Inside and Outside Liquidity*

Patrick Bolton
Columbia University

Tano Santos
Columbia University

Jose Scheinkman
Princeton University

November 2008

Abstract

We consider a model of liquidity demand arising from a possible maturity mismatch between asset revenues and consumption. This liquidity demand can be met with either cash reserves (*inside liquidity*) or via asset sales for cash (*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 and outside liquidity? and (c) if not, what kind of interventions can best restore efficiency? We argue that a key determinant of the optimal liquidity mix is the expected timing of asset sale decisions. An important source of inefficiency in our model is the presence of asymmetric information about asset values, which increases the longer a liquidity trade is delayed. We establish existence of an immediate-trading equilibrium, in which asset trading occurs in anticipation of a liquidity shock, and sometimes also of a delayed-trading equilibrium, in which assets are traded in response to a liquidity shock. We show that, when it exists, the delayed-trading equilibrium is efficient, despite the presence of adverse selection.

*We thank Rafael Repullo and Lasse Pedersen as well as participants at workshops and seminars at several universities and at the 2008 NBER Summer Institute on Risks of Financial Institutions for their comments and suggestions.

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.

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 and inside liquidity? and if not, what kind of interventions can 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 who are only ready to buy these securities at a much higher discount are not explicitly modeled. An important potential source of inefficiency in practice and in our model is asymmetric information about asset values between short and long-horizon investors. That is, even when short-run investors turn to knowledgeable long-run investors 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 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 depends critically on the ex-ante portfolio-composition decisions of both the short and long-run investors. 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 cash 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. 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 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 in a delayed trading equilibrium, one should expect – and we indeed establish – that equilibrium asset prices 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 equilibrium. However, 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.

In the presence of adverse selection, however, outside liquidity involves a dilution of ownership cost so that short-run investors prefer to partially rely on inefficient 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 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 writing a long-term contract for liquidity with long-horizon investors. Such a contract takes the form of an investment fund set up by long-horizon investors in which short-run investors can co-invest and in return receive state-contingent payments. Under complete information such a fund arrangement would always dominate any equilibrium allocation achieved

[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.

through spot trading of assets for cash. However, when the long-horizon investor who manages the fund also has private information about the realized returns on the fund's investments then, as we show, the long-term contract cannot always achieve a more efficient outcome than the delayed-trading equilibrium. Indeed, the fund manager's private information then prevents the fund from making fully efficient state-contingent transfers to the short-run investor, thus raising the cost of providing liquidity.

Given that neither financial markets nor long-term contracts for liquidity can achieve a fully efficient outcome, the question naturally arises whether some form of public intervention may provide an efficiency improvement. There are two market inefficiencies 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–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 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 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 VAR type regulations which require divestitures of assets 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 is a policy of public provision of liquidity by a central bank, which guarantees a minimum price for assets by lending against collateral. Such a policy could improve efficiency simply by inducing intermediaries to rely less on inefficient inside liquidity (and even if short-run investors end up not relying on the lending facility in equilibrium). 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 literatures on banking and liquidity crises, and the limits of arbitrage. Our analysis differs from the main contributions in these 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 literatures. We briefly discuss the most related contributions in each of these literatures.

Consider first the banking literature. Diamond and Dybvig (1983) and Bryant (1980) provide the first models 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 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 models 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, as Jacklin (1987) has shown. 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 by firms 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)[2], 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

[2]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.

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 Dow and Gorton (1993) and Shleifer and Vishny (1997). The typical model in this literature (e.g. Kyle and Xiong, 2001 and Xiong, 2001) also allows for outside liquidity and generates episodes of fire-sale pricing–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 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 mismatch 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 mismatch 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. Our interim trading phase is divided in two distinct periods: at date 1 there is an aggregate, publicly observable, shock that affect both the liquidity and average quality of the risky assets held by short-run investors; at date 2 there are privately observed idiosyncratic shocks affecting risky assets held by short-run investors. Thus asset-trading at this date takes place under asymmetric information.

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 0 and no endowments at subsequent dates. There is also a unit mass of *long-run investors*, each with $\kappa > 0$ units of endowment at $t = 0$ and again no endowments at subsequent dates. Their utility function is simply given by

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

with $C_t \geq 0$.

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

II.B Assets

For simplicity and with almost no loss of generality we assume that the two types of investors have access to different investment opportunity sets. Both types can hold *cash* with a gross per-period rate of return of one. LR investors can also invest in a decreasing-returns-to-scale *long-maturity* asset that returns $\varphi(x)$ at date 3 for an initial investment at date 0 of $x = (\kappa - M)$, where $M \geq 0$ is the LRs' cash holding.

SR investors 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 payoff of the risky asset is the only source of uncertainty in the model and is shown in Figure 1. For simplicity we assume that there is a first aggregate *maturity shock* that affects all 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.

Formally, the risky asset either pays ρ at date 1 (in state $\omega_{1\rho}$), which occurs with probability λ, or it pays a return at a subsequent date, with probability $(1-\lambda)$. In that case the asset yields either a return $\widetilde{\rho}_2 \in \{0, \rho\}$ at date 2, or a late return $\widetilde{\rho}_3 \in \{0, \rho\}$ at date 3. After date 1 shocks are idiosyncratic and are represented by two separate *i.i.d.* random variables: (i) an individual asset can either mature at date 2, with probability θ, or at date 3 with probability $(1 - \theta)$ (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 idiosyncratic shocks is private information to the SR holding the risky asset. We denote by m the amount of cash held by SRs and $1 - m$ 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 LRs. In particular SRs can sell claims to the risky assets in state ω_{1L}, which we refer to as the *immediate trading date*. Alternatively SRs can postpone trading until date 2 and thus get another chance for the asset to pay off before date 3. Recall that the idiosyncratic shocks are private information to SRs so that only SRs they know whether they are in states $\omega_{2\rho}$, ω_{20}, or ω_{2L}. In state $\omega_{2\rho}$ SRs collect the asset payoff and consume it. In state ω_{20} SRs have worthless assets and, in state ω_{2L} they have assets with an expected payoff of $\eta\rho$ realized at date 3. In state ω_{2L} SRs can either liquidate the asset or carry it to date 3. LR investors only know that the assets sold at date 2 could be either be "lemons" (assets sold by SRs in state ω_{20}) or good assets which can still pay at date 3.

II.D Assumptions

We introduce assumptions on payoffs that focus the analysis on the economically interesting outcomes and thus considerably shorten the discussion of the model. We begin with assumptions on the assets. First we assume that outside liquidity is costly:

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

The assumption that $\varphi''(\cdot) < 0$ captures the fact that the opportunities that these long assets represent are scarce and cannot be exploited with limit. We also assume that LRs always want to invest some fraction of their endowment in this long asset:

$$\lim_{x \to 0} \varphi'(x) = +\infty.$$

The key assumption here though is that $\varphi'(\kappa) > 1$. This implies that if LRs carry cash it must be to acquire assets with expected returns at least as high as $\varphi'(\kappa)$. Given our assumption of risk neutrality this can only occur if asset purchases occur at *cash-in-the-market* prices. That is, assets must trade in equilibrium at prices that are below their 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 the risky asset may be more attractive than cash when the asset can be resold at a reasonable price:

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

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.

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

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

II.E Discussion

A central feature of the model is the timing of the aggregate and idiosyncratic shocks. The aggregate shock reveals whether the risky assets pay off ρ at date 1, or whether there is a deterioration of this entire class of assets as well as a postponement of payoffs. The expected payoff of the risky assets in state ω_{1L} is $\eta\rho$, and the payoff may be realized at either dates 2 or 3. Following date 1, additional news accrue only to the holders of the asset, which is why there is an adverse selection premium at that point. In state ω_{1L} all agents are symmetrically informed, which is why asset sales that take place at date 1 do not include an adverse selection premium whereas they do at date 2.

Once the initial portfolios are set, a critical problem agents face is the decision whether to trade assets at dates 1 or 2. The parameter θ plays an important role for this decision. Indeed a high value of θ means that if the risky asset has not paid off at date 1, the probability that it does pay off at date 2 rather than at date 3 is also high. The risky asset is more attractive to SRs when θ is higher, for they care about consumption at date 2 more than they do about consumption at date 3. But at the same time, the higher the value of θ, the more severe the adverse selection problems are at date 2. Similarly for the LRs the question is whether to acquire high quality assets at date 1 at high prices or to trade in a market subject to adverse selection but at better prices. How these tradeoffs affect the (ex-ante) portfolio decisions of both SRs and LRs is the central issue explored in this paper.

We conclude this section with a brief discussion of our assumptions on the asset-opportunity sets of each group of agents. Specifically, we have assumed that SRs cannot invest in the long asset and LRs cannot invest in the risky asset. As $\lim_{x\to 0}\varphi'(x) = +\infty$, 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.

We have also assumed that the LRs cannot invest in the risky asset, but this turns out to be an innocuous assumption. If the returns to holding cash for LRs are lower than the expected return of the risky asset, $\rho\,[\lambda + (1-\lambda)\,\eta]$, then LRs would not want to carry any cash at all 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 higher than the risky asset's expected return then clearly LRs would choose not to invest in it. The first situation is obviously not an interesting one, so that we simply assume that the LRs cannot invest in the risky asset.

III. EQUILIBRIUM

Given that all SRs and LRs are ex-ante identical, we shall restrict attention to symmetric competitive equilibria. We also restrict attention to *pooling equilibria*, for which SR asset trades are the same in states ω_{20} and ω_{2L}.[3] Recall that trade between SRs and LRs can only take place in spot markets at dates 1 and 2, and there are only two information sets at which there are potentially strictly positive gains from trade, ω_{1L} and $\{\omega_{20}, \omega_{2L}\}$. 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_{1L}, ω_{20} and ω_{2L}. LR investors, on the other hand, are unable to distinguish states ω_{20} and ω_{2L} in any pooling equilibrium, and therefore can only condition their trades on their information sets ω_{1L} and $\{\omega_{20}, \omega_{2L}\}$. We denote by $q\,(\omega_{1L})$ the amount of the risky asset supplied by an SR in state ω_{1L} and by $q\,(\omega_{20}, \omega_{2L})$ the amount supplied at date 2. Similarly, we denote by $Q\,(\omega_{1L})$ $(Q\,(\omega_{20}, \omega_{2L}))$ the amount of the risky asset that LR investors acquire in state ω_{1L} $(\{\omega_{20}, \omega_{2L}\})$.

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 date 1 at price P_1 and at date 2 at price P_2. Their objective function is then

$$\pi\,[m, q\,(\omega_{1L})\,, q\,(\omega_{20}, \omega_{2L})] \quad = \quad m + \lambda\,(1-m)\,\rho$$

[3]Pooling equilibria are easily supported by pessimistic out-of-equilibrium beliefs, which attribute any deviation to the lemon type (that is type ω_{20}).

$$
\begin{aligned}
&+ \quad (1-\lambda)\, q\left(\omega_{1L}\right) P_1 \\
&+ \quad (1-\lambda)\, \theta\eta \left[(1-m) - q\left(\omega_{1L}\right)\right] \rho \\
&+ \quad (1-\lambda)\left(\theta\left(1-\eta\right)+1-\theta\right) q\left(\omega_{20},\omega_{2L}\right) P_2 \\
&+ \quad \delta\left(1-\lambda\right)\left(1-\theta\right)\eta\left[(1-m) - q\left(\omega_{1L}\right) - q\left(\omega_{20},\omega_{2L}\right)\right]\rho
\end{aligned} \tag{2}
$$

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.[4] 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 $q\left(\omega_{1L}\right)$ of the risky asset for a unit price P_1.

