Aggregate Risk Perceptions, Adverse Selection, and Securitized Lending: Mitigating Allocational Consequences of Exuberance

Sudipto Bhattacharya
LSE and CEPR

Georgy Chabakauri
LSE

Kjell Gustav Nyborg
ISB, Zurich and CEPR

Preliminary and Incomplete, November 2010

Abstract

We develop a model of securitized ("originate and distribute") lending in which both publicly observed aggregate shocks to values of securitized loan portfolios, and later asymmetrically observed discernment of the qualities of subsets thereof, play crucial roles, as in the recent paper of Bolton, Santos and Scheinkman (2010). Unlike in their framework, we find that originators and potential buyers of such assets may differ in their preferences, over their timing of trades, leading to a reduction in aggregate surplus accruing from securitization. In addition, heterogeneity in agents' selected timing of trades – arising from differences in their ex ante exuberance – may lead to financial crises, via uncoordinated asset liquidation inconsistent with any market equilibrium. We consider mitigating regulatory policy interventions, such as leverage restrictions and asset price guarantees.

I. Introduction

Securitization, via sales of portfolios of long-maturity loans originated by banks to market-based institutions funded with longer maturity liabilities (often aided by implicit governmental support) has been a key part of reality in US and other developed financial markets for a very long time. The presumed benefits arising from such activity are due, in addition to greater cross-sectional diversification in the resulting portfolios backing liabilities, to the "inter-temporal diversification" via which institutions with longer-maturity debt claims are less vulnerable to any short-term aggregate shocks impacting on the current market values of the assets supporting claims on them. Hellwig (1998) was one of the first to emphasize such a role for securitization, in a context of inter-temporal variations in economy-wide interest rates impacting on interim values of long-maturity assets, given fixity of banks' short-maturity liability claims, and the return/interest rates on their loans.

However, it was only in the previous decade, of "financial innovation", that we have witnessed explosive expansion in the securitization of bank-lending based portfolios of credit-backed assets of a far broader quality spectrum, culminating in an even more implosive crash leading to broad-based financial cum economic crises considered to be the worst since the Great Depression of the 1930s. This includes credit card debt-based portfolios of varying qualities, and mortgage loan-backed portfolios with much higher debt to valuation ratios (less borrower income information) all subject to potential losses arising from sectoral shocks - plus their "spillovers" into the broader economy - with origins beyond interest-rate shocks, and their impact on the valuation of payoffs on assets that were largely devoid of default risk, at least in the aggregate. In addition, the structuring of various quasi-independent entities providing funding for such securitization was based on quite complex "tranching" of the payoffs arising from these asset/loan portfolios that backed up their liabilities, creating further non-transparency vis-a-vis their risks.

In essence, this phase of rapid expansion of securitization - of at least ostensibly those lower risk tranches of portfolios backed by (largely) bank-originated loans with far higher degree of heterogeneity, and potentially lower average value than at origination - remained still-born, at or just before the near-closure (flow-wise) of these markets in late 2008. As Adrian and Shin (2009a) have noted, the share of Asset Based Securities (ABS) held by intermediaries with high and short-maturity debt ratios - investment banks, banks and their sponsored investment vehicles - was nearly two-thirds in 2008, with the remainder held by pension funds as well as insurance companies et al. In the process, as securitization markets expanded over 2002-2007 (new issuance sharply slowed in 2007-8, following bad news on some securitized entities which emerged), their funding by the investing firms was provided largely via increases in leverage ratios, either directly as with the investment banks, or within "off the book" special purpose entities sponsored by larger commercial banks, often in the form of short-maturity (overnight) Repos.

Subsequently, declines in market valuations of these underlying asset portfolios, coupled with asymmetric information on their qualities leading to Lemons issues vis-vis acceptable traded prices, led to collapses in these markets. These in turn created the possibility - in some cases reality - of Runs on the investing entities, resulting in both higher spreads on Repo rates, as well as enhanced "haircuts", or margins, on such repo financing. Gorton and Metrick (2009) have documented these crisis-induced phenomena across securities, as well as inter-bank, markets. One of their key findings, elaborated on in Gorton and Metrick (2010), was that these post-crisis effects on spreads and haircuts also occurred, albeit to a lesser degree, in securitization markets other than those backed by sub-prime mortgage backed assets, including even on credit-card receivables based portfolios. On the other hand, the impact on haircuts was much less for high grade corporate bonds, e.g., yield differentials on industrial bonds of differing categories (AAA vs BBB) widened in the financial crisis of 2008-9 to a much lower extent than on ABS.

These circumstances, and findings, have clearly called for a systematic program of research, on the functioning and potential vulnerabilities of a "market based banking" system, in which banks with specialized expertise originate, package and distribute portfolios of securities to other financial market participants. In an initial stage of rapid expansion of such markets, only a few firms may have had the required expertise to evaluate risks associated with such portfolios, to create tranches of these varying in seniority and risk for sale to the ultimate investors, such as pension funds and insurance companies. During this phase, many securities remained in the portfolios of these specialized entities, investment banks and the sponsored investment vehicles and conduits of large commercial banks. This is (or at least was) associated with large increases in their leverage, often of a short-term nature. The resulting increase in funding for the originated assets was often also associated with increase in the prices of such assets in the short run – Adrian and Shin (2009b) – allowing for easy refinancing of loans made to finance these, so that repayment risks pertinent to their associated credit-based portfolios were difficult to judge (as compared to on corporate bonds), by outside rating agencies as well as interim suppliers of short-term funding to the initial portfolio holders. But, ultimately, when these asset price "bubbles" proved to be unsustainable, the resulting aggregate shocks to securities based on loans made to finance such sets of assets collapsed, leading to deleveraging and drastic drops in their prices. Shin (2009) provides an outline of such a process of credit expansion and collapse; on pioneering earlier work on this set of themes, see also Geanakoplos (2010).

Several recent papers have amplified and elaborated on richer micro-economic foundations for bank behaviour and "systemic risk" - of asset price declines and potential bank failures - in these settings. Cifuentes, Ferrucci, and Shin (2005) provide a "reduced form" characterization cum calibration of the impact of inter-bank connections on the contagion of default risks across banks, due to declining asset prices arising from (large) idiosyncratic shocks to some banks' asset values. Acharya, Shin, Yorulmazer (2010), and Stein (2010) have extended this process further, by characterizing banks' ex ante portfolio choices, over risky long-term loans vs risk-less liquid assets. Liquidity for the purchase of long-term assets of banks which are sold to service their debts, in low return states, is provided by a combination of other banks which have surplus liquidity, as well as some pool of outside investors who are less efficient at realizing value from these assets. Both sets of authors emphasize externalities, arising from such inefficient liquidation, that an individual bank may ignore in making its ex ante portfolio choice. Stein focuses on the ostensible liquidity premium (cheaper short-term debt) banks may obtain, with excessive investment in illiquid assets to be sold later at a discount to outside investors in a bad state of nature, whereas Acharya et al emphasize that an originating bank's full return on long-term assets/loans would not be "pledgeable" to facilitate additional interim financing, to stave off asset sales in adverse states.

In contrast to these papers, in which an originating institution sells its longer-term assets/loans only in the low individual or aggregate return state attempting to avert default, Bolton, Santos and Scheikman (2010) develop and analyze another model in which securitization of originated assets to markets is an ongoing, and essential, part of the investment process in longer-maturity and potentially risky assets. The market participants who are potential buyers of these assets ascribe higher values to them than their originators, at least contingent on an aggregate value-reducing shock which leads the originating institutions to consider selling their assets. Their focus is on endogenising the timing of asset sales, by short-run (SR) to long-run (LR) investors, during the time interval following upon an aggregate shock. Over that period, originators (specialized interim holders) of securitized assets come to know more about the qualities, in terms of prospective future payoffs, of subsets within their holdings. Then, if they had not sold all of their holdings at the start of this stage, the asset market price would come to reflect their incentive to sell only those assets about which they have bad news, or at best no (idiosyncratic) news beyond the public aggregate shock. Indeed, Bolton et al (hereafter BSS) make the strong assumption that, for the subset of an SR's assets on which she has received good news, there is no longer any wedge between their (future) values perceived by SR vs LR investors. (Formally they model this by assuming that the return on this subset of assets is realized immediately, while other returns are delayed, and discounted at a higher rate by SR investors). Hence, given that LR investors face an opportunity cost of holding liquidity to buy assets, there can not be gains to be realized via SRs selling good assets to LRs, even under asymmetric information.

Building on the last observation, BSS then show that whenever a Delayed trading equilibrium - in which SRs wait until asymmetric information is (thought to be) prevalent, and then sell only their "bad" and "no new information" assets to LRs - does exist, despite a "lemons discount" in its equilibrium market price, it Pareto dominates an Early trading equilibrium, for both SR and LR agents, in an ex ante sense. It is also associated with relatively higher (equilibrium) origination of the long-maturity asset by SR agents, coupled with higher outside liquidity provision by LR investors. Thus, the overall thrust of their conclusions is in sharp contrast with those reached and emphasised by Acharya et al (2010), and Stein (2010). In discussing policy implications of their model in a companion paper, BSS (2009), they suggest that when a Delayed trading equilibrium may not exist – owing to the opportunity cost of holding liquid assets for LR agents, coupled with prices reflecting asymmetric information about the qualities of assets being sold - the role of government policy should be to provide a price guarantee to restore its existence, rather than, for example, to provide credit based on collateral held on in their books by SR agents/institutions, when private investors refuse to do so. Indeed, they suggest the latter policy may make SR agents more reluctant to sell.

