A Model of Interbank Settlement

Benjamin Lester *
Department of Economics
University of Pennsylvania
lester@econ.upenn.edu

October 12th, 2005

Abstract

A settlement system is a set of rules and procedures that govern when and how funds are transferred between banks. Perhaps the most crucial feature of a settlement system is the frequency with which settlement occurs. On the one hand, a higher frequency of settlement limits the risk of default should a bank be rendered insolvent. On the other hand, a lower frequency of settlement is less costly for banks to operate. We construct a model of the banking sector in which this trade-off between cost and risk arises endogenously. We then complete the economy with a trading sector that has a micro-founded role for credit as a media of exchange. The result is a general equilibrium model that allows for welfare and policy analysis. We parameterize the economy and study the optimal intra-day borrowing policy that the operator of a settlement system should impose on member banks. We also determine conditions under which one settlement system is more appropriate than another.

*I am particularly indebted to Randy Wright for his guidance throughout the writing of this paper, and to Ed Green and Ludo Viscchers for their suggestions. I thank Aleksander Berentsen, Ping He, Ed Nosal, David Mills, Deniz Selman, Skander Van den Heuvel, Alberto Trejos, and seminar participants at the Federal Reserve Bank of Cleveland, the Federal Reserve Bank of New York, the Bank of Canada and the Penn Macro Lunch for helpful commentary. Finally, I'd also like to thank Stephen Millard and Matthew Willison for introducing me to this topic. All errors are my own.

1 Introduction

Many purchases, particularly of large value, are paid for with credit; that is, a buyer transfers to the seller a claim on funds held at his financial institution, which the seller can subsequently redeem at her financial institution. In such a transaction, the exchange of goods and the finalization of payment are separate events, and therefore require a specification for when and how the actual funds are delivered from the buyer's financial institution to the seller's. Such a specification, called a *settlement system*, is a "contractual and operational arrangement that banks and other financial institutions use to transfer funds to each other." (Zhou, 2000)

As a result of recent advances in technology, as well as international financial integration, both the regularity and the value of these transactions have increased dramatically. For example, in the United States the average *daily* value of transfers processed by the two primary settlement systems exceeds three trillion dollars, or approximately one third of *annual* GDP.[1] Therefore, an efficient, stable settlement system is an increasingly essential component of a well-functioning modern economy. While there are many features to a settlement system, perhaps the most crucial is the frequency with which settlement occurs. This paper develops a framework to analyze the costs and risks associated with settlement frequency.

As the frequency of settlement increases, the settlement system becomes more costly to banks. While there are several sources of this costliness, this paper focuses on *financial costs*.[2] If banks do not have on hand sufficient reserves to finalize payments at the time of settlement, they must borrow the necessary funds.

[1] These statistics are similar across most developed countries. For example, in the United Kingdom and Canada the average daily value of transactions equates to approximately 20% and 15%, respectively, of annual GDP.

[2] A second type of costs that arise as settlement frequency increases are *resource costs*. These would include expenditures for the transfer, collection, accounting, and monitoring of payment finalization, among many others. We ignore these here. See Berger, Hancock, and Marquardt (1996) for a more complete exposition of costs inherent in payment systems.

This borrowing is costly, as nearly all systems impose either interest charges or collateral requirements for borrowing, or penalties on banks that are unable to meet their liabilities. Therefore, when a bank is forced to make settlements more often, it chooses to hold a larger amount of idle, liquid reserves. This sub-optimal portfolio allocation bears the opportunity cost of holding assets that could be generating revenue in the loan market.

In a world with perfect commitment between banks, the optimal settlement system would simply be the least costly, and therefore frequent settlement would be undesirable. However, in reality we observe occasional events in which a financial institution is unable to deliver payments to finalize its customers' transactions, and therefore must default on its outstanding liabilities. The settlement system in place will be a crucial determinant of how the economy as a whole is affected by such events.[3] Increased settlement frequency mitigates this risk in two important ways. First, as the frequency of settlement increases, the number of parties awaiting payment decreases. Therefore, a settlement system that reduces the time between the initial transaction and the finalization of payment decreases the potential size of a default. Secondly, since banks retain a higher level of reserves under systems with more frequent settlement, they are more equipped to honor outstanding liabilities in the case of insolvency.

Therefore, there exists tension between the risk associated with large defaults and the costs of controlling that risk. In Section 3, we begin by constructing a partial equilibrium model (in the sense that the price of loans is exogenous) of the banking sector in which a bank's intra-day net liabilities are stochastic, borrowing is costly, and banks are subject to random, exogenous solvency shocks. Within this framework, the risk-cost tradeoff discussed above arises naturally. In Section 4, we introduce a simple example to illustrate the key results from the partial equilibrium model. We proceed with sensitivity analysis, which allows us

[3] There are several papers that assess the performance of a settlement system after a large disruption to the financial sector. See Bernanke (1990) for analysis of clearing and settlement during the 1987 stock market crash, and Fleming and Garbade (2002) for an account of settlement performance after the September 11th terrorist attacks.

to better understand how the magnitude of the risk-cost tradeoff is affected by changes to the parameters of the model economy.

The partial equilibrium example is a useful tool in that it clearly illustrates the tradeoff within the banking sector. However, it takes as given the demand for loans, as well as the willingness of buyers and sellers to trade, and therefore does not consider how their behavior might respond to changes in the costs or risks of banking. As a result, any attempt at welfare analysis within the partial equilibrium setting would require arbitrary weights on the costs and risks of settlement frequency. In Section 5, we complete the economy by introducing a trading sector in which agents conduct transactions using bank credit as the means of payment. The complete economy is a general equilibrium model in the usual sense: there are two sectors, and variables that are endogenously determined in one sector are taken as given in the other. In equilibrium, of course, there is aggregate consistency between the behavior in the two sectors. The complete model allows us to analyze behavioral and welfare implications of the risk-cost tradeoff within the context of a monetary economy.

In section 6, we extend the example introduced earlier to the general equilibrium framework. By considering welfare under alternative settlement frequencies, borrowing policies, and parameter specifications, several insights emerge. First, we find that the determination of the optimal settlement system requires *joint* consideration of both settlement frequency and intra-day borrowing practices. This is a natural result: the frequency of settlement and the costly provision of liquidity are both levers that can be employed to reduce the risks associated with default. However, the method by which they do so is different, as are the financial inefficiencies that they introduce. Therefore, to find the optimal settlement system, one must determine both the proper balance between risk and cost, *and* the most efficient way of achieving that balance. Our welfare analysis also reveals that the optimal use of these two levers will depend heavily on the underlying parameters of the economy. In particular, the factors governing the probability of solvency and liquidity shocks, as well as the costs of recovering from a default, will be the primary determinants of the optimal settlement system. In an economy in which these factors generate a sufficiently high risk, some combination of real-time settlement and costly provision of liquidity may be desirable in order to curb the effects of potential defaults. On the other hand, should these risks be small relative to the financial costs incurred, less frequent settlement and/or free intra-day liquidity may be optimal.

Section 7 concludes, and offers potential extensions of the model.

2 Literature Review

Following Freeman (1996), many economists have assessed settlement system issues with relatively abstract models in which the agents *are* the banks, and the need for a third party to provide liquidity arises from spatial and intertemporal miscoordination. The first generation of these papers, which would include Freeman (1999) and Zhou (2000), introduced default as an exogenous event and concluded that intra-day liquidity should be provided free of charge. In more recent work by Mills (2005) and Koeppl, Monnet, and Temzelides (2005), default has been modeled as the product of a moral hazard; that is, banks *choose* to default on their outstanding debt because it is no longer incentive compatible to repay.[4] In this framework, the authors consider various instruments that the central counter-party may wish to implement in order to limit the risk of default, including collateral requirements, interest charges, and borrowing caps. Their findings suggest that the free provision of intra-day liquidity may not be optimal in the presence of a moral hazard problem.[5]

The approach in these papers is appealing because there is a natural definition of welfare and a transparent role for a welfare-improving monetary authority. However, since these models require a highly stylized timing sequence, they do

[4]Martin (2004) implements a similar model, though he motivates default in a manner more similar to the current paper.

[5]Note that we characterize collateral requirements as being costly, as they generally introduce opportunity costs of foregone investment.

not easily accommodate analysis of one of the key features of a settlement system: the frequency with which settlement occurs. Moreover, by modelling banks and consumers as a single agent, the authors ignore the fact that the two parties have different incentives. As a result, these models will fail to capture the behavior of buyers and sellers in response to a banks choice of risk, and the effect of that behavior on instruments (such as loan prices) that relate consumers and banks.

In response to the first of these limitations, several models have been developed which showcase a relatively sophisticated notion of banks. In Kahn and Roberds (1998), for example, banks must choose the proper portfolio of liquid and non-liquid assets given that intra-day liabilities (to and from other banks) are stochastic.[6] In the presence of a moral hazard problem similar to those discussed above, the model successfully generates the trade-off between cost and risk that is inherent to the choice of settlement frequency. However, as they ignore the behavior of the agents involved in the underlying transactions, Kahn and Roberds require arbitrary weights to conduct welfare and policy analysis.

The current paper hopes to bridge the gap between the two strands of literature by constructing a model that has both a realistic banking environment which allows for the analysis of settlement frequency *and* a trading sector that provides a natural definition of welfare. Yet there remains an important difference between much of the existing literature and the framework developed here. Many of the papers above incorporate risk into a model of payments via *strategic default*; when a bank is *able* to settle its debts, but it *chooses* not to. Policy analysis, therefore, is tantamount to finding the mechanism that properly balances the benefits of (a) maintaining banks incentive compatibility constraints, and (b) ensuring that all welfare-increasing transactions are executed. In contrast, this paper concerns itself solely with solvency shocks; when a bank simply *cannot* settle its outstanding liabilities.[7] As a result, the risk of default arises from the fact that, when choosing the profit-maximizing level of reserves, banks

[6] Also see Kahn and Roberds (2001), and Kahn, McAndrews, and Roberds (2003).

[7] See Rochet and Tirole (1996) for an excellent discussion on why solvency shocks, and not liquidity shocks, should be the primary area of concern.

do not internalize the effects of a default on the economy as a whole. Therefore, when exogenous solvency shocks occur, the size of banks' default is larger than is socially optimal. This motivates the role of central bank policy to promote more conservative reserve management.

Zhou (2000) specifies four main criteria that a model of payment systems should satisfy. First, she proposes that it should directly model the exchange of goods, so that the structure of the system affects the allocation. Secondly, the model should treat separately the debt issued as a means of payment and the debt assumed by the banks for settlement. Finally, she suggests that the model incorporates both liquidity shortage and credit risk, and that these should be generated endogenously by the actions of the agents. In what follows, we construct a model that is able to satisfy each of these four criteria.

3 The Banking Sector

We consider a discrete time, infinite horizon model in which each period, or "day", consists of a finite number of sub-periods. There are a large number of ex-ante identical banks, each with an initial mass of depositors d. Upon receiving a deposit, a bank issues the depositor some form of receipt that can then be used to purchase goods in the trading sector. We loosely refer to this receipt as a "check," though it should be understood that the usual settlement system associated with checks does not necessarily apply. Banks face a fixed cost a to manage each deposit, and charge each depositor a fee ϕ. It is possible, however, that $\phi < 0$, in which case the bank must pay the depositor. This case has the natural interpretation of an interest-bearing bank account.

Before the day's trading begins, banks choose a fraction α of initial deposits to retain as reserves, and loan out the remaining $1-\alpha$ in exchange for an up-front fee ρ. We assume that the probability of an agent redepositing a loan at the lending bank is zero, so that each bank has a mass d of active *buyers* in the market. We also assume that *sellers* are pre-assigned a bank to deposit any payments received, and that sellers are uniformly distributed across banks. Sellers deposit checks immediately after the transaction takes place. Therefore, throughout the day, a bank accrues *due-to*'s and *due-from*'s as its buyers and sellers, respectively, trade with other banks' customers in the marketplace. A bank's net liabilities at any time are equal to the difference between due-to's and due-from's.

Suppose that transactions occur at T discrete intervals throughout each day, and that a bank has cumulative intra-day net liabilities $l_t \in [\underline{L}_t, \bar{L}_t]$ at each $t \in \mathcal{T} \equiv \{1, 2, ..., T\}$, where $\bar{L}_t$ ($\underline{L}_t$) is the maximum (minimum) net position that occurs at time t with strictly positive probability, and $-d \leq \underline{L}_t < 0 < \bar{L}_t \leq d$ for all $t \in \mathcal{T}$. We denote the maximum possible net liability position $\bar{L} = \max_{t \in \mathcal{T}}\{\bar{L}_1, ..., \bar{L}_T\}$. Therefore, a bank's intra-day net liabilities is a T x 1 random variable $l \equiv (l_1, ..., l_T)$ with density $\lambda(\cdot)$, which we assume is continuous. Any outstanding liabilities are settled at the end of each day, so that each period begins with $l_0 = 0$. Moreover, each night deposits are reallocated across banks so that the following morning all banks begin with d deposits.

