Inside and Outside Liquidity[*]

Patrick Bolton   Tano Santos   Jose Scheinkman
Columbia University   Columbia University   Princeton University

June 2008

Abstract

   We consider a model of liquidity demand arising from maturity mismatch on one side of
the market. This demand can be met with either the cash held by those with the liquidity
need, what we refer to as inside liquidity, or alternatively via asset sales. In this latter
case then assets are exchanged for the cash held by agents other than those with these
liquidity needs. We refer to this cash as outside liquidity. The questions we address are:
(a) what determines the mix of inside and outside liquidity in equilibrium? (b) does the
market provide an efficient mix of inside versus inside liquidity? and (c) if not, what kind
of interventions can be proposed to restore efficiency? We argue that a key determinant of
the aforementioned mix is the timing of the liquidation decision. An important source of
inefficiency is the presence of asymmetric information about asset values, which increases
the longer a liquidity trade is delayed. We establish that an immediate-trading equilibrium
always exists, where liquidity trading occurs in anticipation of a liquidity shock. Another,
delayed-trading equilibrium, in which liquidity trading is a response to a liquidity shock,
may also exist. We show that, when it exists, the delayed-trading equilibrium is efficient,
despite the presence of adverse selection.

INTRODUCTION

The main goal of this paper is to propose a tractable model of maturity transformation by financial intermediaries and the resulting liquidity demand arising from the maturity mismatch between assets and liabilities. When financial intermediaries invest in long-term assets but potentially face redemptions before these assets mature they have a need for liquidity. These redemptions can be met either out of cash holdings of the financial intermediaries – what we refer to as *inside liquidity* – or out of the proceeds from asset sales to other investors with a longer horizon–what we refer to as *outside liquidity*. In reality financial intermediaries rely on both forms of liquidity and the purpose of our analysis is to determine the relative importance and efficiency of inside and outside liquidity in a competitive equilibrium of the financial sector.

Our model comprises two different groups of agents that differ in their investment horizons. One class of agents, which we denominate short run investors, prefer early to late payoffs, whereas the second class, which we refer to as long run investors, are indifferent as to the timing of the payoffs associated with their investments. One can think of these long run investors as wealthy individuals, endowments, hedge funds and even sovereign wealth funds and the short run investors as financial intermediaries with short dated liabilities. Short run investors allocate their investments between long-term and liquid assets, or *cash*, which they carry in case their long term investments do not pay in time. If they do not, short run investors can supply some or all of the long term assets in their books in exchange for cash. We refer to the cash carried by the short run investors as inside liquidity. Long-horizon investors directly invest in a portfolio of long-term and liquid assets of their own, which is also cash. These investors carry cash, which we refer to as outside liquidity, precisely to opportunistically acquire the long assets of the short run investors at low prices.

Within this model the key questions we are interested in are: first, what determines the mix of inside and outside liquidity in equilibrium? second, does the market provide an efficient mix of inside versus inside liquidity? and, third, if not, what kind of interventions can be proposed to restore efficiency?

Our model attempts to describe situations in which short run investors hold relatively sophisticated assets or securities, and where long run investors have sufficient expertise with these securities to stand ready to buy them at a relatively good price. Other investors may be ready to buy these securities, but at a much higher discount.

Still, an important potential source of inefficiency in practice and in our model remains asymmetric information about asset values between short and long horizon investors. That is,

1

even when short run investors turn to experienced long run ones to sell claims to their assets, the latter cannot always tell whether the sale is due to a sudden liquidity need or whether the financial intermediary is trying to pass on a *lemon*. This problem is familiar to financial market participants and has been widely studied in the literature in different contexts. The novel aspect our model focuses on is a *timing* dimension. Short run investors learn more about the underlying value of their assets over time. Therefore, when at the onset of a liquidity shock they choose to hold on to their positions – in the hope of riding out a temporary crisis – they run the risk of having to go to the market in a much worse position should the crisis be a prolonged one. The longer they wait the worse is the *lemons* problem and therefore the greater is the risk that they will have to sell assets at fire-sale prices.

This is a common dynamic in liquidity crises, which has not been much analyzed nor previously modeled, and which is a core mechanism in our analysis. We capture the essence of this dynamic unfolding of a liquidity crisis by establishing the existence of two types of rational expectations equilibria: an *immediate trading equilibrium*, where short run investors are rationally expected to trade at the onset of the liquidity shock and a *delayed trading equilibrium*, where they are instead correctly expected to prefer attempting to ride out the crisis and to only trade as a last resort should the crisis be a prolonged one. We show that for some parameter values only the immediate trading equilibrium exists, while for other values both equilibria coexist.

When two different rational expectations equilibria can coexist one naturally wonders how they compare in terms of efficiency. Which is better? Interestingly, the answer to this question turns critically on the implications for ex-ante portfolio-composition decisions of both the short and long run investors of immediate or delayed-trading expectations.

In a nutshell, under the expectation of immediate liquidity-trading, long-run investors expect to obtain the assets of short run investors at close to fair value. In this case the returns of holding outside liquidity are low and thus there is little of it held by long run investors. On the other side of the liquidity trade, short run investors will then expect to be able to sell a relatively small fraction of assets at close to fair value, and therefore respond by relying more heavily on inside liquidity. That is, they tend to hold a larger fraction of their assets in liquid securities or cash. In other words, in an immediate trading equilibrium there is less *cash-in-the-market pricing* (to borrow a term from Allen and Gale, 1998), which reduces the return to outside liquidity and therefore its' supply. The reduced supply of outside liquidity, in turn, causes financial intermediaries to rely more on inside liquidity and, thus, bootstraps

2

the relatively high equilibrium price for the assets held by short run investors under immediate liquidity trading.

In contrast, under the expectation of delayed liquidity trading, short run investors rely more on outside liquidity. Here the bootstrap works in the other direction, as long run investors decide to hold more cash in anticipation of a larger future supply of the assets held by short run investors at more favorable cash-in-the-market pricing. The reason why there is more favorable cash-in-the-market pricing in the delayed trading equilibrium, in spite of the worse lemons problem, is that in this equilibrium the return to investing in the long maturity asset is also higher, due to the lower overall probability of liquidating assets at fire-sale prices.

In sum, immediate trading equilibria are based on a greater reliance on inside liquidity than delayed trading equilibria. And, to the extent that there is a greater reliance on outside liquidity by short run investors in a delayed trading equilibrium, one should expect – and we indeed establish – that equilibrium prices of financial intermediary assets are lower in the delayed-trading than in the immediate-trading equilibrium. In other words, our model predicts a common dynamic of liquidity crises, in which asset prices progressively deteriorate throughout the crisis. Importantly, this predictable pattern in asset prices is consistent with no arbitrage, as short run investors prefer to delay asset sales, despite the deterioration in asset prices, in the hope that they wont have to trade at all at fire sale prices.[1]

Because of this deterioration in asset prices one would expect that welfare is also worse in the delayed-trading than in the immediate-trading equilibrium. However, this is not the case in our model. As it turns out, the Pareto superior equilibrium is in fact the delayed-trading equilibrium. What is the economic logic behind this somewhat surprising result?

The answer is that the fundamental gains from trade in our model are between short horizon investors, who undervalue long term assets, and long horizon investors, who undervalue cash. Thus, the more short horizon investors can be induced to hold long assets and the more long horizon investors can be induced to hold cash, the higher are the gains from trade and therefore the higher is welfare. In other words, the welfare efficient form of liquidity in our model is outside liquidity. Since the delayed trading equilibrium relies more on outside liquidity it is more efficient.

Under complete-information, when fundamental asset values are fully known to both

---

[1]The short run investors' decision to delay trading has all the hallmarks of *gambling for resurrection*. But it is in fact unrelated to the idea of excess risk taking as these financial intermediaries will choose to delay whether or not they are levered.

short run investors and other investors, it is actually efficient to rely exclusively on outside liquidity in our model. In the presence of adverse selection, however, outside liquidity involves too much dilution of ownership and is generally too costly for short run investors, so that they will tend to rely partially on inside liquidity. As the lemons' problem worsens – in particular, as short run investors are less likely to trade for liquidity reasons when they engage in delayed trading – the cost of outside liquidity rises. There is then a point when the cost is so high that short run investors and their investors are better off postponing the redemption of their investments altogether, rather than realize a very low fire-sale price for their valuable long-term assets. At that point the delayed trading equilibrium collapses, as only lemons would get traded for early redemption.

Interestingly, short run investors could reduce the lemons' cost associated with outside liquidity by committing ex-ante to sell their long-term assets for cash in the delayed trading equilibrium whenever these assets mature late. But, as much as they would like to make such a commitment, the subtle state contingent commitment technology that is required will generally be hard to come by. This then naturally raises the question whether some form of public intervention may provide an adequate substitute for the lack of commitment by short run investors.

There are two fundamental market inefficiencies in our model that public policy might mitigate. An ex-post inefficiency, which arises when the delayed-trading equilibrium fails to exist, and an ex-ante inefficiency in the form of an excess reliance on inside liquidity. It is worth emphasizing that a common prescription against banking liquidity crises–namely to require that banks hold cash reserves or excess equity capital–would be counterproductive in our model. Such a requirement would only force short run investors to rely more on inefficient inside liquidity and would undermine the supply of outside liquidity. As we illustrate in an example, such regulations could push the financial sector out of a delayed-trading equilibrium into an immediate-trading equilibrium. Interestingly, in that case short run investors may actually hold liquid assets in equilibrium far in excess of what they are required to hold.

Rather than mandate minimum liquid asset holdings by financial intermediaries, a more effective policy intervention in our model could be to enforce value-at-risk (VAR) type regulations, which require that short run investors sell or reduce their exposure to risky securities when their overall VAR is too high. Such regulations would have the effect in our model of forcing financial intermediaries to rely more on outside liquidity, and could therefore help select a delayed-trading equilibrium. An important novel insight of our analysis is, thus, that

4

VAR type regulations have the unintended positive effect of shifting the overall reliance of the financial system on a more efficient form of liquidity supply: outside liquidity.

Another potentially beneficial intervention in our model is a policy of public provision of liquidity by a central bank, whereby the monetary authority offers to guarantee a minimum price for assets by lend against collateral. Such a policy could improve efficiency in our model even though short run investors may actually never rely on the lending facility in equilibrium, simply by inducing intermediaries to rely less on inefficient inside liquidity in the first place. That is, such a policy would induce long-term investors to hold more cash in the knowledge that short run investors rely less on inside liquidity, and thus help increase the availability of outside liquidity. Far from being a substitute for privately provided liquidity, a commitment to offer public emergency liquidity could be a complement and give rise to positive spillover effects on the provision of outside liquidity. Long-term investors would be prepared to hold more cash in the knowledge that financial intermediaries rely less on inside liquidity, and if the central bank's lending terms are sufficiently punitive, they can then hope to acquire valuable long-term assets at a reasonable price.

Related literature. Our paper is related to the literature on banking crises, on the one hand, and to the literature on liquidity crises, and limits of arbitrage, on the other other. Our analysis differs from the main contributions in these two literatures mainly in two respects: first, our focus on ex-ante efficiency and equilibrium portfolio composition, and second, the endogenous timing of liquidity trading. Still, our analysis shares several important themes and ideas with these two literatures. We briefly discuss the most related contributions in each of these literatures in turn.

Consider first the banking literature. Diamond and Dybvig (1983)[2] provide the first model of investor liquidity demand, maturity transformation, and inside liquidity. In their model a bank run may occur if there is insufficient inside liquidity to meet depositor withdrawals. In contrast to our model, investors are identical ex-ante, and are risk-averse with respect to future liquidity shocks. The role of financial intermediaries is to provide insurance against idiosyncratic investors' liquidity shocks.

Bhattacharya and Gale (1986) provide the first model of both inside and outside liquidity by extending the Diamond and Dybvig framework to allow for multiple banks, which may face different liquidity shocks. In their framework, an individual bank may meet depositor

_____

[2]See also Bryant (1980).

withdrawals with either inside liquidity or outside liquidity by selling claims to long-term assets to other banks who may have excess cash reserves. An important insight of their analysis is that individual banks may free-ride on other banks' liquidity supply and choose to hold too little liquidity in equilibrium.

More recently, Allen and Gale (2000) and Freixas, Parigi, and Rochet (2000) (see also Aghion, Bolton and Dewatripont, 2000) have analyzed a model of liquidity provided through the interbank market, which can give rise to contagious liquidity crises. The main mechanism they highlight is the default on an interbank loan which depresses secondary-market prices and pushes other banks into a liquidity crisis. Subsequently, Acharya (2001) and Acharya and Yorulmazer (2005) have, in turn, introduced optimal bailout policies in a model with multiple banks and cash-in-the-market pricing of loans in the interbank market.

While Diamond and Dybvig considered idiosyncratic liquidity shocks and the risk of panic runs that may arise as a result of banks' attempts to insure depositors against these shocks, Allen and Gale (1998) consider aggregate business-cycle shocks and point to the need for equilibrium banking crises to achieve optimal risk-sharing between depositors. In their model aggregate shocks may trigger the need for asset sales, but their analysis does not allow for the provision of both inside and outside liquidity.

Another strand of the banking literature, following Holmstrom and Tirole (1998) considers liquidity demand on the corporate borrowers' side rather than on depositors' side, and asks how efficiently this liquidity demand can be met through bank lines of credit. This literature emphasizes the need for public liquidity to supplement private liquidity in case of aggregate demand shocks.

Most closely related to our model is the framework considered in Fecht (2004), which itself builds on the related models of Diamond (1997) and Allen and Gale (2000). The models of Diamond (1997) and Fecht (2004) seek to address an important weakness of the Diamond and Dybvig theory, which cannot account for the observed coexistence of financial intermediaries and securities markets. Liquidity trading in secondary markets undermines liquidity provision by banks and obviates the need for any financial intermediation in the Diamond and Dybvig setting. To address this objection, Diamond (1997) introduces a model where banks coexist with securities markets due to the fact that households face costs in switching out of the banking sector and into securities markets. Fecht (2004) extends Diamond (1997) by introducing segmentation on the asset side between financial intermediaries' investments in firms and claims issued directly to investors though securities markets. Also, in his model banks

have local (informational) monopoly power on the asset side, and subsequently can trade their assets in securities markets for cash–a form of outside liquidity. Finally, Fecht (2004) also allows for a contagion mechanism similar to Allen and Gale (2000) and Diamond and Rajan (2005)[3], whereby a liquidity shock at one bank propagates itself through the financial system by depressing asset prices in securities markets.

Two other closely related models are Gorton and Huang (2004) and Parlour and Plantin (2007). Gorton and Huang also consider liquidity supplied in a general equilibrium model and also argue that publicly provided liquidity can be welfare enhancing if the private supply of liquidity involves a high opportunity cost. However, in contrast to our analysis they do not look at the optimal composition of inside and outside liquidity, nor do they consider the dynamics of liquidity trading. Parlour and Plantin (2007) consider a model where banks may securitize loans, and thus obtain access to outside liquidity. As in our setting, the efficiency of outside liquidity is affected by adverse selection. But in the equilibrium they characterize liquidity may be excessive for some banks–as it undermines their loan origination standards–and too low for other banks, who may be perceived as holding excessively risky assets.

The second literature our model is related to is the literature on liquidity and the dynamics of arbitrage by capital or margin-constrained speculators in the line of Shleifer and Vishny (1997). The typical model in this literature (e.g. Kyle and Xiong, 2001 and Xiong, 2001) also analyzes outside liquidity and also obtains episodes of fire-sale pricing–and even destabilizing price dynamics–following negative shocks that tighten speculators' margin constraints. However, most models in this literature do not address the issue of deteriorating adverse selection and the timing of liquidity trading, nor do they explore the question of the optimal mix between inside and outside liquidity. The most closely related articles, besides Kyle and Xiong (2001) and Xiong (2001) are Gromb and Vayanos (2002), Brunnermeier and Pedersen (2007) and Kondor (2007). In particular, Brunnermeier and Pedersen (2007) also focus on the spillover effects of inside and outside liquidity, or what they refer to as funding and market liquidity.

II. THE MODEL

We consider a model with three phases: an investment phase (date 0), an interim trading phase (dates 1 and 2) and an unwinding phase of all long duration assets (date 3). There are

---

[3]Another feature in Diamond and Rajan (2005) in common with our setup is the idea that financial intermediaries possess superior information about their assets, which is another source of illiquidity.

two classes of agents which differ in their investment horizons as well as their investment opportunity sets. In particular, one class of agents is potentially subject to a maturity missmatch during the interim trading phase and this generates a demand for liquidity. This demand can be met with either cash carried by the agents subject to this maturity missmatch or by the sale of assets to the other class of agents, who may also carry cash to acquire these assets opportunistically. We call the cash carried by those who demand liquidity, *inside liquidity*, and the cash supplied by the second class of agents to acquire the assets *outside liquidity*. The novelty of our analysis resides in the timing of these asset sales. As mentioned, our interim trading phase is divided in two distinct periods. In the first, date 1, there is an aggregate shock that determines the average quality of the assets held by one class of agents and that is known to all. In date 2 there are idiosyncratic shocks to the assets and knowledge of the nature of these shocks accrues only to those in possession of the assets. Thus trading in this last date occurs under conditions of asymmetric information. Our purpose is to understand what determines the timing of the asset sales and the distribution of outside versus inside liquidity. We also want to understand the welfare consequences associated with these choices and thus offer regulatory prescriptions in case the market does not produce efficient outcomes. After this brief sketch of the model we now give a more detailed account of the framework.