Despite the richness of its framework, and the elegance of its analysis, the BSS conclusions leave some issues unanswered, and raise other questions. There is, for example, no clear "tipping point" at which a Crisis arises, besides when SR agents discover that there is no delayed trading equilibrium price at which they are willing to trade medium quality assets, about which they have no additional news beyond the initial, and public, value-reducing aggregate shock. In reality, significant doubts about the sustainability of high and safe (flow) returns on sub- prime mortgage-backed securities arose by mid-2007, while the full occurrence of a crisis, with sharply increased haircuts and yields related to credit granted based on such assets, did not materialize until mid-2008, which is a long time for sophisticated (SR) agents to have investigated the qualities of securitized assets in their portfolios. During this interval, there were also reports of some (investment) banks divesting, or curtailing new purchases of, mortgage-backed securities, so uniform co-ordination on the (possibility of a) Delayed Trading equilibrium is far from evident. Rather, it suggests (to us) the possibility of developing differences in opinion among SRs, about the (medium-term) likelihood of the continuation of a benign state for mortgage-backed securities as a whole, leading to their making differing choices on the timing of trades in these assets. But, the BSS framework provides no clear conclusion regarding the impact of such differences, regarding the likelihood of an aggregate value-reducing shock, on its SR agents' preferred trading equilibrium, early versus delayed, when each of these is believed to exist.

In addition, we find a key BSS precept, that there are NO perceived differences between SR and LR investors vis--vis their (ex interim) payoffs accruing from long-term assets whose quality is known to be good (with zero probability of low returns), to be problematic under asymmetric information. Their metaphor, of the return on such an asset accruing immediately as its quality is discerned by an SR, can not be taken literally. In reality, the differential valuation of asset payoffs by SR and LR agents is in most part due to differences in the maturity structure and composition of their liabilities. SR institutions are often funded with low-maturity and demandable (non-renewable) liabilities so that, when their asset values drop, closer to the investment required to originate these, prudence coupled with the "haircuts" demanded by their (Repo) creditors and regulators, necessitates their injecting more equity(-like) Capital, in order to continue holding such assets. LR agents, who have a longer-duration structure to their liabilities, can cope with the risks associated with interim portfolio values with relatively greater flexibility, including the payout-smoothing across low and high return periods of the sort described in Allen and Gale (1995). These equity capital injections also affect the pledgeable (to investors as a whole) values of assets adversely due to tax effects, or greater ability of insider managers of SR institutions to extract more Rents, as noted by Calomiris and Kahn (1990), and Diamond and Rajan (2000). When an SR agent in the BSS model holds onto a subset of its originated/bought assets, which it knows are of good quality, its financiers could still require it to provide an equity capital cushion to reflect the average reduction in asset prices following upon the prior aggregate shock. Indeed, the liability holders of an SR institution need have no idea of the structure of a Delayed trading equilibrium, including the equilibrium extent of selling in it, so that an SR selling just a part of its assets would not provide them a credible signal of the quality of its assets that are held on to.

For these reasons, our beliefs regarding relevant modelling precepts, and our sense that SR agents' (possibly divergent, from 2007) beliefs, regarding the likelihood of an adverse shock to valuation of sub-prime mortgage-backed securities as a whole, had an important impact on their choices of timing of trade on extant holdings of, as well as future investments in these, we develop an alternative analysis otherwise in the spirit of the BSS framework. We assume that the valuation wedge that arises between SR and LR agents, following upon an adverse aggregate shock, applies to all asset subsets, irrespective of their heterogeneous qualities as discerned by SRs; Chari et al (2010) assume the same in a reputation-based secondary market model. We then consider potential existence of delayed and early trading equilibria, as in BSS, and agents' preferences over these. We find, in contrast to their conclusions, that LR agents are always worse off in a delayed trading equilibrium (whenever it exists), relative to an early trading equilibrium for the same exogenous parameters. SR agents, on the other hand, may be better off in the delayed trading equilibrium, but that is the case only if their prior belief regarding the likelihood of the benign aggregate state continuing - the adverse aggregate shock not occurring - is strictly above an interior threshold level; sufficiently "exuberant" ex ante beliefs are essential for delayed trading equilibrium to be preferred by SRs. As in Bolton et al (2010), such SR-preferred delayed trading equilibria are associated with (weakly) higher investment in the long-term risky asset, hence lower holding of inside liquidity by SR. However, overall surplus from origination and trading, summed across SR and LR, is strictly lower in our delayed - as compared to early - trading equilibrium.

We then consider, again consistent with our view of empirical reality, a scenario in which a subset of (optimistic/exuberant) agents, who ascribe a lower likelihood to the adverse aggregate shock arising, make their trading and investment choices based on a delayed trading equilibrium, whereas other SRs and LRs, who are less optimistic, make their trades immediately, before the aggregate shock has arisen. Such immediate trading plays a key role in our model, unlike in BSS (2010). We use this scenario to sketch a plausible process for a "financial crisis", in which the price discovery via immediate trading, by a subset of SR and LR agents, serves to provide a basis for the Leverage choices of the LR agents who plan to trade later in a Delayed trading equilibrium, outlined above. We then show that changes in the beliefs of (at least) the less optimistic LR agents, and its impact on the immediate trading prices, may lead to (Repo) Runs by the short-term creditors of optimistic SRs, even before an adverse aggregate shock has realized, which is a pre-condition for any type of trading in BSS (2010). The resulting asset sales, by the SR agents who had planned to trade a proper subset of their assets in a Delayed equilibrium, leads then to a "market breakdown" (or malfunction, see Hellwig (2008)), prior to to the stage in which idiosyncratic asymmetric information about subsets of their held assets accrues to SRs. The market then collapses, and stays that way. In other words, adverse selection relevant to delayed trading serves to provide a backdrop for, rather than the immediate triggering mechanism in a process of financial crisis.

Our paper is organized as follows. In Section II, we provide an overview of the set up of Bolton et al (2010), emphasizing the departure point for our variation on it. Section III deals with the characterization of (manifolds of) early, then delayed, trading equilibria in our setting. Section IV considers the potential impact of mis-coordination and crises. In Section V we consider policy interventions such as leverage restrictions and guaranteed price supports to mitigate their impact. In section VI, we conclude, with further comparisons with some recent literature.

II. The Model: Agents' Tradeoffs and Detailed Choice Criteria

There are two sets of key players in our set up. Short-run (SR) agents, funded with short-term liabilities, who are uniquely capable of originating long-maturity assets with potentially risky returns, but are less capable of (ascribe lower valuation to) holding such assets when their risks are perceived, as compared to Long-run (LR) investors. The latter are funded with longer-duration liabilities, and hence are less concerned with risks pertaining to interim valuations of assets with long maturities, which may reflect new information about their returns, or macroeconomic variables such as interest rates. As a result, conditional on an aggregate shock that lowers the average valuation of assets originated by SR agents, there are natural trading gains to be obtained by their selling these assets to LR investors. Both sides are aware of these opportunities, and plan their ex ante portfolio choices - over liquid and long-term assets – taking these anticipated trades, and the rationally conjectures market equilibrium prices associated with these, into account. The prospective buyers of long-term assets, LR investors, do however have opportunity costs associated with holding liquid assets in their portfolios, to enable them to buy these SR-originated long-maturity assets when SR agents may wish to sell these. This arises in the form of alternative, and illiquid in an interim period, investments which the LR investors may make. Its resulting payoff is assumed to have diminishing marginal returns.

The novelty and richness added by Bolton et al (2010) to this set up - which they share with several other models of "market-based banking" - is their assumption that, subsequent to a public aggregate shock which reduces the valuation of SR's long-term assets to a level at which they may be motivated to sell to LR investors – when revised valuations are closer to their (short-term) liabilities - additional, idiosyncratic information on subsets of an SR's assets may arise, that affect their values for both types of agents, and which are asymmetrically better known to the originating SR agents. The important question that then arises, regarding trading of long-maturity assets among SR and LR agents, pertains to the timing of such trades: would it occur before (early) or after (delayed) the idiosyncratic, as well as asymmetrically known, shocks to subsets of SR agents' asset portfolios are known (or perceived) to have arisen? In the latter case of delayed trading, the SR agents sell only those subsets of their long-maturity assets about which they do not have better than average news, whereas in early trading they sell all their long-maturity assets to LR investors who value these more. In the BSS set up, interim good news about a subset of an SR's asset portfolio also makes their prospective return risk-free, as well as equal to its anticipated level prior to the aggregate value-reducing shock. As such, and consistent with the nature of SR and LR agents' horizons (cum liability structure), they assume that - unlike with subsets of SR assets about which there is no news, or bad (worse than average) news - SR and LR agents value the (capitalized) returns arising from the "good news" component at the same level.

Given this key assumption, and the fact that LR agents face an opportunity cost for holding liquid assets (reserves) to buy SR assets contingent on the aggregate shock to their valuations, there is a presumption that delayed trading may benefit both SR and LR agents. Indeed, BSS first show that this is always the case when the accrual of new idiosyncratic information about the quality of (subsets of) SR agents' longer-maturity assets is symmetric across SR and LR. Delayed trading equilibrium, based on competitive linear price-taking conjectures, always exists, and Pareto dominates the outcome arising from early trading. It also entails greater investment in longer- maturity assets by SRs, as compared to that arising in an early trading equilibrium.