3.1 Notation, Definitions, and Key Concepts

A settlement system specifies the frequency with which banks must settle net liabilities incurred since the previous settlement. For a given frequency $F \leq T$, banks will settle at intervals of length $S = \frac{T}{F}$.[8] Let $\mathcal{S} \subset \mathcal{T}$ denote the subset of points in time in which banks settle, so that $\mathcal{S} \equiv \{S, 2S, ..., FS\}$, where $FS = T$. Also, for any given point in time t, let s_t denote the time of the previous settlement. That is,

$$s_t = \begin{cases} \max\{t' \in \mathcal{S} : t' < t\}, & \text{for } t > S \\ 0, & \text{for } 0 \leq t \leq S \end{cases}$$

The settlement system is operated by a third party which we will refer to as the central bank, as this is both convenient and consistent with many settlement

[8]For simplicity, we consider frequencies for which S is an integer.

systems world-wide. Each bank has an account at the central bank, where it keeps its reserves. At the time of settlement, the central bank will transfer the appropriate amounts *from* the accounts of banks that have accrued positive net liabilities since the previous settlement *to* the accounts of banks that have accrued negative net liabilities since the previous settlement. Our analysis thus requires that we keep track of a bank's reserves held at the central bank.

Definition 1 *Let a bank's* ***reserves at time*** t *for any realization of net liabilities* l*, reserve ratio* α*, and settlement frequency* F *be denoted* $R_t(l, \alpha; F)$*, and be recursively defined:*

$$R_t(l, \alpha; F) = \begin{cases} \alpha d, & \textit{for } t = 0 \\ \max\{0, R_{s_t}(l, \alpha; F) - (l_t - l_{s_t})\}, & \textit{for } t \in \mathcal{S} \\ R_{s_t}(l, \alpha; F), & \textit{for } t \notin \mathcal{S} \end{cases}$$

As the formula suggests, a bank's account balance at the central bank is non-negative, increases at settlements in which the bank receives payment (when $l_t - l_{s_t} < 0$ for $t \in \mathcal{S}$), and decreases when a bank must settle outstanding liabilities (when $l_t - l_{s_t} > 0$ for $t \in \mathcal{S}$). Notice that for any t that does not correspond to a settlement period, the level of reserves does not change.

Should a bank's obligations exceed its reserves at any settlement time $t \in \mathcal{S}$, it must borrow the difference, $(l_t - l_{s_t}) - R_{s_t}(l, \alpha; F)$, from the central bank.

Definition 2 *Let a bank's* ***total intra-day borrowing requirement*** *for a given realization of net liabilities* l*, reserve ratio* α*, and settlement frequency* F *be denoted* $b(l, \alpha; F)$ *and be defined:*

$$b(l, \alpha; F) = \sum_{t \in \mathcal{S}} \max\{0, (l_t - l_{s_t}) - R_{s_t}(l, \alpha; F)\}.$$

At the end of the trading day, a bank is assessed a fee based on it's total intra-day borrowing requirement given by $\psi\big[b(l, \alpha; F)\big]$. In reality, these fees are imposed through a combination of interest charges, collateral requirements, and borrowing caps (in which a bank is penalized for exceeding a predetermined borrowing limit).[9] Moreover, many systems maintain deductibles for small intra-day overdrafts, while they levy far more severe penalties on large overdrafts. Therefore, the following assumption on the convexity of ψ is not only convenient in that it provides curvature to our problem, but also seems consistent with real world practice.[10]

[9] For a detailed comparison of these various instruments, see Mills(2005).

[10] It is important to note that we are beginning with the assumption that the central bank

Assumption 1 *The cost function for borrowing fees, $\psi(\cdot)$, is C^2, strictly increasing, strictly convex, and satisfies the property $\psi(0) = 0$.*

Since banks choose their reserve levels before the start of the day, they must project borrowing fees for any given level of reserves.

Definition 3 *Let a bank's **expected borrowing fees**, given reserve ratio α and settlement frequency F, be denoted $\Psi(\alpha; F)$ and be defined:*

$$\Psi(\alpha; F) = \mathbb{E}_l \Big[\psi\big[b(l, \alpha; F)\big] \Big] = \int \psi\big[b(l, \alpha; F)\big] \lambda(l) dl.$$

Now that we have defined the structure of the banking sector, we are ready to turn our analysis to the problem facing a profit-maximizing bank.

3.2 The Problem of the Bank

For each deposit, a bank receives a deposit fee (or pays interest) ϕ and earns additional revenue $(1-\alpha)r\rho$ in loans, where r is the rate of time preference of the bank's customers. However, a bank must spend a fixed amount a to manage each account, and expects an additional borrowing fee $\Psi(\alpha; F)$. Therefore, a bank's expected profits are given by

$$\pi(\phi, \alpha; F) = d[\phi + (1-\alpha)r\rho - a] - \Psi(\alpha; F). \tag{1}$$

Assuming perfect competition in the banking sector, we use the zero expected profit condition to get

$$\phi = a - (1-\alpha)r\rho + \frac{\Psi(\alpha; F)}{d}. \tag{2}$$

We are able to characterize $\alpha(F)$, the optimal level of reserves that a profit-maximizing bank will choose under a settlement system with frequency F, as the

only provides liquidity at a cost. We will later relax this assumption to consider the optimality of free provision of intra-day liquidity.

solution to the first order condition[11]

$$-\frac{\partial}{\partial\alpha}\left[\Psi(\alpha;F)\right] = dr\rho. \tag{3}$$

This first order condition has the usual interpretation: the marginal benefit from an increase in the reserve ratio, resulting from a reduction in expected borrowing fees, is equated to the marginal cost of foregone loan revenue.

3.3 Bank Insolvency

In this paper, we take the view that defaults tend to be the result of sudden events that render a bank insolvent. Such events may be caused by poor investments (Continental Illinois Bank or the creditors to Long Term Capital Management), fraud (Barings Bank), or other calamities (such as a hurricane or tsunami). There is a large literature devoted to the analysis of solvency shocks, the difficulties of distinguishing them from liquidity shocks in an environment with imperfect monitoring, and the optimal bail-out policy for the central bank.[12] In addition, there is an equally large literature on bankruptcy laws, and how the priority of claimants to an insolvent institution affects the bank's behavior. However, since the current paper wishes to focus on settlement systems, we will abstract from many of these intricacies and instead motivate insolvency and default by the following simple observations: (1) from time to time, a bank is suddenly rendered insolvent, and (2) an insolvent bank with more available capital (or reserves) is better able to settle outstanding debts than an insolvent bank with fewer available funds. We formalize these ideas below.

3.3.1 Solvency Shocks

Suppose that every day a stochastic fraction Δ of banks are rendered insolvent just prior to the final settlement at $t = T$. We assume that this fraction is i.i.d.

[11] Here we restrict our attention to interior solutions, though we consider corner solutions in the example presented in Section 3. Also, notice that this is a strictly concave function over a convex set, so that we are assured of a unique solution.

[12] See, for example, Aghion, Bolton, and Fries(1999), and Freixas, Parigi, and Rochet(2004).

across days and distributed over the $[0, 1]$ interval according to density $g(\cdot)$. We denote the expected fraction of insolvent banks $\bar{\Delta}$, where

$$\bar{\Delta} = \int \Delta \cdot g(\Delta)\ d\Delta.$$

3.3.2 Default

We make two simplifying assumptions that govern our notion of default. First, we assume that the central bank perfectly observes solvency shocks, and thus will never extend a loan to an insolvent bank.[13] Secondly, we assume that banks awaiting payment are given top priority to an insolvent bank's available assets. Therefore, there are two important statistics to consider when a bank is rendered insolvent: its remaining liabilities $(l_T - l_{s_T})$ and its available assets (R_{s_T}). Should an insolvent bank's available assets be greater than its remaining liabilities (so that $R_{s_T} \geq l_T - l_{s_T}$), the bank is simply liquidated, all outstanding transactions are finalized, and the bank is replaced by an identical, solvent bank. However, an insolvent bank with remaining liabilities greater than available assets (so that $R_{s_T} < l_T - l_{s_T}$) must default on all liabilities in excess of its current liquidation value (i.e. its current reserves). For the ensuing analysis, the following definition is convenient.

Definition 4 *Let a bank's* ***outstanding liabilities in excess of reserves at time*** *t for a given realization of net liabilities l, reserve ratio α, and settlement frequency F be denoted $\hat{l}_t(l, \alpha; F)$ and be recursively defined:*

$$\hat{l}_t(l, \alpha; F) = \max\{0, l_t - l_{s_t} - R_{s_t}(l, \alpha; F)\} \qquad \textit{for all } t \in \mathcal{T},$$

with $\hat{l}_0(l, \alpha; F) = 0$.

[13] That insolvency is perfectly - and in particular instantaneously - observed implies that an insolvent bank would be forced to default immediately, regardless of whether or not there was a scheduled settlement during that interval. Under this assumption, all of the ensuing results are maintained under the more general assumption that there exists a probability of default at every interval $t \in \mathcal{T}$.

Note that, for a solvent bank, this is simply the amount they need to borrow at time t. However, for a bank that is insolvent at time T, $\hat{l}_T(l, \alpha; F)$ is the size of default. Assuming that l is i.i.d. across banks, the law of large numbers implies that the total defaults which occur in a given day can be defined in the following manner.

Definition 5 *Let the* ***aggregate daily default under settlement system*** F *for a given realization of* Δ *be denoted* $\delta(\Delta, F)$ *and be defined*

$$\delta(\Delta, F) = \Delta \int \hat{l}_T(l, \alpha(F); F)\ \lambda(l)\ dl \equiv \Delta\ \mathbb{E}_l\Big[\hat{l}_T(l, \alpha(F); F)\Big]. \tag{4}$$

Moreover, if Δ and l are independently drawn, the expected value of defaults is also easily defined.

Definition 6 *Let the* ***expected aggregate daily default under settlement system*** F *be denoted* $\bar{\delta}(F)$ *and be defined*

$$\bar{\delta}(F) = \int \delta(\Delta, F)\ g(\Delta)\ d\Delta \equiv \bar{\Delta}\ \mathbb{E}_l\Big[\hat{l}_T(l, \alpha(F); F)\Big]. \tag{5}$$

3.4 The Tradeoff Between Cost and Risk

We now turn to the relationship between those variables determined endogenously in the model - namely the reserve ratio α and the subsequent borrowing fee Ψ, banking fee ϕ, and default size δ - and the exogenously specified frequency of default F and borrowing policy ψ. We present our results in the context of the special case of $T = 2$, as this captures the key insights most clearly. However, this is *only* for the ease of exposition; the analagous results for an arbitrary, finite T, along with the necessary proofs, are presented in the appendix.

For $T = 2$, there are only two possible settlement systems. The first, with settlement frequency $F_1 = 1$, is a *deferred* settlement system, as transactions occurring at $t = 1$ are not finalized until the end of the period (at $t = 2$). The second settlement system, with frequency $F_2 = 2$, is a *real-time* settlement system, as transactions are finalized as they take place.

3.4.1 The Costs of Frequent Settlement

We first illustrate how increasing the frequency of settlement will increase the financial costs to banks, and subsequently increase the fees paid (or decrease the interest earned) by depositors. In what follows, we restrict α to the relevant open set $(0, \frac{\bar{L}}{d})$, so that banks hold strictly positive reserves, but not amounts greater than or equal to the maximum net liability position.[14]

Claim 1 *Increasing the reserve ratio decreases the expected borrowing fees at any settlement frequency. That is, for* $F \in \{F_1, F_2\}$,

$$\frac{\partial}{\partial \alpha}\left[\Psi(\alpha; F)\right] < 0.$$

Claim 2 *For any level of reserves, expected borrowing fees are larger under a real-time settlement system than they are under a deferred settlement system. That is,* $\forall\ \alpha \in (0, \frac{\bar{L}}{d})$,

$$\Psi(\alpha; F_1) \leq \Psi(\alpha; F_2).$$

The claims above allow us to formalize the relationship between settlement frequency and the endogenously determined reserve ratio.

Result 1 *Profit-maximizing banks choose to retain a larger fraction of deposits as reserves under a real-time settlement system than they do under a deferred settlement system. That is,*

$$\alpha(F_1) \leq \alpha(F_2).$$

Since banking fees are a function of the reserve ratio and the reserve ratio is endogenously determined as the optimal solution to the banking problem under settlement system F, ϕ can be expressed as a function of F:

$$\phi(F) = a - [1 - \alpha(F)]r\rho + \frac{\Psi(\alpha(F); F)}{d}. \tag{6}$$

[14]The assumption that $\alpha^* > 0$ can be motivated by the assumption that ψ is sufficiently steep, while $\alpha^* < \frac{\bar{L}}{d}$ can be justified by assuming that ψ is not too steep. Therefore, bounding $D\psi$ is a sufficient condition for interiority.

We are now ready to formalize the first major result, relating settlement frequency to the banking fees faced by agents.