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 they still own: $[(1-m) - q\left(\omega_{1L}\right)]\rho$. Another outcome at date 2 is that the asset yields a zero return. This occurs with probability $(1-\lambda)\theta\left(1-\eta\right)$. In that state of nature the SR chooses optimally to sell its pooling position in the risky asset $q\left(\omega_{20},\omega_{2L}\right)$ for the price P_2. Finally, with probability $(1-\lambda)(1-\theta)$ the asset only matures at date 3. The SR again sells its pooling position $q\left(\omega_{20},\omega_{2L}\right)$ at price P_2, as long as this price exceeds what SR gets by holding the asset to maturity at date 3: $\delta\rho(1-\lambda)(1-\theta)\eta$. The SR's optimal investment program is therefore given by:

$$
\max_{m,q(\omega_{1L}),q(\omega_{20},\omega_{2L})} \pi\left[m, q\left(\omega_{1L}\right), q\left(\omega_{20},\omega_{2L}\right)\right] \tag{P_{SR}}
$$

subject to

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

and

$$
0 \leq q\left(\omega_{1L}\right) \leq 1-m
$$

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_{1L} and $\left(\omega_{20},\omega_{2L}\right)$ it cannot sell more than what it holds.

III.B The LR optimization problem

LR investors 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

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

risky assets to purchase at prices P_1 and P_2. 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 risky assets. With this observation in mind we can write the payoff an LR investor that purchases $Q(\omega_{1L})$ at date 1 and $Q(\omega_{20}, \omega_{2L})$ at date 2, as follows:

$$\begin{aligned} \Pi\left[M, Q(\omega_{1L}), Q(\omega_{20}, \omega_{2L})\right] &= M + \varphi(\kappa - M) \\ &+ (1-\lambda)\left[\eta\rho - P_1\right] Q(\omega_{1L}) \\ &+ (1-\lambda) E\left[\tilde{\rho}_3 - P_2 \middle| \mathcal{F}\right] Q(\omega_{20}, \omega_{2L}) \end{aligned} \tag{3}$$

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 $Q(\omega_{1L})$ in risky assets at unit price P_1 at date 1. Indeed, the expected gross return of a risky asset in state ω_{1L} is $\eta\rho$. The last term is the net return from trading in states $(\omega_{20}, \omega_{2L})$. 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 states $(\omega_{20}, \omega_{2L})$ 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 states $(\omega_{20}, \omega_{2L})$ in an equilibrium that prescribes no sales in these states, 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 for sure. In addition, LRs assume that SRs that weakly prefer to sell at price P_2 will sell all their remaining holdings in states $(\omega_{20}, \omega_{2L})$ whereas SRs that prefer not to sell, will not sell any units.

The LR investor's program is thus:

$$\max_{M, Q(\omega_{1L}), Q(\omega_{20}, \omega_{2L})} \Pi\left[M, Q(\omega_{1L}), Q(\omega_{20}, \omega_{2L})\right] \tag{$\mathcal{P}_{LR}$}$$

subject to

$$0 \leq M \leq \kappa \tag{4}$$

and

$$Q(\omega_{1L}) P_1 + Q(\omega_{20}, \omega_{2L}) P_2 \leq M \quad \text{with} \quad Q(\omega_{1L}), Q(\omega_{20}, \omega_{2L}) \geq 0 \tag{5}$$

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 (pooling) rational expectations competitive equilibrium is a vector of portfolio policies $[m^*, M^*]$, supply and demand choices $[q^*(\omega_{1L}), Q^*(\omega_{1L}), Q^*(\omega_{20}, \omega_{2L})]$ and prices $[P_1^*, P_2^*]$ such that (i) at these prices $[m^*, q^*(\omega_{1L})]$ solves $\mathcal{P}_{SR}$ and $[M^*, Q^*(\omega_{1L}), Q^*(\omega_{20}, \omega_{2L})]$ 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

$$M_i^* > 0,\ q^*(\omega_{1L}) = Q^*(\omega_{1L}) = 1 - m_i^* \qquad \text{and} \qquad q^*(\omega_{20}, \omega_{2L}) = Q^*(\omega_{20}, \omega_{2L}) = 0.$$

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

$$P_{1i}^* = \frac{M_i^*}{1 - m_i^*} \geq \frac{1 - \lambda\rho}{1 - \lambda}. \tag{6}$$

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

To gain some intuition notice first that the immediate trading equilibrium has to meet the first order conditions for m and M, which are respectively:

$$P_{1i}^* \geq \frac{1 - \lambda\rho}{1 - \lambda} \qquad \text{and} \qquad \lambda + (1 - \lambda)\frac{\eta\rho}{P_{2i}^*} = \varphi'(\kappa - M_i^*), \tag{7}$$

when $m_i^* < 1$ and $M_i^* > 0$.[5] These expressions follow immediately from the maximization problem $\mathcal{P}_{SR}$ when we set $q^*(\omega_{1L}) = 1 - m_i^*$, and from problem $\mathcal{P}_{LR}$.

[5]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 $m_i^* = 1$.

Next to determine the equilibrium price, take P_{1i}^* to be the unique solution to the equation:

$$\lambda + (1-\lambda)\frac{\eta\rho}{P_{1i}} = \varphi'(\kappa - P_{1i}), \tag{8}$$

and assume first that the solution to (8) is such that

$$P_{1i} > \frac{1-\lambda\rho}{1-\lambda}.$$

In that case we can set $m_i^* = 0$, so that SRs are fully invested in the risky asset, and also $P_{1i}^* = M_i^*$, which by construction satisfies the LR's first order condition. Moreover, by assumption A1 it must also be the case that $M_i^* < \kappa$.

A key step in the construction of the immediate trading equilibrium then, is that the price at date 2, P_{2i}^*, has to be such that both SRs and LRs have incentives to trade at date 1 and not at date 2. That is, it has to be the case that

$$P_{1i}^* \geq \theta\eta\rho + (1-\theta\eta) P_{2i}^* \qquad \text{and} \qquad \frac{\eta\rho}{P_{1i}^*} \geq \frac{\mathbf{E}[\widetilde{\rho}_3|\mathcal{F}]}{P_{2i}^*}. \tag{9}$$

The first expression in (9) states that SRs prefer to sell assets at date 1 for a price P_{1i}^* rather than carrying it to date 2. Indeed if they do the latter, then with probability $\theta\eta$ the risky asset pays off ρ and with probability $(1-\theta\eta)$ they end up in states $(\omega_{2L}, \omega_{20})$ in which the SRs can sell the asset at price P_{2i}^*. If the price P_{2i}^* is low enough then SRs prefer to sell the asset at date 1.[6]

The expression on the right hand side of (9) states that the expected return of acquiring the asset in state ω_{1L} is higher than in states $(\omega_{20}, \omega_{2L})$. To guarantee this outcome it is sufficient to set $P_{2i}^* < \delta\eta\rho$ for in this case SRs in state ω_{2L} would prefer to carry the asset to date 3 rather than selling it for that price. This then only leaves "lemons" in the market at date 2. LRs, anticipating this outcome, set their expectations accordingly, $\mathbf{E}[\widetilde{\rho}_3|\mathcal{F}] = 0$, and therefore for any strictly positive price $P_{2i}^* < \delta\eta\rho$ LRs prefer to acquire assets in state ω_{1L}.

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

$$P_{1i} \leq \frac{1-\lambda\rho}{1-\lambda}, \tag{10}$$

and set P_{1i}^* equal to the right hand side of (10). At this price, SRs are indifferent on the amount of cash carried. Then the solution to the LR's first order condition is such that:

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

[6] See expression (25) in the appendix for a precise upper bound on P_{2i}^* that has to hold to provide incentives for the SRs to sell at date 1 rather than at date 2.

(see the expression (7)).

It is then sufficient to set $m_i^* \in [0,1)$ such that:

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

which is always possible.[7] Finally, the choice of P_{2i}^* can be taken to be the same 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, $P_{1i}^* < \eta\rho$, otherwise there would be no incentive for the LRs to carry cash.

The delayed-trading equilibrium. Proposition 2 establishes the existence of a delayed trading equilibrium..[8]

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

$$q^*(\omega_{1L}) = Q^*(\omega_{1L}) = 0 \qquad \text{and} \qquad q^*(\omega_{20},\omega_{2L}) = Q^*(\omega_{20},\omega_{2L}) = (1-\theta\eta)(1-m_d^*).$$

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

$$P_{2d}^* = \frac{M_d^*}{(1-\theta\eta)\left(1-m_d^*\right)} \geq \frac{1-\rho\left[\lambda+(1-\lambda)\,\theta\eta\right]}{(1-\lambda)\,(1-\theta\eta)}. \tag{12}$$

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

The intuition of how we construct the delayed-trading equilibrium is broadly similar to the one for the immediate-trading equilibrium, with a few 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 date 2, which would destroy the delayed-trading equilibrium, as only lemons would be traded at date 2. Second, a key difference with the immediate-trading equilibrium 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

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

[8]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.

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

date 2.[10] As a result cash-in-the-market pricing under delayed trading is given by:

$$P_{2d}^* = \frac{M_d^*}{(1-\theta\eta)\left(1-m_d^*\right)}.$$

Notice that now the mass of risky assets supplied in the market in states $(\omega_{20}, \omega_{2L})$ is given by $(1-\theta\eta)(1-m_d^*)$. Thus delaying asset liquidation introduces both an adverse selection effect which depresses prices, and a lower supply of the risky asset, which, other things equal, increases prices.

As under the immediate-trading equilibrium, to support a delayed-trading equilibrium requires that both SRs and LRs have incentives to trade at date 2 rather than at date 1, which means that

$$P_{1d}^* \leq \theta\eta\rho + (1-\theta\eta)\, P_{2d}^* \qquad \text{and} \qquad \frac{\eta\rho}{P_{1d}^*} \leq \frac{\mathbf{E}\left[\widetilde{\rho}_3|\mathcal{F}\right]}{P_{2d}^*}, \tag{13}$$

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 state ω_{1L} has to be in the interval

$$P_{1d}^* \in \left[\frac{1-\theta\eta}{1-\theta} P_{2d}^*, \theta\eta\rho + (1-\theta\eta) P_{2d}^*\right].$$

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

It is 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_{2d}^* in (12) reduces to the lower bound in (6) for P_{1i}^*. The only difference between dates 1 and 2 is thus precisely the occurrence of an idiosyncratic shock that reveals to the SRs the true value of the risky asset. When $\theta = 0$ there is no informative idiosyncratic signal to be obtained as at date 1. This feature of our model plays an important role in what follows.