Matters are considerably more delicate when individual SRs become more informed about the qualities of components of her long-maturity asset portfolio, and LR agents realize that when trading. A classic "lemons problem" (Akerlof (1970)) then arises, in that LRs realize that the average value to them of the assets that would be sold in such delayed trading would be strictly lower than that of the average quality asset, after the aggregate valuation shock, because assets about which an SR has received bad news in the interim would be sold for the same price as other assets of average quality. If the resulting lemons discount applied by LRs to the valuation of SR-offered assets is sufficiently high, SRs may stop offering those assets about which they hold unrevised (average) belief about their quality for sale, despite their ascribing lower valuation to these as compared to LRs, who would have offered a better price for such assets had their quality been common knowledge among them. If this is the case, which depends both on the degree of asymmetric information - in the form of both the likelihood of SRs obtaining asymmetric information, and the divergence in quality levels (payoffs) arising thereby - as well as LR agents' opportunity costs of holding liquid reserves, an equilibrium asset price for delayed trading at which SRs are willing to sell their average quality assets, and LR agents are willing to hold sufficient liquidity to buy SR assets at that price, may cease to exist. In contrast, early trading equilibrium will, under mild regularity assumptions.

Our point of departure, from these BSS set of assumptions, is to observe that, under asymmetric information as above, SR agents may no longer value even components of their long-maturity asset portfolios about which they have excellent positive news as above, at the same level as LR agents would have, had they been informed about the asset quality. Our rationale for this is that, just as there is asymmetric information across SR and LR agents about asset quality levels, so there is among SR agents and their liability holders, in particular short-maturity creditors who need to be convinced about asset quality to renew their loans. Hence, after a delayed trading outcome where an SR agent does not sell her best-quality assets (as assumed), she might need to back up the liability structure supporting her portfolio of retained long-maturity assets with more equity, than prior to the aggregate value-reducing shock. As noted above, such equity injections are costly to all investors of an SR as a whole, in part owing to rent- extraction by SR managers, which increases with equity finance. Indeed, we assume that the proportional difference in the valuations of these assets, between SR and LR agents, is the same as that for other long-maturity assets about which SR has no news unknown to LRs, or bad news relative to the average. Thus, in contrast to the BSS set up, there remain some gains from trade to be had via trading of these assets between SRs and LRs, albeit requiring LRs to invest more in liquid assets for a given level of long-maturity asset origination by SRs. In this modified set up, we characterize when early versus delayed trading equilibria would arise, first when agents have common beliefs about structural parameters of the set up, including (especially) the likelihood of the aggregate (average) value-reducing shock - which is deemed necessary for the valuation wedge across SR and LR agents. We then delineate the structure of agents' preferences across these two types of trading equilibrium, and the resulting choices.

We now introduce some notation. SRs can create long-maturity assets with a constant returns to scale technology. We assume that, in the beginning, these generate a gross (capitalized, over a finite horizon) rate of return $\rho > 1$, per unit of investment. Overall SR agents/institutions have an aggregate endowment of unity, which they may invest divisibly in either the long-maturity asset, or in a liquid short-maturity asset with unit rate of return; indeed, as far as SR agents are concerned the liquidity of the latter is of no consequence. This benign aggregate state is expected to continue beyond an initial period with probability λ in $(0, 1)$; with the complementary probability an aggregate value reducing shock arises, which reduces the expected value of these long-maturity assets to $\eta\rho$, η in $(0, 1)$. Some of the assets would ex post provide a return of ρ, with interim probability η, others would provide no return at all. It is assumed that, upon the aggregate shock arising, SR agents value their expected payoff at $\delta\eta\rho$, δ in $(0,1)$ - for the reasons we outlined above - whereas LR agents revise their valuations only to $\eta\rho$, from ρ prior to this shock. Selling, and securitization of such asset (portfolios) for the purpose, is deemed essential for any initial investment by the SRs in these, as:

A1. $[\lambda\rho + (1-\lambda)\delta\eta\rho] < 1$, whereas, in contrast

A2. $[\lambda\rho + (1-\lambda)\eta\rho] > 1$

The first assumption (A1) clearly implies an interior upper bound, $\lambda_u < 1$, on λ. The second assumption would imply an interior lower bound $\lambda_d > 0$, if $\eta\rho$ is weakly below unity. We let SRs choose ex ante portfolio $\{m, 1-m\}$ across short and long-maturity assets, consistent with expected prices for selling (some) long-maturity assets to LRs. LR agents, in contrast, can either invest in a long-maturity asset having no risk at all, under an (aggregate) diminishing returns technology giving gross payoff $F(I)$ given investment level I in it. The function $F(\cdot)$ is assumed to be strictly increasing, concave and to satisfy the Inada conditions, with its first derivative always strictly above unity. LRs in the aggregate are assumed to have an endowment level of $K > 1$, to invest in I this technology as well as to hold a non-negative amount $M = (K-I)$ in a liquid asset paying unit rate of return, which they anticipate using to buy the long-maturity assets of SRs, in the event of the adverse aggregate shock creating their valuation "wedge".

At time $t = 1$, a period after these initial investment portfolio choices by SRs and LRs, an aggregate value-reducing shock to the long-maturity asset originated by SRs either occurs, or not. If not, then SR would not sell its assets to LRs at any unit price below ρ, so there is no gain from such trade for LRs. If it does, SRs expect to be able to sell their long-maturity assets immediately, to LRs, at some unit price $P_e > \delta\eta r$, and/or to wait and sell later, at time $t = 2$, at the unit price P_d. In between $t = 1$ and $t = 2$, SRs acquire private information about the qualities of subsets within portfolios of long- maturity assets. We assume that SRs are well-diversified among such assets, that are subject to conditionally independent idiosyncratic information shocks. An SR obtains perfect information about the future payoffs on a proportion q of her assets, coming to know if their return would turn out to be ρ, or 0, in the sub-proportions η and $(1-\eta)$ respectively. On the remaining proportion $(1-q)$ of her long-maturity assets, an SR obtains no new information, and thus continues to believe that on average returns on these assets would continue to be $\eta\rho$. Thus, assuming (consistent with an equilibrium) $\delta\rho > P_d > \delta\eta\rho$, SR would sell only her "no news" and "bad news" assets at time $t = 2$. We derive SR's ex ante expected payoff as a function of her investments and trades.

If SRs choose ex ante to invest $\{m, 1-m\}$ in the liquid and long-maturity assets, respectively, and then - conditional on the aggregate valuation shock to the latter - to trade amounts X_e at $t = 1$, and X_d at $t = 2$ of now risky asset, at rationally conjectured (equilibrium) prices P_e and

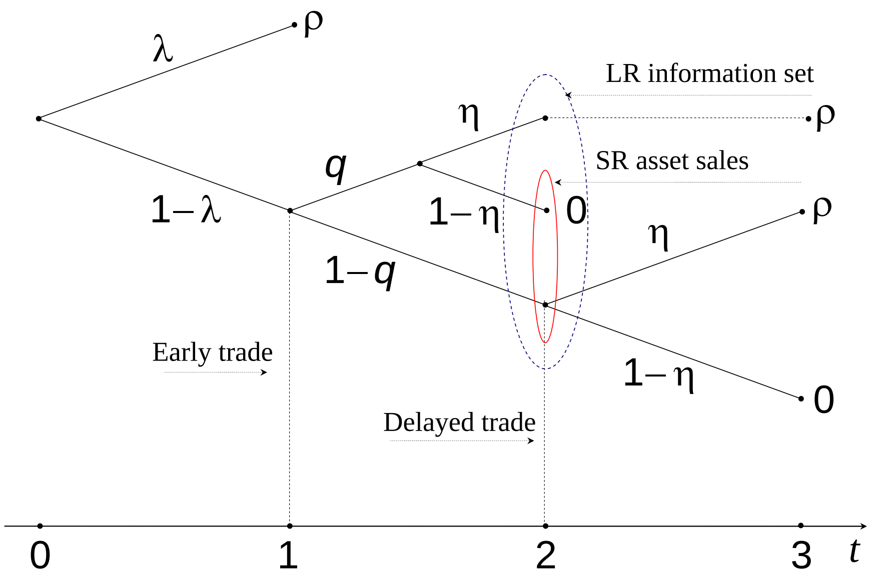


Figure 1: **The Time Line of the Events.**

P_d respectively, her expected payoff equals:

$$\Pi_{SR} = [m + \lambda(1-m)\rho + (1-\lambda)\{X_eP_e + (1-m-X_e)(q\eta\delta\rho + (1-q\eta)P_d)\}], \qquad (1)$$

where we have simplified, by incorporating the assumption, consistent with any non-trivial delayed trading equilibrium, that SR will trade $X_d = [(1-q\eta)(1-m-X_e)]$ at time $t = 2$. Thus, for example, given such a set of conjectures prices, SR would strictly prefer to trade early if $[P_e - \{q\eta\delta\rho + (1-q\eta)P_d\}] > 0$, and wait to trade later if this inequality is strictly reversed. She would choose to invest a strictly positive amount in the liquid asset, pick $m > 0$, given strict preference for early trading in the long-maturity asset if $(1-\lambda\rho) = (1-\lambda)P_e$, and choose $m = 0$ if Pe is strictly higher than the level satisfying this equality. On the other hand, if SR strictly prefers to trade her long-maturity asset late rather than early, she would choose $m > 0$ if $(1-\lambda\rho) = [(1-\lambda)\{q\eta\delta\rho + (1-q\eta)P_d\}]$.