Result 2 *For any loan price ρ, banking fees are higher under a real-time settlement system than they are under a deferred settlement system. That is,*

$$\phi(F_1) \leq \phi(F_2).$$

3.4.2 The Risks of Infrequent Settlement

Though increased settlement frequency is costly, it decreases the risk posed by bank insolvency. In particular, a higher frequency of settlement reduces the size of potential defaults in two important ways.[15] On the one hand, more frequent settlement prohibits the accumulation of large outstanding net liability positions. We refer to this as the *accumulation effect*, and present it formally with the following result.

Result 3 *At any time $t \in \mathcal{T}$ and for any reserve ratio α, the size of outstanding liabilities in excess of reserves is always smaller under a real-time settlement system than under a deferred settlement system. That is, for any t, α, and l,*

$$\hat{l}_t(l, \alpha; F_1) \geq \hat{l}_t(l, \alpha; F_2).$$

The second way in which increased frequency decreases risk, which we will call the *reserve effect*, is illustrated in Result 4. It suggests that the size of a default is decreasing in α, the reserve ratio.

Result 4 *At any time $t \in \mathcal{T}$ and for any settlement frequency $F \in \{F_1, F_2\}$, increasing the reserve ratio decreases the size of outstanding liabilities in excess of reserves. That is, for any t, F, and l,*

$$\alpha_2 > \alpha_1 \quad \Rightarrow \quad \hat{l}_t(l, \alpha_1; F) \geq \hat{l}_t(l, \alpha_2; F).$$

[15] Kahn, McAndrews, and Roberts (2003) point out a third effect of settlement frequency on default that we do not consider. They point out that deferred settlement may decrease the risk of "gridlock", and thus defaults arising from liquidity shortages.

From Results 1, 3, and 4, we trivially arrive at the second major result, relating settlement frequency to total expected defaults.

Result 5 *For any realization of* Δ*, the aggregate daily default is larger under a deferred settlement system than it is under a real-time settlement system. That is, for any* $\Delta \in [0, 1]$*,*

$$\delta(\Delta, F_1) \geq \delta(\Delta, F_2).$$

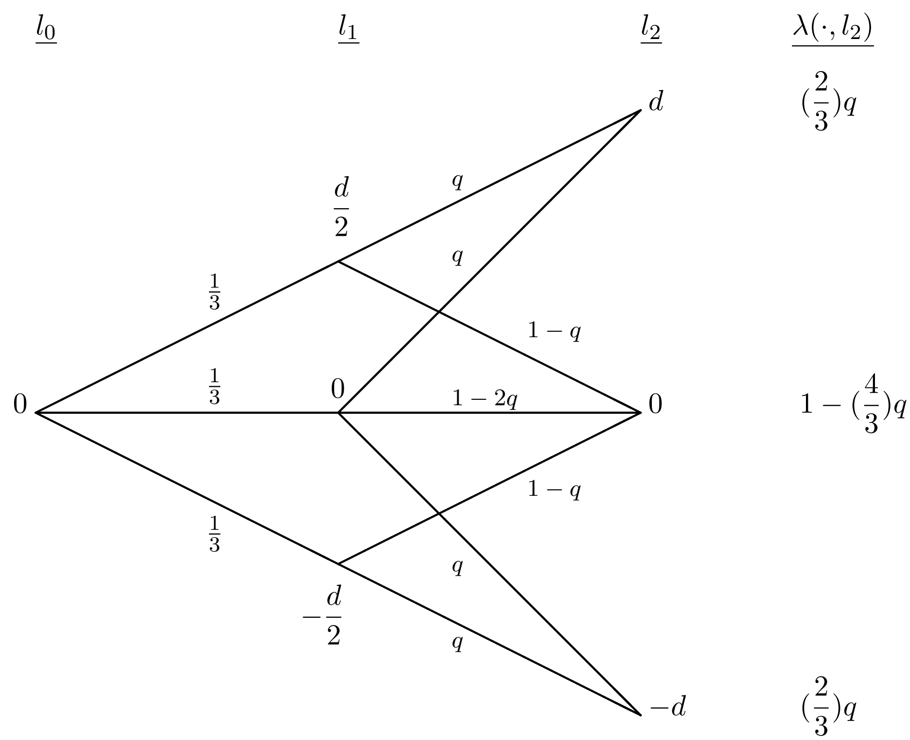
Naturally, it follows that expected aggregate daily defaults are larger under a deferred settlement system as well; that is, $\bar{\delta}(F_1) \geq \bar{\delta}(F_2)$.

4 An Example

We now develop a simple example of the banking sector defined above. One reason for this exercise is to make the results introduced in Section 3 more intuitive and concrete to the reader. Moreover, the example we develop illustrates the type of analysis that the model allows for, and provides a context for the consideration of several important issues often debated in the literature.

Suppose that at $t = 1$, net liabilities increase by $\frac{d}{2}$, stay the same, or decrease by $\frac{d}{2}$ with equal probability $\frac{1}{3}$. If $l_1 = \frac{d}{2}$, then at $t = 2$ net liabilities return to zero with probability $1 - q$ and increase to d with probability q. Similarly, if $l_1 = -\frac{d}{2}$, then at $t = 2$ net liabilities return to zero with probability $1 - q$ and decrease to $-d$ with probability q. Finally, if $l_1 = 0$, then at $t = 2$ net liabilities remain at zero with probability $1 - 2q$, and increase or decrease to d with probability q. Figure 1 below illustrates the distribution of net liabilities.

FIGURE 1: DISTRIBUTION OF l



Suppose further that $\psi : [0, d] \to \mathbb{R}_+$ is of the form $\psi(b) = \eta b^2$ for some $\eta > 0$. This function is clearly consistent with the conditions specified in Assumption 1. We restrict our analysis to parameter values that satisfy the condition $\frac{2}{3} < \frac{r\rho}{dq\eta}$, which guarantees that the optimal reserve ratio under both F_1 and F_2 is strictly less than $\frac{1}{2}$.[16] However, we *do* consider parameter values for which bank's choose the minimum reserve ratio, which we set to zero for simplicity.

4.1 Deferred Settlement

Consider first a deferred settlement system (F_1) in which settlement occurs only once at $t = 2$. Banks face ex-ante expected borrowing costs

$$\Psi(\alpha, F_1) = (\frac{2}{3})q(d - \alpha d)^2\eta. \tag{7}$$

[16]If $\alpha(F_1) \in [\frac{1}{2}, 1]$, then $\alpha(F_1) = \alpha(F_2)$, and we lose two margins of interest: the banking fees are identical under F_1 and F_2 and the reserve effect vanishes.

Therefore, the first order condition for the optimal reserve ratio under settlement system F_1 implies

$$\alpha(F_1) = \begin{cases} 0, & \text{for } \frac{4}{3} \leq \frac{r\rho}{dq\eta} \\ 1 - \frac{3r\rho}{4\eta dq}, & \text{for } \frac{4}{3} > \frac{r\rho}{dq\eta} > \frac{2}{3}. \end{cases} \tag{8}$$

As a result, depositors face a banking fee

$$\phi(F_1) = \begin{cases} a - r\rho + \frac{2\eta dq}{3}, & \text{for } \frac{4}{3} \leq \frac{r\rho}{dq\eta} \\ a - \frac{3}{8}(\frac{(r\rho)^2}{\eta dq}), & \text{for } \frac{4}{3} > \frac{r\rho}{dq\eta} > \frac{2}{3} \end{cases} \tag{9}$$

under a deferred settlement system.

Now consider expected aggregate daily defaults under a deferred settlement system. By construction, there are only two realizations, $l = (\frac{d}{2}, d)$ and $l = (0, d)$, for which a bank has positive net liabilities at $t = 2$, so we can restrict our attention to these outcomes. For either realization, outstanding liabilities in excess of reserves are the same:

$$\hat{l}_2\Big((\tfrac{d}{2}, d), \alpha(F_1); F_1\Big) = \hat{l}_2\Big((0, d), \alpha(F_1); F_1\Big) = \begin{cases} d, & \text{for } \frac{4}{3} \leq \frac{r\rho}{dq\eta} \\ \frac{3r\rho}{4\eta q}, & \text{for } \frac{4}{3} > \frac{r\rho}{dq\eta} > \frac{2}{3}. \end{cases}$$

Therefore, we employ equation (5) to conclude that the expected aggregate daily defaults are

$$\bar{\delta}(F_1) = \begin{cases} \frac{2q\bar{\Delta}d}{3}, & \text{for } \frac{4}{3} \leq \frac{r\rho}{dq\eta} \\ \frac{\bar{\Delta}r\rho}{2\eta}, & \text{for } \frac{4}{3} > \frac{r\rho}{dq\eta} > \frac{2}{3}. \end{cases} \tag{10}$$

4.2 Real-Time Settlement

Now consider a real-time settlement system with frequency $F_2 = 2$, so that settlement occurs at both $t = 1$ and $t = 2$. We first observe that the expected borrowing fees are increased under a real-time system, because banks now must make payments for the realizations $l = (\frac{d}{2}, d)$, $l = (0, d)$, ***and*** $l = (\frac{d}{2}, 0)$. Therefore,

$$\Psi(\alpha, F_2) = (\frac{2}{3})q(d - \alpha d)^2\eta + (\frac{1}{3})(1-q)(\frac{d}{2} - \alpha d)^2\eta, \tag{11}$$

or, equivalently,

$$\Psi(\alpha, F_2) = \Psi(\alpha, F_1) + (\frac{1}{3})(1-q)(\frac{d}{2} - \alpha d)^2\eta.$$

A profit-maximizing bank will find the optimal reserve ratio under F_2 to be

$$\alpha(F_2) = \begin{cases} 0, & \text{for } 1 + \frac{1}{3q} \leq \frac{r\rho}{dq\eta} \\ \frac{3}{2(1+q)}[\frac{1}{3} + q - \frac{r\rho}{\eta d}], & \text{for } 1 + \frac{1}{3q} > \frac{r\rho}{dq\eta} > \frac{2}{3}. \end{cases} \tag{12}$$

Simple algebra confirms that $\alpha(F_2) \geq \alpha(F_1)$. Given the optimal reserve ratio $\alpha(F_2)$, the banking fees imposed on depositors are uniquely determined by (6):

$$\phi(F_2) = \begin{cases} a - r\rho + \frac{3\eta dq}{4}, & \text{for } 1 + \frac{1}{3q} \leq \frac{r\rho}{dq\eta} \\ a - \frac{1}{2(1+q)}[(1-q)(r\rho - \frac{\eta dq}{3}) + \frac{3(r\rho)^2}{2\eta d}], & \text{for } 1 + \frac{1}{3q} > \frac{r\rho}{dq\eta} > \frac{2}{3}. \end{cases} \tag{13}$$

Though it is not immediately obvious from the equation above, again it can be easily shown that $\phi(F_2) \geq \phi(F_1)$.

Finally, we consider expected aggregate daily defaults under a real-time system. Again, we restrict our attention to $l = (\frac{d}{2}, d)$ and $l = (0, d)$. However, under frequency F_2, outstanding liabilities in excess of reserves will not be the same for these two realizations. For $l = (\frac{d}{2}, d)$, the bank is forced to settle at $t = 1$ so that, for $\frac{r\rho}{dq\eta} > \frac{2}{3}$,

$$\hat{l}_2\Big((\frac{d}{2}, d), \alpha(F_1); F_1\Big) = \frac{d}{2}.$$

That a higher frequency of settlement limits the potential default for $l = (\frac{d}{2}, d)$ to $\frac{d}{2}$, as opposed to d under deferred settlement, is precisely the *accumulation* effect presented in result 3. For $l = (0, d)$, on the other hand,

$$\hat{l}_2\Big((0,d),\alpha(F_1);F_1\Big) = \begin{cases} d, & \text{for } 1+\frac{1}{3q} \leq \frac{r\rho}{dq\eta} \\ \frac{d}{2(1+q)}[1-q+\frac{3r\rho}{\eta d}], & \text{for } 1+\frac{1}{3q} > \frac{r\rho}{dq\eta} > \frac{2}{3}. \end{cases}$$

Here, the size of a potential default has been reduced via the *reserve effect*; a real-time system has encouraged more conservative reserve management, and therefore insolvent banks are better able to meet outstanding liabilities. We calculate expected aggregate daily defaults under frequency F_2,

$$\bar{\delta}(F_2) = \begin{cases} \frac{q\bar{\Delta}d}{2}, & \text{for } 1+\frac{1}{3q} \leq \frac{r\rho}{dq\eta} \\ \frac{q\bar{\Delta}d}{6} + \frac{q\bar{\Delta}}{3}[\frac{d}{2(1+q)}[1-q+\frac{3r\rho}{\eta d}]], & \text{for } 1+\frac{1}{3q} > \frac{r\rho}{dq\eta} > \frac{2}{3}, \end{cases} \tag{14}$$

and conclude that, due to *both* the accumulation and the reserve effect, $\bar{\delta}(F_1) \geq \bar{\delta}(F_2)$.