II.A Agents

There are two types of agents, short and long run investors with preferences over periods $t = 1, 2, 3$. *Short run investors*, of which there is a unit mass, have preferences

$$u(C_1, C_2, C_3) = C_1 + C_2 + \delta C_3, \tag{1}$$

where $C_t \geq 0$ denotes consumption at dates $t = 1, 2, 3$ and $\delta \in (0, 1)$. These investors have one unit of endowment at date $t = 0$ and 0 in every other date. There are also *long run investors*: There is a unit mass of long run investors, each with $\kappa > 0$ units of endowment at $t = 0$ and again no endowment at subsequent dates. Their utility function is simply given by

$$\hat{u}(C_1, C_2, C_3) = \sum_{t=1}^{3} C_t,$$

provided $C_t \geq 0$.

In what follows we refer to short and long run investors as SRs and LRs respectively.

II.B Assets

For simplicity we assume that the two types of investors have access to different investment opportunity sets and comment on this assumption further below. First, both types can hold *cash* with a gross per-period rate of return of one. LRs in addition can invest in a *long* asset. Specifically, we assume that each LR has access to a decreasing-returns-to-scale long maturity asset that returns $\varphi(x)$ at date $t = 3$ for an initial investment at date 0 of $x = (\kappa - M)$, where $M \geq 0$ is the LRs' cash holding and our definition of outside liquidity.

SRs can invest in a *risky* asset, which is a constant returns to scale technology, that pays an amount $\widetilde{\rho}_t$ at dates $t = 1, 2, 3$ where $\widetilde{\rho}_t \in \{0, \rho\}$ and $\rho > 1$. The payoffs of the risky asset is the only source of uncertainty in the model and are shown in Figure 1. For simplicity we assume that there is a first aggregate *maturity shock* that affects risky assets. That is, agents learn first whether all risky assets mature at date 1, or at some later date. Subsequently, the realized value of a risky asset and whether it matures at date 2 or 3 is determined by an idiosyncratic shock. More specifically, at date $t = 1$ the risky asset is subject to an aggregate shock. The asset either pays ρ (in state $\omega_{1\rho}$), which occurs with probability λ, or with probability $(1 - \lambda)$ state ω_{1L} occurs, in which asset yields either a payoff $\widetilde{\rho}_2 \in \{0, \rho\}$ at date 2 or a late payoff $\widetilde{\rho}_3 \in \{0, \rho\}$ at date 3. After date $t = 1$ all shocks are idiosyncratic. The idiosyncratic shocks are represented by two separate *i.i.d.* random variables: (i) conditional on a delayed aggregate maturity shock (past date 1), an individual asset matures with probability θ at date 2 and with probability $(1 - \theta)$ at date 3 (in state ω_{2L}); (ii) when the asset matures at either dates $t = 2, 3$ it yields $\widetilde{\rho}_t = \rho$ with probability η (in states $\omega_{2\rho}$ and $\omega_{3\rho}$, respectively) and $\widetilde{\rho}_t = 0$ with probability $(1 - \eta)$ for $t = 2, 3$ (in states ω_{20} and ω_{30}.) The realization of the idiosyncratic shocks is private information to the SR holding the risky asset. We denote by m the amount of cash held by SRs and then $1 - m$ is the amount invested in the risky asset. m is thus our measure of inside liquidity.

II.C Financial markets

At dates 1 and 2 a secondary market opens where claims on the SR's risky asset can be traded for the cash held by the LRs. In particular SRs can liquidate the risky assets in state ω_{1L}, which we refer to as the *immediate trading date*. To emphasize this immediate aspect we denote the state as $\omega_i = \omega_{1L}$ and the price at which the risky asset trades as $P(\omega_i)$. Alternatively SRs can instead postpone decisions until $t = 2$ and thus "give another chance" for the asset to pay off. Recall that knowledge of the idiosyncratic shocks accrues only to SRs

and thus they know whether they are in states $\omega_{2\rho}$, ω_{20}, or ω_{2L}. SRs in state $\omega_{2\rho}$ collect the payoff and consume accordingly. Agents in state ω_{20} have worthless assets whereas agents in ω_{2L} have assets with an expected payoff of $\eta\rho$, a payoff that is realized at $t = 3$. SRs in ω_{2L} can either liquidate the asset or carry it to date $t = 3$. LRs instead only know that the assets sold at date $t = 2$ can either be "lemons," those sold by SRs in state ω_{20}, or good assets which can still pay at $t = 3$. The nature of the adverse selection problem is thus that the LR's information set is $\omega_d = \{\omega_{20}, \omega_{2L}\}$, where "$d$" stand for *delayed*, which is how we refer to the $t = 2$ trading date. Accordingly we let $P(\omega_d)$ be the price at which assets can be sold at ω_d.

II.D Assumptions

We introduce some minimal assumptions that will focus the analysis on the economically interesting questions and simplify considerably the presentation. We start with assumptions on the different technologies. First we assume that outside liquidity is costly:

$$\varphi'(\kappa) > 1 \qquad \text{with} \qquad \varphi''(x) < 0 \quad \text{and} \quad \lim_{x \longrightarrow 0} \varphi'(x) = +\infty \tag{A1}$$

The assumption that $\varphi''(\cdot) < 0$ captures the the fact that the opportunities that these long assets represent are scarce and cannot be exploited limitlessly. To simplify the presentation we also assume that LRs always want to invest in this long asset, that is, that $\lim_{x \to 0} \varphi'(x) = +\infty$. The key assumption here though is that $\varphi'(\kappa) > 1$. Thus if LRs carry cash it must be to acquire, in some states of nature, assets with high expected returns. Given the assumption of risk neutrality this can only occur if asset purchases occur at cash-in-the-market prices. That is, as we will show more formally below, assets trade at prices that are below the expected payoff, for otherwise LRs would have no incentive to carry cash.

Our second assumption says that SRs would not invest in the risky asset in autarchy, though investment in it is more attractive than cash when the asset can be resold:

$$\rho\left[\lambda + (1 - \lambda)\eta\right] > 1 \qquad \text{and} \qquad \lambda\rho + (1 - \lambda)\left[\theta + (1 - \theta)\delta\right]\eta\rho < 1 \tag{A2}$$

Assumption A2 is needed to get the economically interesting situation where the liquidity of secondary markets at dates 1 and 2 affects asset allocation decisions at date 0. If instead we assumed that

$$\lambda\rho + (1 - \lambda)\left[\theta + (1 - \theta)\delta\right]\eta\rho \geq 1$$

then SRs would always choose to put all their funds in a risky asset irrespective of the liquidity of the secondary market at date 1.

10

Finally we assume that there are gains from trade, at least, in the immediate trading date, at $t = 1$. That is, $\varphi'(\kappa)$ is not as high as to rule out the possibility of the LRs carrying cash to trade at $t = 1$ altogether. As it will become clear below assumption A3 says that the agents' isoprofit lines cross in the "right way:"

$$\frac{\varphi'(\kappa) - \lambda}{(1 - \lambda)\eta\rho} < \frac{1 - \lambda}{1 - \lambda\rho} \tag{A3}$$

II.E Discussion

A central feature of the model is the timing of the aggregate and idiosyncratic shocks. The aggregate shock reveals news about the entire class of risky assets, whether they pay off, as it occurs in state $\omega_{1\rho}$, or instead whether there is a deterioration of the entire class as well as postponement of the actual payoff, as it occurs in state ω_{1L}. Indeed, the expected payoff of the risky assets in state ω_{1L} is $\eta\rho$, and the payoff may or not be realized at date $t = 2$ or $t = 3$. After $t = 1$, additional news accrue only to the holder of the asset and that is what generates the adverse selection premium. In state ω_{1L} all agents are informed about problems in a particular financial market, which results in a drop in prices. The key is that some of the holders of the risky asset are indeed affected and others are not, but they themselves only learn of which state they are in at $t = 2$. It is for this reason that if transactions take place in state ω_{1L} prices do not include an adverse selection premium whereas they do in state ω_d.

Once the initial portfolios are set, a critical trade-off the agents face in our framework is the decision of whether to liquidate and acquire the assets at ω_i, the immediate-trading state, or ω_d, the delayed-trading state. The parameter θ plays an important role in determining this decision. Indeed a high value of θ means that if the risky asset has not paid off at date $t = 1$, the probability that it does so at date $t = 2$ rather than at $t = 3$ is also high. This makes the risky asset attractive to the SRs for they care about consumption at date $t = 2$ more than they do about consumption at $t = 3$. But at the same time, the higher the value of θ, the more severe the adverse selection problems are at date $t = 2$ because the higher the θ the higher the probability that the asset acquired is a lemon. This translates into an adverse selection premium that discourages the SRs from carrying the asset to $t = 2$. Similarly for the LRs the trade-off is between acquiring high quality assets at date $t = 1$ at high prices or trade in a market subject to adverse selection but potentially at better prices. How these trade-offs affect the (ex-ante) portfolio decisions of both SRs and LRs is the central issue explored in this paper.

We finish this section with some comments about our particular assumptions on the contractual framework the agents face. First we assume that SRs cannot invest in the long asset and LRs cannot invest in the risky asset. Clearly, because $\lim_{x \to 0} \varphi'(x) = +\infty$, SRs would want to invest in the long asset at least a small amount, but recall that their marginal utility of consumption is given by δ, which we assume to be "small," in a sense we make precise below, and thus SRs would still be faced with a trade-off between the risky asset and cash. Also we assume that the LRs cannot invest in the risky asset, but this turns out to be less relevant for our purposes than it may appear at first. Indeed the risky asset is a constant returns to scale technology. If the returns to holding cash, what we refer to as outside liquidity, are below the expected return of the risky asset for the LRs, which is $\rho[\lambda + (1 - \lambda)\eta]$, then the LRs would not want to carry any cash and by assumption A2, the only resulting equilibrium would be one where the SRs would not invest in the risky asset. If instead the returns to holding cash are above the risky asset's expected return then clearly the LRs would not invest in it. Clearly the first situation is not an interesting one in which to analyze the determinants of outside versus inside liquidity and it is for this reason that we simply assume that the LRs cannot invest in the risky asset.[4]

Finally a word about the endogeneity of the outside capital that can potentially absorb the risky asset sales by part of SRs. In principle the entire supply of capital in the economy should stand ready to absorb these sales the moment returns are attractive enough, which realistically should yield small quantitative effects for the types of problems discussed in this paper. But we argue that, as the literature on limits to arbitrage postulates, sometimes the capital and the knowledge of the particular good opportunities in a market are unbundled, either because the capital of those with knowledge has been depleted due to some adverse shocks or simply because moral hazard and adverse selection problems are preventing capital from flowing to the market with the apparent attractive opportunities. Understanding the frictions that lead to this unbundling is essential to have a complete picture of financial markets and it is the subject of our current research. In this paper we simply take this amount of outside capital, κ, that can potentially absorb the SR's sales to be a parameter, rather than a variable.

---

[4]An issue we do not consider is the possibility that SRs and LRs enter into ex-ante contracts that could somehow improve the efficiency of the transactions. Our paper is concerned with the role of the market in providing liquidity to those facing any type of maturity missmatch. Clearly if we allowed SRs and LRs to enter into any type of contract at date $t = 0$ it is difficult to see what could be the residual role of markets. Still, understanding how asymmetric information problems may prevent complete ex-ante contracts so as to leave some role for market provided liquidity is one of the issues in our current research.

III. EQUILIBRIUM

Given that all SRs and LRs are ex-ante identical, we shall restrict attention to symmetric competitive equilibria. Recall that trade between SRs and LRs can only occur in spot markets at date 1 and/or 2, and there are only two states of nature where there are potentially strictly positive gains from trade, $\omega_i = \omega_{1L}$ and $\omega_d = \{\omega_{20}, \omega_{2L}\}$, where "$i$" and "$d$" stand for $i$mmediate and $d$elayed trading. Given that SRs have private information about realized returns on their risky asset at date 2, they can condition their trading policy on states ω_i, ω_{20} and ω_{2L}. We denote by $q(\omega_i)$ be the amount of the risky asset supplied by an SR in state ω_i and similarly for the other states. LR investors on the other hand are unable to distinguish states ω_{20} and ω_{2L}, and therefore can only condition their trades on states ω_i and ω_d. We write $Q(\omega_i)$ $(Q(\omega_d))$ for the amount of the risky asset that LR investors acquire in state ω_i (ω_d). In these equilibria we write prices as $P(\omega_i)$ and $P(\omega_d)$.

III.A The SR optimization problem

SRs must determine first how much of an investor's savings to invest in cash and how much in a risky asset. Second, they must decide how much of the risky asset to trade at respective prices $P(\omega_i)$ and $P(\omega_d)$. Their objective function is then

$$
\begin{aligned}
\pi\left[m, q\left(\omega_i\right), q\left(\omega_d\right)\right] = \ & m + \lambda\left(1-m\right)\rho \\
& + \left(1-\lambda\right)q\left(\omega_i\right)P\left(\omega_i\right) \\
& + \left(1-\lambda\right)\theta\eta\left[\left(1-m\right)-q\left(\omega_i\right)\right]\rho \\
& + \left(1-\lambda\right)\theta\left(1-\eta\right)\left[\left(1-m\right)-q\left(\omega_i\right)\right]P\left(\omega_d\right) \\
& + \left(1-\lambda\right)\left(1-\theta\right)q\left(\omega_{2L}\right)P\left(\omega_d\right) \\
& + \delta\left(1-\lambda\right)\left(1-\theta\right)\eta\left[\left(1-m\right)-q\left(\omega_i\right)-q\left(\omega_{2L}\right)\right]\rho
\end{aligned}
\tag{2}
$$

Notice that implicit in this objective function is the fact that SRs don't trade in states $\omega_{1\rho}$, $\omega_{2\rho}$ and $\omega_{3\rho}$. As we have noted above, given SRs' preferences and the LRs' objective function below there is actually no gain from trading assets in these states of nature.[5] In state $\omega_{1\rho}$, which occurs with probability λ, the risky asset pays in full at date 1 and SRs consume all the proceeds. In contrast, in state ω_{1L} the risky asset matures at a later date, and SRs may choose to sell an amount $q(\omega_i)$ of the risky asset for a unit price $P(\omega_i)$.

---

[5] Recall that the marginal utility of consumption at date $t = 3$ is $\delta \in (0, 1)$.

13

The risky asset then matures with ex-ante probability $(1 - \lambda)\theta\eta$ at date 2, in which case SRs consume the share of the proceeds of the asset it still owns: $[(1 - m) - q(\omega_i)]\rho$. Another outcome at date 2 is that the asset yields a zero return. This occurs with probability $(1 - \lambda)\theta(1 - \eta)$. In that state of nature the SR chooses optimally to sell its full position in the risky asset (which is now a lemon) for the price $P(\omega_d)$. Finally, with probability $(1 - \lambda)(1 - \theta)$ the asset only matures at date 3. The SR may then sell an additional amount of the risky asset $q(\omega_{2L})$ at unit price $P(\omega_d)$ rather than holding it to maturity at date 3, when the asset yields a return ρ with ex-ante probability $(1 - \lambda)(1 - \theta)\eta$. Given this the SRs program is:

$$\max_{m, q(\omega_i), q(\omega_{2L})} \pi[m, q(\omega_i), q(\omega_{2L})] \qquad (\mathcal{P}_{SR})$$

subject to

$$m \in [0, 1]$$

and

$$q(\omega_i) + q(\omega_{2L}) \leq 1 - m \quad \text{with} \quad q(\omega_i), \ q(\omega_{2L}) \geq 0$$

The constraints simply state that the SR cannot invest more in the risky asset than the funds at its disposal and that in states ω_i and ω_d it cannot sell more than what it holds.

III.B The LR optimization problem

LRs must first determine how much of their savings to hold in cash, M, and how much in long term assets, $\kappa - M$. They must then decide at dates 1 and 2 how much of the risky assets to purchase at prices $P(\omega_i)$ and $P(\omega_d)$. Recall that, given assumption 2 cash is costly to carry for LRs and thus they never carry more cash than they expect they will need to purchase risky assets from SRs at dates 1 and 2. In other words, in the states of nature where trade occurs LR investors completely exhaust their cash reserves to purchase the available supply of SR long assets. With this observation in mind we can write the payoff an LR investor that purchases $Q(\omega_i)$ in period 1 and $Q(\omega_d)$ in period 2, as follows:

$$
\begin{aligned}
\Pi[M, Q(\omega_i), Q(\omega_d)] \ = \ & M + \varphi(\kappa - M) \\
& + \ (1 - \lambda)[\eta\rho - P(\omega_i)]Q(\omega_i) \\
& + \ (1 - \lambda)E[\tilde{\rho}_3 - P(\omega_d)|\mathcal{F}]Q(\omega_d)
\end{aligned}
\qquad (3)
$$

The first term in the above expression is simply what the LR investor gets by holding an amount of cash M until date 3 without ever trading in secondary markets at dates 1 and 2.

The second term is the net return from acquiring a position $Q(\omega_i)$ in risky assets at unit price $P(\omega_i)$ in state ω_i. Indeed, the expected gross return of a risky asset in state ω_i is $\eta\rho$. The last term is the net return from trading in state ω_d. This net return depends on the payoff of the risky asset at date 3 and in particular on the quality of assets purchased at date 2. As we postulate rational expectations, the LR investor's information set, $\mathcal{F}$, will include the particular equilibrium that is being played. In computing conditional expectations the LRs assume that the mix of assets offered in state ω_d corresponds to the one observed in equilibrium.