Before we close this section we introduce the example that we use throughout to illustrate our results. As the parameter θ plays a critical role in our analysis it is the focus of our comparative statics, and we parameterize the set of economies that we consider throughout by the different values that θ 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.

[10]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 $t = 2$.

Example. In this example the parameter values are:

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

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 hold, as well as assumption A4 below. All the Figures in our paper refer to this example. A summary of the main results is as 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 that δ is small enough fails to hold as we explain below. $\square$

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

How does the composition of inside and outside liquidity vary across equilibria? To build some intuition on this question it is useful to illustrate the immediate and delayed-trading equilibria that obtain in our example when $\theta = .35$. Figure 2 represents the immediate and delayed-trading equilibria 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*.[11] In the figure we display the isoprofit lines for both the immediate and delayed-trading equilibrium.[12] 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 SR and LR isoprofit curves.

Consider first the immediate trading equilibrium, located at the point marked $(M_i^*, m_i^*) = (.0169, .9358))$. There are two isoprofit curves going through that point; the straight line corresponds to the SR, and the dashed-dotted line corresponds to the LR isoprofit curve. In fact the

[11] Assumption A3 simply says that the slope of the isoprofit lines at $M = 0$ at date 1 are such that there are gains from trade: the LR isoprofit curve is “flatter” than the SR isoprofit line.

[12] 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.

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 entirely 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 states $(\omega_{20}, \omega_{2L})$ is lower than in state ω_{1L} 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_{1L}, so that SRs must be compensated with a high price relative to the price in states $(\omega_{20}, \omega_{2L})$, 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,

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

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 $\theta' \in (0, \overline{\theta}]$ 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.[13]

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_{1L} or $(\omega_{20}, \omega_{2L})$, 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_{1L}, 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_{1L}.

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 states $(\omega_{20}, \omega_{2L})$. 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.

[13]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 $\theta' < \overline{\theta}$.

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 states $(\omega_{20}, \omega_{2L})$, as measured by changes in θ, affect equilibrium outcomes. In particular, we are interested in understanding how equilibrium cash 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 states $(\omega_{20}, \omega_{2L})$ is worsened and therefore equilibrium prices P_{2d}^* are likely to be lower. These lower prices that SRs face in states $(\omega_{20}, \omega_{2L})$ 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 states $(\omega_{20}, \omega_{2L})$,

$$s_d^*(\omega_{20}, \omega_{2L}) \equiv (1 - m_d^*(\theta))(1 - \theta\eta) \tag{15}$$

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 2 in the delayed-trading equilibrium, which is defined as

$$R_d^*(\omega_{20}, \omega_{2L}) \equiv \frac{(1-\theta)\eta\rho}{(1-\theta\eta)P_{2d}^*} \tag{16}$$

The next proposition establishes how these countervailing effects net out and how M_d^*, m_d^*, $s_d^*(\omega_{20}, \omega_{2L})$, and $R_d^*(\omega_{20}, \omega_{2L})$ vary with θ. Throughout we assume, of course, that $\theta \leq \overline{\theta}$.

Proposition 4. (*Comparative statics.*) 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 there exists a unique $\widehat{\theta} \in [0, \overline{\theta}]$, possibly $\widehat{\theta} = \overline{\theta}$, such that:

I. 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 $\theta \in [\widehat{\theta}, \overline{\theta}]$ 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 $\theta \in [\widehat{\theta}, \overline{\theta}]$.

II. 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 $\theta \in (\widehat{\theta}, \overline{\theta}]$.

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

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 in states $(\omega_{20}, \omega_{2L})$ ($R_d^*(\omega_{20}, \omega_{2L})$ and P_{2d}^*), 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 states $(\omega_{20}, \omega_{2L})$. 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_{2d}^* changes with θ. As can be seen, P_{2d}^* 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 states $(\omega_{20}, \omega_{2L})$. 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 $\theta < \widehat{\theta}$, $R_d^*(\omega_{20}, \omega_{2L})$ is an increasing function of θ: The expected payoff of the risky asset in the delayed-trading date

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

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 $R_d^*(\omega_{20}, \omega_{2L})$ as a function of θ in this range.

In conclusion then, for $\theta \in [0, \widehat{\theta}]$ the more severe the adverse selection problem, 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 is due to the drop in prices, which makes the risky asset more attractive to the LRs in states $(\omega_{20}, \omega_{2L})$. The larger the liquidity correction at date 2, the more attractive it is for LRs to carry cash and 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 states $(\omega_{20}, \omega_{2L})$, 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.[14]

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 $\theta \leq \widehat{\theta} = .4196$ 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 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 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 states $(\omega_{20}, \omega_{2L})$. The planner's

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

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 states $(\omega_{20}, \omega_{2L})$. 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 the value of commitment to trade

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 3 rather than trading it for P_{2d}^{*} at date 2.[15] 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 date 3 whenever

$$\widetilde{P}_d\left(\omega_{20}, \omega_{2L}\right) < \delta\eta\rho, \tag{18}$$

where $\widetilde{P}_d\left(\omega_{20}, \omega_{2L}\right)$ is the candidate price for the risky asset at date 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 $\widetilde{P}_d(\omega_{20}, \omega_{2L})$ in state ω_{2L}. In this case, the LRs anticipating that the pool of

[15]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.

assets supplied in states $(\omega_{20}, \omega_{2L})$ would also include assets of higher quality would be willing to bring more outside liquidity than in the immediate trading equilibrium. As shown above, this is always Pareto improving in our framework because it substitutes inside with outside liquidity.

IV. LONG-TERM CONTRACTS FOR LIQUIDITY

So far we have only allowed SR and LR investors to trade assets for cash at dates 1 and 2, after they have each made their portfolio investment decisions at date 0. Given that SR investors could reduce the lemons' cost associated with outside liquidity by committing to sell their long-term assets at date 2, the question naturally arises whether some form of long term contract between an SR and LR at date 0 may improve on the outcome obtained in the immediate and delayed-trading equilibria. We address this question in this section and allow LRs to write long-term contracts with SRs at date 0, whereby an SR invests her endowment in a fund managed by an LR and in return LR promises state-contingent payments to SR.

More precisely, the fund allocates the total endowment $(1 + \kappa)$ of the two investors in a portfolio that may comprise the LR long-run asset, the SR risky asset, and cash. The LR manager in turn promises state contingent payments

$$\{C_1(\omega_{1\rho}), C_1(\omega_{1L}), C_2(\omega_{2\rho}), C_2(\omega_{20}), C_2(\omega_{2L}), C_3(\omega_{3\rho}), C_3(\omega_{30})\}$$

to SR for investing her unit of endowment in the fund. If the LR investor chooses to invest part of the endowments in the SR risky asset then it is the LR investor who privately observes the realization of the idiosyncratic shocks affecting the risky asset at dates 2 and 3. Accordingly, an LR investor may also face incentive compatibility constraints, which limit the efficiency of the long-term contract.[16]

Throughout this section we assume that $\delta\varphi'(k) < 1$, so that the LR investor would not find it optimal to simply invest the whole endowment $(1 + \kappa)$ in the long asset and repay the SR at date 3. Incentive compatibility for the LR fund manager then requires first that:

$$C_3(\omega_{3\rho}) = C_3(\omega_{30})$$

[16]Note that if the SR investor can also observe the realization of idiosyncratic shocks then the asymmetric information problem in the delayed-trading equilibrium would not be present, so that the long-term contract at date 0 would clearly yields a superior outcome. The more consistent and interesting case, however, is when the observation of idiosyncratic shocks is private information to the manager of the risky asset.

at date 3, for otherwise the LR simply announces the state which involves the lower payment to SR, and second that at date 2,

$$C_2(\omega_{2L}) + C_3(\omega_{30}) = C_2(\omega_{20}) + C_3(\omega_{30}) = C_2(\omega_{2\rho}) + C_3(\omega_{3\rho})$$

Otherwise, again, LR would always announce the state

$$(\hat{\omega}_2, \hat{\omega}_3) \in \{(\omega_{2L}, \omega_{30}), (\omega_{2L}, \omega_{3\rho}), (\omega_{20}, \omega_{30}), (\omega_{2\rho}, \omega_{3\rho})\}$$

which involves the lowest total payout $C_2(\hat{\omega}_2) + C_3(\hat{\omega}_3)$.

The long term contract between LR and SR must also satisfy the SR and LR participation constraints. An individual SR investor has the choice of entering into a long-term contract at date 0 with an LR or investing in a financial intermediary. The return offered by financial intermediaries, in turn, depends on which equilibrium they expect. If they expect the immediate-trading equilibrium then SR participation in a long-term requires that LR offer at least the same payoff as financial intermediaries in that equilibrium. On the other hand, if they expect the delayed-trading equilibrium the LR must offer at least the payoff in the delayed-trading equilibrium.

As is easy to see, when SR investors expect the immediate-trading equilibrium, then any pair of LR and SR investors are weakly better off writing a long-term contract at date 0. Indeed, at worst the contract simply replicates the allocation under immediate trading. But, the contract can also implement other allocations that are not feasible under the immediate trading equilibrium, in particular by investing part of the SR endowment in the LR project and by specifying payments to SR that (weakly) dominate the allocation under immediate trading:

$$C_1(\omega_{1\rho}) \geq m_i^* + \lambda(1 - m_i^*)\rho$$

$$C_1(\omega_{1L}) \geq m_i^* + (1 - m_i^*)P_i^*(\omega_{1L}).$$

It follows that:

Proposition 6. (*Ranking of long-term contract and immediate-trading equilibrium*) The optimal long-term contract weakly (and sometimes strictly) dominates the equilibrium allocation under immediate trading.

In contrast, when SR investors expect the delayed-trading equilibrium, then the long-term contract cannot always replicate the allocation under delayed trading. The reason is that under delayed trading, SRs face different incentive constraints at date 2 from those faced by LRs under the long-term contract.

Under delayed trading, the SR investor seeking to trade assets for cash at date 2 must trade assets at the same price in states ω_{20} and ω_{2L} to be induced to truthfully reveal these two states to LRs. However, in state $\omega_{2\rho}$ there is no trade between SR and LR, and therefore also no need to reveal that state truthfully to LR. In other words, no incentive constraint applies in this state of nature.

Under the long-term contract, however, LR promises transfers to SR in four different realizations of the idiosyncratic shocks $\{(\omega_{20}, \omega_{30}); (\omega_{2\rho}, \omega_{3\rho}); (\omega_{2L}; \omega_{30}); (\omega_{2L}; \omega_{3\rho})\}$ (see Figure X). As we have argued above, the total transfers in each of these four states must be the same to satisfy incentive compatibility. Therefore, LR simply cannot replicate the allocation under the delayed trading equilibrium with a suitable long-term contract. Given that the delayed-trading equilibrium allocation is not in the feasible set for the long-term contract it is not obvious a priori which allocation is superior. To be able to answer this question we must first characterize the optimal long-term contract (when SR requires a payoff at least as high as under the delayed trading equilibrium).