In stating the above, we are assuming existence of an equilibrium with at least one of $\{X_e, X_d\}$ strictly positive, so that $m < 1$. An assumption to follow, made also in the Bolton et al (2010) paper, ensures that this is true for some early trading equilibrium. It amounts to stating that, at the maximum level of LR investment in its long-maturity illiquid technology, via setting $I = K$ and $M = 0$, the marginal product of investment $dF(I)/dI$ is strictly less than the expected returns to increasing M at the margin, to buy long-maturity assets from SRs at the unit price $Pe(\lambda) = (1-\lambda\rho)/(1-\lambda)$. Given such a unit price for purchasing the SR-originated asset after an aggregate shock, LR obtains:

$$\Pi_{LR} = F(I = K - M) + \lambda M + (1-\lambda)\frac{M}{P_e}\eta\rho \qquad (2)$$

as expected payoff. She optimizes by choosing M to satisfy the first order condition:

$$\frac{dF(I)}{dI} = \{\lambda + (1-\lambda)\frac{\eta\rho}{P_e}\}. \qquad (3)$$

The assumption below ensures that the level of M chosen by LR is strictly positive:

A3. $\frac{F'(K)-\lambda}{(1-\lambda)\eta\rho} < \frac{1-\lambda}{1-\lambda\rho}$, where, $F'(K) = \frac{dF(I)}{dI}$ at $I = K$.

In what follows, we proceed to describe both early and delayed trading equilibrium, and characterize the conditions under which one or the other should be expected to arise, depending on agents' preferences over these. Our characterization of the set of early trading equilibrium, as a function of λ, is essentially the same as in Bolton et al (2010). It involves two segments, in the first of which $m^* > 0$ for SRs, and a second in which $m^* = 0$, implying $M^* = P_e$. It is in the characterization of delayed trading equilibrium that the difference between our setup and theirs emerges in a stark way. We show that, unlike in their model, even if a delayed trading equilibrium exists in ours, it is never preferred to the early trading equilibrium by both SR and LR agents, even weakly. Further, SR agents prefer to trade in a delayed equilibrium if and only if their belief regarding the likelihood of an adverse aggregate shock $(1-\lambda)$ is sufficiently low.

III. Early vs Delayed Equilibrium: Descriptions and Comparisons

Early Trading Equilibrium

Proposition 1: (**Bolton et al**) *For all λ in $[\lambda_d, \lambda_u)$, an early trading equilibrium exists, with unit trading prices P_e, and liquidity holding levels $\{m, M_e^*\}$satisfying:*

(i) *For $\lambda < \lambda_c$, $m^* > 0$, $P_e(\lambda) = \frac{1-\lambda\rho}{1-\lambda}$, $M_e^* = (1-m^*)P_e$, satisfying equation (3);*

(ii) *For $\lambda_c < \lambda < \lambda_u$, $m^* = 0$, and $M^* = P_e(\lambda)$, again satisfying equation (3)*

Remarks: Note that, for λ in $[\lambda_d, \lambda_c]$, $\Pi_{SR} = 1$; all of any strictly positive surplus, resulting from the origination of long-maturity assets by SRs, accrues only to LRs.

Corollary 1: *$M_e^*(\lambda)$ is strictly increasing in λ at all λ in $[\lambda_d, \lambda_c)$, whereas it strictly decreases in λ for λ in $[\lambda_c, \lambda_u)$; that is also the case for LRs' expected payoff $\Pi_{LR}(\lambda)$.*

Proof: Consider LR's expected return on money holding M as a function of λ, $R(\lambda)$:

$$R(\lambda) = \lambda + (1-\lambda)\frac{\eta\rho}{P_e(\lambda)}, \tag{4}$$

which in the segment λ in $[\lambda_d, \lambda_c)$, in part (i) of the Proposition, implies that:

$$R(\lambda) = \lambda + \eta\rho\frac{(1-\lambda)^2}{1-\lambda\rho}. \tag{5}$$

It is straightforward to show, via differentiation, that the right hand side of (5) is strictly increasing in λ, which using LR's optimality condition (3) yields the result.[1]

For the second part of the corollary, concerning the region $[\lambda_c, \lambda_u)$ in which money holdings m^* of SRs equals zero, so that $M_e^*(\lambda) = P_e(\lambda)$, suppose to the contrary that $M_e^*(\lambda)$, and thus $P_e(\lambda)$ are (weakly) increasing in λ. But, then $R(\lambda)$ would be strictly decreasing in λ, which would contradict LRs' optimality condition, in equation (3).

Finally, the statement that LRs' overall expected payoff is increasing (vs decreasing) in λ whenever $M_e^*(\lambda)$ is increasing (vs decreasing) in λ, is implied by the axioms of Revealed Preference, applied to LR's objective function, described in equation (2). *Q.E.D.*

Remark: The co-movement of the unit asset prices $P_e(\lambda)$, and LR money holdings $M_e^*(\lambda)$, across the set of early trading equilibria when λ is in $[\lambda_d, \lambda_c)$, may well be thought of as the inverse of "cash in the market pricing" – see Shin (2009) for its exposition - in that unit asset prices, and external (LR) liquidity holdings held in the anticipation of buying these assets following on an aggregate shock to their value, move in opposite directions as a function $(1-\lambda)$, the probability of such a shock. The reason, of course, is that m^* decreases, hence the quantity of the long-maturity asset supplied by SRs, $(1-m^*)$, increases strictly in λ, i.e., as the probability of the adverse aggregate shock decreases. However, SRs gain nothing from that enhanced surplus!

Delayed Trading Equilibrium

First we note that the nature of any delayed trading equilibrium when it exists, despite the "lemons" discount in the price of assets sold in it - which excludes the subset of an SR's portfolio about which she has received better than average information - is very different in our setting, as compared to that of Bolton et al (2010). In particular, there exist no set of commonly conjectured prices $\{P_e, P_d\}$, for trading assets early vs late, that are consistent with a non-trivial equilibrium, and thereby (weakly) exceed $\delta\eta\rho$, such that both selling (SR) and buying (LR) agents would prefer delayed over early trading, even weakly. Indeed if, as we shall assume, it is the selling side's choice which determines when trading will occur, when prices are determined

[1] It is easily shown that:

$$\frac{dR(\lambda)}{d\lambda} = \Big[1 - \eta\Big\{\frac{(1-\lambda\rho)^2 - (1-\rho)^2}{(1-\lambda\rho)^2}\Big\}\Big] > 0.$$

in competitive centralized markets, LR agents become discretely worse off at the switching point at which SR agents just prefer to trade later rather than early. We show that, given the other parameters, such a shift in SR preferences regarding timing (and nature) of their trading arises at a point where λ is sufficiently high, and strictly above zero if $\eta\rho$ is at least weakly below unity. In other words, a sufficiently high value associated with the net Surplus arising from SR-originated assets is essential for them to strictly prefer to trade a subset of their assets later, despite the cost coefficient $(1-\delta)$ that is associated with not selling their better-quality assets. In contrast, in the Bolton et al (2010) paper there is a range of examples, involving SRs choosing strictly positive money holdings $m^* > 0$ in both early and delayed trading equilibrium - and thus being indifferent vis- a-vis their payoffs across the two - in which the LR agents strictly prefer to trade late, benefiting from being able to buy a subset of a greater quantity of SR investment in the long-maturity assets in the delayed equilibrium, with lower money holdings M_d^*.

Lemma 1: *There does not exist a tuple of prices $\{P_e, P_d\}$, consistent with non-trivial early and late trading equilibrium, thus (weakly) exceeding $\delta\eta\rho$, such that both SR and LR agents prefer, even weakly, their delayed over their early trading payoff outcomes.*

Proof: For SRs to prefer trading late, it must be the case that $P_e < [q\eta\delta\rho + (1-q\eta)P_d]$, at least weakly. For LRs to do so as well, it must be that the expected return on their money holding is (weakly) higher in delayed trading, $(\eta\rho/P_e) < (1-q)\eta\rho/(1-q\eta)P_d$. The last expression recognizes that, in delayed trading the assets sold by any SR, the proportion $(1-q\eta)$ of her long-maturity investments, has the overall expected payoff of only $(1-q)\eta\rho$, since a subset $q(1-\eta)$ of these are "lemons", with zero payoffs.

From the second (LRs') inequality, we obtain that $(1-q\eta)P_d < (1-q)P_e$, weakly at least. This can be consistent with the first inequality, for SR, even weakly, only if P_e is lower than or equal to $\delta\eta\rho$. But this would imply that $P_d < \delta\eta\rho$, for all $\eta < 1$, and that contradicts P_d being consistent with a non-trivial delayed trading equilibrium. Indeed, $P_e = \delta\eta\rho$ is also inconsistent with initial investment by SR, given (A1) above. *Q.E.D.*

Given the observation above, the following corollary is an immediate consequence.