We require a standard, and weak, rationality condition from LRs, that if they succeed in purchasing a unit in state ω_d in an equilibrium that prescribes no sales in state ω_d, and furthermore at a price for which SRs strictly prefer to hold the asset until date 3 to selling it in state ω_{2L}, then LR assumes that state ω_{20} (where SRs always prefer to sell) has occurred. In addition, LRs assume that SRs that weakly prefer to sell at price $P(\omega_d)$ will sell all their remaining holdings in state ω_d whereas SRs that prefer not to sell, will not sell any units.

The LR investor's program is thus:

$$\max_{M, Q(\omega_i), Q(\omega_d)} \Pi\left[M, Q(\omega_i), Q(\omega_d)\right] \qquad (\mathcal{P}_{LR})$$

subject to

$$0 \leq M \leq \kappa \qquad (4)$$

and

$$Q(\omega_i) P(\omega_i) + Q(\omega_d) P(\omega_d) \leq M \quad \text{with} \quad Q(\omega_i), Q(\omega_d) \geq 0 \qquad (5)$$

The first constraint (4) is simply the LR investor's wealth constraint: LRs' cannot carry more cash than their initial capital κ and they cannot borrow. The second constraint (5) says that LRs' cannot purchase more SR long-assets than their money, M, can buy. In our model M is, thus, the *supply of outside liquidity* by LRs.

III.C Definition of equilibrium

A (non-revealing) rational expectations competitive equilibrium is a vector of portfolio policies $[m^*, M^*]$, supply and demand choices $[q^*(\omega_i), q^*(\omega_{2L}), Q^*(\omega_i), Q^*(\omega_d)]$ and prices $[P^*(\omega_i), P^*(\omega_d)]$ such that (i) at these prices $[m^*, q^*(\omega_i), q^*(\omega_{2L})]$ solves $\mathcal{P}_{SR}$ and $[M^*, Q^*(\omega_i), Q^*(\omega_d)]$ solves $\mathcal{P}_{LR}$ and (ii) markets clear in all states of nature.

III.D Characterization of equilibria

III.D.1 Immediate and delayed-trading equilibria

The immediate-trading equilibrium. Under our stated assumptions we are able to establish first that there always exists an immediate trading equilibrium.

Proposition 1. (*The immediate-trading equilibrium*) Assume A1-A3 hold then there always exists an immediate trading equilibrium, where

$$M_i^* > 0, \ q^* (\omega_i) = Q^* (\omega_i) = 1 - m_i^* \qquad \text{and} \qquad q^* (\omega_{2L}) = Q^* (\omega_d) = 0.$$

In this equilibrium cash-in-the-market pricing obtains and

$$P_i^* (\omega_i) = \frac{M_i^*}{1 - m_i^*} \geq \frac{1 - \lambda \rho}{1 - \lambda}. \tag{6}$$

Moreover the cash positions m_i^* and M_i^* are unique.

To gain some intuition start by noticing that the immediate trading equilibrium has to meet the corresponding first order conditions for m and M, which are

$$P_i^* (\omega_i) \geq \frac{1 - \lambda \rho}{1 - \lambda} \qquad \text{and} \qquad \lambda + (1 - \lambda) \frac{\eta \rho}{P_i^* (\omega_d)} = \varphi' (\kappa - M_i^*), \tag{7}$$

when $m_i^* < 1$ and $M_i^* > 0$,[6] as it follows from immediate inspection of the maximization problems $\mathcal{P}_{SR}$, once we assume $q^*(\omega_i) = 1 - m_i^*$, and $\mathcal{P}_{LR}$, respectively. The question is how to determine $P_i^* (\omega_i)$. For this consider $P_i (\omega_i)$ to be the unique solution to:

$$\lambda + (1 - \lambda) \frac{\eta \rho}{P_i (\omega_i)} = \varphi' (\kappa - P_i (\omega_i)). \tag{8}$$

Assume first that the solution to (8) is such that

$$P_i (\omega_i) > \frac{1 - \lambda \rho}{1 - \lambda},$$

then we can set $m_i^* = 0$, and thus the SR is fully invested in the risky asset, and also set $P_i^* (\omega_i) = M_i^* = P_i (\omega)$, which by construction satisfies the LR's first order condition and because of A1 has to be such that $M_i^* < \kappa$. A key step in the construction of the immediate

---

[6]The proof of Proposition 1 establishes that assumption A3 rules out the possibility of a "no trade" immediate trading equilibrium in which $M_i^* = 0$ and $m_i^* = 1$.

trading equilibrium is that the price at date $t = 2$, $P_i^*(\omega_d)$, has to be such that both SRs and LRs have incentives to trade at ω_i and not ω_d. That is, it has to be that

$$P_i^*(\omega_i) \geq \theta\eta\rho + (1 - \theta\eta) P_i^*(\omega_d) \qquad \text{and} \qquad \frac{\eta\rho}{P_i^*(\omega_i)} \geq \frac{\mathbf{E}[\widetilde{\rho}_3|\mathcal{F}]}{P_i^*(\omega_d)}. \tag{9}$$

The first expression in (9) says that SRs prefer to sell assets at date 1 for a price $P_i^*(\omega_i)$ rather than carrying it to $t = 2$. If they do the latter, with probability $\theta\eta$ the risky asset pays off ρ. With probability $(1 - \theta\eta)$ instead they are in state $\omega_d = \{\omega_{2L}, \omega_{20}\}$ and the SRs can sell the asset for a price $P_i^*(\omega_d)$. If the price $P_i^*(\omega_d)$ is low enough,[7] then indeed the SRs would prefer to liquidate at date $t = 1$.

The expression on the right hand side of (9) says that the expected return of acquiring the asset in state ω_i is higher than in state ω_d. For this it is enough to set $P_i^*(\omega_d) < \delta\eta\rho$ for in this case agents in state ω_{2L} would prefer to carry the asset to $t = 3$ rather than selling it for that price. This leaves only "lemons" in the market. The LRs, anticipating this, set their expectations accordingly, $\mathbf{E}[\widetilde{\rho}_3|\mathcal{F}] = 0$, and thus for any strictly positive price $P_i^*(\omega_d)$ LRs prefer to acquire assets in state ω_i.

Assume now that instead the solution to (8) is such that

$$P_i(\omega_i) \leq \frac{1 - \lambda\rho}{1 - \lambda}, \tag{10}$$

then set $P_i^*(\omega_i)$ equal to the right hand side of (10). At this price, SRs are indifferent on the amount of cash carried. Then with this the solution to the LR's first order condition, see expression (7), is such that

$$M_i^* < \frac{1 - \lambda\rho}{1 - \lambda}.$$

Then it is enough to set $m_i^* \in [0, 1)$ such that

$$\frac{M_i^*}{1 - m_i^*} = \frac{1 - \lambda\rho}{1 - \lambda},^{8} \tag{11}$$

which can always be done. The choice of $P_i^*(\omega_d)$ is as above.

Notice that in our framework, and by assumption A1, cash in the market has to obtain and prices are lower than their discounted expected payoff, $P_i^*(\omega_i) < \eta\rho$, otherwise there would be no incentive for the LRs to carry cash.

---

[7]See expression (25) in the appendix for a precise upper bound on $P_i^*(\omega_d)$ that has to hold to provide incentives for the SRs to sell at date 1 rather than at $t = 2$.

[8]Notice that assumption A2 implies that $1 - \lambda\rho > 0$

The delayed-trading equilibrium. Proposition 2 establishes the existence of the delayed trading equilibrium.

Proposition 2.[9] (*The delayed-trading equilibrium*) Assume A1-A3 hold and that δ is small enough[10] then there always exists an delayed-trading equilibrium, where $m_d^* \in [0, 1)$, $M_d^* \in (0, \kappa)$,

$$q^* (\omega_i) = Q^* (\omega_i) = 0 \qquad \text{and} \qquad q^* (\omega_{2L}) = Q^* (\omega_d) = (1 - \theta\eta) (1 - m_d^*).$$

In this equilibrium cash-in-the-market pricing obtains and

$$P_d^* (\omega_d) = \frac{M_d^*}{(1 - \theta\eta) (1 - m_d^*)} \geq \frac{1 - \rho [\lambda + (1 - \lambda) \theta\eta]}{(1 - \lambda) (1 - \theta\eta)}. \tag{12}$$

Moreover the cash positions m_d^* and M_d^* are unique.

The intuition of the construction of the delayed-trading equilibrium is very similar to the immediate-trading one but there are still some differences that we emphasize next. First as stated in the proposition δ needs to be small enough. Otherwise SRs in state ω_{2L} prefer to carry the asset to date 3 rather than selling it at $t = 2$, which would destroy the delayed-trading equilibrium as it would only leave lemons in the market. We postpone a discussion of the existence problems that arise when δ is relatively high until later in the paper.

Second, a key difference with immediate trading is that the supply of risky assets by SRs is reduced under delayed trading by an amount $\theta\eta$, which is the proportion of risky assets that pay off at date $t = 2$.[11] Cash-in-the-market pricing is now:

$$P_d^* (\omega_d) = \frac{M_d^*}{(1 - \theta\eta) (1 - m_d^*)}.$$

Notice that now the mass of risky assets supplied in the market in state ω_d is given by $(1 - \theta\eta) (1 - m_d^*)$. Thus delaying asset liquidation introduces an adverse selection effect, which depresses prices, and a lower supply of the risky asset, which, other things equal, increases prices.

---

[9]There may also be a third equilibrium, which involves positive asset trading at both dates 1 and 2. We do not focus on this equilibrium as it is unstable.

[10]The proof of the proposition clarifies the upper bound on δ that guarantees existence, see expression (36) in the Appendix.

[11]This is one of the key differences that arises when the shocks at date $t = 2$ are aggregate rather than idiosyncratic. In this case the supply of risky assets is always the same. The difference is that there is one aggregate state of nature, $\omega_{1\rho}$ where there is no mare for the risky asset at date $t = 2$.

As before supporting a delayed-trading equilibrium requires that both SRs and LRs have incentives to trade at $t = 2$ rather than at $t = 1$, which means

$$P_d^* (\omega_i) \leq \theta\eta\rho + (1 - \theta\eta) P_d^* (\omega_d) \qquad \text{and} \qquad \frac{\eta\rho}{P_d^* (\omega_i)} \leq \frac{\mathbf{E} \left[\widetilde{\rho}_3|\mathcal{F}\right]}{P_d^* (\omega_d)}, \tag{13}$$

where now the expected payoff is given by

$$\mathbf{E} \left[\widetilde{\rho}_3|\mathcal{F}\right] = \frac{(1 - \theta)\eta\rho}{1 - \theta\eta}.$$

If (13) is to be met, the price in ω_i has to be in the interval,

$$P_d^* (\omega_i) \in \left[\frac{1 - \theta\eta}{1 - \theta} P_d^* (\omega_d), \theta\eta\rho + (1 - \theta\eta)P_d^* (\omega_d)\right].$$

The key step of the proof of Proposition 2 is that this interval is non empty.

It is perhaps worth emphasizing that the delayed trading equilibrium collapses to the immediate trading equilibrium when $\theta = 0$. Indeed notice, for instance, that the lower bound in the price $P_d^* (\omega_d)$ in (12) reduces to the lower bound in (6) for $P_i^* (\omega_i)$. The only difference between dates 1 and 2 is precisely the occurrence of an idiosyncratic shock that reveals to the SRs, the holders of the risky asset, its true value. When $\theta = 0$ the idiosyncratic signal does not occur and thus all the information is common knowledge as it is always the case at $t = 1$. This feature of our model will play an important role in what follows.

Before we close this section we introduce the example that we use throughout to illustrate the results. Because as already explained, the parameter θ plays a critical role in our analysis it is the focus of the comparative statics below and thus we parameterize the set of economies that we considered throughout by the different values that it takes. In light of assumption A2 it is then convenient to define $\overline{\theta}$ as the value for which

$$1 = \rho \left[\lambda + (1 - \lambda)\eta\rho \left(\overline{\theta} + (1 - \overline{\theta})\delta\right)\right], \tag{14}$$

for a given λ, δ, η, and ρ.

Example. In this example the parameter values are:

$$\lambda = .85 \qquad \eta = .4 \qquad \rho = 1.13 \qquad \kappa = .2 \qquad \delta = .1920 \qquad \varphi\left(x\right) = x^\gamma \quad \text{with} \quad \gamma = .4$$

Having fixed the value of δ, we need to restrict the values of our only free parameter θ to $\theta \leq \overline{\theta} = .4834$ so as to ensure that assumption A2 holds. It is immediate to check that in this example assumptions A1-A3 hod, as well as assumption A4 below. Figures and numbers throughout the paper refer to this example. A summary of the main features follows.

- Both the immediate and delayed trading equilibria exist for $\theta \in [0, .4196)$ and moreover in the delayed trading equilibrium we have $m_d^* > 0$.

- For $\theta \in [.4196, .4628]$ both equilibria exist and the delayed trading equilibrium is such that $m_d^* = 0$.

- For $\theta \in (.4628, .4834]$ the delayed trading equilibrium fails to exists. It is for this range that the assumption of δ being "small enough" fails to hold and we elaborate on this case below. □

III.D.2 Inside and and outside liquidity in the immediate and delayed trading equilibria

How does the composition of inside versus outside liquidity vary across equilibria? To build some initial intuition on this question it is useful to illustrate the immediate and delayed-trading equilibrium that obtain in our example when $\theta = .35$. Figure 2 represents the immediate and delayed-trading in a diagram where the x-axis measures the amount of cash carried by LRs, M, and the y-axis the amount of cash carried by the SRs, m. The dashed lines are the *isoprofit curves* of the LRs and the straight (continuous) lines are the SR *isoprofit lines*.[12] In the figure we have both the isoprofit lines for both the immediate and delayed exchange, which of course, correspond to exchange of different assets as the one at $t = 2$ has different payoffs than the one at $t = 1$.[13] It is for this reason that isoprofit lines appear to cross in the plot: They simply correspond to different dates. Equilibria are located at the tangency points between the isoprofit lines of the SRs and LRs.

Start with the immediate trading equilibrium, located at the point marked $(M_i^*, m_i^*) = (.0169, .9358))$. There are two isoprofit lines going through that point; the straight line corresponds to the SR and the dashed-dotted line corresponds to the LR's isoprofit line. In fact the straight line corresponds to the SR's reservation utility, $\overline{\pi} = 1$. Thus whatever gains from trade there are in the immediate trading equilibrium they accrue to the LRs. Turn next to the delayed-trading equilibrium, which is marked $(M_d^*, m_d^*) = (.0540, 4860)$ and features a mix of outside versus inside liquidity that is tilted towards the former relative to the latter when compared to the immediate-trading equilibrium. Also, observe that the SR's isoprofit line has shifted down, reflecting the fact that the perceived "quality" of SR assets in state ω_d is lower

---

[12]Assumption A3 simply says that the slope of the isoprofit lines at $M = 0$ in the immediate-trading date are such that there ganis from trade: The LR's isoprofit line is "flatter" than the SR's isoprofit line.

[13]Figure 2 is simply the result of superimposing two Edgeworth Boxes, the one corresponding to the immediate exchange and the one corresponding to the delayed exchange.

than in state ω_i due to adverse selection, so that one should expect that the SRs would have to settle for a lower price in that state. The SR's isoprofit line remains that associated with it's reservation value.

One way of understanding the portfolio choices in the immediate-trading equilibrium is that the risky asset is of high quality in state ω_i, so that SRs must be compensated with a high price relative to the price in state ω_d, which also includes an *adverse selection discount*, to be willing to sell the asset at that point. This observation is reflected in the slope of the isoprofit lines in Figure 2: The SRs' isoprofit line in the immediate trading equilibrium is *flatter* suggesting that SRs require a higher price per unit of risky asset sold at that date. But this higher price can only come at the expense of lower returns to holding cash for LRs. The latter are thus induced to cut back on their cash holdings. This, in turn, makes it less attractive for SRs to invest in the risky asset, and so on. The outcome is that in the immediate trading equilibrium most of the liquidity is *inside liquidity* held by SRs, whereas the delayed-trading equilibrium features relatively more *outside liquidity* than inside liquidity.

The next proposition formalizes this discussion, specifically, it characterizes the mix of inside versus outside liquidity across the two types of equilibria. For this we make one additional assumption that allows for a particularly clean characterization of the aforementioned mix,

$$\frac{1 - \lambda\rho}{1 - \lambda} > \kappa \tag{A4}$$

As the Result in the Appendix shows under assumption A4 the immediate-trading equilibrium is such that $m_i^* \in (0, 1)$, that is the SRs is carrying a strictly positive amount of cash. Roughly, we need to guarantee that $m_i^* > 0$ in order to obtain non trivial cash allocation decisions for the SRs, which otherwise would be equal to 0 for both the immediate and the delayed-trading equilibria, as will become clear in Proposition 4. The present paper is concerned with the ex-ante efficiency costs associated with portfolio choices that result in the particular timing of the liquidation decisions and thus the most economically interesting case is the one where the economy is not "at a corner," that is $m_i^* = 0$, at the immediate-trading date. Armed with this new assumption we can prove the following

Proposition 3. (*Inside and outside liquidity across equilibria.*) Assume that A1-A4 hold and that δ is small enough so that a delayed trading equilibrium exists for all $\theta \in (0, \overline{\theta}]$ then there exists a $\theta' \in \left(0, \overline{\theta}\right]$ such that $m_i^* > m_d^*$ and $M_i^* < M_d^*$ for all $\theta \in (0, \theta')$.