4.3 The Magnitude of the Risk-Cost Tradeoff

We've established above the existence of a tradeoff between risk and cost that arises within the analysis of settlement frequency. However, the magnitude of this tradeoff depends heavily on the parameters of the economy. In an effort to better understand the key forces at play, we now parameterize the economy, and consider the sensitivity of $\alpha(F_i)$, $\phi(F_i)$, and $\bar{\delta}(F_i)$ to two key parameters: η and q.[17]

Let $a = 0$, $d = .15$, $r = .02$, and $\rho = .1$, and the distribution of shocks, $g(\cdot)$, be such that

$$\Delta = \begin{cases} 0, & \text{with probability } 1-p \\ .1, & \text{with probability } p \end{cases} \tag{15}$$

This distribution is meant to reflect that there is a high probability of zero solvency shocks, and a small probability of a solvency shock affecting a fraction of banks in the settlement system. We set $p = .05$, so that $\bar{\Delta} = .005$.

First, we will fix $q = .1$ and focus on how the policy parameter η, representing the cost of borrowing from the central bank, affects a bank's choice of reserves,

[17] Note that this is not a calibration, but rather a purely illustrative exercise.

along with the resulting fees and defaults. Figure 2A illustrates that, for small η, banks choose to retain the minimum level of reserves (zero) under either settlement system. However, notice in Figure 2B that even when banks hold zero reserves, the banking fees are larger under F_2 than they are under F_1, as the increased intra-day borrowing costs are passed along to the depositors. Moreover, due to the increased liquidity requirements under a real-time system, banks begin to hold strictly positive reserves for smaller values of η under F_2 than they do under F_1; at these values of η, the difference between banking fees under the two systems is most pronounced. As η increases towards its upper bound, the difference in banking fees diminishes. Figure 2C illustrates the two key effects of settlement frequency on total defaults: the accumulation effect and the reserve effect. As in the previous graphs, the solid line and the dotted line represent the equilibrium expected aggregate daily defaults under frequencies F_1 and F_2, respectively. However, here we've added a third component: the intermediate dashed line represents the expected aggregate daily default under frequency F_1 for a bank that has chosen the optimal reserve ratio for frequency F_2 (i.e. $\alpha(F_2)$). Therefore, the distance between the dashed and dotted lines represents the accumulation effect - the decreased size of expected defaults resulting from a shorter interval between settlements. This effect is most pronounced for small values of η, when there is no reserve effect, and vanishes as $\alpha(F_2)$ approaches $\frac{1}{2}$. The distance between the solid and dashed lines represents the reserve effect, which accounts for the downward slope of the curves corresponding to regions where they retain strictly positive reserves. Naturally, the size of this effect corresponds directly to the difference between $\alpha(F_1)$ and $\alpha(F_2)$.

[INSERT FIGURE 2A - 2C HERE]

We now fix $\eta = .1$, and concentrate on the magnitude of the risk-cost tradeoff for different values of q. Recall that q represents the conditional probability that a bank's intra-day liabilities will increase to d , leaving a bank with positive outstanding net liabilities at $t = 2$. In other words, q is the probability of a bank receiving a liquidity shock. For small values of q, banks in a deferred settlement system will not retain any reserves, since the probability of needing them for settlement at $t = 2$ is small. However, as is evident in Figure 3A, banks in a real-time settlement system with $\eta = .1$ will always retain positive reserves to meet liabilities accrued at $t = 1$. Furthermore, notice that for small q, $\phi(F_2)$ is noticeably larger than $\phi(F_1)$, while $\bar{\delta}(F_2)$ and $\bar{\delta}(F_1)$ appear similar. One might conclude that for small q, a deferred settlement system would be preferred. As q increases, the tradeoff seems to skew toward a real-time settlement system, as the fees become similar while the difference in default sizes becomes more pronounced.

[INSERT FIGURE 3A - 3C HERE]

Overall, the partial equilibrium example above is a useful tool for illustrating the primary forces at work within the banking sector. However, since it takes the underlying transactions as exogenous, it ignores the behavior of buyers and sellers; and therefore does not consider how their behavior might respond to changes in the banking fees or default risks. Moreover, we have thus far ignored any profits or losses assumed by the central bank. As such, any definition of welfare (and thus optimality) would require arbitrary weights on the risks and costs of settlement. We now introduce a trading sector to complete the economy. In this general equilibrium model, we have a more clear definition of welfare and a framework to discuss optimal settlement policies.

5 The Trading Sector

5.1 The Basic Model

The trading sector is modeled in the spirit of He, Huang, and Wright (2005), which we find appropriate for several reasons. First, the use of credit as a medium of exchange is micro-founded, in the sense that agents *choose* whether or not to use the banking system given the costs and risks. This feature provides some discipline as we consider various policy choices, as all equilibria must satisfy the incentive constraints of buyers and sellers. Moreover, this framework generates a simple demand for money that responds to changes in interest rates, loan prices, and the risks of failed settlement. Finally, the model assumes that bilateral transactions occur randomly across time, which seems consistent with the stochastic evolution of intra-day net liabilities that we introduced in section 3.[18]

There is a $[0, 1]$ continuum of agents, and a fraction M are initially endowed with one unit of fiat money. Money is indivisible and agents hold either zero or one unit. Before entering the market, an agent with money has the option to deposit his money at the bank in exchange for a "check." Again, this service has a cost ϕ, where $\phi > 0$ can be interpreted as a banking fee, and $\phi < 0$ as interest earned on deposits. An agent with money will trivially choose to use banking if $\phi \leq 0$ and not if $\phi > 0$. Since this paper is concerned with payment systems, we will restrict our attention to equilibria in which $\phi \leq 0$ and sellers always accept checks as payment. An agent without money can take out a loan for an up-front fee ρ.[19] Therefore, when all agents choose to deposit their money at the bank, a fraction $M_1 = \frac{M}{\alpha}$ will enter the market with checks (these are *buyers*), while the remaining $1 - M_1$ enter the market without money (these are *sellers*). Naturally, M_1 is the total money supply. To ensure that $M_1 \leq 1$, we will impose a required reserve ratio $\underline{\alpha} \geq M$.

Agents trade in a random, bilateral matching (or *decentralized*) market in which meetings occur at T discrete points in time throughout the day. When a consumer meets a producer, there is probability x that he likes the producer's specialized good. In this case, the seller produces the good at cost c in exchange

[18] The key distinction here is that we do not introduce theft to motivate banking, as in He, Huang, and Wright (2005). Instead, we will concentrate on the case of interest-bearing bank accounts. Moreover, we do not require a Lagos-Wright (2005) framework, in which agents alternate between a centralized and decentralized market, because we limit our analysis to the case of $\{0, 1\}$ money holdings.

[19] As He, Huang, and Wright (2005) point out, this fee can be interpreted as a coupon payment of $r\rho$ in each period.

for the buyer's check, and the buyer immediately consumes the good and receives utility u. The probability of a double-coincidence of wants is zero. Goods are indivisible, non-storable, and they are produced, traded, and consumed immediately. After an agent sells her good, she leaves the market and deposits the check at her bank. When the bank receives payment finalization, her account is credited one unit and she becomes a buyer the following day.

5.2 The Costs of Default

As we have seen in Section 3, each day a fraction Δ of banks are rendered insolvent at $t = T$. We assume that a deposit at an insolvent bank will be honored at the solvent bank that is created in its place the following morning. Therefore, buyers who have not spent their check are fully insured. However, a seller who has accepted a check does not necessarily enjoy such insurance. On any given day, a fraction $\delta(\Delta, F)$ of agents will fail to receive payment finalization for goods they've produced and sold. Recovering these payments is a costly endeavor; in general, the process proves much more costly than simply the size of the default itself. For one thing, a default by one bank can trigger defaults by other banks within the settlement system. Moreover, defaults erode confidence in the banking sector, increasing the perceived risk of deposits and potentially sparking bank runs. Finally, the default process is generally costly in terms of both time and legal fees. Though we don't model any of these effects directly, we would like to capture the notion that recovering from a default has potentially large social costs. To this end, we assume that there is a cost $\Omega(\delta)$ of recovering from a default of size δ.

Definition 7 *Let the* ***expected costs of default under settlement system*** *F be denoted $\omega(F)$ and be defined*

$$\omega(F) = \int \Omega\Big(\delta(\Delta, F)\Big)\ g(\Delta)\ d\Delta. \tag{16}$$

Depending on the specifications of the settlement system, these costs may be born (ultimately) by either the buyers or the sellers. One common practice, known as *survivors pay*, requires that the remaining solvent banks pay the recovery costs. Under the assumption of perfect competition in the banking sector, this implies that the buyers will absorb the burden of default through increased banking fees. On the other hand, if the system does not require the participants of the settlement system to pay for recovery, then the costs may be born by the sellers. In such a case, we assume the sellers will pay an (actuarially fair) insurance premium to protect themselves from payment default. This practice, the so-called *third party pays*, is also fairly common. Let us assume that a fraction of the recovery costs, $\xi \in [0, 1]$, are paid by the buyers and the remaining $1 - \xi$ by the sellers. Such generality nests the two most common practices discussed above as special cases ($\xi = 1$ and $\xi = 0$, respectively).

5.3 The Central Bank and Transfers

A general equilibrium model also requires that we account for the revenue and operating expenses of the central bank. In particular, any profits or losses should be incorporated into a welfare comparison of settlement frequency. To this end, suppose that the central bank issues a lump-sum transfer to the consumers equal to the revenue generated from providing intra-day liquidity less operating expenses. Since we have assumed for simplicity that the central bank can perform the necessary tasks costlessly, the expected lump-sum transfer for a particular policy (ψ, F) is given by

$$\tau(\psi, F) = (1 - \bar{\Delta})\Psi(\alpha(F), F). \tag{17}$$

Notice that $\tau(\psi, F)$ is simply the borrowing fees collected from solvent banks at the end of the day.

5.4 Banking Equilibria

We denote the value function of an agent with zero money holdings (prior to paying the costs of default) V_0, and the value function of an agent with one unit of money (prior to paying banking fees or the costs of default) V_1. With this notation, we derive the flow Bellman equations:

$$
\begin{aligned}
rV_0 &= M_1x(V_1 - V_0 - c) - (1+r)[(1-\xi)\omega] + \tau && (18)\\
rV_1 &= (1-M_1)x(u + V_0 - V_1) - (1+r)[\phi + \xi\omega] + \tau && (19)
\end{aligned}
$$

where we've simplified notation by suppressing the arguments of functions such as ϕ, ω, and τ.

An equilibrium in which agents use banks must satisfy the following conditions:

1. *Individual Rationality*: neither producers nor consumers would rather drop out of the economy and live in autarchy.

$$V_0 \geq 0 \tag{20}$$

$$V_1 \geq 0 \tag{21}$$

2. *Incentive Compatibility*: both producers and consumers will trade in a bilateral match.

$$V_1 - c \geq V_0 \tag{22}$$

$$u + V_0 \geq V_1 \tag{23}$$

3. *Banking*: agents with money choose to use banks.

$$\phi \leq 0 \tag{24}$$

These simplify to the following conditions, expressed as constraints on the banking fee ϕ:[20]

$$
\begin{aligned}
\phi &\leq 0 && (25)\\
\phi &\leq \frac{(1-M_1)x(u-c) - rc}{1+r} + \omega\Big[1 - 2\xi - \frac{(r+x)(1-\xi)}{M_1x}\Big] + \frac{(r+x)\tau}{(1+r)M_1x} && (26)\\
\phi &\geq \frac{M_1xc + (1+r)(1-2\xi)\omega - (r+M_1x)u}{1+r} && (27)
\end{aligned}
$$

[20] It turns out that $V_0 \geq 0 \Rightarrow V_1 - V_0 - c \geq 0$, and that this further implies that $V_1 \geq 0$.

The first and second conditions, respectively, ensure that banks pay interest on deposits, and that this interest is sufficiently large that a seller is willing to incur the costs of production in order to become a buyer. The third condition ensures that the interest earned on deposits is not so high that buyers are unwilling to withdraw their deposits from the bank.