Lemma 2: *SRs would never strictly prefer a Delayed trading equilibrium in which $m^* > 0$, over any early trading equilibrium. Such a delayed equilibrium would also make LR agents strictly worse off than in early trading - unlike as in Bolton et al (2010).*

Given Lemmas 1 and 2 above, the only case in which a delayed trading equilibrium could arise in our setup is one where, given other parameters, SR agents perceive that they will be strictly better off in such an equilibrium, as compared to an early trading equilibrium. As a result, they withhold their supply of the long-maturity asset from its market, until it is common belief that they have asymmetric information about subsets of their portfolio, and would only be selling their average and bad quality assets. For example, consider a set of parameters such that early trading equilibrium, described in Proposition 1 above, entails money holdings $m^* > 0$ by SR agents, whereas delayed equilibrium entails $m^* = 0$ for SRs. As noted in remarks above, SR agents' payoff in such an early equilibrium would be no more than if she had invested only in the liquid asset, setting $m = 1$. In contrast, in a delayed equilibrium with $m^* = 0$, in which SRs invest all of their endowment in the long-maturity asset, their expected payoff from so doing, $[\lambda\rho + (1-\lambda)\{q\eta\delta\rho + (1-q\eta)P_d\}]$, must necessarily strictly exceed the unit payoff from just holding the liquid asset, despite gains from trade given up (to the detriment of LR agents' payoffs) by SRs planning not to trade their better quality asset subsets.

The issue is, when would a tuple of such a Delayed and an Early trading equilibrium exist? To answer this, we proceed in two steps. First, we describe the conditions under a non-trivial delayed trading equilibrium could conceivably exist, given the "lemons discount" which LRs would take into account for their returns when buying assets, as well as the opportunity cost that LRs face at the margin for additional liquidity M. In this respect, our conditions are the same as those derived in Bolton et al (2010). Then we derive the necessary and sufficient conditions for a delayed trading equilibrium in which SRs hold $m^* = 0$ to exist in our setup. SRs would strictly prefer to trade in this delayed trading equilibrium, as compared to any early equilibrium involving $m^* > 0$.

In such a delayed trading equilibrium, SRs would be trading a proportion $(1 - q\eta)$ of their long-maturity assets, invested in unit measure, and having an overall expected payoff of $(1 - q)\eta\rho$, arising from only those assets within this mix about which SR does not have bad information, predictive of null payoffs. To buy these assets at the market clearing price P_d, LR investors would have to hold $M_d = (1 - q\eta)P_d$ in liquid assets, on which they obtain the expected return of $[\lambda + (1 - \lambda)(1 - q)\eta\rho/(1 - q\eta)P_d]$. For such a scenario to be an equilibrium, we would need to have $P_d > \delta\eta\rho$, weakly at least, so that SR agents would be wiling to sell their average quality assets, implying:

$$1 < \frac{dF(I = K - M_d)}{dI} = \lambda + (1 - \lambda)\frac{(1 - q)\eta\rho}{(1 - q\eta)P_d}, \tag{6}$$

which is the LR agents' optimality condition for their choice to hold Md in liquid assets. Combining these two conditions, we see that for any λ it must be true that:

$$\delta < \frac{1 - q}{1 - q\eta} < 1. \tag{7}$$

Of course, in addition a consistent equilibrium price P_d must be such that SR agents strictly prefer to trade in this equilibrium, rather than coordinating on an early one, or

$$P_e(\lambda) = \frac{1 - \lambda\rho}{1 - \lambda} < q\eta\delta\rho + (1 - q\eta)P_d(\lambda), \tag{8}$$

where we have assumed that $\lambda < \lambda_c$, so that early trading equilibrium entails $m^* = 0$. Combining the conditions (7) and (8) above, we can derive the following necessary condition for the existence of a delayed trading equilibrium with $m^* = 0$ held by SRs:

Lemma 3: *Define the "social surplus" per unit of the SR-created long-maturity asset,*

$$S(\lambda) = [\lambda\rho + (1-\lambda)\eta\rho - 1]. \tag{9}$$

A necessary condition for the existence of a delayed trading equilibrium with $m^* = 0$ *is*

$$S(\lambda) > (1-\lambda)q^2 \frac{1-\eta}{1-q\eta}\eta\rho. \tag{10}$$

Proof: Conditions (8) and $P_d > \delta\eta\rho$, required for a delayed equilibrium, together imply

$$S > (1-\lambda)[(1-\delta)\eta\rho - q(1-\eta)\delta\eta\rho]. \tag{11}$$

Which, upon substitution for $\{\delta, (1-\delta)\}$ from the inequality (7), implies inequality (10).

Remarks: Under the maintained hypothesis that $\lambda < \lambda_c$, this creates the possibility of a lower bound λ^*, $0 < \lambda^* < \lambda_c$, such that the selected equilibrium would entail early trading for all $\lambda < \lambda^*$, and delayed trading for $\lambda > \lambda^*$. Such direct dependence of the (SR-)selected timing of trading, hence implied equilibrium investment $(1-m^*)$ in the SR-originated long-maturity asset, is absent in the Bolton et al (2010) paper, in which a delayed trading equilibrium, provided it exists given the parameters as well the $F(I)$ function, is Pareto preferred by both agent types, albeit weakly by SRs if $m^* > 0$ in it.

Given an $F(I)$ function, denoting LR agents' opportunity cost of holding liquid assets for trading with SRs, and assuming $\lambda < \lambda_c$ vis-a-vis the early equilibrium, we have:

Proposition 2: *Condition* (10) *above, together with the condition in inequality* (13) *below, are necessary and sufficient for the existence of a delayed trading equilibrium in which* m^*, *the liquid asset holdings of the selling SR agents, equals zero. Defining:*

$$P_{\min} = \frac{P_e(\lambda)}{1+q(1-\eta)}, \tag{12}$$

$$\frac{dF(I = K - (1-q\eta)P_{\min})}{dI} < \left[\lambda + (1-\lambda)\frac{(1-q)\eta\rho}{(1-q\eta)P_{\min}}\right]. \tag{13}$$

Moreover, there exist upper and lower bounds on δ, *given by:*

$$\delta^*(\lambda) = \frac{x}{\eta\rho}, \qquad \delta_*(\lambda) = \max\left\{\frac{x}{\rho}, \frac{P_e(\lambda) - (1-q\eta)x}{q\eta\rho}\right\}, \tag{14}$$

where x *solves a nonlinear equation*

$$\frac{dF(I = K - (1-q\eta)x)}{dI} = \lambda + (1-\lambda)\frac{\eta\rho(1-q)}{(1-q\eta)x}, \tag{15}$$

such that for all pairs $\{\lambda, \delta\} \in \Big\{\{\lambda, \delta\} : \delta_*(\lambda) \leq \delta \leq \delta^*(\lambda)\Big\}$ *there exists a unique delayed equilibrium with* $m^* = 0$ *and price* $P_d = x \geq P_{\min}$ *which SRs prefer to an early equilibrium with price* $P_e(\lambda)$. *Furthermore, the length of the equilibrium existence interval on* δ *satisfies the following inequality:*

$$\delta^*(\lambda) - \delta_*(\lambda) < \min\left\{1-\eta, \frac{1-q}{q}\right\}. \tag{16}$$

Proof: See the Appendix.

Proposition 2 establishes necessary and sufficient conditions for the existence of a unique delayed trading equilibrium and provides a tractable characterization of the equilibrium existence regions. It also establishes a lower bound on the equilibrium price P_d, given by (12), which guarantees that the equilibrium price is high enough to induce SRs to choose to trade late and supply not only the lemons but also average quality assets. The existence region is characterized in terms of upper and lower bounds (14) on the discount parameter δ. Intuitively, on one hand, parameter δ should be sufficiently high to induce the SRs to trade at $t = 2$, so that they get higher expected discounted payoff after receiving good news after $t = 1$. On the other hand, it cannot be too high since otherwise $P_d \geq \delta\eta\rho$ is violated and hence only lemons are traded in the market. Consequently, the equilibrium exists only for δ in the medium range, bounded by some δ_* and δ^*.

The results of Proposition 2 indicate that the bounds on parameter δ become tighter as η or q increases. To understand the intuition we note that as η increases a good outcome becomes more likely in the no-news state at $t = 2$. Therefore, for the delayed trade to be an equilibrium outcome, SRs with no news should be more impatient to be willing to sell the asset at time $t = 2$. Consequently, the upper bound δ^* should decrease leading to the shrinkage of the interval for δ supporting the delayed equilibrium. Furthermore, the interval for δ shrinks as q increases. The reason is that higher q makes the no-news state less likely, increasing the proportion of lemons traded at $t = 2$. Consequently, price P_d decreases, and the no-news SRs should be more impatient (as measured by their δ) to sell assets at $t = 2$, and hence δ^* should decrease reaching zero in the limit.

From the results of Proposition 2 it can additionally be demonstrated that SRs prefer a delayed equilibrium with $m_d^* = 0$ to an early one with $m_e^* = 0$ or $m_e^* > 0$, so that $\pi_d \geq \pi_e$, where expected payoffs π_d and π_e are given by:

$$\pi_d = \lambda\rho + (1-\lambda)(q\eta\rho\delta + (1-q\eta)P_d), \tag{17}$$

$$\pi_e = m_e^* + (1-m_e^*)(\lambda\rho + (1-\lambda)P_e). \tag{18}$$

Consequently, the SRs choose to trade late, enforcing the delayed equilibrium.