Thus for the range $\theta \in [0, \theta']$ the delayed-trading equilibrium features more outside

liquidity and less inside liquidity than the immediate-trading equilibrium. In our example $\theta' = \overline{\theta}$ so that Proposition 3 holds for the entire range of admissible θs.[14]

We close this section by making two additional comments. First, note that all equilibria are *interim efficient*. That is, conditional on trade occurring in either states ω_i or ω_d, there is no additional reallocation of the risky asset that would make both sides better off. As can be seen immediately in Figure 2, it is not possible to improve the ex-post efficiency of either equilibrium, as in each case the equilibrium allocation is located at the tangency point of the isoprofit curves. As we shall further explore below, in our model inefficiencies arise through distortions in the *ex-ante* portfolio decisions of SRs and LRs and through the particular timing of liquidity trades they give rise to. When agents anticipate trade in state ω_i, SRs lower their investment in the risky asset and carry more *inside liquidity* m_i. In contrast LRs, carry less liquidity M_i as they anticipate fewer units of the risky asset to be supplied in state ω_i.

A second observation is that A4, which implies that the immediate trading equilibrium is such that $m_i^* > 0$, does not necessarily imply that $m_d^* > 0$. Indeed, Figure 3 shows the immediate and delayed-trading equilibrium when θ is increased from $\theta = .35$, as it was the case in Figure 2, to $\theta = .45$ and thus the adverse selection problem is relatively worse than in the previous case. The delayed-trading equilibrium is located in $(M_d^*, m_d^*) = (.0716, 0)$, the immediate-trading equilibrium being unaffected as it is independent of θ. Clearly the equilibrium is ex-post efficient, but now, unlike in the case considered in Figure 2, gains from trade do not solely accrue to the LRs but also to the SRs. In Figure 3 the isoprofit line marked IP_{SR} corresponds to the profit level $\overline{\pi} = 1$ for the SR, which is the same as under autarky. The isoprofit line through the delayed trading equilibrium lies strictly to the right of IP_{SR}, which implies that FIs now command strictly positive profits. The reason is that at the corner when $m = 0$, FIs are at full capacity in supplying the risky asset in state ω_d. They may then earn *scarcity rents*, as LRs compete for the limited supply of the risky asset supplied by the SRs by increasing their bids for these assets.

III.D.3 Adverse selection and the delayed trading equilibrium

We now turn to the relevant comparative statics in our analysis, namely how changes in the adverse selection problem SRs face in state ω_d, as measured by changes in θ, affect equilibrium outcomes. In particular, we are interested in understanding how equilibrium cash

---

[14]In fact though we have been unable to prove it formally, we have not found an example of an economy that meets assumptions A1-A4 for which $\theta' < \overline{\theta}$.

holdings and equilibrium prices vary with θ.

Several important effects are at work as θ changes, some of which we have already mentioned. First, the incentives of both SRs and LRs to hold cash are affected by changes in θ. In addition, SRs' incentives to hold onto their asset position until date 2 (when the risky asset does not mature at date 1) are affected. As θ rises the risky asset is more likely to mature at date 2 and thus becomes more attractive to SRs. Other things equal, SRs are then both more likely to invest in the risky asset and to carry the asset from date 1 to date 2.

However, as θ rises the adverse selection problem in state ω_d is worsened and therefore equilibrium prices $P_d^*(\omega_d)$ are likely to be lower. These lower prices that SRs face in state ω_d in turn reduce their incentives to invest in the risky asset and to carry it to date 2. An additional effect that complicates the analysis is that as θ increases the supply of the risky asset in state ω_d,

$$s_d^*(\omega_d) = (1 - m_d^*(\theta))(1 - \theta\eta) \tag{15}$$

diminishes on account of the fact that a larger share of the available risky assets pay off and thus are not liquidated.

We are interested also in the expected return on acquiring the risky asset at date $t = 2$ in the delayed-trading equilibrium, which is defined as

$$R_d^*(\omega_d) = \frac{(1 - \theta)\eta\rho}{(1 - \theta\eta)P_d^*(\omega_d)} \tag{16}$$

The next proposition establishes how these countervailing effects net out and how M_d^*, m_d^*, $s_d^*(\omega_d)$, and $R_d^*(\omega_d)$ vary with θ. Throughout of course we assume that $\theta \leq \bar{\theta}$.

Proposition 4. (*Comparative statics.*) Assume that A1-A4 hold and that δ is small enough for all $\theta \in [0, \bar{\theta}]$ so that a delayed trading equilibrium always exists, then there exists a unique $\widehat{\theta} \in [0, \bar{\theta}]$, possibly $\widehat{\theta} = \bar{\theta}$, such that:

1. The SR's cash position: (a) m_d^* is a (weakly) decreasing function of θ, (b) $m_d^* > 0$ for all $\theta \in [0, \widehat{\theta})$ $m_d^* = 0$ for all $\theta \in [\widehat{\theta}, \bar{\theta}]$ and (c) s_d^* is a strictly increasing function of θ for $\theta \in [0, \widehat{\theta})$ and a strictly decreasing function of θ for $\theta \in [\widehat{\theta}, \bar{\theta}]$.

2. The LR's cash position: M_d^* is a strictly increasing function of θ for $\theta \in [0, \widehat{\theta})$ and a strictly decreasing function of θ for $\theta \in (\widehat{\theta}, \bar{\theta}]$.

3. Expected returns at $t = 2$: R_d^* is an increasing function of θ for $\theta \in [0, \widehat{\theta})$ and a decreasing function of θ for $\theta \in (\widehat{\theta}, \bar{\theta}]$.

23

We illustrate the comparative statics described in Proposition 3 in our example, for which, given our parametric assumption, it can be shown that $\widehat{\theta} = .4196$. Figures 4 and 5 exhibit the comparative statics with respect to θ for the cash positions, m_d^* and M_d^*, and the expected return and the price of the risky asset at ω_d ($R_d^*(\omega_d)$ and $P_d^*(\omega_d)$), respectively.

Consider first Figure 4. As we would expect, based on our discussion above, the amount of cash carried by the SR is a decreasing function of θ, and $m_d^* = 0$ for $\theta \geq \widehat{\theta} = .4196$. It is less obvious how the amount of cash carried by LR investors varies with θ. Consider first the case where $\theta \leq \widehat{\theta}$. The amount of cash carried by LR investors is then an increasing function of θ. This is surprising: the more severe the adverse selection problem the more cash LR investors bring to state ω_d. What is the logic behind this result?

In this range of θs there are actually two effects at work. First, an increase in θ does indeed worsen the adverse selection problem and would other things equal result in LRs reducing their supply of liquidity. But there is a countervailing effect, which is that an increase in θ also results in a higher investment in the risky asset by the SRs. Indeed as shown in 1-(c) in the proposition, s_d^* is an increasing function of θ in this range. It is this higher supply of the risky asset that in turn increases the supply of outside liquidity. The latter effect dominates and thus results in an increasing M_d^* as a function of θ when $\theta \leq \widehat{\theta}$. Instead, when $\theta > \widehat{\theta}$ the supply effect gets reversed and s_d^* is a decreasing function of θ. Both the supply side and the adverse selection effect decrease the incentives of the LRs to carry cash and it is for this reason that M_d^* is now a decreasing function of θ.

Figure 5 illustrates how the price $P_d^*(\omega_d)$ changes with θ. As can be seen, $P_d^*(\omega_d)$ is a decreasing function of θ. But note that the decline is more pronounced when $\theta < \widehat{\theta}$. The reason has already been mentioned. As long as $\theta < \widehat{\theta}$ an increase in θ has a double effect. A higher θ worsens adverse selection concerns and thus the drop in prices. In addition, a higher θ increases investment in the risky asset, which gets liquidated in state ω_d. This supply effect produces a further decline in prices that is absent when $\theta \geq \widehat{\theta}$ for then $m_d^* = 0$ and there can be no further investment in the risky asset. Notice that for $\theta > \widehat{\theta}$ prices keep dropping but a lower rate for now the supply is decreasing and thus the competition for the risky asset amongst the LRs dampens the adverse selection effect on prices.

The pattern of returns is also revealing about the incentives of the LRs to carry outside liquidity to the delayed-trading equilibrium. For $\theta < \widehat{\theta}$, $R_d^*(\omega_d)$ is an increasing function of θ: The expected payoff of the risky asset in the delayed-trading date

$$\mathbf{E}\left[\widetilde{\rho}_3 | \mathcal{F}\right] = \frac{(1-\theta)\,\eta\rho}{(1-\theta\eta)}, \tag{17}$$

24

is a decreasing function of θ. But the price is dropping *faster* than returns on account both of the adverse selection effect and the supply effect. This produces the increasing pattern in returns. It is for this reason that the incentives of the LRs to carry outside liquidity are also increasing in θ. Instead when $\theta > \widehat{\theta}$, the expected payoff is still decreasing but the decrease in supply makes for a slow drop in prices as a function of θ, as we just saw, and thus the negative slope of $R_d^*(\omega_d)$ as a function of θ in this range.

In conclusion then, for $\theta \in [0, \widehat{\theta}]$ the more severe the adverse selection, as measured by θ, the higher the amount of outside liquidity brought to the market and the lower the amount of inside liquidity carried by those holding the risky asset. This counterintuitive result springs from the fact that the drop in prices makes the risky asset progressively more attractive to the LRs in state ω_d. Roughly, for the LRs the more pronounced the liquidity correction at $t = 2$ the more attractive it is to carry cash to trade opportunistically. Armed with these insights we turn to the question of the Pareto ranking of the two equilibria.

III.D.4 Pareto ranking of the immediate and delayed-trading equilibria

Given these differences in *ex-ante* portfolio allocations an obvious question is whether there a clear ranking of the two equilibria in terms of Pareto efficiency when they coexist? Interestingly, the answer to this question is *yes* and, somewhat surprisingly, it is the delayed trading equilibrium that *Pareto dominates* the immediate trading equilibrium. This is surprising, as delayed trade is hampered by the information asymmetry that arises in state ω_d, and therefore will take place at lower equilibrium prices.

Proposition 5. (*Pareto ranking of equilibria.*) Assume that A1-A4 hold and that δ is small enough for all $\theta \in [0, \overline{\theta}]$ so that a delayed trading equilibrium always exists, then then there exists a $\theta' \in (0, \overline{\theta}]$ such that $\pi_i^* \leq \pi_d^*$ and $\Pi_i^* < \Pi_d^*$ for all $\theta \in (0, \theta')$.

In our example $\theta' = \overline{\theta}$ and though we have not been able to prove a tighter characterization of Proposition 5, we have been unable to find an example that meeting assumptions A1-A4, features $\theta' < \overline{\theta}$. Thus in our example the delayed-trading equilibrium Pareto dominates the immediate-trading equilibrium for all $\theta \in (0, \overline{\theta}]$. This is illustrated in Figure 6, where the expected profits of both the SRs and LRs are plotted for a particular range of θs.[15]

Figure 6 shows the expected profits of SRs and LRs as a function of θ for the delayed trading equilibrium. The top panel shows the SRs' expected profits. Notice that for $\theta \leq \widehat{\theta} =$

---

[15]The starting $\theta = .35$ is simply chosen to show the figures in a convenient scale.

.4196 the SRs are left at their reservation profits, which obtain if they were to be fully invested in cash. Indeed, the SRs' risky asset is a constant returns to scale technology and, a shown in Proposition 4, in this range they are not fully invested in the risky asset. Figure 2 offered a preview of this result. In that particular case $\theta = .35 < \widehat{\theta}$ and thus the delayed-trading equilibrium was located at the tangency point of the SR's isoprofit line which corresponds to its reservation value of $\overline{\pi} = 1$ and the LR's isoprofit line. The lower panel shows the LR's expected profit. The flat line corresponds to the LR's expected profit in the immediate-trading equilibrium, which is everywhere below the expected profit in the delayed-trading equilibrium. What may be at first surprising is that the LR's expected profits are, in this range, an increasing function of θ: The higher the adverse selection the higher the LR's expected profit. This again is due to the fact that as we increase θ the asset is more likely to pay in date $t = 2$, when SRs care the most for payoffs, and thus it becomes more attractive to them. This leads them to invest more in the risky asset and carry less inside liquidity, which translates into more goods for the LR in the event that the market opens at date $t = 2$. The liquidity premium associated with the adverse selection problem combined with the increased supply of assets translates, as we saw in Proposition 4, into an improvement of the investment opportunities available to the LRs at the interim stage, which can only make them better off.

For $\theta > \widehat{\theta}$ the SRs are fully invested in the risky asset and because of this fact they now acquire some rents. Indeed for this range $\pi_d^* > 1$ and increasing with θ, whereas for the LRs the expected profits are a decreasing function. Notice though that $\Pi_d^* > \Pi_i^*$ throughout. A particular example was depicted in Figure 3, where an example was shown where $\theta = .45 > \widehat{\theta}$. As could be seen there, the delayed trading equilibrium was located strictly in the interior of the lens formed by the two reservation isoprofit lines.

In Section IV below we explore the efficiency of equilibria more broadly and compare equilibrium allocations to the allocations chosen by a planner at date 0 who seeks to maximize a weighted sum of LR and SR payoffs subject to participation and incentive compatibility constraints. Consistent with our observation on Pareto ranked equilibria, the planner's optimal allocation is to enforce trade of the risky asset at date 2 in state ω_d. The planner's overall objective is to maximize total surplus and total gains from trade. This requires both limiting total cash reserves at date 0 and the trading of assets at date 2 in state ω_d. Interestingly, the planner's optimal choice of cash holdings is generally strictly smaller than in either of our two equilibria.

III.E Non-existence of the delayed-trading equilibrium and "commitment"

A maintained assumption throughout our analysis in section III.D is that the δ is small enough so as not to compromise the existence of the delayed trading equilibrium. Here we explore more deeply this assumption. Indeed, an important feature of our model is precisely the fact that the SRs in state ω_{2L} may prefer to carry the asset to date $t = 3$ rather than trading it for $P_d^*(\omega_d)$ at $t = 2$.[16] When this happens the delayed-trading exchange cannot be supported as a competitive equilibrium for only lemons would appear in the market. SRs in state ω_{2L} would have an incentive to retain the asset and carry it to $t = 3$ whenever

$$\widetilde{P}_d(\omega_d) < \delta\eta\rho, \tag{18}$$

where $\widetilde{P}_d(\omega_d)$ is the candidate price for the risky asset at $t = 2$ constructed as in Proposition 2. In our example this occurs for the range economies for which

$$\theta \in (.4628, .4834)$$

To illustrate graphically the welfare costs associated with this lack of existence, Figure 7 plots the expected profits for the SRs and the LRs as a function of θ where we have selected the Pareto superior equilibrium whenever there are two of them. There are three regions in the plot. The first two correspond to the cases already discussed. In region A, $\theta \in (0, .4196)$ the delayed trading equilibrium is Pareto superior and is such that $m_d^{ast} > 0$. Region B, $\theta \in [.4196, .4628)$, also features the delayed trading equilibrium as the Pareto superior one and in that range of θs, $m_d^* = 0$. Region C is one where, though assumptions A1-A4 are met, a delayed-trading equilibrium cannot be supported and thus the immediate trading equilibrium gets selected.

The dashed line in both panels of Figure 7 shows the additional expected profits that would accrue to SRs and LRs if the former could commit ex-ante to liquidate their assets at the candidate price $\widetilde{P}_d(\omega_d)$ in state ω_{2L}. In this case, the LRs anticipating that the pool of assets supplied in ω_d would also include assets of higher quality would be willing to bring more outside liquidity than the amount they are willing to bring in the immediate trading equilibrium. As shown above, this is always Pareto improving in our framework because it substitutes inside liquidity with outside one.

Clearly it is difficult to design such a contingent commitment technology but we elaborate on some policy implications below, were we also discuss how a monopolistic supplier of outside

---

[16]It is easy to check that A2 implies that it is never optimal to retain the asset in an immediate trading equilibrium constructed as above.

liquidity can also improve efficiency by internalizing the effect that its portfolio policy has on the prices of the risky asset, and thus on the quality of the pool, at the different dates.

IV. WELFARE OPTIMUM

We shall take the planner's objective to be the maximization of a weighted average of investor and LR payoffs, where the weights are $\gamma \geq \frac{1}{2}$ for short-run investors (SRs) and $(1-\gamma)$ for LRs. Moreover, since all agents are risk neutral and since LR and SR investors have the same unit mass, we can reduce the planner's problem to the maximization of the weighted average payoffs of a pair of individual representative LR and SR investors.

The planner's choice variables at date 0 are the investment allocations (m, M). And at dates $1, 2$ and 3 the planner's controls are consumption allocations to the two representative SR and LR investors in the seven different states of nature, implemented through asset allocations and cash transfers. It is helpful to begin our analysis of the planner's problem by imposing only resource constraints on the planner. We shall refer to this as the unconstrained welfare optimum.