In equilibrium, the loan price (or interest rate) that clears the market must satisfy $\rho = V_1 - V_0$. From (18) and (19), it follows that

$$\rho = \frac{(1 - M_1)xu - M_1xc - (1 + r)\phi + (1 + r)(1 - 2\xi)\omega}{r + x}. \tag{28}$$

Definition 8 *Consider an economy* $\{a, d, r, u, c, x, \xi, T, M, \tau, \lambda(\cdot), g(\cdot), \Omega(\cdot)\}$*. Given a settlement system* $\{F, \psi\}$*, a* ***banking equilibrium*** *is a fixed point* $\{\alpha^*, \rho^*\}$ *such that*

- *Given* ρ^**,* α^* *satisfies the bank's optimality condition in equation (3). Moreover,* α^* *uniquely determines banking fees* ϕ^* *according to (6), the expected costs of default* ω^* *according to (16), and the total money supply* $M_1^* = \frac{M}{\alpha}$*.*
- *Given* $\{\phi^*, \omega^*, M_1^*\}$*, the constraints (25) - (27) are satisfied and the price of loans* ρ^* *is determined as in equation (28).*
- *Deposits equal the supply of fiat money, so that* $d = M$*.*
- *Transfers* τ *are determined as in (17).*

5.5 Welfare

Assuming equal weight on all agents, we can define equilibrium welfare

$$W^* = (1 - M_1^*)V_0^* + M_1^*V_1^*. \tag{29}$$

Substituting (18) and (19) into (29) implies that

$$W^* \propto M_1^*\Big[(1-M_1^*)x(u-c)-(1+r)\phi^*\Big]-(1+r)\Big[M_1^*(2\xi-1)+1-\xi\Big]\omega^*+\tau. \tag{30}$$

Armed with a clear definition of welfare, we are now able to identify the *optimal* settlement system as the set (ψ, F) that maximizes W^*. In the next section, we extend the partial equilibrium example from Section 4 in order to consider how settlement frequency affects equilibrium prices, behavior, and welfare; and pinpoint which features of an economy might make one settlement system more appropriate than another.

6 Extending the Example

Let $\lambda(\cdot)$ remain as depicted in Figure 1, and $g(\cdot)$ as described in (15). Let the cost of default be specified as $\Omega(\delta) = \kappa\delta^2$. We begin with benchmark parameter values $p = .05$, $q = .1$, and $\kappa = 150$, though we will later analyze how the equilibria change as these values vary. Finally, we set $\xi = .5$, $u = .6$, $c = .05$, $x = .4$, $\underline{\alpha} = M$, and maintain the remaining parameter values from section 4.

A *demand curve* for each F_i can be derived by substituting the appropriate equations for $\phi(F_i)$ (equations (9) and (13), respectively) and $\delta(F_i)$ (equations (10) and (14), respectively) into equation (28). This demand curve relates the price of money, ρ, to the supply of money $\frac{M}{\alpha}$, and is downward sloping. The *supply curve*, given by the banks optimality condition (equations (8) and (12), respectively), implies an upward-sloping relationship between the price of loans and a bank's willingness to supply them. We compute the intersection to find $\{\alpha^*, \rho^*\}$, which can be used to pin down the banking fees ϕ^*, the costs of default ω^*, and equilibrium welfare W^*. Though closed form solutions are not practical, we continue our analysis graphically.[21]

We present equilibria under two different assumptions on the initial supply of money, M. First, we assume that M is fixed. Under this assumption, the banks' choice of reserves has real effects on the allocation of goods, as $\alpha(F_i)$ determines the ratio of buyers to sellers $\frac{M_1}{1-M_1}$. Our analysis under a fixed money supply suggests that general equilibrium effects are important, in the sense that some

[21] All parameter values considered below support equilibria that satisfy the constraints (25) - (27), as well as the example-specific constraint $\frac{2}{3} < \frac{r\rho}{\eta dq}$.

previous conclusions from the partial equilibrium analysis are not necessarily sustained.

Unfortunately, the effect of F_i on the ratio of buyers to sellers when M is fixed is the dominating factor in the analysis of welfare (particularly for corner solutions); and while it is important to be aware of this effect, there are several reasons - both practical and model-specific - to believe that it may be more marginal than the results suggest. For one, the choice of a settlement system is a relatively long-run decision, and affords the central bank ample time to adjust the money supply to any large changes in the reserve ratio. Therefore, the money supply *is* somewhat flexible over this sort of time range. Moreover, that this effect is the dominating factor in welfare analysis is more an artifact of our modeling strategy than a true concern in the choice of settlement systems. In particular, the assumption that agents are *either* buyers *or* sellers implies that M_1 completely determines the probability of a trade. Therefore, as M_1 approaches either 0 or 1, trade comes to a halt and welfare is drastically reduced, regardless of the banking fees or default costs. In response to these concerns, we also consider equilibria under the assumption that M is flexible, and can be adjusted by the central bank to achieve the optimal M_1.

6.1 Equilibrium analysis with a fixed money supply

Figure 4A illustrates the optimal reserve ratio under F_1 and F_2 as a function of η, the cost of intra-day borrowing. As expected, we find that a bank will choose to retain more reserves under a real-time system than they will under a deferred system for all values of η in the relevant parameter space. Figure 4C also illustrates an expected result: the costs of default are less under F_2 than F_1.

[INSERT FIGURE 4A - 4E HERE]

On the contrary, the results presented in Figure 4B are not as expected; after all, in section 3 we prove that for any ρ, it must be that $\phi(F_1) \leq \phi(F_2)$. However, in the general equilibrium framework, ρ is an endogenous variable and will depend on the settlement frequency. At low levels of η, banks choose to retain few reserves and flood the market with loans, putting downward pressure on the price ρ. Since banks are more inclined to hold reserves under a real-time system, we see that the price of loans is greater under F_2 than it is under F_1. The effect of increased revenue from larger values of ρ is sufficiently strong that it dominates the effect of decreased revenue from larger intra-day liquidity requirements, and the result is in fact *decreased* banking fees (or *increased* interest payments) under F_2. Therefore, in equilibrium, the costs of increased settlement frequency are passed to agents in the form of higher *loan* fees, and not banking fees.

Turning to welfare analysis, Figure 4E illustrates the concerns expressed above. As η approaches zero, $\alpha(F_i) \rightarrow \underline{\alpha}$ and $M_1 \rightarrow 1$, thereby drastically reducing welfare. Similarly, as η tends toward its upper bound, $\alpha(F_i) \rightarrow 1$ and $M_1 \rightarrow M$, also reducing welfare substantially. To shut down this effect and concentrate on the costs (ϕ) and risks (ω) of settlement frequency, we now move to the case of a flexible money supply.

6.2 Equilibrium analysis with a flexible money supply

Suppose now that M can be adjusted by the central bank so that $M_1 = .5$ for any frequency F. Under this assumption, the equilibrium values of α, ϕ, and ω depicted in Figures 5A - 5C are similar to those presented in section 4. Moreover, since ρ is independent of M_1, it is primarily determined by ϕ. Therefore, since deposits earn higher interest under a deferred system, loans are more attractive and therefore priced higher under F_1 than they are under F_2. As η increases, interest rates decline and the price of loans follows. Finally, we consider Figure 5E, which illustrates welfare under the two settlement frequencies for all possible intra-day borrowing policies. For the benchmark values, the optimal settlement system is a real-time system with free intra-day liquidity. The risk posed by default is substantial enough that the accumulation effect is a factor, which explains the gap between welfare under F_2 and F_1 for $\eta = 0$. However, as η increases the reduction in risk from the reserve effect is small relative to the increase in costs, so that welfare decreases as liquidity becomes more costly.

[INSERT FIGURE 5A - 5E HERE]

Of course, the optimal settlement system (F, ψ) will depend heavily on the values of several key parameters. Perhaps the most obvious such parameter is p, the probability of a solvency shock. In Figure 6A, we consider equilibrium welfare across borrowing policies for the case of no solvency shocks (i.e. $p = 0$), holding all other parameters at their benchmark values. As expected, in an economy without risk, a deferred settlement system dominates a real-time system. Moreover, there is no reason for costly intra-day liquidity, as it serves only a negative, distortionary role in this economy. On the opposite extreme, Figure 6B plots welfare when solvency shocks occur at a high rate of 10 %. Here, the risks associated with default are sufficiently high that a real-time system is warranted. However, again the reserve effect is not sufficiently strong that costly liquidity would improve welfare.

[INSERT FIGURE 6A - 6B HERE]

Another variable of interest is q, the probability of a bank receiving a liquidity shock and accruing large intra-day net liabilities. Figures 7A - 7B illustrate welfare when q assumes two extreme values: .05 and .5, respectively. When there is a small probability of a liquidity shock, again the risks of default are small and a deferred system with free liquidity is preferred. However, when liquidity shocks are more likely, a real-time system with costly intra-day borrowing becomes optimal. As is evident in Figure 7B, both the accumulation effect and the reserve effect are pertinent: the accumulation effect is evident at $\eta = 0$, and the reserve effect is illustrated by the upward slope of the welfare graph at values of η such that banks retain reserves strictly above the required level.

[INSERT FIGURE 7A - 7B HERE]

Finally, we consider the optimal settlement system as the cost of recovering a default, κ, varies. Consistent with the results above, when κ assumes a small value a deferred settlement system with free liquidity is optimal. However, when the costs of recovery become sufficiently large, the reduction in risk from high frequency settlement outweighs the additional costs, and a real-time system is optimal.

[INSERT FIGURE 8A - 8B HERE]

6.3 Discussion

There are two primary topics of debate in the literature on settlement systems. The first is whether the system should be a real-time, a deferred, or (more recently) a hybrid settlement system; and the second is whether the central bank should provide intra-day liquidity at a cost or for free. However, much of the literature has addressed these questions *separately.* That is, they have assumed a particular settlement frequency and concentrated on the optimal borrowing policy, or vice-versa.[22] The current analysis suggests that these questions need to be addressed *jointly.* As Figure 7B illustrates, the optimal borrowing policy varies across settlement frequency; for this case, the best policy under a real-time system is to allow banks to borrow free of charge, while a deferred system would benefit from the costly provision of intra-day liquidity. Therefore, we conclude that any discussion of borrowing policies must be made within the context of a specific settlement frequency.

A second insight that arises from the analysis above is that the optimal settlement system will depend very heavily on the economy in which it operates. An economy in which solvency shocks, liquidity shocks, or recovery costs are small will likely find a deferred settlement system with free liquidity to be optimal. On the other hand, an economy with volatile intra-day net liabilities, or one that is

[22] A noteable exception is Kahn and Roberds (1998), who consider the optimal use of borrowing caps in the context of both real-time and deferred settlement systems.

vulnerable to solvency shocks or contagion can benefit from the risk reduction of a real-time settlement system. From the analysis above, it seems clear that any blanket statement claiming the superiority of one policy over another, irrespective of the particulars of the economy of interest, would be erroneous.

7 Conclusion

This paper first develops a partial equilibrium model of reserve management in the presence of stochastic realizations of intra-day liabilities and costly borrowing from the central bank. Within this framework, increased settlement frequency implies that banks face greater financial costs. In a perfectly competitive banking sector, these costs are transferred to the customers via higher banking fees. However, in the presence of exogenous solvency shocks, increased settlement frequency also mitigates the risk to the economy by decreasing the size of potential defaults. In the context of a simple example, we illustrate this tradeoff between cost and risk, and examine its sensitivity to key model parameters.

We then complete the model by introducing a trading sector in which agents endogenously choose to use credit to conduct bilateral transactions. In this general equilibrium framework, the effects of banking fees and default risk on the behavior of both buyers and sellers are taken into account, and a natural definition of welfare is established. Two key insights emerge from our analysis. First, we find that determination of the optimal settlement system requires *joint* consideration of both settlement frequency and intra-day borrowing practices. Secondly, we conclude that the optimal use of these two policy levers will depend heavily on the underlying parameters of the economy. In particular, the factors governing the probability of solvency and liquidity shocks, as well as the costs of recovering from a default, will be the primary determinants of the optimal settlement system. Economies that face larger costs of recovery from a default, or are more prone to solvency or liquidity shocks, prefer a real-time settlement system. Conversely, when these risks are less substantial, a less expensive, deferred settlement system may be welfare-improving.

The primary contribution of this paper is the development and synthesis of a partial equilibrium model of a banking sector with an existing model of trade. The result is a general equilibrium model that captures key elements of *both* bank's and consumer's behavior. The policy experiments and sensitivity analysis we perform represent only a small fraction of the potential topics that this model could address. More careful consideration of the informational issues that accompany solvency shocks may provide new insights into settlement systems and the effects of central bank bail-out policies and bankruptcy procedures. Alternatively, heterogeneity with respect to the costs that banks (the variable a in the current model) face may provide an interesting mechanism to explore *why* certain banks choose to join a settlement system and others do not. Finally, by considering the case of divisible money within the trading sector, we might find that equilibrium prices differ for agents that pay with credit as opposed to cash (when both are circulating). These are areas for future work.

References

Phillipe Aghion, Patrick Bolton, and Steven Fries (1999), "Optimal Design of Bank Bailouts: The Case of Transition Economies." *Journal of Institutional and Theoretical Economics* 155, 51 - 70.

Allen Berger, Diana Hancock, and Jeffrey Marquardt (1996), "A Framework for Analyzing Efficiency, Risks, Costs, and Innovations in the Payments Systems." *Journal of Money, Credit and Banking* 28, 696 - 732.

Ben S. Bernanke (1990), "Clearing and Settlement during the Crash." *The Review of Financial Studies* 3 (1), 133 - 151.

Committee on Payment and Settlement Systems (2003), "The Role of Central Bank Money in Payment Systems." Bank for International Settlements.

Committee on Payment and Settlement Systems (2005), "Statistics on Payment and Settlement Systems in Selected Countries - Figures for 2003." Bank for International Settlements.