To facilitate the numerical analysis, from Proposition 1 we observe that the early equilibrium price P_e (required for the construction of bounds δ_* and δ^*) can conveniently be written as follows:

$$P_e(\lambda) = \max\left\{\frac{1-\lambda\rho}{1-\lambda}, y\right\}, \tag{19}$$

where y solves a nonlinear equation:

$$\frac{dF(I = K - y)}{dI} = \left\{\lambda + (1-\lambda)\frac{\eta\rho}{y}\right\}. \tag{20}$$

Indeed, it follows from Proposition 1 that:

$$P_e(\lambda) = \begin{cases} \dfrac{1-\lambda\rho}{1-\lambda}, & \text{if} \quad m_e^* > 0, \\ y(\lambda), & \text{if} \quad m_e^* = 0. \end{cases} \tag{21}$$

Moreover, from Proposition 1, $M_e^* < P_e$ when $m_e^* > 0$, and hence from the first order condition (2) and concavity of function $F(\cdot)$ it follows that $F'(K-P_e) \geq \lambda + (1-\lambda)\eta\rho/P_e$. Consequently, in the early equilibrium with $m_e^* > 0$ it can easily be demonstrated that $P_e = (1-\lambda\rho)/(1-\lambda) \geq y$, giving rise to expression (19). Expressions (14) for the bounds δ_* and δ^* along with expression (19) for the price in the early equilibrium allow for an efficient numerical computation of the existence regions for delayed and early equilibria, which we describe in the next subsection.

Numerical Analysis

In this subsection we numerically explore the existence regions for different equilibria in $\{\lambda, \delta\}$-space, aggregate welfare across different equilibria and other relevant economic parameters. In particular, we are interested in the regions where the delayed equilibrium with $m_d^* = 0$ coexists with early equilibrium with either $m_e^* > 0$ or $m_e^* = 0$. Our construction of these regions is based on the bounds for discount parameter δ derived in Proposition 2. In addition to bounds δ_* and δ^*, we also note that assumption (A2) imposes the following upper bound on δ:

$$\delta \leq \bar{\delta}(\lambda) = \frac{1-\lambda\rho}{1-\lambda}\frac{1}{\eta\rho}. \tag{22}$$

From the results of Proposition 1 we note that (22) along with assumption (A3) are enough to guarantee the existence of an early equilibrium.

For our numerical analysis we pick the following specification for LR investment technology, satisfying all the conditions in Section II:

$$F(I) = \frac{K^{1-\alpha}I^{\alpha}}{\alpha}, \tag{23}$$

where $\alpha \in (0,1)$. Given the concavity of $F(\cdot)$ it can easily be demonstrated that the nonlinear equations (15) and (20) have unique solutions x and y in terms of which the early P_e and delayed P_d equilibrium prices are derived. We calculate x and y numerically and by substituting them into expressions (14) obtain the upper and lower bounds for δ as functions of λ.

The characterization of the existence regions in Proposition 2 allows us to calculate the lower bound λ_* for the benign state probability λ, such that the delayed equilibrium with $m^* = 0$ exists (for some δ) whenever $\lambda \geq \lambda_*$. The discussion in Proposition 2 implies that λ_* can be obtained as a solution to equation $\delta_*(\lambda_*) = \delta^*(\lambda_*)$. Similarly, the expression for the early equilibrium price P_e in (19) can be employed to characterize the "switching point" λ_c, introduced in Section II, which separates early equilibria with $m^* > 0$ (when $\lambda < \lambda_c$) and early equilibria with $m^* = 0$

(when $\lambda \geq \lambda_c$). In particular, it can easily be demonstrated that parameters λ_* and λ_c solve the following equations:

$$\begin{aligned} x(\lambda_*) &= \frac{P_e(\lambda_*)}{1+q(1-\eta)}, \\ y(\lambda_c) &= \frac{1-\lambda_c\rho}{1-\lambda_c}, \end{aligned} \tag{24}$$

where x and y in turn solve equations (15) and (20), and price $P_e(\lambda)$ is given by (19).

Figure 2 shows the existence regions for delayed and early equilibria in $\{\lambda, \delta\}$-space. For the numerical calculations we use the following set of parameters: $K = 2$, $\rho = 1.2$, $\eta = 1/\rho$, $q = 0.3$, and $\alpha = 0.75$. The existence region for the delayed equilibrium with $m_d^* = 0$ is the region bounded from above by $\delta^*(\lambda)$ and $\bar{\delta}(\lambda)$ and from below by $\delta_*(\lambda)$. The early equilibrium exists for all parameters λ and δ such that $\delta \leq \bar{\delta}(\lambda)$, and λ_c separates the equilibria with $m_e^* > 0$ (when $\lambda < \lambda_c$) and the equilibria with $m_e^* = 0$ (when $\lambda \geq \lambda_c$).

The numerical calculations demonstrate that the existence regions for the delayed and early equilibria overlap, and $\lambda_* < \lambda_c$. Moreover, bounds $\delta_*(\lambda)$ and $\delta^*(\lambda)$ turn out to be decreasing functions of the good state probability λ. To explain this result, we note that the delayed price $P_d = x$, where x solves equation (15), is a decreasing function of λ, which can be established by differentiating equation (15) and showing that $\partial x/\partial\lambda < 0$. Intuitively, as probability λ increases, a bad shock at $t = 1$ becomes less likely. Since the SRs trade only conditional on observing the bad state at $t = 1$, ex ante at $t = 0$ the probability of trade after $t = 0$ goes down. Therefore, LRs face higher opportunity cost of holding liquidity M, preferring to invest more in their riskless technology. Consequently, conditional on a bad shock at $t = 1$, SRs will face lower demand for their assets both in early and delayed equilibria, and hence, the delayed and early prices are decreasing functions of λ, which translates into decreasing bounds for δ.

The delayed equilibrium coexists with the early equilibrium with $m_e^* > 0$ when $\lambda \in [\lambda_*, \lambda_c]$ and with the early equilibrium with $m_e^* = 0$ when $\lambda \geq \lambda_c$. The size of $[\lambda_*, \lambda_c]$ interval depends on the curvature of the technology function, $-F''(I)I/F'(I) = 1 - \alpha$, parameterized by α. To investigate the sensitivity of the size of this region with respect to α we numerically calculate λ_* and λ_c as functions of α. Figure 3 presents the results of the calculations and demonstrates that the size of the region decreases as parameter α goes up.

We now investigate the welfare implications of our analysis. Given the significant overlap of the existence regions it becomes important to compare the aggregate welfare across different equilibria. We quantify the aggregate welfare by an expected total payoff defined as the sum of the expected payoffs of LRs and SRs, denoted by Π and π, respectively. The expected payoffs of SRs are given by expressions (17) and (18) whereas for LRs the expected payoffs in delayed and early equilibria take the following form:

$$\Pi_d = F(K - M_d) + \lambda M_d + (1-\lambda)\frac{M}{P_d}\frac{1-q}{1-q\eta}\eta\rho, \quad (25)$$

$$\Pi_e = F(K - M_e) + \lambda M_d + (1-\lambda)\frac{M}{P_e}\eta\rho. \quad (26)$$

Figure 3 shows aggregate welfare in delayed and early equilibria for the model parameters $K = 2$, $\rho = 1.2$, $\eta = 1/\rho$, $q = 0.3$, $\alpha = 0.75$, and $\delta = 0.74$.

The aggregate welfare functions are increasing in probability λ. Moreover, the aggregate welfare in the early equilibrium exceeds that in the delayed equilibrium for each parameter λ. To understand the economic intuition we note that SRs have higher expected payoff in the delayed equilibrium with $m_d^* = 0$ than in the early equilibrium. However, according to Lemma 2 they do not have strict preference for a delayed equilibrium with $m_d^* > 0$ over an early equilibrium. Therefore, in the neighborhood of λ_* their welfare will be almost unchanged by switching from a delayed to an early equilibrium. Furthermore, according to Lemma 2 LRs are always strictly better off in the early trading equilibrium, and hence, at least in the neighborhood of λ_* the aggregate welfare should be higher in the early equilibrium.

The fact that the aggregate welfare is higher in the early equilibrium can also rigorously be demonstrated for $\lambda \geq \lambda_*$. For the ease of exposition we demonstrate this only for the case when $m_e^* = 0$. In this case, from the expressions for SR and LR expected payoffs in (17), (18), (25), and (25), as well as market clearing conditions $M_e = P_e$ and $M_d = (1 - q\eta)P_d$ it follows that:

$$\Pi_d + \pi_d = \lambda\rho + (1-\lambda)\eta\rho(1 - q + q\delta) + F(K - M_d) + M_d, \quad (27)$$

$$\Pi_e + \pi_e = \lambda\rho + (1-\lambda)\eta\rho + F(K - M_e) + M_e. \quad (28)$$

Comparing the expressions in (27) and (28) we observe that the sufficient condition for the aggregate welfare to be higher in early equilibrium is that $F(K - M_e) + M_e > F(K - M_d) + M_d$, or equivalently, given that $F(K - M) + M$ is an increasing function, the sufficient condition becomes $M_e > M_d$. To prove that $M_e > M_d$ we note that the first order condition for M_d in (6) along with the market clearing condition $M_d = (1 - q\eta)P_d$ imply the following inequality:

$$\frac{dF(I = K - M_d)}{dI} < \lambda + (1-\lambda)\frac{\eta\rho}{M_d}. \quad (29)$$

From the comparison of the above equation with the first order condition for M_e in (3) and the properties of technology function $F(\cdot)$ it can easily be demonstrated that $M_e > M_d$, and hence the aggregate welfare is higher in an early equilibrium.