Solving the long-term contracting problem is a somewhat involved constrained optimization problem, as it involves two investment variables (α, M) and seven state-contingent transfers to SR. This problem can be simplified to some extent, as the next lemma establishes, since the combination of all the incentive and feasibility constraints reduce the long-term contracting problem to the determination of optimal values for only: i) the amount $\alpha \in [0, 1]$ invested in the risky SR project, ii) the amount M of cash held by the fund, and iii) payments to SR in states $\omega_{1\rho}$, $\omega_{2\rho}$ and ω_{30}.

Lemma 7. Without loss of generality, any feasible, incentive-compatible long-term contract between LR and SR takes the form:

	$C_2(\omega)$	$C_3(\omega)$
$\omega_{1\rho}$	$M + \alpha\rho$	$C_3(\omega_{1\rho})$
$\omega_{2\rho}$	$C_2(\omega_{2\rho})$	$C_3(\omega_{2\rho})$
ω_{20}	M	$C_3(\omega_{30})$
ω_{2L}, ω_{30}	M	$C_3(\omega_{30})$
$\omega_{2L}, \omega_{3\rho}$	M	$C_3(\omega_{3\rho})$

Given that SR discounts date 3 consumption by δ it seems inefficient to offer any date 3 consumption to SR. Still, we cannot rule out that $C_3(\omega) > 0$ for either $\omega \in \{\omega_{1\rho}, \omega_{2\rho}, \omega_{30}, \omega_{3\rho}\}$ since a date 3 transfer in one state may be required for LR to satisfy all incentive constraints he faces. That is, to be able to credibly disclose that the realized state is ω_{20}, for example, LR may have to promise a high transfer $C_3(\omega_{30})$ at date 3. Nevertheless, intuition suggests that if δ is very small, λ sufficiently large, and the opportunity cost of holding cash for LR is bounded, then the optimal contract ought to specify $C_3(\omega_{1\rho}) = C_3(\omega_{2\rho}) = 0$. This is indeed the case as the following Proposition establishes.

Proposition 8. Suppose that δ is close to zero and that

$$\eta(1-\lambda)\rho + \varphi(0) \leq \varphi(\kappa), \tag{\textbf{A5}}$$

then the optimal long-term contract is such that $C_3(\omega_{1\rho}) = C_3(\omega_{2\rho}) = 0$.

With this characterization of the optimal long-term contract we are able to numerically solve for the optimal contract and to compare LR payoffs under the contract to LR equilibrium payoffs under the delayed trading equilibrium. We then show that for some parameter values, the long-term contract is dominated by the delayed trading equilibrium outcome for high values of θ.

The economic logic behind this result is that when θ is high then the SR risky asset already matures most of the time at dates 1 or 2. The added value of additional liquidity offered by LR through a long-term contract is then not that high. In addition, when θ is high LR also faces high costs of meeting incentive constraints under the long-term contract. To be able to credibly claim that the risky asset did not yield a return ρ at either dates 1 or 2, LR must commit to wasteful date 3 payments $C_3(\omega_{3\rho}) = C_3(\omega_{30}) = \alpha\rho$ which SR does not value much. The deadweight cost of these distortions exceeds the benefit of extra liquidity insurance, which is why the delayed-trading equilibrium outcome is superior.

Example. In our example we keep θ as a free parameter and fix the other parameters to the following values:

$$\lambda = .7 \qquad \eta = .4 \qquad \rho = 1.25 \qquad k = .12 \qquad \delta = .1 \quad \text{and} \quad \varphi(x) = x^{\gamma} \quad \text{with} \quad \gamma = .19.$$

Note that all our assumptions are then met as long as $\theta \leq .8148$, which is the first value of θ for which the assumption $[\lambda + (1-\lambda)\eta[\theta + (1-\theta)\delta]] \leq 1$ is violated. Accordingly our plots below are restricted to the interval $\theta \in [0, .8148]$.

The payoffs of SR and LR under the long-term contract are given by respectively:

$$
\begin{aligned}
\pi_x^* &= \lambda[M + \alpha\rho + \delta C_3(\omega_{1\rho})] + \\
&\qquad (1-\lambda)[\theta\eta(C_2(\omega_{2\rho}) + \delta C_3(\omega_{2\rho})) + (1-\theta\eta)(M + \delta C_3(\omega_{30}))]
\end{aligned}
$$

and

$$
\begin{aligned}
\Pi_x^* &= M + \varphi[\kappa + (1-\alpha) - M] + \\
&\quad \lambda[\alpha\rho - (C_2(\omega_{1\rho}) + C_3(\omega_{1\rho}))] + \\
&\quad (1-\lambda)[\eta\alpha\rho - (M + C_3(\omega_{30}))].
\end{aligned}
$$

In our numerical example we set $\pi_x^* = \pi_d^*$ the SR payoff in the delayed trading equilibrium to satisfy the SR participation constraint. Numerical computations show that for the chosen parameter values the optimal long-term contract is such that $C_3(\omega_{1\rho}) = C_3(\omega_{2\rho}) = C_3(\omega_{30}) = 0$, and therefore that $C_2(\omega_{2\rho}) = M$.

Note that unlike in the previous example, a delayed trading equilibrium always exists here. In the top panel of Figure 9 we graph the expected utility of the SR in the delayed trading equilibrium as a function of θ whereas the bottom panel shows the expected utility of the LR in the delayed trading equilibrium, Π_d^*, as well as the LR's expected payoff under the long-term contract, Π_x^*. For $\theta > \tilde{\theta}$ this payoff is less than what LR gets in the delayed trading equilibrium. The bottom panel of Figure 10 shows that when θ increases, the amount of cash carried by the LR to fulfill his commitments under the long-term contract increases, making the contract less efficient, in sharp contrast with the total amount of cash $m_d^* + M_d^*$ carried by both the LRs and SRs in the delayed trading equilibrium, shown in the top panel of the same figure. This increase in cash follows because the expected payoff of SR in the delayed trading equilibrium increases with θ. Incentive constraints limit the difference in payments in states $\omega_{2\rho}$ and ω_{20}, and since payments at date 3 are very inefficient, the contract specifies higher payments at date 2, which requires carrying more cash.

V. REGULATORY IMPLICATIONS

V.A Minimum reserve requirements and the scope of regulation

The source of liquidity demand is a potential maturity mismatch problem for SR investors: they prefer to consume at dates 1 and 2, but the risky asset may only pay at date 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 common regulatory response is the imposition of minimum reserve or liquidity requirements that compel banks to maintain a certain percentage of their assets in "cash" or other liquid securities such as treasury bills.

To analyze the effect of these minimum reserve requirements it is useful to turn to the example illustrated in Figure 3, where the efficient delayed trading equilibrium is such that SRs are fully invested in the risky asset, $m_d^* = 0$. Clearly *any* reserve requirement in this situation risks undermining the delayed-trading equilibrium and to leave as the only equilibrium the immediate-trading equilibrium. The general equilibrium effects are such that when SRs are required to carry a minimum amount of cash the returns of cash holdings in the unregulated sector go down, as SRs now invest less and sell fewer risky assets. This in turn makes the risky assets less attractive to SRs. Also, the resulting equilibrium outcome could be one where the minimum cash position exceeds the regulatory requirement in equilibrium.

V.B Value-at-Risk (VAR) regulation and mandated asset sales

As we have shown, SR investors's option to postpone consumption until date 3 may undermine existence of the efficient delayed-trading equilibrium when $P_{2d}^* \leq \delta\eta\rho$. A regulatory intervention mandating SRs to liquidate assets at date 2 could thus conceivably improve efficiency by creating an environment that helps support a delayed-trading equilibrium. A VAR regulation requiring financial institutions to sell assets when too much value is at risk may be seen as achieving such an intervention.[17] Although a surprise increase in asset sales at date 2 only results in lower prices, the anticipation of such sales would encourage LRs to hold more cash and SRs less, thus increasing ex-ante efficiency. Once again the benefits of such an intervention obtain through a 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 assets, the lemons as well as the good assets.

V.C Public and private provision of liquidity

V.C.1 Inefficient private supply of outside liquidity

So far we have assumed that outside liquidity is supplied competitively, so that LRs

[17] However, VAR regulation that results in financial institutions holding on to their assets and raising more regulatory capital would be counterproductive.

do not take into account the effect that their choices have on the equilibrium price P^*_{2d}. A monopoly LR, on the other hand, would internalize the effect of its supply of liquidity on price. The obvious question then arises whether a monopoly LR might be more efficient in situations where the competitive delayed-trading equilibrium fails to exist?

When $\theta < \widehat{\theta}$, SRs carry a strictly positive amount of inside liquidity $m^*_d > 0$ and make zero profits. All the surplus goes to the LRs, who cannot obtain higher profits by changing their holdings of liquidity. It follows then that in this range the competitive and monopoly solutions are identical. In contrast, when $\theta \geq \widehat{\theta}$, the level of inside liquidity in the competitive equilibrium is $m^*_d = 0$, LRs compete for a fixed supply of the risky asset, and SRs obtain some of the surplus from trade. In this situation, a monopoly LR would gain by restricting its supply of outside liquidity and thereby raising the price P^*_{2d}. 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 states $(\omega_{20}, \omega_{2L})$ are plotted against θ.

Notice first that in region **A**, which corresponds to the case $\theta < \widehat{\theta}$, the prices and profits in a monopoly are identical to those under perfect competition. In region **B**, SRs set the level of inside liquidity to $m^*_d = 0$ and the monopoly LR restricts the supply of outside liquidity so as to capture fully all the gains from trade. This explains why the price of the risky asset in states $(\omega_{20}, \omega_{2L})$ under a monopoly is below the competitive price.

But once θ exceeds a certain threshold the competitive delayed-trading equilibrium no longer exists. The monopoly LR in this case has to set the price for the risky asset to $\delta\eta\rho$ to guarantee a profitable trade at date 2. In this parameter region, region **C** in Figure 8, a monopoly LR will indeed improve efficiency. By carrying and supplying enough outside liquidity the monopolist elicits the supply of risky assets by SRs in state ω_{2L}, thus avoiding the break down of the delayed exchange. Notice that as shown in the top panel of Figure 8, the monopoly's profits are, in this region, above those that obtain in the immediate-trading equilibrium, which is the only one that exists with competitive LRs.[18]

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

In the more plausible situation where outside liquidity is supplied competitively, the price support function played by a monopoly LR has to be taken up the public sector. 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

[18] It is worth emphasizing than in this region SR profits 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}.

up to $\delta\eta\rho$ in region **C**, LRs have an incentive to also carry outside liquidity and purchase the risky assets supplied in states $(\omega_{20}, \omega_{2L})$. In contrast, if the public sector supplies liquidity in regions **A** or **B** it will crowd out liquidity provided by LRs. In this parameter region 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 restore efficiency? A novel dimension of our model is the cross sectional supply of liquidity, which seems to be a core feature of modern financial markets, where different actors outside the regulated financial intermediary sectors that stand ready to absorb asset sales by distressed financial intermediaries.[19] The incentives of the different parties to carry liquidity in our model are driven by their different opportunity costs and different investment horizons. An important question we address is whether a competitive price mechanism would elicit the optimal cross-sectional cash reserve decisions by all the different actors.