IV.A Unconstrained Welfare Optimum

The resource constraints the planner faces are, first at date 0, $m \in [0, 1]$ and $M \in [0, \kappa]$. In other words, all the planner can do at date 0 is pick the SR and LR portfolio weights between cash and the respective risky assets. Second, at dates $1, 2$ and 3 the planner can enforce state-contingent asset allocations $\phi(\omega)$, where $\phi(\omega) \in [0, 1]$ denotes the share of risky assets given to SR investors, and cash transfers $T(\omega)$ from LR to SR, that, however, cannot exceed the total cash-holdings of LR (if $T(\omega)$ is positive) or SR (if $T(\omega)$ is negative).

For a given allocation $\{m, M, \phi(\omega), T(\omega)\}$ SR and LR respective utilities in each state are then as follows. When risky assets mature early, they get respectively

$$m + (1-m)\rho + T(\omega_{1\rho}) \qquad \text{and} \qquad M + \varphi(\kappa - M) - T(\omega_{1\rho}),$$

where $T(\omega_{1\rho}) \in [-(m + (1-m)\rho), M]$. Note that we have suppressed asset allocations $\phi(\omega_{1\rho})$, without loss of generality, as any asset allocation can be undone through a transfer $T(\omega_{1\rho})$.

Similarly, when an SR's risky asset matures at date 2, and the realized state of nature is $\omega_{2\rho}$, they get

$$m + (1-m)\rho + T(\omega_{2\rho}) \text{ and } M + \varphi(\kappa - M) - T(\omega_{2\rho}).^{17}$$

And when the realized state is ω_{20} they only get:

$$m + T(\omega_{20}) \text{ and } M + \varphi(\kappa - M) - T(\omega_{20}).$$

Consider next the interim state ω_{1L}, where the planner learns that a risky asset matures at some subsequent date, but does not yet know the realized return. In this state the planner can specify asset allocations ϕ and transfers T, but he cannot perfectly determine final consumption allocations. Observe also that the planner can subsequently "undo" the allocations $\phi(\omega_{1L})$ and transfer $T(\omega_{1L})$ in states $\omega_{2\rho}, \omega_{20}$ and ω_{2L}. Thus, without loss of generality we can set $\phi(\omega_{1L}) = 1$ and $T(\omega_{1L}) = 0$.

In contrast, T set in state ω_{2L} cannot be undone in period 3. Therefore, when SR risky assets only mature at date 3 SR and LR respective utilities are now partly determined by the transfer $T(\omega_{2L})$. Thus, in state $\omega_{3\rho}$ SR and LR respectively get:

$$m + T(\omega_{2L}) + \delta[\phi(\omega_{3\rho})(1 - m)\rho + T(\omega_{3\rho})]$$

and

$$M - T(\omega_{2L}) + \varphi(\kappa - M) + (1 - \phi(\omega_{3\rho}))(1 - m)\rho - T(\omega_{3\rho}).$$

(Note that SR can always consume its cash holdings $m + T(\omega_{2L})$ at date 2. As, for date 3 returns on a risky asset and transfers $T(\omega_{3\rho})$, those can only be consumed at date 3 and therefore only provide marginal utility δ). Similarly, when the realized state is ω_{30} the respective utilities are:

$$m + T(\omega_{2L}) + \delta T(\omega_{30}) \text{ and } M - T(\omega_{2L}) + \varphi(\kappa - M) - T(\omega_{30}).$$

Combining these terms, we therefore obtain the following expression for the expected date 0 SR payoff, which we denote by $V_{SR}(m, M)$:

$$V_{SR}(m, M) = \lambda(m + (1 - m)\rho + T(\omega_{1\rho})) +$$

$$(1 - \lambda)\{\theta[\eta(m + (1 - m)\rho + T(\omega_{2\rho})) + (1 - \eta)(m + T(\omega_{20}))] +$$

$$(1 - \theta)[m + T(\omega_{2L}) + \eta\delta(\phi(\omega_{3\rho})(1 - m)\rho + T(\omega_{3\rho})) + (1 - \eta)\delta T(\omega_{30})]\}$$

Similarly, the expression for $V_{LR}(m, M)$, the expected long run investor payoff at date 0 is:

$$V_{LR}(m, M) = \lambda(M + \varphi(\kappa - M) - T(\omega_{1\rho})) +$$

$$(1 - \lambda)[M + \varphi(\kappa - M) - \theta(\eta T(\omega_{2\rho}) + (1 - \eta)T(\omega_{20})) +$$

$$(1 - \theta)(\eta(1 - \phi(\omega_{3\rho}))(1 - m)\rho - T(\omega_{2L}) - (\eta T(\omega_{3\rho}) + (1 - \eta)T(\omega_{30})]$$

Thus, the planner's optimization problem (subject to meeting resource constraints) reduces to:

$$\max_{m; M; \phi(\omega); T(\omega)} \{\gamma V_{SR}(m, M) + (1 - \gamma)V_{LR}(m, M)\}$$

(where we restrict the welfare weights to be such that $\gamma \geq \frac{1}{2}$).

This problem is linear in all the controls $\phi(\omega)$ and $T(\omega)$, and since $\gamma \geq \frac{1}{2}$ we immediately obtain that $T^*(\omega_{1\rho}) = M$, $T^*(\omega_{2\rho}) = M$, $T^*(\omega_{20}) = M$, and $T^*(\omega_{2L}) = M$. As for $\phi(\omega_{3\rho})$ and $T(\omega_{3\rho})$, we obtain that:

$$\phi^*(\omega_{3\rho}) = 0 \text{ and } T^*(\omega_{3\rho}) = T^*(\omega_{30}) = 0 \text{ if and only if } (1 - \gamma) \geq \delta\gamma.$$

Henceforth we shall always assume that:

$$(1 - \gamma) \geq \delta\gamma \tag{A5}$$

When assumption A5 holds the *outside* value of risky assets held by SRs is always higher than the *inside* value. Under this assumption, the unconstrained welfare optimum for any given portfolio allocation (m, M) is given by:

$$W(m, M) = \gamma[m + M + (\lambda + (1 - \lambda)\theta\eta)(1 - m)\rho] +$$

$$(1 - \gamma)[\varphi(\kappa - M) + (1 - \lambda)(1 - \theta)\eta(1 - m)\rho]$$

Differentiating $W(m, M)$ with respect to m and M we obtain:

$$W_m = \gamma[1 - (\lambda + (1 - \lambda)\theta\eta)\rho] - (1 - \gamma)(1 - \lambda)(1 - \theta)\eta\rho,$$

and

$$W_M = \gamma - (1 - \gamma)\varphi'(\kappa - M).$$

Next, consider the socially optimal asset allocations m^* and M^* at the two extremes for the range of welfare weights, $\gamma \longrightarrow \frac{1}{2}$–when the planner puts equal weight on SR and LR payoffs–and $\gamma \longrightarrow \frac{1}{1+\delta}$–when the planner puts almost all the weight on SR payoffs.

1. When $\gamma \longrightarrow \frac{1}{2}$ we have:

$$W_m \longrightarrow \frac{1}{2}[1 - (\lambda + (1-\lambda)\theta\eta)\rho - (1-\lambda)(1-\theta)\eta\rho]$$

$$= \frac{1}{2}[1 - (\lambda + (1-\lambda)\eta)\rho],$$

and

$$W_M \longrightarrow \frac{1}{2}[1 - \varphi'(\kappa - M)].$$

By assumption A2, $1 < (\lambda + (1-\lambda)\eta)\rho$ so that the unconstrained welfare optimum is $m^* = 0$ when $\gamma \longrightarrow \frac{1}{2}$. In words, the unconstrained social optimum for γ close to $\frac{1}{2}$ is to have financial intermediaries invest only in risky assets. As for long run investors, their unconstrained socially optimal asset allocation turns out to be the same as the long run investors' preferred asset composition under *autarky*, $M^* = 0$. Indeed $W_M = -\varphi'(\kappa - M) + 1 < 0$ for all $M \geq 0$ by assumption, so that the socially optimal long run investor cash-holding is 0.

This unconstrained welfare optimum always satisfies the LR participation constraint as $0 < (1-\lambda)(1-\theta)\eta\rho$. But it does not satisfy SR's participation constraint as $(\lambda + (1-\lambda)\theta\eta)\rho < 1$ by assumption A2.

2. When $\gamma \longrightarrow \frac{1}{1+\delta}$ we have:

$$W_m \longrightarrow \frac{1}{1+\delta}[1 - (\lambda + (1-\lambda)\eta - (1-\delta)(1-\lambda)(1-\theta)\eta)\rho],$$

and

$$W_M \longrightarrow \frac{1}{1+\delta} - \frac{\delta}{1+\delta}\varphi'(\kappa - M).$$

Thus, as long as θ is large enough that

$$1 - (\lambda + (1-\lambda)\eta)\rho + (1-\delta)(1-\lambda)(1-\theta)\eta\rho \leq 0$$

$m^* = 0$ remains the socially optimal allocation. However, the socially optimal cash holdings for LR may now be substantially higher than under autarky and is given by:

$$\varphi'(\kappa - M^*) = \frac{1}{\delta}.$$

This unconstrained welfare optimum is less likely to satisfy the LR participation constraint, but more likely to satisfy the SR participation constraint.

To summarize, the unconstrained welfare optimum is extremely simple and intuitive. First, it involves ex-post asset allocations to the agents who value the asset the most. That is, $\phi(\omega) = 1$ in all states ω except state $\omega_{3\rho}$ when $\phi(\omega_{3\rho}) = 0$, so that the fruits of the risky asset are obtained by SR in all states except $\omega_{3\rho}$. Second, to the extent that cash in SR hands is worth more than in LR hands at dates 1 and 2, all cash holdings (if there are any) are transferred to SR. Third, given that the social welfare optimum implements an efficient trade at dates 1 and 2 between LR and SR, it is welfare efficient for SR to hold no cash reserves ($m^* = 0$). That is, cash holdings by SR are wasteful and therefore it is efficient to minimize these holdings. Finally, LR cash holdings are set so that: *i)* $M^* = 0$ when γ is close to $\frac{1}{2}$. That is, since there is a positive social opportunity cost of holding cash the welfare maximizing portfolio allocation is to have no cash holdings at all; *ii)* when $\gamma \longrightarrow \frac{1}{1+\delta}$, the marginal return on cash and *discounted* LR risky assets–at SR discount factor δ–are just equalized.

IV.B Constrained Welfare Optimum

We have ignored two sets of constraints in the planner's optimization problem, incentive compatibility and ex-ante participation constraints. As is the case for LR, the planner cannot distinguish between states ω_{20}–when the asset matures at date 2 but yields no dividends–and ω_{2L}–when the risky asset is known to mature at date 3. The planner must therefore rely on truthful reporting of these states by the financial intermediaries. In other words, the following incentive compatibility constraints must hold:

$$T(\omega_{20}) \geq T(\omega_{2L}) + (1 - \eta)\delta T(\omega_{30})$$

and

$$T(\omega_{2L}) + \eta\delta(\phi(\omega_{3\rho})(1 - m)\rho + T(\omega_{3\rho})) + (1 - \eta)\delta T(\omega_{30}) \geq T(\omega_{20}).$$

The first constraint guarantees that SR prefers to truthfully announce state ω_{20} rather than lie and pretend that state ω_{2L} has occurred. If SR lies, it knows that the dividend at date 3 will be equal to zero, so that it obtains at most $T(\omega_{2L}) + (1 - \eta)\delta T(\omega_{30})$. The second constraint guarantees that SR prefers to truthfully report ω_{2L} rather than lie and announce state ω_{20}.

It is interesting to observe that the unconstrained planning solution, where $\phi^*(\omega_{20}) = \phi^*(\omega_{2L}) = 1$, $T^*(\omega_{20}) = T^*(\omega_{2L}) = M$, and $\phi^*(\omega_{3\rho}) = 0$ with $T^*(\omega_{3\rho}) = T^*(\omega_{30}) = 0$, satisfies both constraints.

Thus, the only constraints that may be violated are the ex-ante participation constraints. From the preceding analysis of the unconstrained problem, the SR participation constraint is

violated when $\gamma \longrightarrow \frac{1}{2}$, and the LR constraint may be violated when $\gamma \longrightarrow \frac{1}{1+\delta}$. We consider each constrained welfare optimum in turn.

1. When $\gamma \longrightarrow \frac{1}{2}$ we have:

$$\max_{m,M} W(m,M) \;=\; \gamma[m + M + (\lambda + (1-\lambda)\theta\eta)(1-m)\rho] +$$

$$(1-\gamma)[\varphi(\kappa - M) + (1-\lambda)(1-\theta)\eta(1-m)\rho]$$

subject to:

$$m + M + (\lambda + (1-\lambda)\theta\eta)(1-m)\rho = 1.$$

Substituting for M we obtain the unconstrained problem:

$$\max_m W(m) = \gamma + (1-\gamma)[\varphi(\kappa - M) + (1-\lambda)(1-\theta)\eta(1-m)\rho]$$

It immediately follows that the constrained welfare optimum solution is to set $m^* = 0$ and

$$M^* = 1 - (\lambda + (1-\lambda)\theta\eta)(1-m)\rho.$$

2. When $\gamma \longrightarrow \frac{1}{1+\delta}$ we have:

$$\max_{m,M} W(m,M) \;=\; \gamma[m + M + (\lambda + (1-\lambda)\theta\eta)(1-m)\rho] +$$

$$(1-\gamma)[\varphi(\kappa - M) + (1-\lambda)(1-\theta)\eta(1-m)\rho]$$

subject to:

$$\varphi(\kappa) - \varphi(\kappa - M) = (1-\lambda)(1-\theta)\eta(1-m)\rho.$$

In this situation the constrained optimal solution may possibly involve an interior value for the SR cash holdings, $m^* > 0$.

To summarize, when $\gamma \longrightarrow \frac{1}{2}$ the constrained welfare optimum can only be obtained by requiring LR to hold a sufficient amount of cash. In a competitive equilibrium situation with an unregulated long run investor sector this allocation is generally not attainable as long run investors would only hold positive cash amounts if there is 'cash-in-the-market' pricing of SR assets at dates 1 or 2. When $\gamma \longrightarrow \frac{1}{1+\delta}$ the welfare optimum favors investors even more and may be even further removed from what is attainable in a competitive equilibrium with an unregulated long run investor sector.

V. REGULATORY IMPLICATIONS

V.A Minimum reserve requirements and the scope of regulation

In our framework the SRs face a potential maturity mismatch problem: Their investors strictly prefer to consume at dates $t = 1$ and $t = 2$, but the risky asset may only pay at date $t = 3$. Financial intermediaries that engage in maturity transformation, such as banks and insurance companies, are heavily regulated in most jurisdictions in an effort to protect claimants from potential losses on the asset side of these institutions' balance sheets. One potential regulatory response to these concerns is the imposition of minimum reserve or liquidity requirements that forces the bank to maintain a certain percentage of its assets in "cash" or other liquid securities such as treasuries.

To illustrate the effect of these minimum cash positions it is useful to turn to the example illustrated in Figure 3, where recall the efficient delayed trading equilibrium was associated with a SR sector that was fully invested in the risky asset, $m_d^* = 0$. Clearly *any* reserve requirement in this situation risks undermining the delayed trading equilibrium and to leaves as the only equilibrium the immediate trading equilibrium. The general equilibrium effects are such that when forcing the SR to carry a minimum amount of cash the returns associated with cash holding for the unregulated sector go down, as now the SR invests less in the risky asset. This in turn makes the risky asset less attractive to the SR. Also, the resulting equilibrium outcome could be one where the minimum cash position is a non binding constraint.

It is here that the interpretation of our model is of great importance. We argue that the way to understand the capital carried by the LR investor sector is not as the total amount of cash (very liquid securities) available to absorb the SR's firesales but rather as the capital that is bundled with the knowledge to absorb the particular risky asset held by the SR. Other cash is either not available to absorb the firesale or available at a much higher premium. By biasing the SR's portfolio towards more cash, less of a particular risky asset is held, which in turn decreases the returns of holding cash for the subset of the LR investors that specialize in that particular market.

V.B VAR

An important feature of our model is that SR's investors can consume at date $t = 3$, albeit at a much lower marginal utility of consumption. This, as we discussed, can undermine

the existence of the efficient delayed trading equilibrium when $P^*(\omega_d) \leq \delta\eta\rho$. As shown in our example, for sufficiently high θ the SRs in state ω_{2L} prefer to retain the asset and consume the proceeds in state $\omega_{3\rho}$, which occurs with probability η, rather than selling the risky asset at date $t = 2$ for a price $P^*(\omega_d)$. The LR investors, anticipating this, reduce the amount of liquidity they carry and the SRs reduce in turn the share of their assets invested in the long one. The only equilibrium that survives is the immediate trading equilibrium, which is dominated by the allocation that would result if the SRs were *forced* to sell in state ω_{2L}.

With some abuse of terminology, the SRs would be better off if they could commit to a sale policy at date $t = 2$ for in that case LRs would be certain to carry the cash to absorb these firesales even in the presence of adverse selection problems. The problem of course is that the SRs in state ω_{2L} cannot commit to such a policy: Given the low price in state ω_d they prefer to carry the risky asset to $t = 3$. SRs in state ω_{20} prefer of course to liquidate the now worthless asset.

It is here that a regulation that forces SRs to liquidate assets improves efficiency by yielding a pooling equilibrium that benefits all parties ex-ante. As usual in our set up the benefits accrue through the better portfolios that the different agents carry. Perhaps surprisingly a VAR like regulation that requires the SRs to liquidate their positions in the risky asset could produce a larger position in that risky asset even when the price at which the asset can be liquidated drops as a result. Once again our result obtains through the general equilibrium effect. In the presence of a VAR regulation the returns of carrying outside liquidity increase for the LRs for they now know that the *all* SRs will be forced to liquidate, the lemons as well as the good assets.