Finally, Figure 5 shows $P_e(\lambda_*)$ as a function of the parameter α for different levels of return ρ, while the other parameters are as for the previous graphs. It turns out that this function is an increasing function of the parameter α, as well as the return ρ. As demonstrated in the subsequent part of the paper, $P_e(\lambda_*)$ can be thought of as a government intervention price that induces the SRs to switch to early trading equilibrium.

Figure 2: **Existence Regions for Early and Delayed Equilibria.**

This Figure shows the existence regions for early and delayed trading equilibria for parameters $K = 2$, $\rho = 1.2$, $\eta = 1/\rho$, $q = 0.3$, and $\alpha = 0.75$. The delayed equilibrium with $m^* = 0$ exists for all $\{\lambda, \delta\}$ such that $\delta_* \leq \delta \leq \delta^*$, $\lambda_* \leq \lambda \leq 1/\rho$. The early equilibrium with $m^* > 0$ exists for all $\{\lambda, \delta\}$ such that $\delta \leq \delta_e^*$ and $0 < \lambda \leq \lambda_c$, and the early equilibrium with $m^* > 0$ exists for all $\{\lambda, \delta\}$ such that $\delta \leq \delta_e^*$ and $\lambda_c \leq \lambda \leq 1/\rho$.

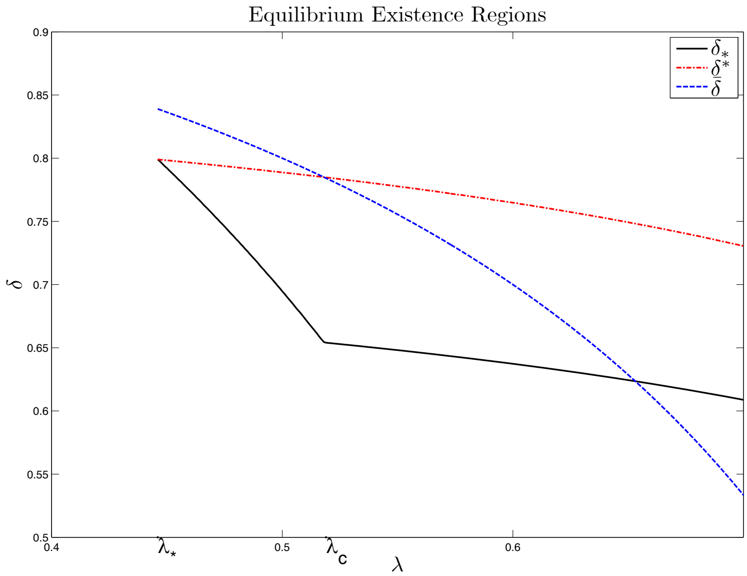


Figure 3: **Equilibrium λ_* and λ_c as Functions of Curvature Parameter α.**

This Figure plots parameters λ_* and λ_c as functions of curvature parameter α for parameters $K = 2$, $\rho = 1.2$, $\eta = 1/\rho$, $q = 0.3$. Delayed equilibria with $m^* = 0$ and early equilibria with $m^* > 0$ coexist if $\lambda_* < \lambda < \lambda_c$, while delayed equilibria with $m^* = 0$ and early equilibria with $m^* = 0$ coexist if $\lambda_c \leq \lambda \leq 1/\rho$.

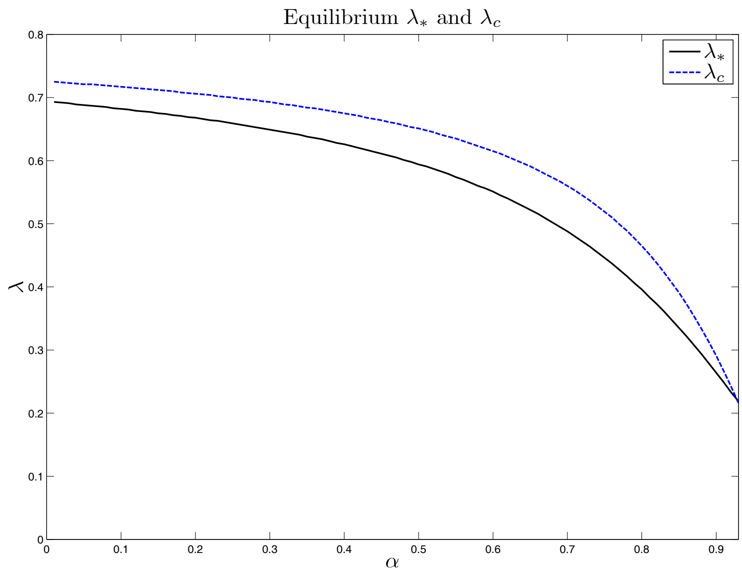


Figure 4: **Aggregate Welfare Across Early and Delayed Equilibria.**

This Figure shows the aggregate welfare in delayed and early equilibria for parameters $K = 2$, $\rho = 1.2$, $\eta = 1/\rho$, $q = 0.3$, and $\delta = 0.74$. $\Pi_e + \pi_e$ is the aggregate welfare of LRs and SRs in early equilibrium while $\Pi_d + \pi_d$ is the aggregate welfare of LRs and SRs in delayed equilibrium.

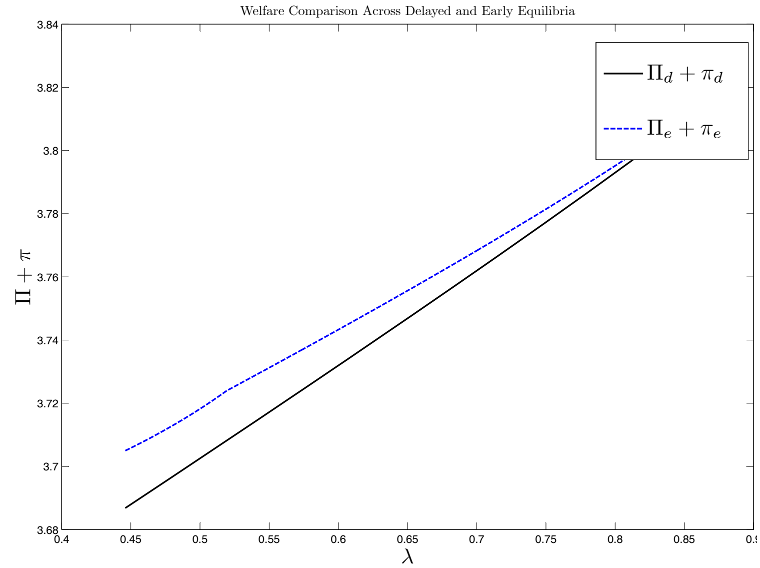
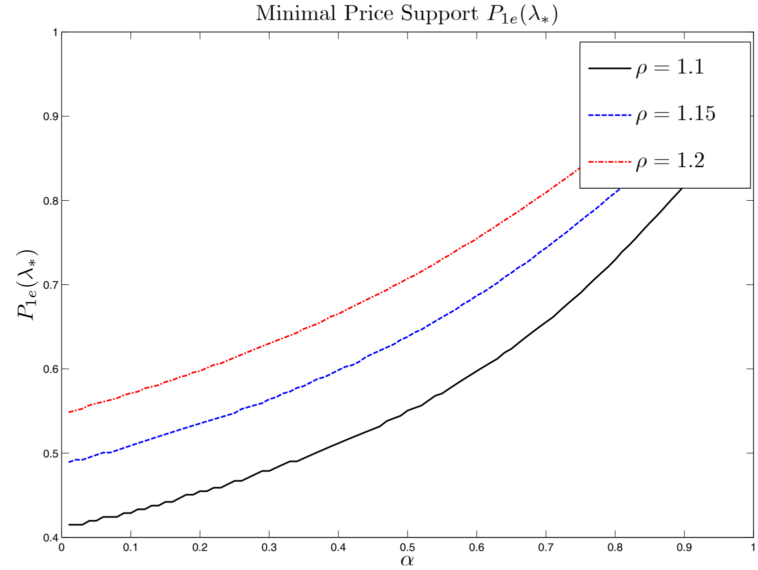


Figure 5: **Price Support $P_e(\lambda_*)$ as Function of Curvature Parameter α.**

This Figure shows the price support function $P_e(\lambda_*)$ for parameters $K = 2$, $\rho = 1.2$, $\eta = 1/\rho$, $q = 0.3$.