A second element in the model that departs from the existing literature is the endogenous timing of asset sales and the deterioration of adverse selection problems over time. Financial intermediaries face the choice of raising liquidity early, or in anticipation of a crisis, before adverse selection problems set in, or in the midst of a crisis at more depressed prices. The benefit of delaying asset sales and attempting to ride through the crisis is that the intermediary may be able to entirely avoid any sale of assets at distressed prices should the crisis be short and mild. We show that when the adverse selection problem is not too severe there are multiple

[19]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.

equilibria, an immediate-trading and a delayed-trading equilibrium. In the first equilibrium, intermediaries liquidate their positions in exchange for cash early in the liquidity crisis. In the second equilibrium, liquidation takes place late in the liquidity event and in the presence of adverse selection problems.

We show that surprisingly the latter equilibrium Pareto-dominates the former because it saves on cash reserves, which are costly to carry. However, the delayed-trading equilibrium does not 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 they were able to do so, intermediaries would be better off committing ex-ante to liquidating their assets at these depressed prices in the distressed states. We also show, perhaps more surprisingly, that a monopoly supplier of liquidity may be able to improve welfare.

Thus, we argue 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. We show, in particular that public provision of liquidity can act as a complement for private liquidity in situations where lemon's problems are so severe that the market would break down without any public price support. An important assumption in our analysis, however, is that the public liquidity provider is able to determine which state of nature the economy is in (ω_{1L} or $(\omega_{20}, \omega_{2L})$) and to adequately adapt its response to circumstances. An important remaining task is thus to analyze the benefits of public policy in our model under the assumption that the public agency may be ignorant about the true state of nature in which it is intervening.

Another central theme in our analysis 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 off. This trade-off is 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 question to explore in future research.

Finally, in our model long-run investors are those with sufficient knowledge to be able to value and absorb the risky assets for sale by financial intermediaries. Only their capital and liquid reserves matter for equilibrium pricing to the extent that they are the only participants with the knowledge to perform an adequate valuation. Other, less knowledge intensive, capital will only step in at steeper discounts for which there may be no market. Our current work focuses precisely on understanding how different knowledge-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.

Dow, James and Gary Gorton (1993), "Arbitrage Chains", NBER Working Paper No. W4314. Available at SSRN: http://ssrn.com/abstract=478730

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.

Jacklin, Charles J. (1987) "Demand Deposits, Trading Restrictions, and Risk Sharing." In E.C. Prescott, and N. Wallace (ed.), *Contractual Arrangements for Intertemporal Trade*, University of Minnesota Press.

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

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

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_{1L} is such that

$$P_{1i}^* \leq \frac{1-\lambda\rho}{1-\lambda}$$

but by assumption A3 this implies

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

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 $P_i(\omega_{1L})$

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

and define

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

as positive number by assumption A2.

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

$$P_{1i}^* = \frac{1-\lambda\rho}{1-\lambda} = \frac{M_i^*}{1-m_i^*} \tag{24}$$

As for prices in states $(\omega_{20}, \omega_{2L})$ they have to be such that both the SRs and the LRs prefer to trade at ω_{1L}. For this set

$$P_i^*(\omega_{20}, \omega_{2L}) < \min\left\{\delta\eta\rho, \frac{1-\rho[\lambda + (1-\lambda)\theta\eta]}{(1-\lambda)(1-\theta\eta)}\right\}. \tag{25}$$

Given this price the LR investors expect only lemons (assets with zero payoff) in the market at ω_{20}, ω_{2L} and thus the demand is equal to zero $Q^*(\omega_{20}, \omega_{2L}) = 0$. As for the SRs notice that if they wait to liquidate at ω_{20}, ω_{2L} they obtain

$$\theta\eta\rho + (1-\theta\eta)\,P_i^*(\omega_{20}, \omega_{2L}) < \theta\eta\rho + \frac{1-\rho[\lambda + (1-\lambda)\theta\eta]}{1-\lambda} = P_{1i}^*,$$

and thus SRs set $q^*(\omega_{20}, \omega_{2L}) = 0$ and $q^*(\omega_{1L}) = 1 - m_i^* = Q^*(\omega_{1L})$. □

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

$$M_d^* = P_{2d}^* (1 - \theta\eta)(1 - m) \tag{26}$$

Define

$$\lambda + (1 - \lambda) \frac{(1 - \theta)\eta\rho}{(1 - \theta\eta) P_d(\omega_{20}, \omega_{2L})} = \varphi' [\kappa - (1 - \theta\eta) P_d(\omega_{20}, \omega_{2L})] \tag{27}$$

This equation always a unique solution which in addition satisfies

$$P_d(\omega_{20}, \omega_{2L}) \in \left(0, \frac{\kappa}{1 - \theta\eta}\right). \tag{28}$$

There are two cases to consider:

- Case 1: $P_d(\omega_{20}, \omega_{2L})$ is such that

$$P_d(\omega_{20}, \omega_{2L}) < \frac{1 - \rho[\lambda + (1 - \lambda)\theta\eta]}{(1 - \lambda)(1 - \theta\eta)} = \underline{P}. \tag{29}$$

In this case set

$$P_{2d}^* = \underline{P}, \tag{30}$$

and set M_d^* to be the solution of

$$\lambda + (1 - \lambda) \frac{(1 - \theta)\eta\rho}{(1 - \theta\eta) P_{2d}^*} = \varphi'(\kappa - M_d^*), \tag{31}$$

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

$$M_d^* < (1 - \theta\eta) P_{2d}^*. \tag{32}$$

Choose m_d^* such that

$$P_{2d}^* = \frac{M_d^*}{(1 - \theta\eta)(1 - m_d^*)} \tag{33}$$

Notice that because $P_{2d}^* = \underline{P}$ the SRs are indifferent in the level of cash held. Both types of traders would prefer to wait to trade at date 2 provided that $P_d^*(\omega_{1L})$ is in the interval

$$\left[\frac{(1 - \theta\eta) P_{2d}^*}{1 - \theta}, \theta\eta\rho + (1 - \theta\eta) P_{2d}^*\right], \tag{34}$$

which is non empty if and only if

$$P_{2d}^* \leq \frac{(1 - \theta)\eta\rho}{1 - \theta\eta} = \overline{P}. \tag{35}$$

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_{2d}^* is independent of δ and for $\delta \leq \overline{\delta}$, where

$$\overline{\delta} = \frac{P_{2d}^*}{\eta\rho}, \tag{36}$$

the SR (weakly) prefers to trade at date 2 for a price P_{2d}^* than carrying the asset to date 3.

- Case 2: $P_d(\omega_{20}, \omega_{2L}) \geq \underline{P}$, then choose

$$P_{2d}^* = P_d(\omega_{20}, \omega_{2L}) \qquad M_d^* = P_{2d}^*(1-\theta\eta) \qquad \text{and} \qquad m_d^* = 0. \tag{37}$$

Except for establishing inequality (35), the remainder of the proof follows as in the previous case. To establish that P_{2d}^* 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 1 is given by

$$P_{1i}^* = \frac{1-\lambda\rho}{1-\lambda}$$

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

$$R_d^* = \frac{(1-\theta)\eta\rho}{(1-\theta\eta) P_{2d}^*},$$

where P_{2d}^* 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).

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

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

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

After some algebra:

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

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 $M_{d,\theta}^{a\,st} < 0$ for $\theta > \widehat{\theta}$, it follows that $R_{d,\theta}^{*} < 0$ for that range.

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

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

Tedious algebra shows that (43) is equal to

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

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. □

Proof of Lemma 7. Given a choice M of cash carried by LR and α invested in the SR risky project ($0 \leq \alpha \leq 1$), the feasibility constraints on transfers to SR are given by:

$$\begin{aligned} C_1(\omega_{1\rho}) &\leq \alpha\rho + M, \\ C_1(\omega_{1\rho}) + C_3(\omega_{1\rho}) &\leq \alpha\rho + M + \varphi\left[\kappa + (1-\alpha) - M\right], \\ C_1(\omega_{1L}) + C_2(\omega_{20}) &\leq M, \\ C_1(\omega_{1L}) + C_2(\omega_{2L}) &\leq M, \\ C_1(\omega_{1L}) + C_2(\omega_{2\rho}) &\leq \alpha\rho + M, \\ C_1(\omega_{1L}) + C_2(\omega_{2\rho}) + C_3(\omega_{2\rho}) &\leq \alpha\rho + M + \varphi\left[\kappa + (1-\alpha) - M\right], \\ C_1(\omega_{1L}) + C_2(\omega_{20}) + C_3(\omega_{20}) &\leq M + \varphi\left[\kappa + (1-\alpha) - M\right], \\ C_1(\omega_{1L}) + C_2(\omega_{2L}) + C_3(\omega_{30}) &\leq M + \varphi\left[\kappa + (1-\alpha) - M\right]. \end{aligned}$$

Consider next the following observations concerning equilibrium contracts:

I. State $\omega_{1\rho}$ is observable and since there is no discounting between periods 1 and 2 we may assume without any loss of generality that $C_1(\omega_{1\rho}) = C_1(\omega_{1L}) = 0$.

II. If $C_3(\omega_{1\rho}) > 0$ then $C_2(\omega_{1\rho}) = \alpha\rho + M$. For if $C_2(\omega_{1\rho}) < \alpha\rho + M$, both agents can be made better off by increasing $C_2(\omega_{1\rho})$ and decreasing $C_3(\omega_{1\rho})$.

III. Incentive compatibility requires that $C_3(w_{30}) = C_3(w_{3\rho})$. Hence any feasible and incentive compatible payment in histories that follow from ω_{2L} is also feasible in histories that follow ω_{20}. Incentive compatibility also requires that

$$C_2(\omega_{2L}) + C_3(\omega_{30}) = C_2(\omega_{20}) + C_3(\omega_{20}).$$

Therefore any payment prescribed for the histories starting at ω_{2L} must also be prescribed for histories starting at ω_{20}:

$$C_2(\omega_{2L}) = C_2(\omega_{20}),$$

and

$$C_3(\omega_{30}) = C_3(\omega_{20}).$$

IV. If $C_3(\omega_{30}) > 0$ then $C_2(\omega_{2L}) = M$. For if $C_2(\omega_{2L}) < M$ SR can be made better off, while keeping LR indifferent, by increasing the payment at date 2 and decreasing by the same amount the payments in states ω_{30} and $\omega_{3\rho}$ at date 3. The same reason, together with observation 3, implies that if $C_3(\omega_{20}) > 0$ then $C_2(\omega_{20}) = M$. We can also use the same reasoning to show that if $C_3(\omega_{2\rho}) > 0$ then $C_2(\omega_{2\rho}) = M + \alpha\rho$.