V.C Public and private provision of liquidity

V.C.1 The source of the inefficiency in the private supply of outside liquidity

As shown in Section III, when the adverse selection problem is severe enough the delayed-trading equilibrium fails to exists, which leaves only the immediate-trading equilibrium. This occurs whenever $P_d^*(\omega_d) < \delta\eta\rho$. This results in a loss of efficiency, which could be remedied if the SRs could commit to liquidate the assets in state ω_d for that would relieve the lemons problem that arises at date $t = 2$. In the absence of such a commitment technology it is natural to consider whether the provision of public liquidity can sufficiently raise the price of the risky asset in state ω_d so as to elicit the supply of the good assets which may in turn elicit

35

the provision of private liquidity. But, why is there a role for public liquidity in our model? What is the source of the inefficiency?

An important feature of our model is that outside liquidity is supplied competitively: In choosing the level of cash carried to absorb the firesales, the LRs do not take into account the effect that their choices have on the price of the risky asset in state ω_d and thus on the particular mix of assets supplied by the SRs. Clearly the monopolist, by internalizing what its choice of inside liquidity means for the price of the risky asset in state ω_d can always guarantee the existence of the delayed trading equilibrium.

To illustrate this it may be helpful to consider the case where the outside liquidity is supplied by a monopolist in the context of our example. Clearly the monopolist prefers to supply the liquidity in state ω_d, where it is "cheaper," than in state ω_i. Recall also that whenever $\theta < \widehat{\theta}$ the SRs carry a strictly positive amount of inside liquidity $m_d^* > 0$ and thus they make zero profits; all the surplus accrues to the LRs, who cannot obtain higher profits by any other choice of inside liquidity than the one that obtains in the competitive case. It follows then that in this range the competitive and monopolistic solutions are identical. When $\theta \geq \widehat{\theta}$, the level of inside liquidity is drawn to $m_d^* = 0$. Recall that in this case some of the surplus goes to the SRs because the LRs now compete for a fixed supply of the risky asset. The monopolist instead would restrict the supply of outside liquidity in this situation in order to extract all the surplus from the SRs. This can be seen in Figure 8, where, in the top panel, the profits of the monopolist are plotted together with the competitive ones and in the bottom panel the prices in state ω_d are plotted against θ.

First notice that, as just discussed, in region **A**, which corresponds to the case $\theta < \widehat{\theta}$ the prices and profits of the monopolist are identical to the competitive ones. In region **B**, the FIs set the level of inside liquidity to $m_d^* = 0$ and the monopolist LR restricts the supply of outside liquidity so as to capture fully all the gains from trade. It is for this reason that now the price of the risky asset in state ω_d is below the competitive one, $P_d^*(\omega_d)$. But, after a certain value of θ, the monopolist has to fix the price for the risky asset to $\delta\eta\rho$, for below this level SRs in state ω_{2L} prefer to carry it to date $t = 3$ rather than liquidating it in state ω_d.

The crucial gain on efficiency occurs in region **C**. There when the supply of outside liquidity is competitive the adverse selection premium is severe enough as to make the price in state ω_d such that $P_d^*(\omega_d) < \delta\eta\rho$ and thus the delayed trading equilibrium breaks down. The LRs do not internalize that their choice of outside liquidity affects the pool of assets being supplied in the market. Instead now when there is a single LR, it internalizes the effect of

its choice of outside liquidity on the SR's decision to supply the risky asset: By carrying and supplying enough outside liquidity the monopolist elicits the supply of the risky asset by those SRs in state ω_{2L}, avoiding the break down of the delayed exchange. Notice that as shown in the top panel of Figure 8, the monopolist's profits are, in this region, above the one that obtains in the immediate exchange equilibrium, which is the only one that could be supported in the competitive case.[21]

V.C.1 Public liquidity as complementary to private outside liquidity

In summary then the monopolist improves the efficiency by the supporting the delayed exchange in the regions where the competitive case cannot. Clearly in the more realistic case where the outside liquidity is supplied competitively this price support can only come from the public sector, which can finance the public provision of liquidity by taxing the parties appropriately as there is surplus relative to the one that obtains when only the immediate exchange equilibrium survives. Notice that in this case the provision of public liquidity is **complementary** to the private provision of liquidity. Indeed, when the public provision of liquidity increases the price of the risky asset up to $\delta\eta\rho$ in region C, the LRs have now an incentive to carry outside liquidity to absorb the risky assets supplied in state ω_d, which now include risky assets in state ω_{2L}. Instead, if the public sector supplies liquidity in regions A or B it can only come at the expense of lowering the returns of the LRs, which then decrease the level of outside liquidity they carry. In this case then public and private liquidity are **substitutes** rather than complements. It follows then that whether the public provision of liquidity is beneficial or not depends critically on the degree of adverse selection in the market: If it is high enough to prevent the delayed exchange from occurring then there is an efficiency improving role for public liquidity which encourages in turn the private supply of outside liquidity.

VI. CONCLUSIONS

This paper is concerned with three questions. First, what determines the distribution of liquidity across market participants? Second, is this distribution (constrained) efficient? Finally, if it is not efficient, what are the regulatory remedies that can brought to bear to restore efficiency? A novel aspect of our paper is the emphasis on the cross sectional aspects

---

[21]It is worth emphasizing than in this region the profits of the SRs are such that $\pi > 1$. The reason is that the monopolist has to "leave some rents" to the SRs precisely to elicit the supply of the assets in state ω_{2L}.

of liquidity provision because it seems a fixture of today's financial markets the presence of different actors that stand ready to absorb sales in distressed states. The incentives of the different parties in financial markets to carry liquidity is driven by their different opportunity costs and an important question is whether a competitive price mechanism would elicit the right cash reserve decisions. A second element in the model that departs from the existing literature is our assumption on the timing of the liquidity event. An important feature of our framework is precisely the fact that adverse selection problems worsen as the liquidity event progresses. Intermediaries face the choice of raising liquidity early on the liquidity crisis, before adverse selection issues set in, or at further depressed prices later in the crisis. The benefit of riding the crisis, of course, is that the intermediary may avoid the sale of assets at distressed prices altogether.

We show that, first, when the adverse selection problem is not "too severe," there are multiple equilibria. In the first, what we termed the immediate trading equilibrium, the intermediaries liquidate their positions in exchange for cash early on the liquidity crisis. In the second, what we termed the delayed trading equilibrium, liquidation takes place late in the liquidity event and in the presence of adverse selection problems. We show that the latter Pareto dominates the former because it saves on cash, which is costly to carry but that it sometimes fails to exist when the adverse selection problem is severe enough. The reason is that in this case prices are so depressed as to make it profitable for the agents holding good assets to carry them to maturity even when it is very costly to do so. We show that if a state contingent commitment technology were to be available to the intermediaries they would be better off committing ex-ante to liquidate at these depressed prices in the distressed states. We also showed, perhaps more surprisingly, that a monopolistic supplier of liquidity can dramatically improve welfare.

Our emphasis on the cross section of liquidity holdings responds to the present state of financial markets, where a myriad of different participants stand ready to step in and provide liquidity in a way that did not seem the norm twenty years ago.[22]

---

[22]For instance, a recent article in the Wall Street Journal of Friday May 9th 2008 by Lingling WeI and Jennifer S. Forsyht emphasized that the discounts on commercial real estate debt are less pronounced than in the previous real-estate collapse of the early 1990s. As the authors point out "[t]oday there are at least 55 active or planned commercial real-estate debt funds seeking to raise $33.8 billion, according to Real Estate Alert, a trade publication. And many have begun to do deals." In the recent period of distress that started in the summer of 2007, the role of sovereign funds has been notorious and major source of recapitalization for institutions such as Citi.

Clearly an important message of this paper is that the role of the public sector as a provider of liquidity has to be understood in the context of a competitive provision of liquidity by the private sector. In particular, the issue of whether public provision of liquidity is a complement or a substitute to private liquidity is an important one. Here a key assumption is that the public liquidity provider knows which state of nature the economy is in (ω_i or ω_d) for the economic impact on the incentives to supply private liquidity of one dollar of public money in one state or the other is rather different. An important remaining task is thus to develop models that account for the fact that the public provision of liquidity occurs under conditions of ignorance about the true state of nature in which it is being provided.

Another important theme in our model is the particular timing of the liquidity crisis that we propose. Liquidity crisis have, in our opinion, always real origins, small as they may be. In our framework the onset of the liquidity event starts with a real deterioration of the quality of the risky asset held by financial intermediaries. It is only later that the adverse selection sets in. Our purpose here was to capture the fact that the market may realize slowly the extent to which many participants are affected by the deterioration in asset values. Financial institutions here face a choice of whether to liquidate early or ride out the crisis in the hope that the asset may ultimately pay. This trade-off is, in our set up, unrelated to the incentives that may force institutions to liquidate at particular times due to accounting and credit quality restrictions in the assets they can hold that have featured more prominently in the literature. Understanding the effect that these restrictions have on the portfolio decisions of the different intermediaries remains an important extension of the present model.

Finally, in our model long run investors are those with sufficient knowledge to absorb the sales of the risky, sophisticated, asset held by the financial intermediaries. In this sense only their capital matters for valuation; other, less knowledge intensive, capital will only step in at sufficient discounts. Our current work focuses precisely on understanding how different capital gets "earmarked" to specific markets. What arises is a theory of market segmentation and contagion that, we believe may shed light on the behavior of financial markets in states of crisis.

REFERENCES

Acharya, Viral (2001) "A Theory of Systemic Risk and Design of Prudential Bank Regulation" *manuscript*, New York University and London Business School.

Acharya, Viral and Tanju Yorulmazer (2005) "Cash-in-the-market Pricing and Optimal Bank Bailout Policy," *manuscript*, CEPR.

Aghion, Philippe, Patrick Bolton, and Mathias Dewatripont (2000) "Contagious Bank Failures in a Free Banking System" *European Economic Review*, May.

Allen, Franklin and Douglas Gale (1998) "Optimal Financial Crisis," *Journal of Finance*, 53, 1245-1284.

Bhattacharya, Sudipto and Douglas Gale (1986) "Preference Shocks, Liquidity, and Central Bank Policy," *manuscript*, CARESS Working Paper #86-01, University of Pennsylvania.

Brunnermeier, Markus and Lasse Pedersen (2008) "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, forthcoming.

Bryant, John (1980) "A Model of Reserves, Bank Runs, and Deposit Insurance" *Journal of Banking and Finance*, 3, 335-344.

Diamond, Douglas W. and Philip H. Dybvig (1983) "Bank Runs, Deposit Insurance, and Liquidity" *Journal of Political Economy*, vol. 91 no. 3, 401-419.

Diamond, Douglas W. (1997) "Liquidity, Banks, and Markets" *Journal of Political Economy*, vol. 105 no. 5, 928-956.

Diamond, Douglas W. and Raghuram Rajan (2005) "Liquidity Shortage and Banking Crisis," *Journal of Finance,* April.

Fecht, Falko (2004) "On the Stability of Different Financial Systems," *Journal of the European Economic Association,* vol. 2 no. 6, 969-1014.

Freixas, Xavier, Bruno Parigi and Jean-Charles Rochet (2000) "Systemic Risk, Interbank Relations, and Liquidity Provision by the Central Bank" *Journal of Money, Credit and Banking*, vol. 32 (2), 611-638.

Gorton, Gary and Lixin Huang (2004) "Liquidity, Efficiency, and Bank Bailouts," *American Economic Review,* vol. 94 no. 3, 455-483.

Gromb, Dennis and Dimitri Vayanos (2002) "Financially Constrained Arbitrage and the Cross-section of Market Liquidity," *manuscript*, London Business School.

Holmstron, Bengt and Jean Tirole (1998) "Private and Public Supply of Liquidity" *Journal of Political Economy,* vol. 106 no. 1, 1-40.

Kondor, Peter (2008) "Risk in Dynamic Arbitrage: The Price Effects of Convergence," *Journal of Finance,* forthcoming.

Kyle, Peter and Wei Xiong (2001) "Contagion as a Wealth Effect," *Journal of Finance*, vol. 56, 1401-1440.

Parlour, Christine and Guillaume Plantin (2008) "Loan Sales and Relationship Banking," *Journal of Finance*, forthcoming.

Xiong, Wei (2001) "Convergence Trading with Wealth Effects," *Journal of Financial Economics,* vol. 62, 247-292.

APPENDIX

Proof of Proposition 1. The first order condition of the LR is

$$\lambda + (1-\lambda)\frac{\eta\rho}{P_i(\omega_i)} \leq \varphi'(\kappa - M). \tag{19}$$

First we establish that it is not possible to support an equilibrium with $M_i^* = 0$ and $m_i^* = 1$. Indeed if $m_i^* = 1$ it has to be the case that the price in state ω_i is such that

$$P_i^*(\omega_i) \leq \frac{1-\lambda\rho}{1-\lambda}$$

but by assumption A3 this implies

$$\lambda + (1-\lambda)\frac{\eta\rho}{P_i(\omega_i)} > \varphi'(\kappa), \tag{20}$$

and thus $M_i^* > 0$ a contradiction.

Having ruled the no trade immediate trading equilibrium we proceed next as follows. Start by solving the following equation in $P_i(\omega_i)$

$$\lambda + (1-\lambda)\frac{\eta\rho}{P_i(\omega_i)} = \varphi'(\kappa - P_i(\omega_i)), \tag{21}$$

and define

$$\underline{P} = \frac{1-\lambda\rho}{1-\lambda}, \tag{22}$$

as positive number by assumption A2.

- Case 1: Assume first that $P_i(\omega_i) \geq \underline{P}$, then set $P_i^*(\omega_i) = M_i^* = P_i(\omega_i)$ and $m_i^* = 0$, which meets the first order condition of the SRs as can be checked by inspection of expression (2).

- Case 2: Assume next that $P_i(\omega_i) < \underline{P}$, then set $P_i^*(\omega_i) = \underline{P}$ and M_i^* to be the solution to

$$\lambda + (1-\lambda)\frac{\eta\rho}{\underline{P}} \leq \varphi'(\kappa - M_i^*), \tag{23}$$

  which by assumption A3 is such that $M_i^* > 0$ and clearly it has to be such that $M_i^* < \underline{P}$. Because, given these prices, the SRs are indifferent on the level of cash carried set m_i^* so that

$$P_i^*(\omega_i) = \frac{1-\lambda\rho}{1-\lambda} = \frac{M_i^*}{1-m_i^*} \tag{24}$$

As for prices in state ω_d they have to be such that both the SRs and the LRs prefer to trade at ω_i. For this set

$$P_i^*(\omega_d) < \min\left\{\delta\eta\rho, \frac{1-\rho[\lambda + (1-\lambda)\theta\eta]}{(1-\lambda)(1-\theta\eta)}\right\}. \tag{25}$$

Given this price the LR investors expect only lemons (assets with zero payoff) in the market at ω_d and thus the demand is equal to zero $Q^*(\omega_d) = 0$. As for the SRs notice that if they wait to liquidate at ω_d they obtain

$$\theta\eta\rho + (1-\theta\eta)P_i^*(\omega_d) < \theta\eta\rho + \frac{1-\rho[\lambda + (1-\lambda)\theta\eta]}{1-\lambda} = P_i^*(\omega_i),$$

and thus SRs set $q^*(\omega_d) = 0$ and $q^*(\omega_i) = 1 - m_i^* = Q^*(\omega_i)$. $\qquad\square$

Proof of Proposition 2. First notice that since $\varphi'(\kappa) > 1$ in any delayed trading equilibrium there must be cash-in-the-market pricing thus

$$M_d^* = P_d^*(\omega_d)(1 - \theta\eta)(1 - m) \tag{26}$$

Define

$$\lambda + (1 - \lambda)\frac{(1 - \theta)\eta\rho}{(1 - \theta\eta)P_d(\omega_d)} = \varphi'[\kappa - (1 - \theta\eta)P_d(\omega_d)] \tag{27}$$

This equation always a unique solution which in addition satisfies

$$P_d(\omega_d) \in \left(0, \frac{\kappa}{1 - \theta\eta}\right). \tag{28}$$

There are two cases to consider:

- Case 1: $P_d(\omega_d)$ is such that

$$P_d(\omega_d) < \frac{1 - \rho[\lambda + (1 - \lambda)\theta\eta]}{(1 - \lambda)(1 - \theta\eta)} = \underline{P}. \tag{29}$$

  In this case set

$$P_d^*(\omega_d) = \underline{P}, \tag{30}$$

  and set M_d^* to be the solution of

$$\lambda + (1 - \lambda)\frac{(1 - \theta)\eta\rho}{(1 - \theta\eta)P_d^*(\omega_d)} = \varphi'(\kappa - M_d^*), \tag{31}$$

  which from the strict concavity of $\varphi(\cdot)$ is

$$M_d^* < (1 - \theta\eta)P_d^*(\omega_d). \tag{32}$$

  Choose m_d^* such that

$$P_d^*(\omega_d) = \frac{M_d^*}{(1 - \theta\eta)(1 - m_d^*)} \tag{33}$$

  Notice that because $P_d^*(\omega_d) = \underline{P}$ the SRs are indifferent in the level of cash held. Both types of traders would prefer to wait to trade in period $t = 2$ provided that $P_d^*(\omega_i)$ is in the interval

$$\left[\frac{(1 - \theta\eta)P_d^*(\omega_d)}{1 - \theta}, \theta\eta\rho + (1 - \theta\eta)P_d^*(\omega_d)\right], \tag{34}$$

  which is non empty if and only if

$$P_d^*(\omega_d) \le \frac{(1 - \theta)\eta\rho}{1 - \theta\eta} = \overline{P}. \tag{35}$$

  Clearly, given assumption A1, specifically the fact that $\varphi'(\kappa) > 1$, and equation (31), equation (35) is trivially met. Clearly, given assumption A1, specifically the fact that $\varphi'(\kappa) > 1$, and equation (31), equation (35) is trivially met.