Michael J. Fleming and Kenneth D. Garbade (2002), "When the Back Office Moved to the Front Burner: Settlement Fails in the Treasury Market After 9/11." Federal Reserve Bank of New York Economic Policy Review, November Issue, 35-57.

Scott Freeman (1996), "The Payment System, Liquidity, and Rediscounting." *American Economic Review* 86, 1126 - 1138.

Scott Freeman (1999), "Rediscounting Under Aggregate Risk." *Journal of Monetary Economics* 43, 197 - 216.

Xavier Freixas, Bruno Parigi, and Jean-Charles Rochet (2004), "The Lender of Last Resort: A 21st Century Approach." *Journal of the European Economic Association* 2 (6), 1085 - 1115.

Edward Green (1999). "Money and Debt in the Structure of Payments." *Quarterly Review*, Federal Reserve Bank of Minneapolis.

Ping He, Lixin Huang, and Randall Wright (2005), "Money and Banking in Search Equilibrium." *International Economic Review* 46 (2), 637 - 670.

Charles Kahn and William Roberds (1998), "Payments System Settlements and Bank Incentives." *Review of Financial Studies* 11, 845 - 870.

Charles Kahn and William Roberds (2001), "Real-time Gross Settlement and the Costs of Immediacy." *Journal of Monetary Economics* 47, 299 - 319.

Charles Kahn, James McAndrews, and William Roberds (1999), "Settlement Risk Under Gross and Net Settlement." *Journal of Money, Credit, and Banking* 35 (4), 591 - 608.

Thorsten Koeppl, Cyril Monnet, and Ted Temzelides (2005), "Mechanism Design and Payments. " *mimeo.*

Ricardo Lagos and Randall Wright (2005), "A Unified Framework for Monetary Theory and Policy Analysis." *Journal of Political Economy* 113 (3), 463 - 484.

Benjamin Lester, Stephen Millard, and Matthew Willison (2005), "Settlement Systems in Search Equilibria with Bank Default." *mimeo*, Bank of England.

Antoine Martin (2004), "Optimal Pricing of Intraday Liquidity." *Journal of Monetary Economics* 51 (2), 401 - 424.

David C. Mills, Jr. (2005), "Alternative Central Bank Credit Policies for Liquidity Provision in a Model of Payments." *Journal of Monetary Economics*, forthcoming.

Jean-Charles Rochet and Jean Tirole (1996), "Controlling Risk in Payment Systems." *Journal of Money, Credit and Banking* 28 (4), 832-862.

Ruilin Zhou (2000) "Understanding Intraday Credit in Large-Value Payment Systems." *Economic Perspectives*, Federal Reserve Bank of Chicago, third quarter, 29-44.

A Results and proofs for an arbitrary, finite T.

In Section 3, we presented several claims and results for the special case of $T = 2$. We now present and prove the analogous claims and results for an arbitrary, finite T. Of course, the proofs below encapsulate the special case of $T = 2$ with $F_1 = 1$ and $F_2 = 2$.

We require a slightly narrow definition of *increased settlement frequency*. In words, for any settlement system which specifies a subset of intervals at which settlement occurs, we say that increasing the settlement frequency implies adding at least one additional interval to the existing subset. We formalize this now.

Definition 9 *Consider an arbitrary, finite T, and settlement frequency F' with an associated $\mathcal{S}' \subset \mathcal{T}$ such that $\mathcal{S}' \equiv \{t'_1, t'_2, ..., t'_F\}$ denotes the subset of intervals in which settlement occurs. We say that F'', with associated set $\mathcal{S}'' \equiv \{t''_1, t''_2, ..., t''_F\}$, has a **higher frequency of settlement** than F' if and only if $\mathcal{S}'$ is a strict subset of $\mathcal{S}''$. That is, $\forall\ t'_i \in \mathcal{S}'$, $t'_i \in \mathcal{S}''$ and $\exists\ t''_i \in \mathcal{S}''$ such that $t''_i \notin \mathcal{S}'$.*

The following two results will be helpful in the ensuing proofs.

Result 6 *For any l, α, $t \in \mathcal{T}$, and F with associated $\mathcal{S} = \{t_1, \ldots, t_F\}$,*

$$\hat{l}_t(l, \alpha; F) = \max\Big\{0, \min\{l_t - l_{s_t}, l_t - l_{s_t - 1}, \ldots, l_t - l_{t_1}, l_t - \alpha d\}\Big\}, \tag{31}$$

where s_t corresponds to the period in which the most recent settlement occurred prior to t, s_{t-1} the settlement prior to s_t, and so on. The proof follows immediately from recursive substitution.

Result 7 *For any l, α, and F with associated $\mathcal{S} = \{t_1, \ldots, t_F\}$,*

$$b(l, \alpha; F) = \max\{0, l_{t_1} - \alpha d, l_{t_2} - \alpha d, \ldots, l_{t_F} - \alpha d, \}. \tag{32}$$

Proof. We sketch the proof, which proceeds by induction. Recall that

$$b(l, \alpha; F) = \hat{l}_{t_1} + \hat{l}_{t_2} + \cdots + \hat{l}_{t_F},$$

so that it is sufficient to show, for any finite $N \in \mathbb{N}$,

$$\hat{l}_{t_1} + \hat{l}_{t_2} + \cdots + \hat{l}_{t_N} = \max\{0, l_{t_1} - \alpha d, l_{t_2} - \alpha d, \ldots, l_{t_N} - \alpha d, \}. \tag{33}$$

<u>Step 1:</u> By definition, this is true for $N = 1$.

<u>Step 2:</u> Assuming that this is true for an arbitrary N, we show that it is then

true for $N+1$.

If
$$\hat{l}_{t_1} + \hat{l}_{t_2} + \cdots + \hat{l}_{t_N} = \max\{0, l_{t_1} - \alpha d, l_{t_2} - \alpha d, \ldots, l_{t_N} - \alpha d, \}$$
then using equation (31)
$$\hat{l}_{t_1} + \cdots + \hat{l}_{t_N} + \hat{l}_{t_{N+1}} = \max\{0, l_{t_1} - \alpha d, \ldots, l_{t_N} - \alpha d\} + \max\Big\{0, \min\{l_{t_{N+1}} - l_{t_N}, l_{t_{N+1}} - l_{t_{N-1}}, \ldots, l_{t_{N+1}} - l_{t_1}, l_{t_{N+1}} - \alpha d\}\Big\}.$$
Suppose $\exists\ t_i \in \{t_1, \ldots, t_N\}$ such that $l_{t_i} \geq l_{t_{N+1}}$, then
$$\max\Big\{0, \min\{l_{t_{N+1}} - l_{t_N}, l_{t_{N+1}} - l_{t_{N-1}}, \ldots, l_{t_{N+1}} - l_{t_1}, l_{t_{N+1}} - \alpha d\}\Big\} = 0$$
and
$$\begin{aligned} \hat{l}_{t_1} + \cdots + \hat{l}_{t_N} + \hat{l}_{t_{N+1}} &= \max\{0, l_{t_1} - \alpha d, \ldots, l_{t_N} - \alpha d\} \\ &= \max\{0, l_{t_1} - \alpha d, \ldots, l_{t_N} - \alpha d, l_{t_{N+1}} - \alpha d\}. \end{aligned}$$
Conversely, suppose $l_{t_{N+1}} > l_{t_i}\ \forall t_i \in \{t_1, \ldots, t_N\}$ and let $l_{t^*} = \max\{l_{t_1}, \ldots, l_{t_N}\}$. Then
$$\begin{aligned} \hat{l}_{t_1} + \cdots + \hat{l}_{t_N} + \hat{l}_{t_{N+1}} &= \max\{0, l_{t^*} - \alpha d\} + \max\Big\{0, \min\{l_{t_{N+1}} - l_{t^*}, l_{t_{N+1}} - \alpha d\}\Big\} \\ &= \max\{0, l_{t_{N+1}} - \alpha d\} \\ &= \max\{0, l_{t_1} - \alpha d, \ldots, l_{t_N} - \alpha d, l_{t_{N+1}} - \alpha d\}. \end{aligned}$$

■

We now present the claims and results for a generic T along with the necessary proofs. Claims 3 and 4 are analogous to claims 1 and 2, respectively.

Claim 3 *Increasing the reserve ratio decreases the expected borrowing fee at any frequency. That is, for any F and $\alpha \in (0, \frac{\bar{L}}{d})$,*
$$\frac{\partial}{\partial \alpha}\Big[\Psi(\alpha; F)\Big] < 0.$$

Proof. From (32), it is clear that $\alpha_2 > \alpha_1$ implies
$$\begin{aligned} b(l, \alpha_1; F) &\geq b(l, \alpha_2; F) \text{ for all } l \\ b(l, \alpha_1; F) &> b(l, \alpha_2; F) \text{ for } l = (\bar{L}_t,\ t \in \mathcal{S}). \end{aligned}$$
Since ψ is strictly increasing, this implies that
$$\begin{aligned} \psi[b(l; \alpha_1, F)] &\geq \psi[b(l; \alpha_2, F)] \qquad \text{for all } l \\ \psi[b(l; \alpha_1, F)] &> \psi[b(l; \alpha_2, F)] \qquad \text{for } l = (\bar{L}_t, t \in \mathcal{S}). \end{aligned}$$

We conclude that

$$\begin{aligned}\alpha_2 > \alpha_1 \quad &\Rightarrow \quad \mathbb{E}_l\Big[\psi[b(l;\alpha_1,F)]\Big] > \mathbb{E}_l\Big[\psi[b(l;\alpha_2,F)]\Big] \\ &\Rightarrow \quad \Psi(\alpha_1;F) > \Psi(\alpha_2;F).\end{aligned}$$

■

Claim 4 *For any level of reserves, expected borrowing fees are larger under a system with a higher frequency of settlement. That is, for* F', F'' *with associated* $\mathcal{S}' \subset \mathcal{S}''$*,* $\forall\ \alpha \in (0, \frac{\bar{L}}{d})$*,*

$$\Psi(\alpha; F') \leq \Psi(\alpha; F'').$$

Proof. This result follows directly from (32). ■

We now present the generalized form of result 1.

Result 8 *Profit-maximizing banks choose to retain a larger fraction of deposits as reserves under a system with a higher frequency of settlement. That is, for* F', F'' *with associated* $\mathcal{S}' \subset \mathcal{S}''$*,*

$$\alpha(F') \leq \alpha(F'').$$

Toward a Proof of Result 8. We require two intermediate results to prove Result 8:

Result 9 *Expected borrowing fees exhibit convexity in* α *at any settlement frequency. That is, for any* $F \leq T$ *and any* $\alpha_1, \alpha_2 \in (0, \frac{\bar{L}}{d})$*,*

$$\alpha_1 > \alpha_2 \quad \Rightarrow \quad \frac{\partial}{\partial \alpha}\Big[\Psi(\alpha_1;F)\Big] > \frac{\partial}{\partial \alpha}\Big[\Psi(\alpha_2;F)\Big].$$

Proof. Suppose that $\alpha_1 > \alpha_2$. Since ψ is strictly increasing and strictly convex, we know that

$$b_2 > b_1 \quad \Rightarrow \quad \frac{\psi(b_2+\Delta b) - \psi(b_2)}{\Delta b} > \frac{\psi(b_1+\Delta b) - \psi(b_1)}{\Delta b}$$

for $\Delta b > 0$ sufficiently small. It follows that

$$b_2 > b_1 \quad \Rightarrow \quad \psi(b_2+\Delta b_2) - \psi(b_2) > \psi(b_1+\Delta b_1) - \psi(b_1) \tag{34}$$

if $\Delta b_2 \geq \Delta b_1 > 0$. Recall that a positive perturbation in α is a negative perturbation in b. Therefore, let

$$\begin{aligned} b_1 &= b(l, \alpha_1 + \Delta\alpha; F) = \max\{0, l_{t_1} - (\alpha_1 + \Delta\alpha)d, \ldots, l_{t_F} - (\alpha_1+\Delta\alpha)d\} \\ b_1 + \Delta b_1 &= b(l, \alpha_1; F) = \max\{0, l_{t_1} - \alpha_1 d, \ldots, l_{t_F} - \alpha_1 d\} \\ b_2 &= b(l, \alpha_2 + \Delta\alpha; F) = \max\{0, l_{t_1} - (\alpha_2 + \Delta\alpha)d, \ldots, l_{t_F} - (\alpha_2+\Delta\alpha)d\} \\ b_2 + \Delta b_2 &= b(l, \alpha_2; F) = \max\{0, l_{t_1} - \alpha_2 d, \ldots, l_{t_F} - \alpha_2 d\} \end{aligned}$$

for sufficiently small $\Delta\alpha > 0$. We know that $b_2 \geq b_1$ for all l, and strictly greater for some l that occurs with positive probability. Further, let $l'_{t^*} = \max_{t \in \mathcal{S}'} l_t$ and $l''_{t^*} = \max_{t \in \mathcal{S}''} l_t$. Then

$$\begin{aligned}
\Delta b_1 &= (b_1 + \Delta b_1) - b_1 = \max\{0, l'_{t^*} - \alpha_1 d\} - \max\{0, l'_{t^*} - (\alpha_1 + \Delta\alpha) d\} \\
\Delta b_2 &= (b_2 + \Delta b_2) - b_2 = \max\{0, l''_{t^*} - \alpha_2 d\} - \max\{0, l''_{t^*} - (\alpha_2 + \Delta\alpha) d\}
\end{aligned}$$