V. Since $(\lambda + (1-\lambda))\,\rho > 1$ and $\varphi'(\kappa) > 1$, if cash is carried by the LR it must be distributed in some state (at either dates 1 or 2). Hence either $C_2(\omega_{1\rho}) = M + \alpha\rho$, or $C_2(\omega_{2\rho}) = M + \alpha\rho$ or $C_2(\omega_{20}) = C_2(\omega_{2L}) = M$. Note furthermore from observation 4 and incentive compatibility that we must have $C_2(\omega_{2\rho}) > 0$ and $C_2(\omega_{20}) > 0$ unless SR consumption is zero in all histories starting at ω_{1L}. However, in the latter case, because of discounting and $\delta\phi'(k) < 1$ the ex-ante contract is dominated by autarky. Hence we may assume that $C_2(\omega_{2\rho}) > 0$ and $C_2(\omega_{20}) > 0$. In an analogous fashion we can establish that $C_2(\omega_{1\rho}) > 0$.

VI. Suppose, that $C_2(\omega_{1\rho}) \leq M + \alpha\rho - \mu$ for some $\mu > 0$, and let $\gamma > 0$ be small enough that $\gamma < \frac{\mu}{2}$ and $\gamma\frac{\lambda}{1-\lambda} < \min\{C_2(\omega_{2\rho}); C_2(\omega_{20})\}$. Consider the payment

$$\hat{C}_2(\omega_{1\rho}) = C_2(\omega_{1\rho}) + \gamma$$

and lower date 2 payments for all realizations following ω_{1L} by $\gamma\frac{\lambda}{1-\lambda}$. This new contract, leaves SR indifferent and economizes in cash. This cash can be invested in the LR project, which has a marginal product above one, and yield extra utility for LR at date 3. Hence the initial contract cannot be optimal.

VII. Suppose that $C_2(\omega_{2L}) < M$, then from observation 4, $C_3(\omega_{30}) = 0$. Hence $C_2(\omega_{20}) < M$ and $C_2(\omega_{2\rho}) < M + \alpha\rho$. Using the same logic as in observation 6 we may then show that this contract is not optimal.

VIII. Incentive compatibility requires that

$$C_2(\omega_{2\rho}) + C_3(\omega_{2\rho}) = M + C_3(\omega_{30}).$$

Since $C_2(\omega_{2\rho}) = M$ satisfies the LR budget constraint, it follows that

$$C_3(\omega_{2\rho}) \leq C_3(\omega_{30}).$$

This concludes the proof. □

Proof of Proposition 8. Under assumption (**A5**) LR's opportunity cost of holding cash, $\varphi'(\kappa + (1-\alpha) - M)$, is bounded. To see this, note first from lemma 7 that LR must pay SR at least M following the realization of state ω_{1L}. LR's date 0 expected payoff therefore cannot exceed:

$$\eta(1-\lambda)\alpha\rho + \varphi(\kappa + (1-\alpha) - M).$$

Since participation by LR requires that

$$\eta(1-\lambda)\alpha\rho + \varphi(\kappa + (1-\alpha) - M) \geq \varphi(\kappa),$$

we must have $\kappa + (1-\alpha) - M > 0$, by assumption **A5**. It follows that

$$\varphi'(\kappa + (1-\alpha) - M) < B, \quad \text{for some } B > 0.$$

Now, suppose by contradiction that $C_3(\omega_{1\rho}) \geq \epsilon > 0$. Then lowering $C_3(\omega_{1\rho})$ by ϵ and increasing M by $\delta\lambda\epsilon$ keeps SR indifferent, but makes LR strictly better off if $B\delta < 1$. Similarly, if $\min\{C_3(\omega_{2\rho}); C_3(\omega_{30})\} = C_3(\omega_{2\rho}) \geq \epsilon$, a decrease of $C_3(\omega_{30})$ and $C_3(\omega_{2\rho})$ by ϵ and an increase of M by $(1-\lambda)\delta\epsilon$, again keeps SR indifferent but makes LR better off (provided that $B\delta < 1$). $\square$

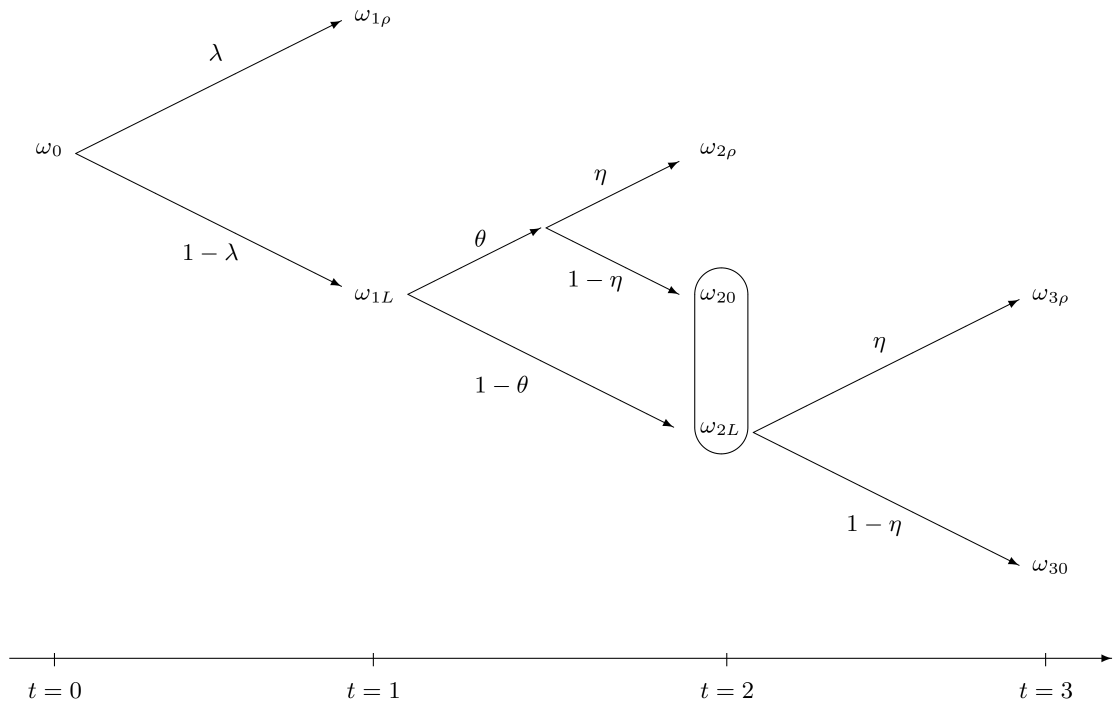

Figure 1. The risky asset. There are four dates. Investment in the risky asset occurs at date0. At date 1 there is an *aggregate* shock. Specifically there are two possible aggregate states, $\omega_{1\rho}$, which occurs with probability λ, and ω_{1L}, which occurs with probability $1-\lambda$. In $\omega_{1\rho}$ 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 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 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 2 is given by $\{\omega_{20}, \omega_{2L}\}$, which generates the adverse selection problem that is key in the analysis. At date 1, agents are symmetrically informed.

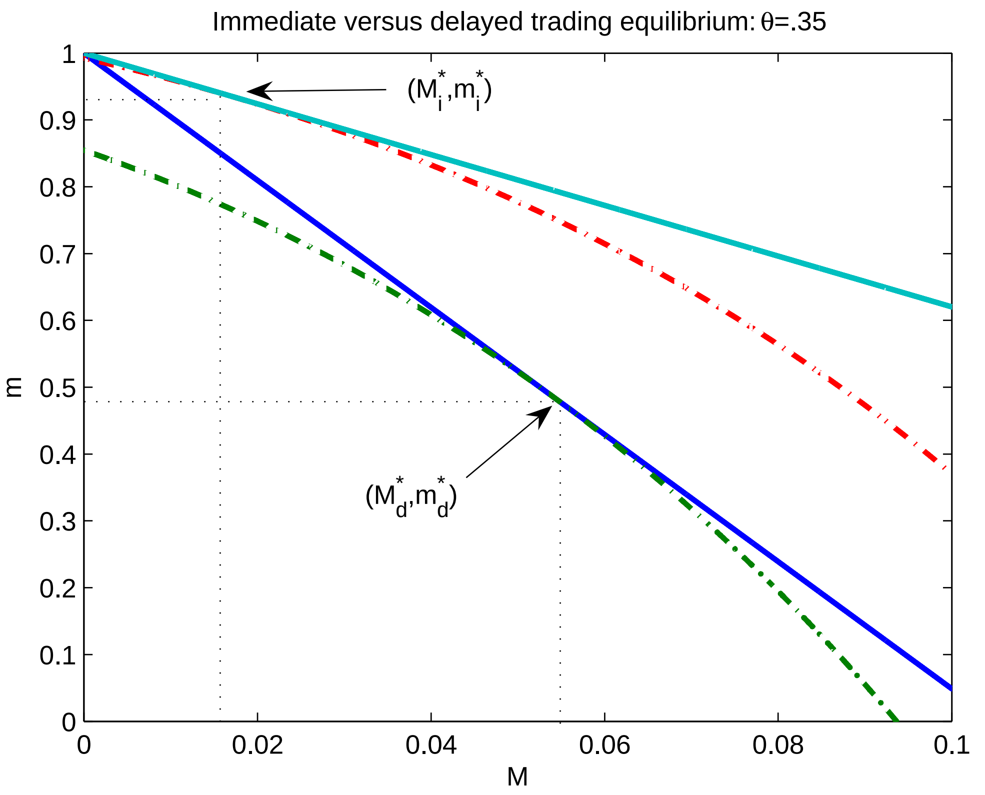

Figure 2. Immediate and delayed-exchange equilibria in the Example for the case $\theta = .35$. The graph represents cash holdings, with the cash holdings of the LRs in the x-axis and the cash holdings of the SRs in the y-axis. The dashed curves represent isoprofit lines for the LR and the straight continuous lines represent the SR's isoprofit lines, for both when the exchange occurs in state ω_{1L} and in states $(\omega_{20}, \omega_{2L})$. 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.