  Notice that $P_d^*(\omega_d)$ is independent of δ and for $\delta \le \overline{\delta}$, where

$$\overline{\delta} = \frac{P_d^*(\omega_d)}{\eta\rho}, \tag{36}$$

  the SR (weakly) prefers to trade in period $t = 2$ for a price $P_d^*(\omega_d)$ than carrying the asset to period $t = 3$.

- Case 2: $P_d(\omega_d) \geq \underline{P}$, then choose

$$P_d^*(\omega_d) = P_d(\omega_d) \qquad M_d^* = P_d^*(\omega_d)(1 - \theta\eta) \qquad \text{and} \qquad m_d^* = 0. \tag{37}$$

Except for establishing inequality (35), the remainder of the proof follows as in the previous case. To establish that $P_d^*(\omega_d)$ meets (35) it is enough to substitute $\underline{P}$ in (35) and appeal to assumption A2. □

Before proving Propositions 3, 4, and 5, it is useful to establish the following

Result. Assume A1-A4 hold. Then the immediate trading equilibrium is such that $m_i^* \in (0, 1)$.

Proof. By the SR's first order condition if the price at date $t = 1$ is given by

$$P_i^*(\omega_i) = \frac{1 - \lambda\rho}{1 - \lambda}$$

then the SR investor is indifferent about the cash position carried. Let M_i^* be the solution to

$$\lambda + (1 - \lambda)^2 \frac{\eta\rho}{1 - \lambda\rho} = \varphi'(\kappa - M_d^*),$$

which by assumption A3 exists and is unique. By assumption A4,

$$\frac{1 - \lambda\rho}{1 - \lambda} > \kappa > M_i^*.$$

Then set $m_i^* \in (0, 1)$ so that

$$\frac{1 - \lambda\rho}{1 - \lambda} = \frac{M_i^*}{1 - m_i^*}.$$

The construction now of the immediate trading equilibrium follows as in the proof of Proposition 1. □

We prove Proposition 4 first. The proof of Proposition 3 following trivially after that.

Proof of Proposition 4 First notice that by the result above, the immediate trading equilibrium is such that $m_i^* > 0$ (and, obviously, $M_i^* > 0$). Thus because the delayed trading equilibrium specializes to the immediate trading equilibrium when $\theta = 0$, it follows that there exists a neighborhood $(0, \tilde{\theta})$ such that $m_d^* > 0$. Then from the LR's and SR's first order conditions, combined with cash in the market pricing, M_d^* and m_d^* are fully determined by

$$\psi^{(M)} = \lambda + (1 - \lambda) R_d^*(\theta) - \varphi'(\kappa - M_d^*) = 0 \tag{38}$$

$$\psi^{(m)} = (1 - m_d^*)(1 - \rho(\lambda + (1-)\theta\eta)) - (1 - \lambda) M_d^* = 0 \tag{39}$$

Expression (38) is the LR's first order condition. Expression (39) is the SR's first order condition combined with the cash-in-the-market pricing equation. These two equations determine M_d^* and m_d^*. In the above expression

$$R_d^* = \frac{(1 - \theta)\eta\rho}{(1 - \theta\eta) P_d^*(\omega_d)},$$

where $P_d^*(\omega_d)$ is given by $\underline{P}$ (see expression (29). Then basic algebra shows that

$$R_{d,\theta}^* = \frac{\partial R_d^*}{\partial \theta} \propto \rho[\lambda + (1 - \lambda)\eta] - 1 > 0, \tag{40}$$

by assumption (A1).

43

$$\partial_x \psi = \begin{pmatrix} \psi_M^{(M)} & \psi_m^{(M)} \\ \psi_M^{(m)} & \psi_m^{(m)} \end{pmatrix} \qquad \text{and} \qquad \partial_\theta \psi = \begin{pmatrix} \psi_\theta^{(M)} \\ \psi_\theta^{(m)} \end{pmatrix}, \tag{41}$$

where

$$
\begin{aligned}
\psi_M^{(M)} &= \varphi''\left(\kappa - M_d^*\right) < 0 \\
\psi_m^{(M)} &= 0 \\
\psi_m^{(m)} &= -\left[1 - \rho\left(\lambda + (1-\lambda)\theta\eta\right)\right] < 0 \\
\psi_M^{(m)} &= -(1-\lambda) \\
\psi_\theta^{(M)} &= (1-\lambda)R_{d,\theta}^* > 0 \\
\psi_\theta^{(m)} &= -(1-m_d^*)(1-\lambda)\eta\rho < 0,
\end{aligned}
$$

First,

$$|\partial_x \psi| = -\left[1 - \rho\left(\lambda + (1-\lambda)\theta\eta\right)\right]\varphi''\left(\kappa - M_d^*\right) > 0$$

Second, by an application of the implicit function theorem

$$M_{d,\theta}^* = \frac{\partial M_d^*}{\partial \theta} = -\begin{bmatrix} 1 & 0 \end{bmatrix}(\partial_x \psi)^{-1}\partial_\theta \psi \qquad \text{and} \qquad m_{d,\theta}^* = \frac{\partial m_d^*}{\partial \theta} = -\begin{bmatrix} 0 & 1 \end{bmatrix}(\partial_x \psi)^{-1}\partial_\theta \psi.$$

After some algebra:

$$
\begin{aligned}
m_{d,\theta}^* &= -|\partial_x \psi|^{-1}\left[-\psi_M^{(m)}(1-\lambda)R_{d,\theta}^* - \psi_M^{(M)}(1-m)(1-\lambda)\eta\rho\right] \\
&= -|\partial_x \psi|^{-1}\left[(1-\lambda)^2 R_{d,\theta}^* - \varphi''\left(\kappa - M_d^*\right)(1-m_d^*)(1-\lambda)\eta\rho\right] \tag{42} \\
&< 0
\end{aligned}
$$

and

$$
\begin{aligned}
M_{d,\theta}^* &= -|\partial_x \psi|\left[\psi_m^{(m)}\psi_\theta^{(M)} - \psi_m^{(M)}\psi_\theta^{(m)}\right] \\
&= -|\partial_x \psi|\psi_m^{(m)}\psi_\theta^{(M)} \\
&> 0
\end{aligned}
$$

Because m_d^* is strictly decreasing in θ if $m_d^* = 0$ for some $\widehat{\theta}$, then $m_d^* = 0$ for all $\theta \geq \widehat{\theta}$. For $\theta \geq \widehat{\theta}$ the LR's first order condition is given by

$$\lambda + (1-\lambda)\frac{(1-\theta)\eta\rho}{M_d^*} = \varphi'\left(\kappa - M_d^*\right),$$

where we have made use of the fact that cash-in-the-market pricing obtains and $m_d^* = 0$. Then a basic application of the implicit function theorem shows that $M_{d,\theta}^* < 0$ for $\theta > \widehat{\theta}$. As for the behavior of expected returns when $\theta > \widehat{\theta}$, notice that the LR's first order condition is written as

$$\lambda + (1-\lambda)R_d^* = \varphi'\left(\kappa - M_d^*\right),$$

and thus given that $M_{d,\theta}^{ast} < 0$ for $\theta > \widehat{\theta}$, it follows that $R_{d,\theta}^* < 0$ for that range.

44

We turn now to the properties of the aggregate supply of the risky asset at $t = 2$ in the delayed trading equilibrium s_d^*. Using (42),

$$
\begin{aligned}
s_{d,\theta}^* &= |\partial_x \psi|^{-1} (1-\lambda)^2 R_{d,\theta}^* (1 - \theta\eta) \\
&\quad - |\partial_x \psi|^{-1} \varphi'' (\kappa - m_d^*) (1 - m_d^*) (1-\lambda) \eta\rho (1-\theta\eta) - \eta(1 - m_d^*)
\end{aligned}
\tag{43}
$$

Tedious algebra shows that (43) is equal to

$$
(1 - m_d^*) \eta \left[\frac{\rho - 1}{1 - \rho(\lambda + (1-\lambda)\theta\eta)} \right],
$$

which is positive by assumption A2. This completes the proof of Proposition 3. □

Proof of Proposition 4. That $m_i^* > m_d^*$ follows immediately from the fact that $m_i^* = m_d^*(\theta = 0)$ and Proposition 3. Clearly for $\theta \leq \widehat{\theta}$ where $\widehat{\theta}$ was defined in the proof of Proposition 3, $M_i^* < M_d^*$. For $\theta > \widehat{\theta}$ M_d^* is a decreasing function of θ and thus, by continuity there exists a (unique) θ', possibly higher than $\overline{\theta}$, for which $M_d^*(\theta') = M_i^*$; for any $\theta < \theta'$, $M_i^* < M_d^*$. □

Proof of Proposition 5. Under assumption A4, $m_i^* > 0$ and thus $\pi_i^* = 1 \leq \pi_d^*$. As for the expected profits of the LR investors, firs notice that

$$
\frac{\partial \Pi_d^*}{\partial \theta} = \Pi_{d,\theta}^* = (1-\lambda) R_{d,\theta}^* M_d^*.
$$

Given that $\Pi_i^* = \Pi_d^*(\theta = 0)$ and the characterization of expected returns in Proposition 3 the result follows immediately. □
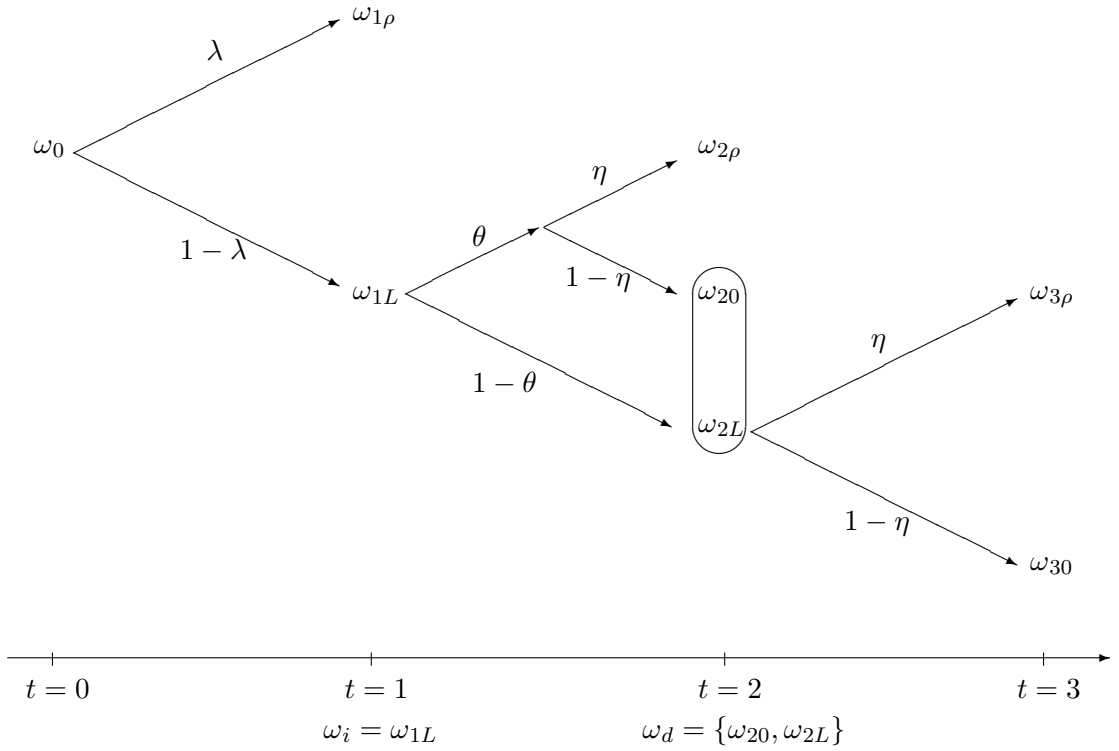
Figure 1. The risky asset. There are four dates. Investment in the risky asset occurs at date $t = 0$. At date $t = 1$ there is an *aggregate* shock. Specifically there are two possible aggregate states, $\omega_{1\rho}$, which occurs with probability λ, and ω_{1L}, which occurs with probability $1 - \lambda$. In $\omega_{1\rho}$ the risky asset matures at date 1 and yields a cash dividend ρ. ω_{1L} is the state when the long duration asset matures later than date 1 (either at dates 2 or 3). At date $t = 2$, there are three *idiosyncratic* states of nature, $\omega_{2\rho}$, which occurs with probability $\theta\eta$, ω_{20}, which occurs with probability $\theta(1-\eta)$, and ω_{2L}, which occurs with probability $1 - \theta$. $\omega_{2\rho}$ is the state when the asset matures at date 2 and yields dividend ρ. Thus the probabilities also denote the mass of SRs which are in the corresponding states of nature. ω_{20} is the state when the asset matures at date 2 but yields no dividends. In ω_{2L} the risky asset is known to mature at date 3. Finally in date $t = 3$ there are again two states, $\omega_{3\rho}$, which occurs with probability η, and ω_{30}, which occurs with probability $1 - \eta$. In state $\omega_{3\rho}$ the asset matures at date 3 and yields dividend ρ and in state ω_{30} the asset matures at date 3 and yields zero dividends. The information set of the LRs at date $t = 2$ is given by $\omega_d = \{\omega_{20}, \omega_{2L}\}$, which generates the adverse selection problem that is key in the analysis. At date $t = 1$, agents are symmetrically informed and $\omega_i = \omega_{1L}$.
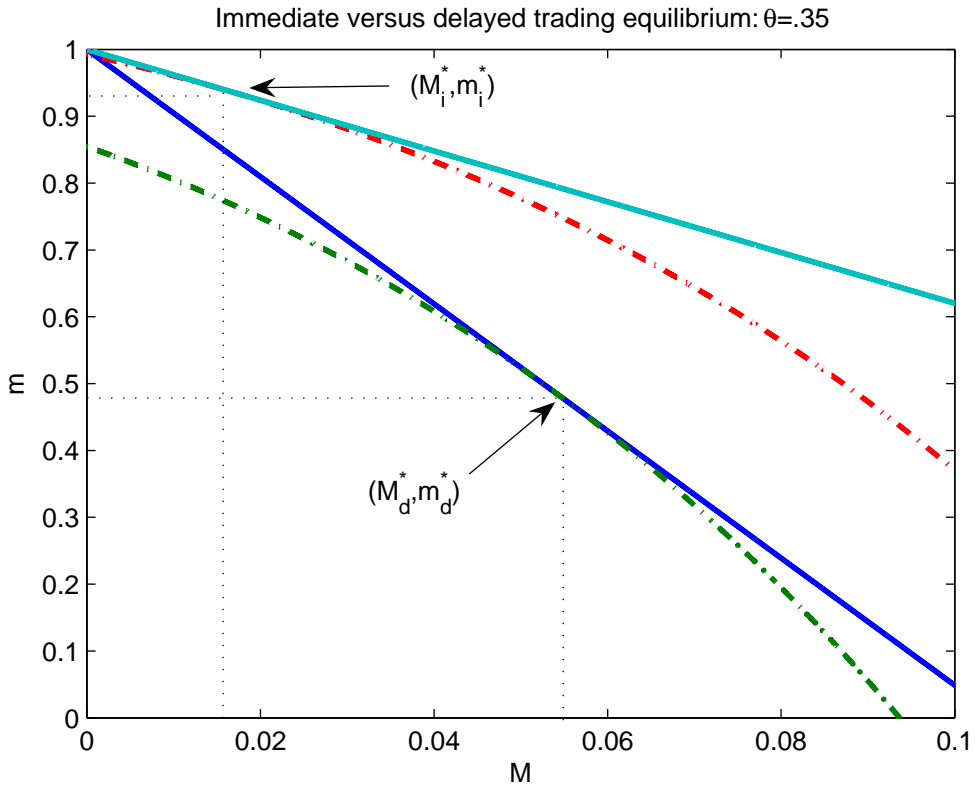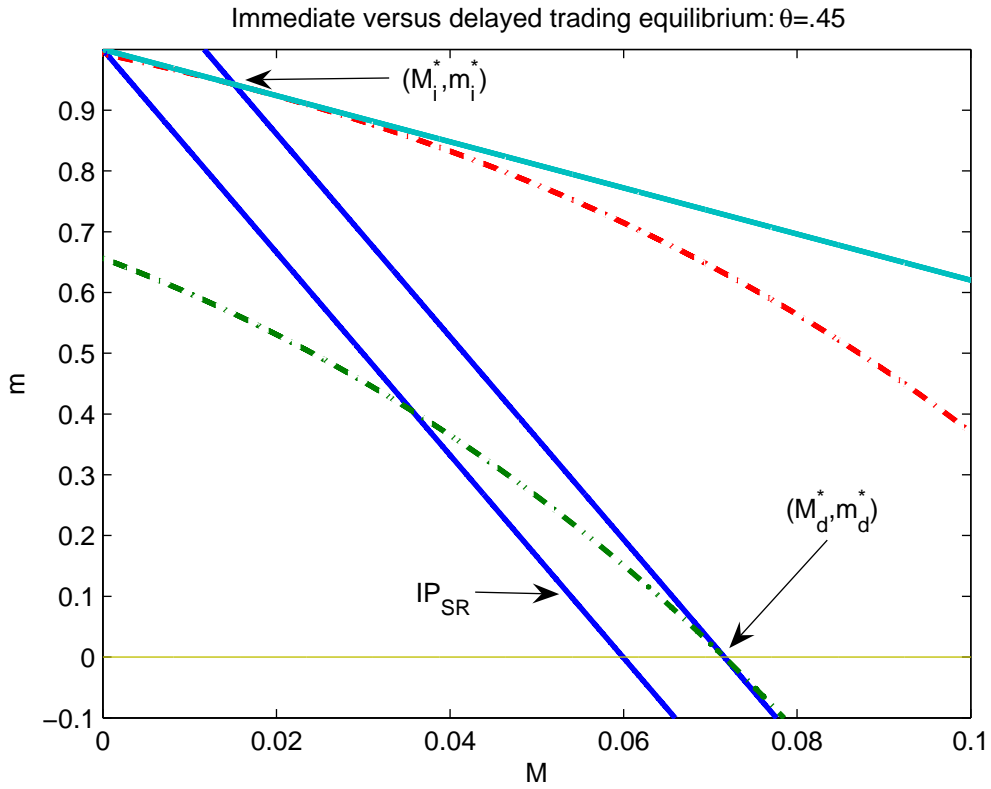
46

Figure 2. Immediate and delayed-exchange equilibria in the Example for the case $\theta = .35$. The graph represents cash holdings, with the cash holdings of the LRs in the x-axis and the cash holdings of the SRs in the y-axis. The dashed curves represent isoprofit lines for the LR and the straight continuous lines represent the SR's isoprofit lines, for both when the exchange occurs in state ω_i and in state ω_d. The isoprofit lines for the SR correspond to its reservation profits $\pi_i^* = \pi_d^* = 1$. The immediate and delayed-trading equilibrium cash holdings are marked (M_i^*, m_i^*) and (M_d^*, m_d^*), respectively.