Consider the three possible cases:

$$\begin{array}{lll}
l''_{t^*} - \alpha_2 d \ > l'_{t^*} - \alpha_1 d \ > 0 & \Rightarrow & \Delta b_1 = \Delta b_2 = \Delta\alpha d. \\
l''_{t^*} - \alpha_2 d \ > 0 > l'_{t^*} - \alpha_1 d & \Rightarrow & \Delta b_1 = 0 < \Delta\alpha d = \Delta b_2. \\
0 > l''_{t^*} - \alpha_2 d \ > l'_{t^*} - \alpha_1 d & \Rightarrow & \Delta b_1 = \Delta b_2 = 0.
\end{array}$$

Clearly, $\Delta b_1 \leq \Delta b_2$ for all l, and moreover $\Delta b_1 > 0$ for some l with $\lambda(l) > 0$. So by (34) we have that $\alpha_1 > \alpha_2$ implies

$$\psi[b(l, \alpha_1 + \Delta\alpha; F)] - \psi[b(l, \alpha_1; F)] \geq \psi[b(l, \alpha_2 + \Delta\alpha; F)] - \psi[b(l, \alpha_2; F)]$$

for all l and strictly greater than for some l that occurs with strictly positive probability. Therefore, taking the expectation across l, we find

$$\begin{array}{lll}
\alpha_1 > \alpha_2 & \Rightarrow & \dfrac{\Psi(\alpha_1 + \Delta\alpha; F) - \Psi(\alpha_1; F)}{\Delta\alpha} > \dfrac{\Psi(\alpha_2 + \Delta\alpha; F) - \Psi(\alpha_2; F)}{\Delta\alpha}, \\
 & \Rightarrow & \dfrac{\partial}{\partial\alpha}\Big[\Psi(\alpha_2; F)\Big] > \dfrac{\partial}{\partial\alpha}\Big[\Psi(\alpha_1; F)\Big],
\end{array}$$

the desired result. ■

Result 10 *The impact of a marginal increase in the reserve ratio is larger (more negative) under a higher frequency of settlement. That is, for F', F'' with associated $\mathcal{S}' \subset \mathcal{S}''$,*

$$\frac{\partial}{\partial\alpha}\Big[\Psi(\alpha; F')\Big] > \frac{\partial}{\partial\alpha}\Big[\Psi(\alpha; F'')\Big].$$

Proof. We follow the same logic employed in the proof of Result 1. Let

$$\begin{aligned}
b_1 &= b(l, \alpha + \Delta\alpha; F') = \max\{0, l_{t'_1} - (\alpha + \Delta\alpha)d, \ldots, l_{t'_F} - (\alpha + \Delta\alpha)d\} \\
b_1 + \Delta b_1 &= b(l; \alpha, F') = \max\{0, l_{t'_1} - \alpha d, \ldots, l_{t'_F} - \alpha d\} \\
b_2 &= b(l; \alpha + \Delta\alpha, F'') = \max\{0, l_{t''_1} - (\alpha + \Delta\alpha)d, \ldots, l_{t''_F} - (\alpha + \Delta\alpha)d\} \\
b_2 + \Delta b_2 &= b(l; \alpha, F'') = \max\{0, l_{t''_1} - \alpha d, \ldots, l_{t''_F} - \alpha d\}
\end{aligned}$$

for a sufficiently small $\Delta\alpha > 0$. We have shown above that $b_2 \geq b_1$ for all l and strictly greater for some l with $\lambda(l) > 0$. Again, let $l'_{t^*} = \max_{t \in \mathcal{S}'} l_t$ and $l''_{t^*} = \max_{t \in \mathcal{S}''} l_t$, and consider Δb_1 and Δb_2:

$$\begin{aligned}
\Delta b_1 &= \max\{0, l'_{t^*} - \alpha d\} - \max\{0, l'_{t^*} - (\alpha + \Delta\alpha) d\} \\
\Delta b_2 &= \max\{0, l''_{t^*} - \alpha d\} - \max\{0, l''_{t^*} - (\alpha + \Delta\alpha) d\}
\end{aligned}$$

We know that $l''_{t^*} \geq l'_{t^*}$. If $l'_{t^*} = l''_{t^*}$, we have that $b_1 = b_2$ and $\Delta b_1 = \Delta b_2$ so that $\psi(b_1 + \Delta b_1) - \psi(b_1) = \psi(b_2 + \Delta b_2) - \psi(b_2)$. For the case of $l''_{t^*} > l'_{t^*}$, we have three subcases:

$$\begin{array}{lll} l''_{t^*} - \alpha_2 d > l'_{t^*} - \alpha_1 d > 0 & \Rightarrow & \Delta b_1 = \Delta b_2 = \Delta\alpha d. \\ l''_{t^*} - \alpha_2 d > 0 > l'_{t^*} - \alpha_1 d & \Rightarrow & \Delta b_1 = 0 < \Delta\alpha d = \Delta b_2. \\ 0 > l''_{t^*} - \alpha_2 d > l'_{t^*} - \alpha_1 d & \Rightarrow & \Delta b_1 = \Delta b_2 = 0. \end{array}$$

Therefore, by (34), we conclude that

$$\psi[b(l,\alpha;F'')] - \psi[b(l,\alpha+\Delta\alpha;F'')] \geq \psi[b(l,\alpha;F')] - \psi[b(l,\alpha+\Delta\alpha;F')]$$

for all l and strictly greater than for some l which occurs with strictly positive probability. Therefore, we can rearrange to conclude that

$$\frac{\Psi(\alpha+\Delta\alpha;F') - \Psi(\alpha;F')}{\Delta\alpha} > \frac{\Psi(\alpha+\Delta\alpha;F'') - \Psi(\alpha;F'')}{\Delta\alpha},$$

the desired result. ■

Proof of Result 8. By the bank's first order condition, $\alpha(F')$ and $\alpha(F'')$ satisfy

$$\frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F');F')\Big] = \frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F'');F'')\Big] = dr\rho.$$

Suppose, towards a contradiction, that $\alpha(F') > \alpha(F'')$. By Result 9,

$$\frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F');F')\Big] > \frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F'');F')\Big].$$

But by Result 10, we have

$$\frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F'');F')\Big] > \frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F'');F'')\Big],$$

implying that

$$\frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F');F')\Big] > \frac{\partial}{\partial\alpha}\Big[\Psi(\alpha(F'');F'')\Big],$$

a contradiction. ■

Similarly, the following result generalizes result 2 in section 3.

Result 11 *For any loan price ρ, banking fees are larger under a higher frequency of settlement. That is,for F', F'' with associated $\mathcal{S}' \subset \mathcal{S}''$,*

$$\phi(F') \leq \phi(F'').$$

Proof. Since $\alpha(F')$ is the optimal solution to the bank's problem,

$$[1-\alpha(F')]r\rho - a - \frac{\Psi(\alpha(F');F')}{d} \geq [1-\alpha(F'')]r\rho - a - \frac{\Psi(\alpha(F'');F')}{d}.$$

This simplifies to

$$[\alpha(F'')-\alpha(F')]r\rho \geq \frac{\Psi(\alpha(F');F') - \Psi(\alpha(F'');F')}{d}.$$

We have shown that $\Psi(\alpha;F'') \geq \Psi(\alpha;F')$, so that

$$[\alpha(F'')-\alpha(F')]r\rho \geq \frac{\Psi(\alpha(F'),F') - \Psi(\alpha(F''),F'')}{d}.$$

This means that

$$a - [1-\alpha(F'')]r\rho + \frac{\Psi(\alpha(F''),F'')}{d} \geq a - [1-\alpha(F')]r\rho + \frac{\Psi(\alpha(F'),F')}{d},$$

or, equivalently,

$$\phi(F'') \geq \phi(F').$$

■

We conclude with the the generalized results regarding the risks of infrequent settlement (i.e. analogues to results 3, 4, and 5).

Result 12 *At any time $t \in \mathcal{T}$ and for any reserve ratio α, the size of outstanding liabilities in excess of reserves is always smaller under higher frequency of settlement. That is, for any t, α, l, and F', F'' with associated $\mathcal{S}' \subset \mathcal{S}''$,*

$$\hat{l}_t(l,\alpha;F') \geq \hat{l}_t(l,\alpha;F'').$$

Proof. Since $t_i' \in \mathcal{S}' \Rightarrow t_i' \in \mathcal{S}''$ and $\exists\ t_i'' \in \mathcal{S}''$ such that $t_i'' \notin \mathcal{S}'$, we know that

$$\min\{l_t - l_{s_t'}, \ldots, l_t - l_{t_1'}, ;_t - \alpha d\} \geq \min\{l_t - l_{s_t''}, \ldots, l_t - l_{t_1''}, ;_t - \alpha d\},$$

where s_t' (s_t'') corresponds to the period in which the most recent settlement occurred prior to t under frequency F' (F''). From (31), it follows directly that $\hat{l}_t(l,\alpha;F') \geq \hat{l}_t(l,\alpha;F'')$. ■

Result 13 *At any time $t \in \mathcal{T}$ and for any settlement frequency F, increasing the reserve ratio decreases the size of outstanding liabilities in excess of reserves. That is, for any t, F, and l,*

$$\alpha_2 > \alpha_1 \quad \Rightarrow \quad \hat{l}_t(l,\alpha_1;F) \geq \hat{l}_t(l,\alpha_2;F).$$

We omit the proof, as it is a trivial implication of (31). From these two results, we arrive at the following conclusion.

Result 14 *For any realization of Δ, the aggregate daily default decreases under increased settlement frequency. That is, for any $\Delta \in [0,1]$, and F', F'' with associated $\mathcal{S}' \subset \mathcal{S}''$,*

$$\delta(\Delta, F') \geq \delta(\Delta, F'').$$

Figure 2A - Reserves (eta)

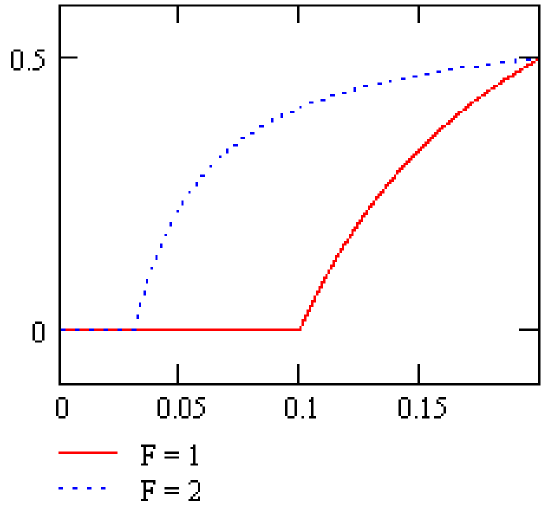


Figure 3A - Reserves (q)

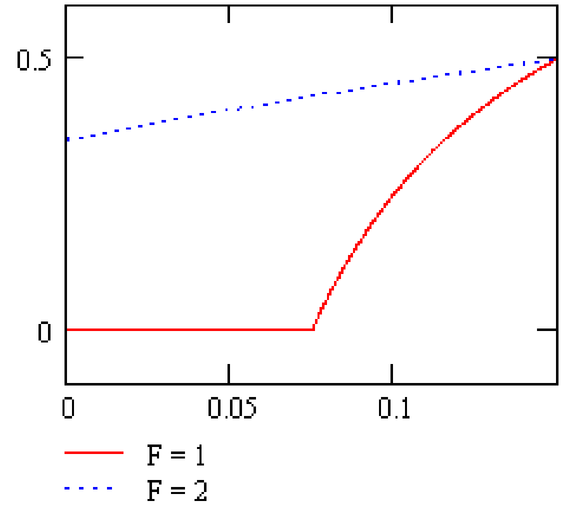


Figure 2B - Banking Fees (eta)

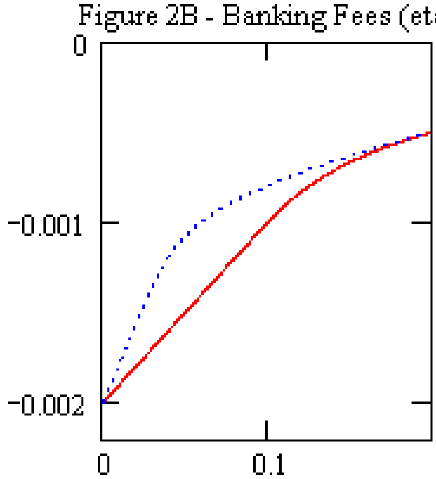


Figure 3B - Banking Fees (q)