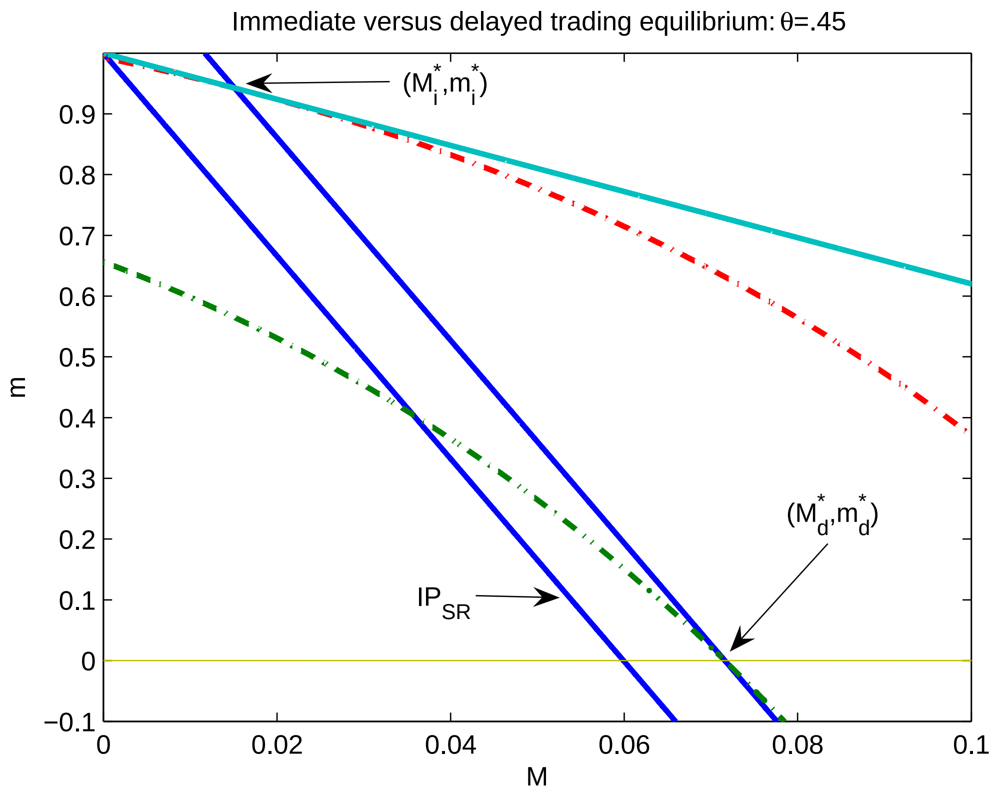

Figure 3. Immediate and delayed-exchange equilibria in the example when $\theta = .45$. The graph represents cash holdings, with the cash holdings of the LRs in the x-axis and the cash holdings of the SRs in the y-axis. The dashed curves represent isoprofit lines for the LR and the straight continuous lines represent the SR's isoprofit lines, for both when the exchange occurs in state ω_{1L} and in states $(\omega_{20}, \omega_{2L})$. 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 states $(\omega_{20}, \omega_{2L})$.

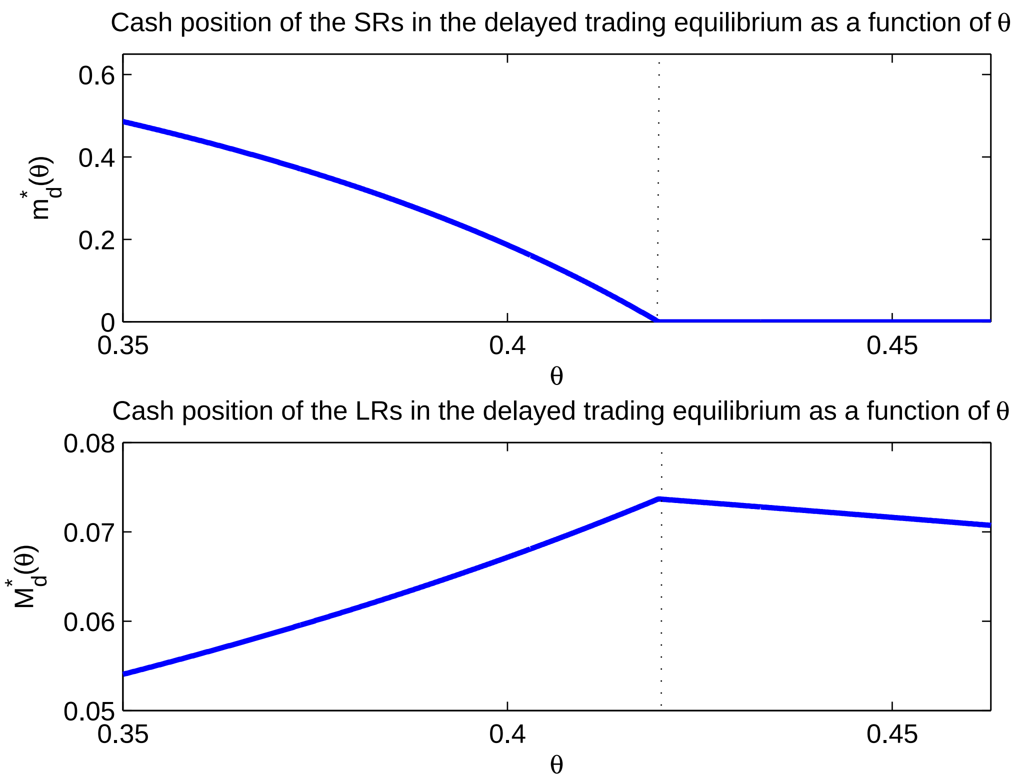

Figure 4. Cash holdings as a function of θ for the Example. Panel A represents the SR's cash holdings in the delayed-trading equilibrium, m_d^* as a function of θ and Panel B does the same for the LR, M_d^*. The dashed vertical line, which sits at $\widehat{\theta} = .4196$ delimits the set of θs for which $m_d^* > 0$ and the one for which $m_d^* = 0$.

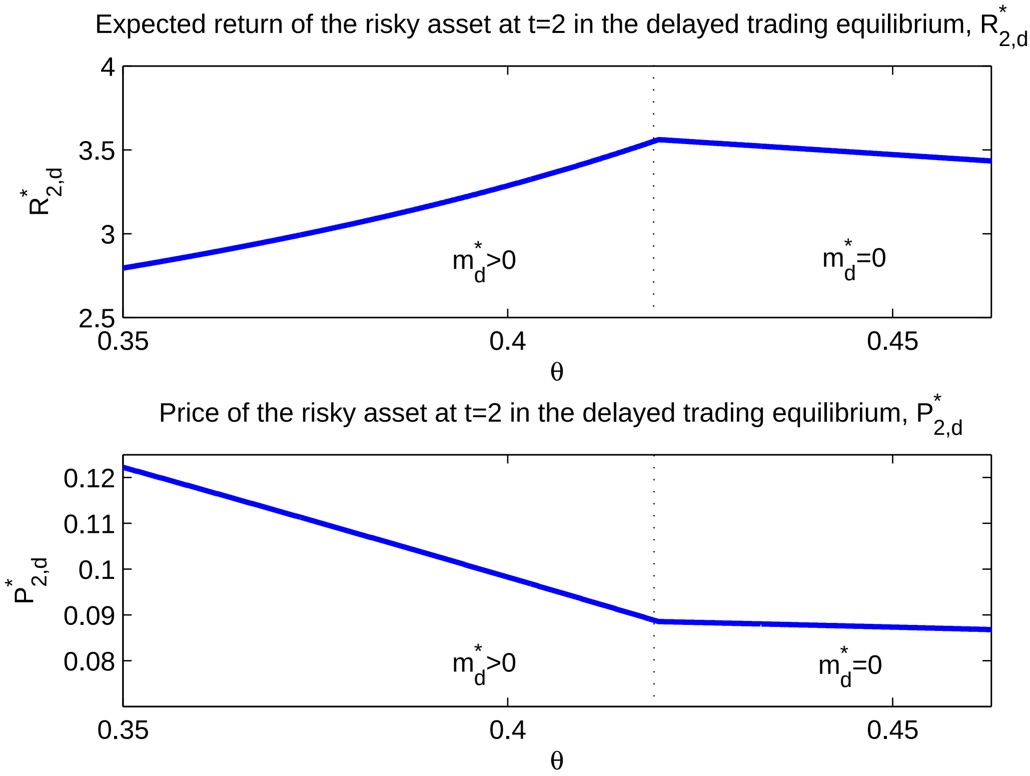

Figure 5. The top panel shows the expected return of the risky asset, $R_d^*(\omega_{20}, \omega_{2L})$, as a function of θ at date 2 in the delayed trading equilibrium. The bottom panel shows the price of the risky asset in states $(\omega_{20}, \omega_{2L})$, P_{2d}^*, as a function of θ at date 2 in the delayed trading equilibrium. The dashed vertical line corresponds to $\widehat{\theta} = .4196$.

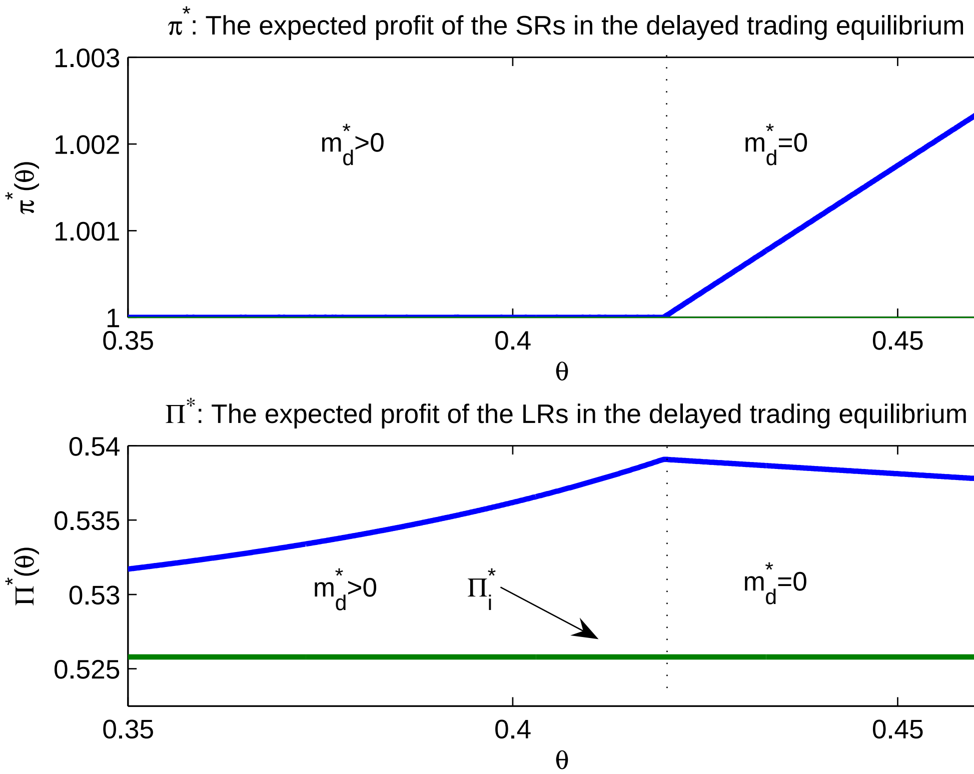

Figure 6. Expected profits for the SR, π^*, (top panel) and the LR (bottom panel), Π^*, as a function of θ in the delayed trading equilibrium. The dashed vertical line corresponds to $\widehat{\theta} = .4196$.