Figure 3. Immediate and delayed-exchange equilibria in the example when $\theta = .45$. The graph represents cash holdings, with the cash holdings of the LRs in the x-axis and the cash holdings of the SRs in the y-axis. The dashed curves represent isoprofit lines for the LR and the straight continuous lines represent the SR's isoprofit lines, for both when the exchange occurs in state ω_i and in state ω_d. As opposed to the case in Figure 2 now the delayed-treading equilibrium, marked (M_d^*, m_d^*), has the SRs commanding strictly positive profits, $\pi_d^* > 1$. The line marked IP_{SR} denotes the SR's reservation isoprofit line in state ω_d.
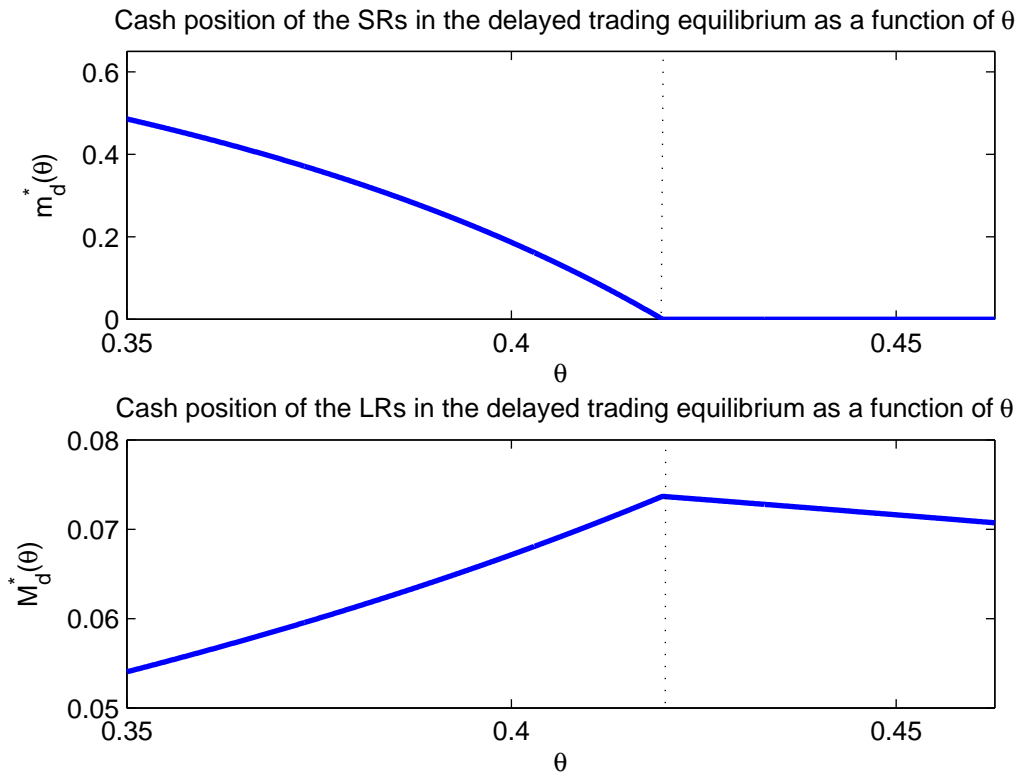
Figure 4. Cash holdings as a function of θ for the Example. Panel A represents the SR's cash holdings in the delayed-trading equilibrium, m_d^* as a function of θ and Panel B does the same for the LR, M_d^*. The dashed vertical line, which sits at $\widehat{\theta} = .4196$ delimits the set of θs for which $m_d^* > 0$ and the one for which $m_d^* = 0$.
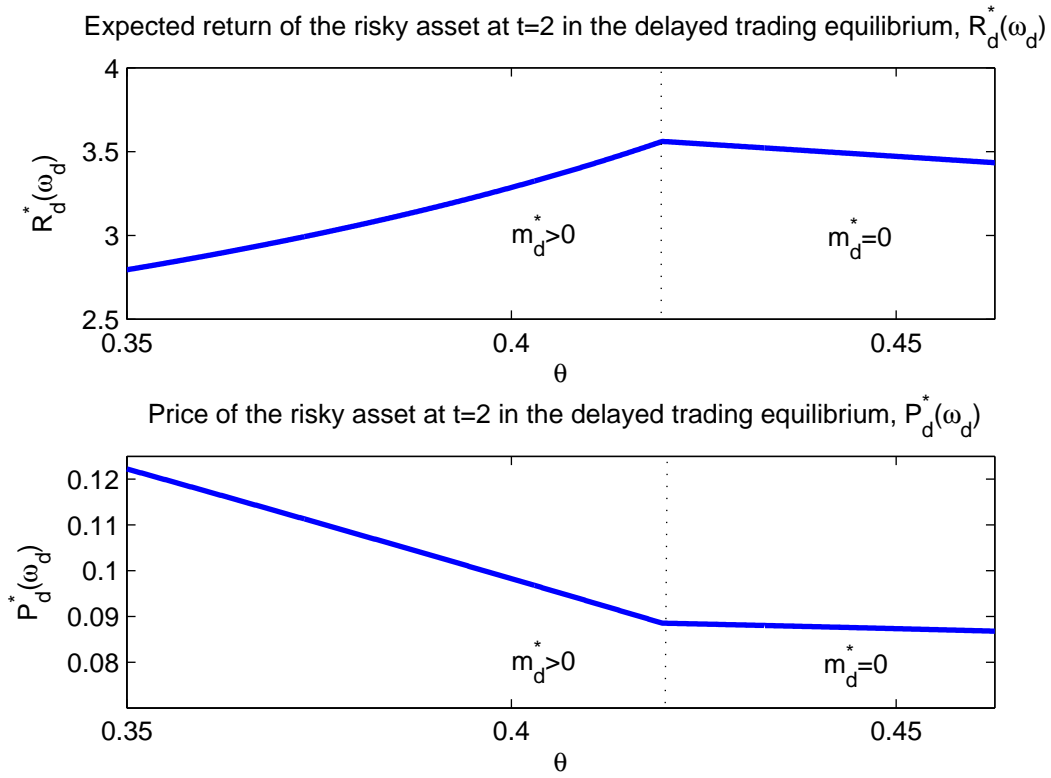
Figure 5. The top panel shows the expected return of the risky asset, $R_d^*(\omega_d)$, as a function of θ at $t = 2$ in the delayed trading equilibrium. The bottom panel shows the price of the risky asset in state ω_d, $P_d^*(\omega_d)$, as a function of θ at $t = 2$ in the delayed trading equilibrium. The dashed vertical line corresponds to $\widehat{\theta} = .4196$.
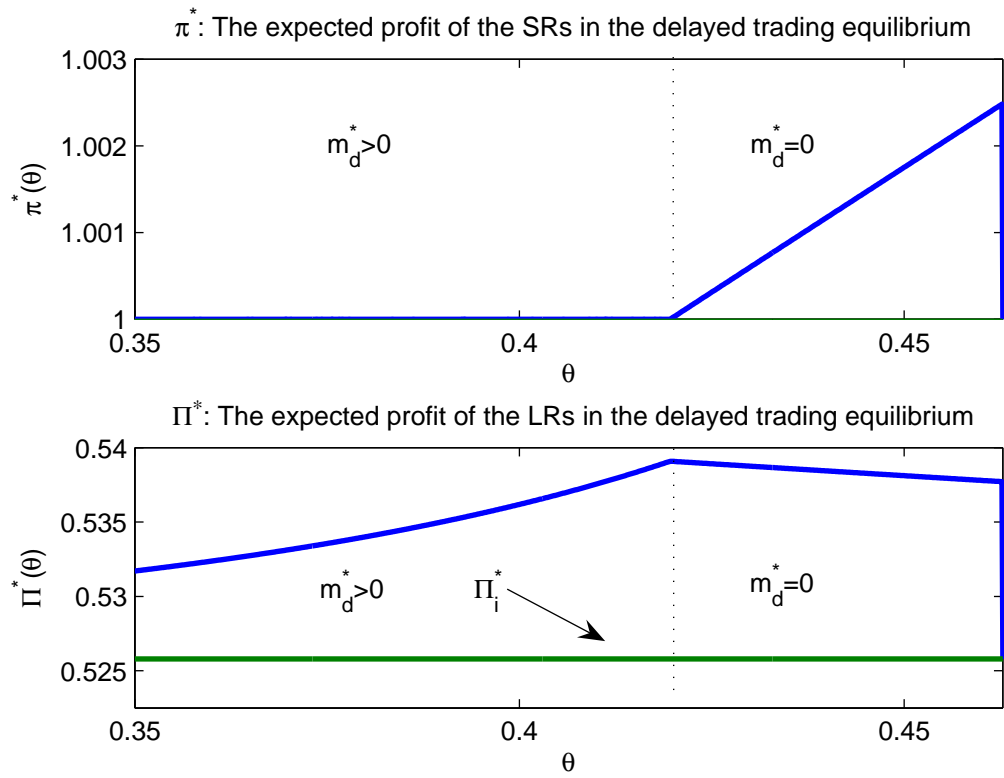
Figure 6. Expected profits for the SR, π^*, (top panel) and the LR (bottom panel), Π^*, as a function of θ in the delayed trading equilibrium. The dashed vertical line corresponds to $\widehat{\theta} = .4196$.
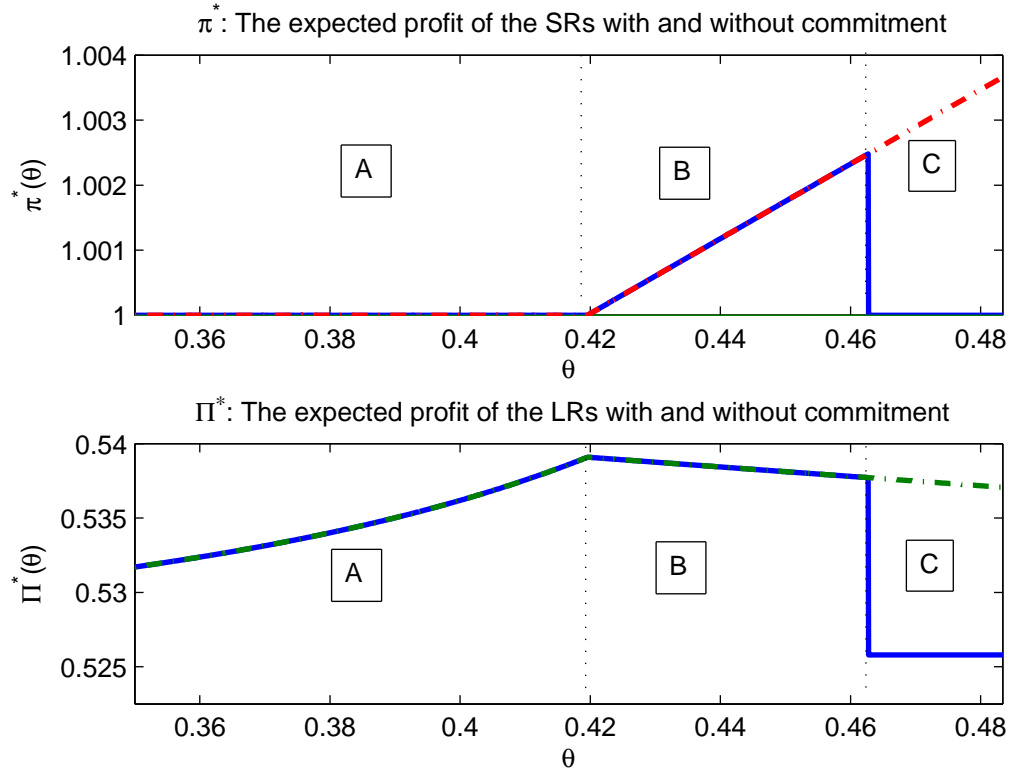
Figure 7. Expected profits for the SR, π^*, (top panel) and the LR (bottom panel), Π^*, as a function of θ. The first dashed vertical line corresponds to $\widehat{\theta} = .4196$. The continuous line plots the expected profits when the Pareto superior equilibrium is chosen. In regions A and B, the delayed trading equilibrium exists and it is the Pareto superior equilibrium. In region C, which corresponds to $\theta \in (.4628, .4834]$, the delayed trading equilibrium no longer exists as $P_d^* < \delta\eta\rho$ and the sole equilibrium is the immediate trading equilibrium. The dashed line corresponds to the expected profits when the SRs can commit to liquidate assets in state ω_{2L}.
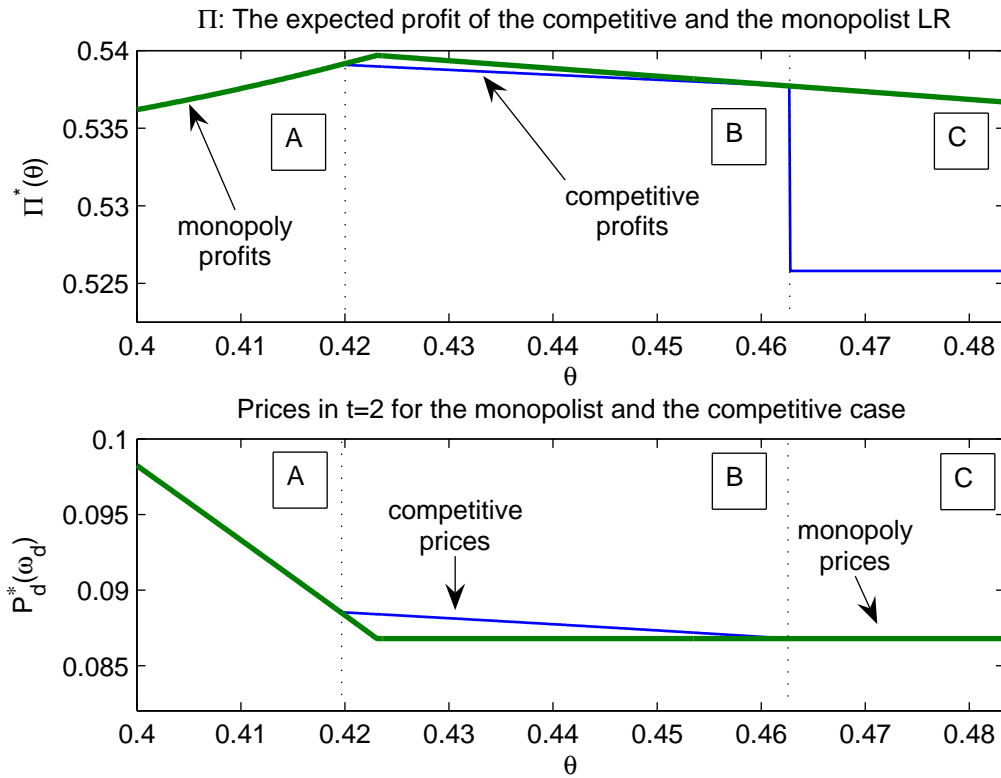
Figure 8. Top panel: Expected profits of the monopolist (the thick line) and the competitive LR (the thin line) as a function of θ. Bottom panel: Prices in state ω_d, $P_d^*(\omega_d)$, in the monopolist (the thick line) and the competitive (the thin line) LR case.