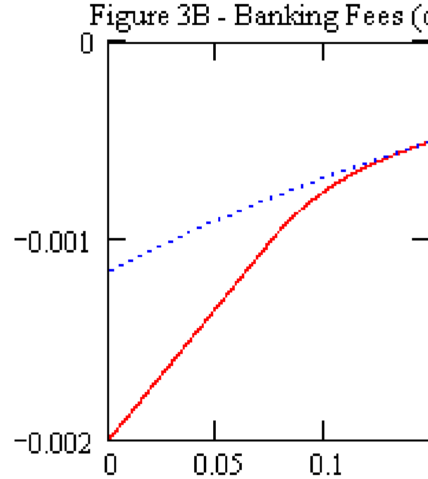


Figure 2C - Default (eta)

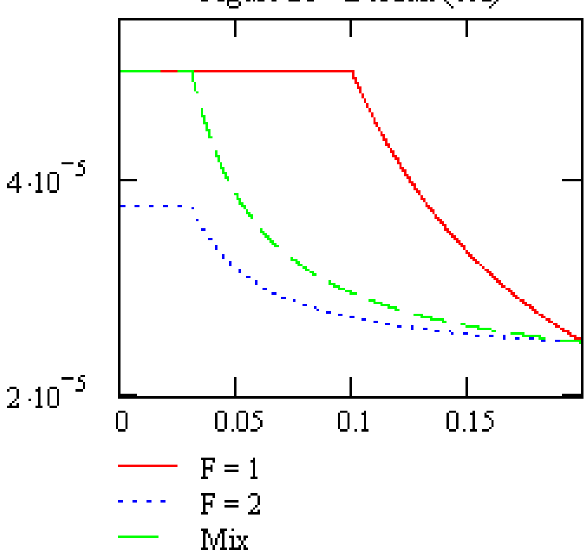
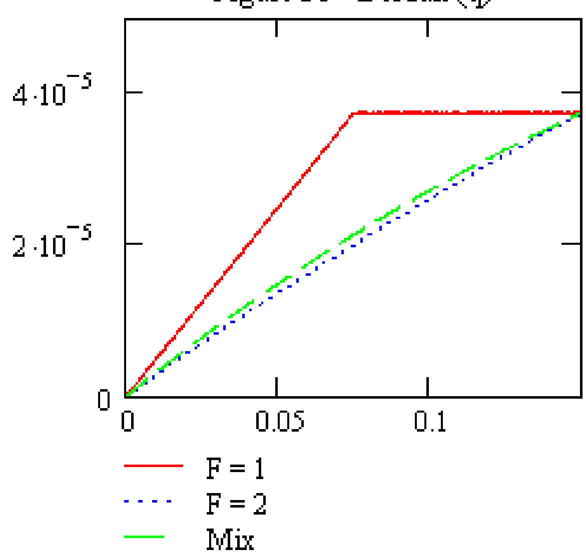


Figure 3C - Default (q)



Fixed *M*, Benchmark Values: $p=.05$, $q=.1$, $\kappa=150$

Figure 4A - Reserves

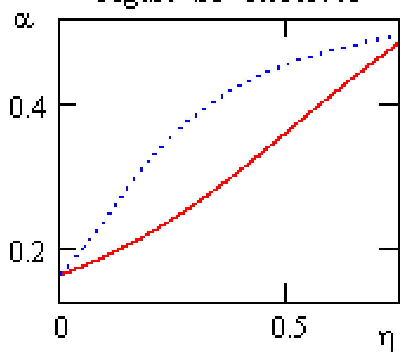


Figure 4B - Banking Fees

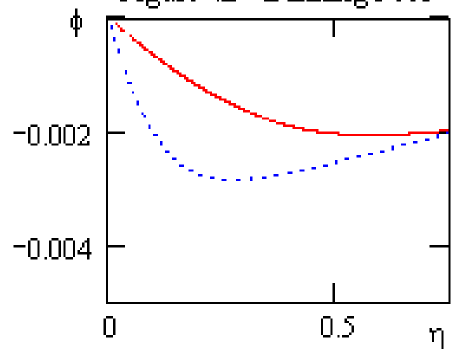


Figure 4C - Default Costs

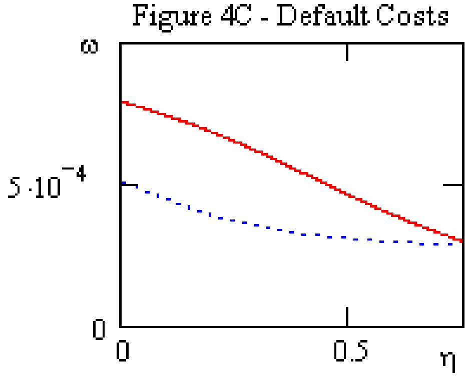


Figure 4D - Price of Loans

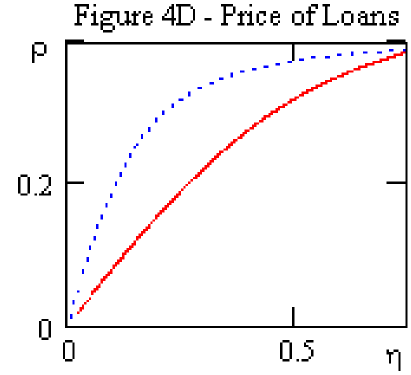
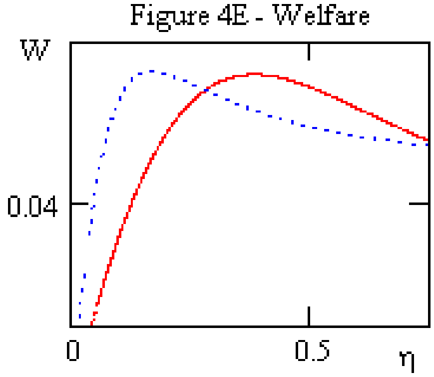
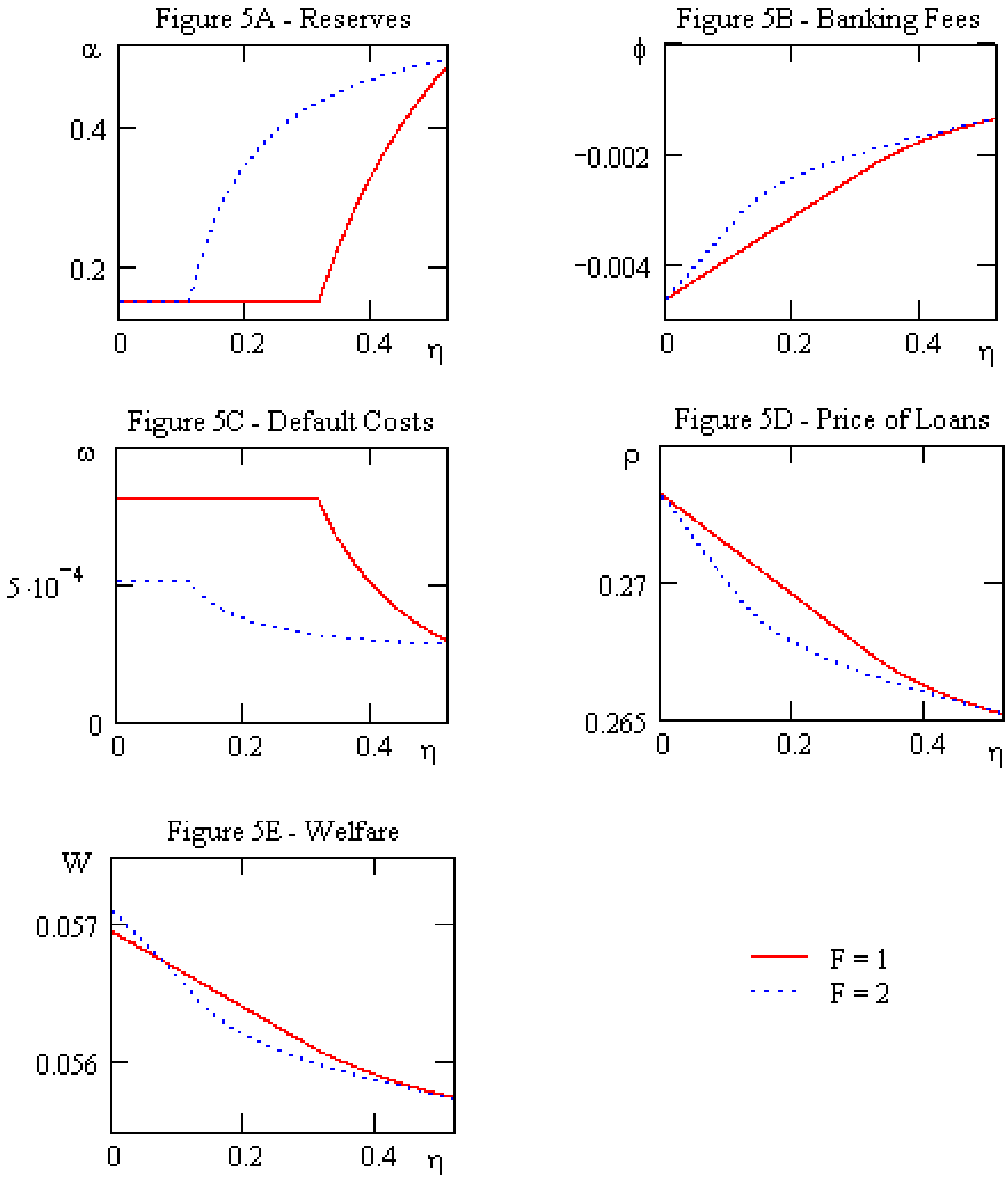


Figure 4E - Welfare

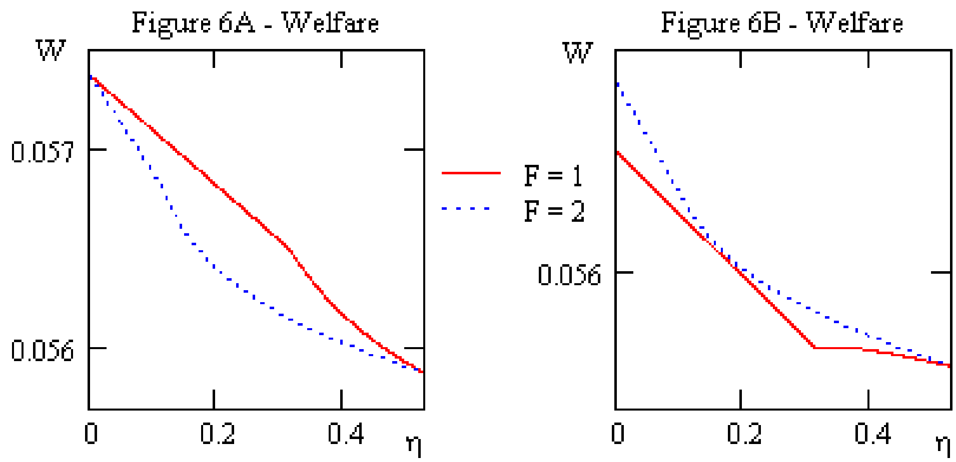


F = 1
F = 2

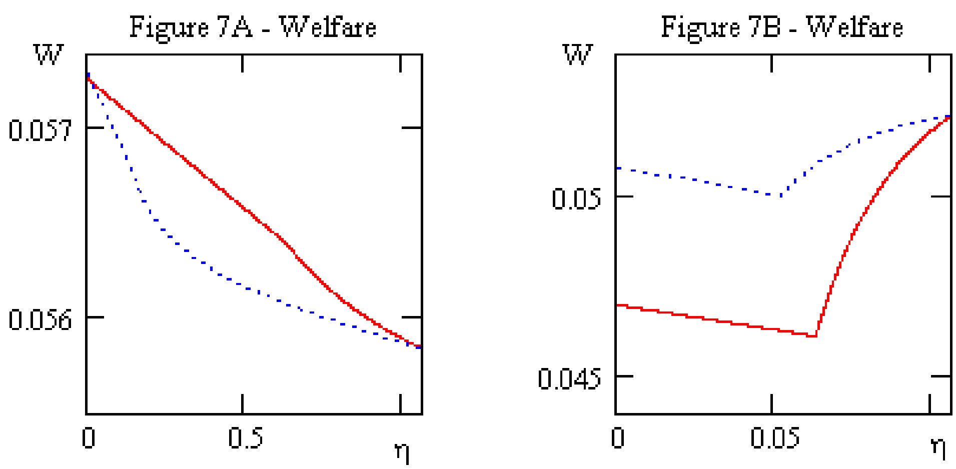
Flexible M, Benchmark Values: $p = .05$, $q = .1$, $\kappa = 150$



Flexible M, Solvency Shocks: $p=0$ (A) & $.1$ (B), $q = .1$, $\kappa=150$



Flexible M, Liquidity Shocks: $p=.05$, $q = .05$ (A) & $.5$ (B), $\kappa=150$



Flexible M, Recover Costs: $p=.05$, $q = .1$, $\kappa=50$ (A) & 450 (B)