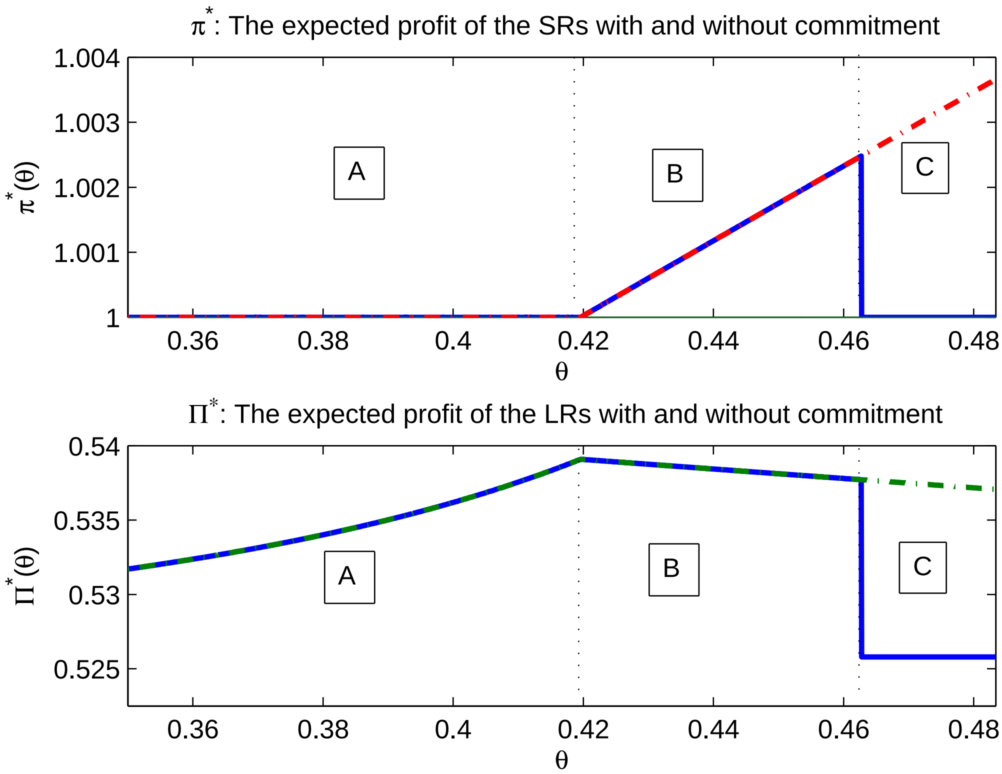

Figure 7. Expected profits for the SR, π^*, (top panel) and the LR (bottom panel), Π^*, as a function of θ. The first dashed vertical line corresponds to $\widehat{\theta} = .4196$. The continuous line plots the expected profits when the Pareto superior equilibrium is chosen. In regions A and B, the 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}.

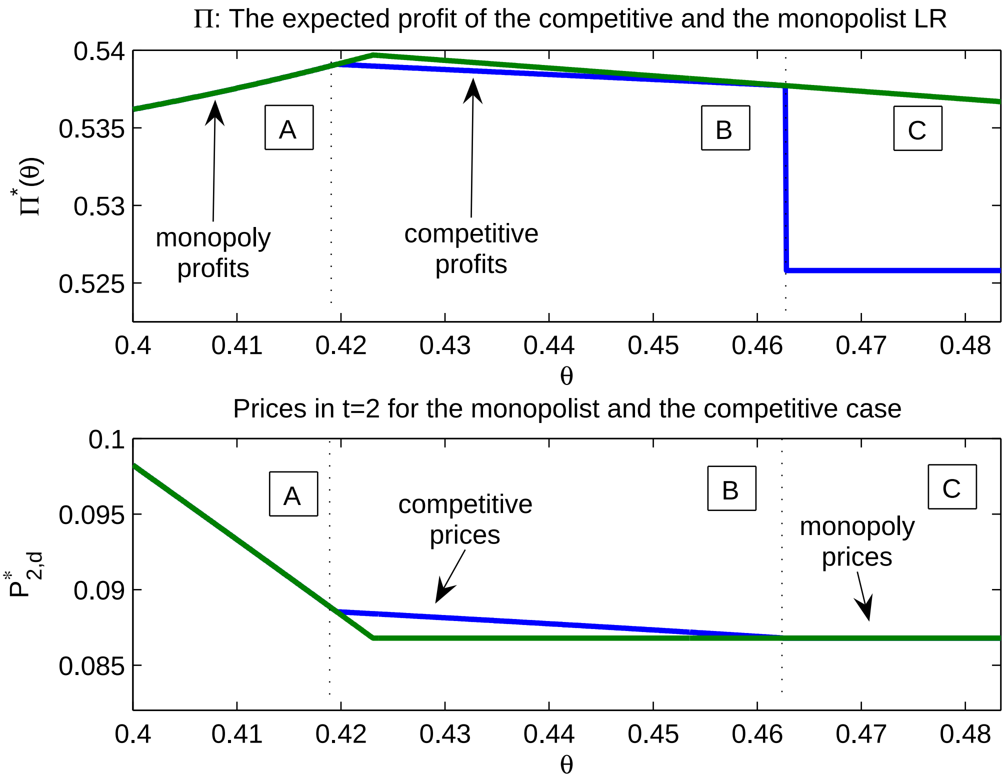

Figure 8. Top panel: Expected profits of the monopolist (the thick line) and the competitive LR (the thin line) as a function of θ. Bottom panel: Prices in states $(\omega_{20}, \omega_{2L})$, P^*_{2d}, in the monopolist (the thick line) and the competitive (the thin line) LR case.

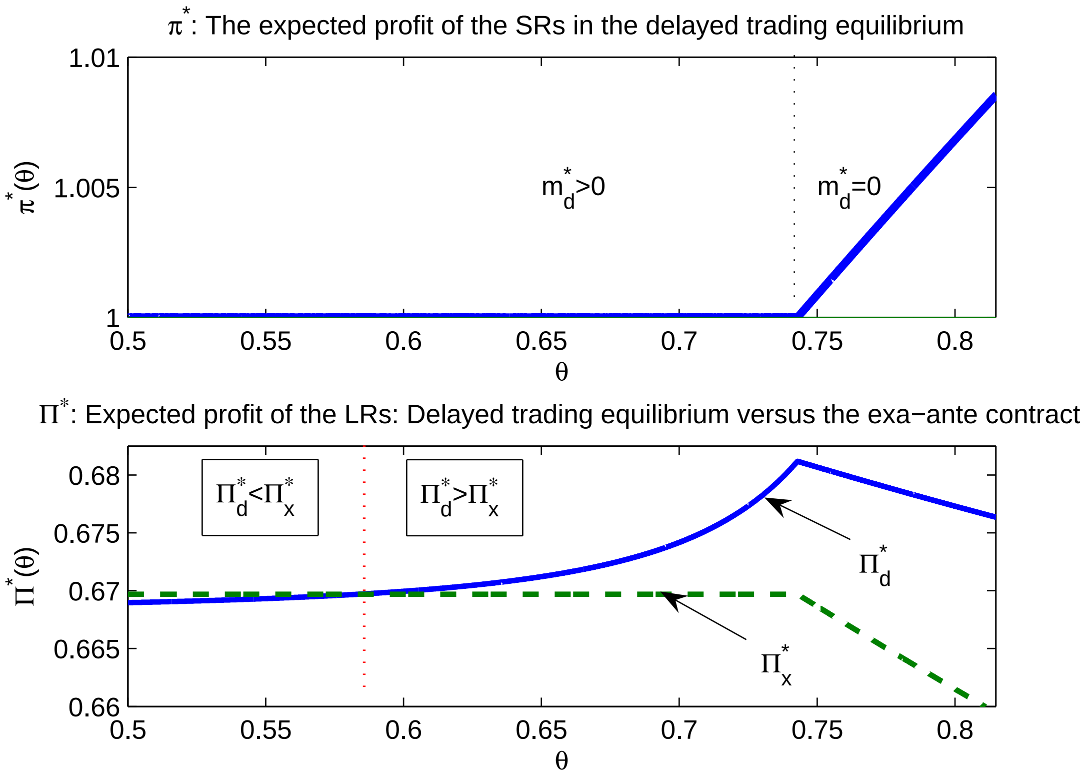

Figure 9. Top panel: Expected profits for the SR in the ex-ante contract when the outside value is the expected profit associated with the delayed trading equilibrium. Bottom panel: Expected profits of the LR in the ex-ante contract, Π_x^*, when the outside value of the SR is the expected profit associated with the delayed trading equilibrium (π_d^*). Also included are the expected profit of the LR in the delayed trading equilibrium, Π_d^*, and in the immediate trading equilibrium, Π_i^*.

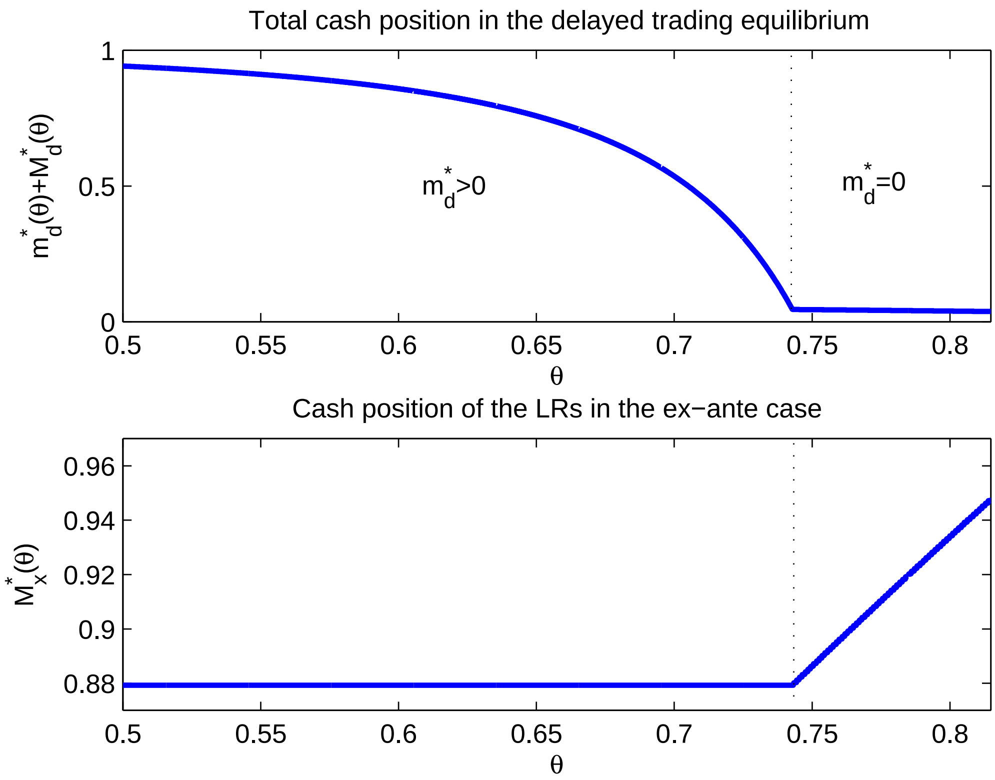

Figure 10. Top panel: Total cash position, $m_d^* + M_d^*$ in the delayed trading equilibrium as a function of θ. Bottom panel: Cash position of the LR in the ex-ante contract case as a function of θ when the outside opportunity of the SR is the expected profit in the delayed trading equilibrium.