.

Under the unified system, all of the technology B investments will have to be liquidated in state $\bar{\alpha}$, or else higher welfare can be achieved by investing more in technology A , without affecting the consumption of the impatient. In the special case of deterministic α , technology B investments will be used exclusively for impatient consumers and will be fully liquidated. When α is random, rather than rationing the impatient consumers, the optimal contract will likely have period-1 consumption tapering off as more consumers arrive, exhibiting the “partial suspension of convertibility” obtained by Wallace (1988, 1990), Green and Lin (2003), and Peck and Shell (2003). We conjecture that run equilibria are impossible under the optimal contract, at least for reasonable parameters in which period-2 consumption demand is substantial.

Under the separated system, the typical case will be overinvestment in technology B . For example, if α is deterministic, then the unified system harvests all of the technology B investments in period 1, leaving nothing for the patient consumers.³⁸ To satisfy incentive compatibility, the optimal contract in the separated system will invest more in technology B and/or reduce consumption in period 1. When α is random, but not too volatile, we anticipate the same conclusion.

We conjecture that run equilibria are possible at the optimal contract in the separated system. Our basic intuition remains: Investment in technology B must satisfy incentive compatibility for each patient consumer, based on the other patient consumers waiting. A patient consumer is indifferent between choosing to arrive at the bank during period 1 and choosing to arrive during period 2, weighing all possible values of α . State $\bar{\alpha}$ is likely to be one of the states in which a patient consumer is better off choosing period 1, because this is the state in which the most investments are harvested in period 1 and the fewest remain for period 2. If so, then there is a run equilibrium.

In our model, as in Wallace (1996), the banking restrictions under consideration cannot improve welfare and can cause harm. Removing the restrictions will improve welfare. We find that the optimal contract in the separated financial system tends to exhibit overinvestment in the liquid technology. In our model, there is always a bank run equilibrium in the separated financial system, but bank runs will occur with probability zero in the unified financial system.³⁹ This is in contrast to the widespread view that narrow-banking-type restrictions provide stability. In response to the world's current financial crisis, we hope that proposals to

³⁸As in our setup, the patient consumers are compensated by a bigger share of the technology A investments in the unified system. Notice that it is difficult even to talk about the separated system without the impatient caring about future consumption. If the impatient were finished with consumption in period 1, what would happen to their investments in technology A ?

³⁹On the other hand, the unified system is subject to non-run rationing in period 1 if the realization of α is sufficiently large.

restrict the activities of firms, and claims that restrictions promote stability, are scrutinized carefully.

References

Allen, Franklin and Gale, Douglas. "Financial Contagion." *Journal of Political Economy*, 108(1) (2000): 1-33.

Calomiris, Charles W. and Kahn, Charles M. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review*, 81(3) (1991): 497-513.

Champ, Bruce; Smith, Bruce; and Williamson, Stephen. "Currency Elasticity and Banking Panics: Theory and Evidence," *Canadian Journal of Economics*, 29 (1996): 828-864.

Cooper, Russell and Ross, Thomas W. "Bank Runs: Liquidity Costs and Investment Decisions." *Journal of Monetary Economics* 41 (1998): 27-38.

Diamond, Douglas W. "Liquidity, Banks and Markets." *Journal of Political Economy*, 105(5) (1997): 928-56.

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

Diamond, Douglas W. and Rajan, R. G. "Liquidity Risk, liquidity creation, and financial fragility: A theory of banking," *Journal of Political Economy*, 109(2) (2001): 287-327.

Ennis, Huberto M. and Keister, Todd. "Economic Growth, Liquidity, and Bank Runs," *Journal of Economic Theory*, 109(2) (2003): 220-45.

Ennis, Huberto M. and Keister, Todd. "Commitment and Equilibrium Bank Runs," March 2006 Working Paper.

Ennis, Huberto M. and Keister, Todd. "Run Equilibria in a Model of Financial Intermediation," December 2007 Working Paper.

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

Haubrich, Joseph G. “Optimal Financial Structure in Exchange Economies,” *International Economic Review*, 29(2) (1988), 217-235.

Jacklin, Charles J. “Demand Deposits, Trading Restrictions, and Risk Sharing.” In *Contractual Arrangements for Intertemporal Trade*, edited by Edward J. Prescott and Neil Wallace. Minneapolis: University of Minnesota Press, 1987.

Kiyotaki, Nobuhiro and Moore, John. “Credit-Cycles.” *Journal of Political Economy*, 105(2) (1997): 211-48.

Lacker, Jeffrey M., “Financial Stability and Central Banks,” Distinguished Speakers Seminar, European Economics and Financial Centre, London, England, June 5, 2008.

Lagunoff, Roger and Schreft, Stacey. “A Model of Financial Fragility.” *Journal of Economic Theory*, 99(1-2) (2001): 220-64.

Peck, James. “Demand Uncertainty, Incomplete Markets, and the Optimality of Rationing.” *Journal of Economic Theory*, 70(2) (1996): 342-63.

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

Postlewaite, Andrew and Vives, Xavier. “Bank Runs as an Equilibrium Phenomenon.” *Journal of Political Economy*, 95(3) (1987): 485-491.

Wallace, Neil. “Another Attempt to Explain an Illiquid Banking System: The Diamond-Dybvig Model with Sequential Service Taken Seriously.” *Quarterly Review of the Federal Reserve Bank of Minneapolis*, 12(4) (1988): 3-16.

Wallace, Neil. “A Banking Model in Which Partial Suspension is Best.” *Quarterly Review of the Federal Reserve Bank of Minneapolis*, 14(4) (1990): 11-23.

Wallace, Neil. “Narrow Banking Meets the Diamond-Dybvig Model.” *Quarterly Review of the Federal Reserve Bank of Minneapolis*, 20(1) (1996): 3-13.