IV. Strategy-Proofness, Immediate Trading, and Exuberant Priors

As we noted in the Introduction, in early 2008, even after some aggregate valuation shocks to the mortgage backed securities market had occurred, in the opinion of the majority of participants in it, highly levered institutions such as banks and investment banks continued to hold nearly two-thirds in value of these assets on balance sheets. This suggests strongly that the SR agents were not coordinating their planned trading of these assets, with LR agents such as insurance firms and pension funds, based on an Early trading equilibrium. At the same time, it is also appears to be the case that such LR agents had acquired quite significant (nearly one third by value) proportional stakes in such assets - or tranches thereof - from SR originators cum/and distributors, over the period 2002-7, before the (fully perceived) accrual of an aggregate shock in housing markets, in turn leading to value declines and risk recognition on mortgage- backed securities, culminating in significant lowering of credit ratings on these by the relevant agencies by mid-2007. Such Immediate trading in these assets plays no role in the Bolton et al (2010) model; indeed it is strictly sub-optimal for SRs and LRs to engage in such trades in their setting, as they note. Their reasoning is simple: relative to an Early trading equilibrium, Immediate trading, at a set of prices satisfying $\Pi(\lambda) = [\lambda\rho + (1-\lambda)P_e]$ for SRs to be indifferent between trading at $t = 0$ and $t = 1$, would simply serve to make LR agents worse off, by having to hold a strictly higher amount of liquidity $M_d(\lambda) > M_e(\lambda)$. As a result, any immediate trading equilibrium would result in strictly lower origination of the tradable asset by SRs, leading to a (weakly) Pareto inferior outcome. A similar argument applies vis-a-vis comparing a Delayed vs an Immediate trading equilibrium in their model, as in it Delayed trading equilibrium outcomes Pareto dominate those arising from Early trading, when the former exists.

This latter argument no longer applies, vis a vis a Delayed trading equilibrium, in the modified setting of our model and, as a result, there is a clear possibility for a role for Immediate trading in our set up. However, at least given heterogeneous prior beliefs regarding the likelihood of an (adverse) aggregate valuation shock across agents, not all SR agents would choose to engage in immediate trading either, leading to the possibility of "segmented markets", in which more optimistic SR agents, along with LR agents with higher marginal liquidity holding costs, would wait to trade assets in a Delayed trading equilibrium instead. The reason such a possibility arises in our setting is the following. Unlike in the BSS model, in which their delayed trading equilibrium exhausts all feasible gains from trade across SR and LR agents, and hence is Pareto-preferred by them to the Early trading equilibrium, in our modified set up LR agents would have strictly preferred trading early instead. Indeed, essentially because of this feature of our analysis, it is easily shown that – faced with the prospect of engaging in trade with an SR agent at a delayed equilibrium price P_d – at the margin an LR agent could make herself, and her SR trading partner, strictly better off, by making an offer to buy an unit of the latter's assets early, at time $t = 1$, as well as initially at time $t = 0$. In other words, our Delayed trading equilibrium notion is simply not "strategy-proof".

Lemma 4: *Given a Delayed trading equilibrium price P_d, there is always an Early trade price offer by an LR of $P_o > q\eta\delta\rho + (1 - q\eta)P_d$ - that makes both her and her SR trading partner strictly better off, via exchanging an unit of the asset at this price.*

Proof: Given a P_d, the commonly held conjecture of SR and LR agents about Delayed trading equilibrium price, at the margin an LR agents would be indifferent between reducing her planned trade at $t = 2$ by an unit, and making an offer to use the liquidity freed up to buy P_d/P_o units of an SR's asset early, at the unit price P_o equalling;

$$P_o = P_d \frac{(1 - q\eta)}{(1 - q)}. \tag{30}$$

The SR agent who is offered this price would be strictly better off by selling early if:

$$P_o > q\eta\delta\rho + (1 - q\eta)P_d, \tag{31}$$

which holds provided

$$\eta\delta\rho < P_d \frac{1 - q\eta}{1 - q} \tag{32}$$

Inequality (32) holds for $P_d > \delta\eta\rho$, which is true of any non-trivial Delayed trading equilibrium. In contrast, in the BSS model, the analogue of (32) would require that:

$$\eta\rho < P_d \frac{(1 - q\eta)}{(1 - q)} \tag{33}$$

which contradicts the LR agent's First Order Condition for optimality, in equation (6).

At first sight, this key "weakness" of our Delayed equilibrium notion may appear to be devastating to its validity, leading to the conclusion that the only valid competitive price-taking equilibrium outcomes in our set up could be those which are associated with some Early trading equilibrium. We take a more eclectic view, by bringing in the consideration of the possibility of Immediate trading offers, based on the same idea as in Lemma 4 above. We then show, via an extended Example, that if the LR agents' offers are based on a lower estimate of λ than that of a subset of SR agents, then the latter may not find it profitable to sell their asset immediately, as compared to waiting to trade these at their (conjectured) Delayed trading equilibrium price P_d. What this example does not accomplish, however, is the task of full integration of Immediate trading based on bilateral offers by some SR and LR agents, with others trading in a Delayed, and price-taking equilibrium (with a Lemons discount in prices) later on.

Example: Consider a scenario where $\rho = 1.20$, $\eta\rho = 1$, $\alpha = 0.87$, $\delta = 0.84$, $q = 0.3$, and P_d is such that $q\eta\delta\rho + (1 - q\eta)P_d$ is between .892 and .9. LR agents, and some SRs as well, believe that the ex ante probability of the benign state continuing is $\lambda_p = .35$, whereas as other "exuberant" SR agents believe that it is $\lambda_o = .45$. Both beliefs are consistent with the conjecture that SR agents would prefer to trade in a price-taking Delayed equilibrium over an Early trading one, as $P_e(\lambda_p) = [1 - (1.2)(.35)/(1 - .35)] = 0.892 < .9$. Suppose that LR agents are willing to offer SR agents the equivalent of an early trading price of $P_o = ..92$ in their immediate offers, amounting to offers of $\Pi = (.35)(1.2) + (.65)(.92) = 1.02$. The exuberant SR agents would prefer not to sell immediately at this price, as they conjecture that if they wait and then trade in a Delayed equilibrium, at the price P_d, if and when the aggregate shock would occur, they would obtain the ex ante (at $t = 0$) expected payoff of $(.55)(1.20) + (.45)(.892) = 1.03 > 1.02$, their offered immediate trading price. This gives rise to a market segmentation whereby assets are traded at both $t = 0, 2$. Indeed, one may think of the post aggregate but pre idiosyncratic private information state $t = 1$, as a conceptual rather than a "real time" state, when trading is carried out.

V. Excessive Leverage, Crises, and Alternative Regulatory Policies

It is, of course, well known that the explosion of securitization of (potentially) lower quality and riskier (more heterogeneous) loan products, especially over the years of 2002-7, was funded with much higher (compared to that in other banking activities) and short-term uninsured debt, in the form Repo financing for example. As market doubts about the qualities of the underlying loans, and their implications for valuation of even higher grade securities (tranches) based on these, stated accruing in from late 2006 on wards, leading to significant downgrades by credit rating agencies starting in early 2007, both the rates paid and "haircuts" (margin requirements) demanded on the repo financing increased, as Gorton and Metrick (2009, 2010) have documented. But, the process was slow in the beginning, in the sense that while sales of new securities based on newly originated (by now realized to be lower-grade) loans, to be funded via a lower extent of repo finance, essentially ceased by mid-2007, the process of higher haircuts and rates on such repo financing crept upwards from mid-2007 until the first quarter of 2008, before accelerating to basically full-fledged systemic bank/repo Runs during the summer of 2008. In a magisterial review of the consequences and possible causes of the great financial crisis of (culminating in) 2008, Hellwig (2008) suggests that "we must distinguish between the contribution to systemic risk that came from excessive maturity transformation through SIVs and Conduits (used by levered banks to park their holdings of securitized products) and the contribution to systemic risk that came from the interplay of market malfunctioning, fair value accounting, and the insufficiency of bank equity". What precisely did he mean by "market malfunctioning"?

To us, based in part on our reading of Hellwig (2008), "market malfunctioning" might imply at least the two following aspects of agents' behaviour and its consequences for the "unforeseen nature" of some of the market trades. The first, which Hellwig (2010) discusses under the heading of "excessive confidence in quantitative models", could account for basing leverage choices on currently prevailing levels of the Immediate trading prices of the assets, in a time such as that captured in our example above, with a cushion for potential adverse shocks to prices based largely on very recent historical changes (volatilities) in these. For example, in the context of our Example above, an "exuberant" SR who intends not to sell her asset immediately, at the offered price of 1.02, may take on a leverage level of 0.99 per unit of the asset, even if she reckons that - contingent on the aggregate shock (fully) realizing - her overall Delayed payoff (supportable with a mix of asset sales, and lower leverage on the good subset held on to), is only 0.892. Implicit in such behaviour is her belief or hubris that she would have the capability to sell enough of her asset, prior to an aggregate shock fully manifesting itself in its immediate trading price, to reduce her leverage ratio to .892, from .99. As Rajan (2010) remarks, even Charlie Prince of Citicorp, to whom the by now notorious statement about "keeping on dancing as long as the music is playing" is attributed, expressed a caveat regarding what might happen "if liquidity dried up" in secondary markets for securitized assets that Citi was holding onto, including those it had carried on the books of its SIVs and Conduits, with implicit promises of Citi supporting their debt liabilities via equity injections, if required. This brings us to what we believe is a second important dimension of market malfunctioning and its interactions with debt.

Suppose that, as say over the last two quarters of 2007, the less exuberant LR agents had lowered their estimated likelihood of the benign aggregate state continuing, from 0.35 to 0.25, so that their maximal price offer for immediate trading now decreased to $(.25)(1.2) + (.75)(.92) = .99$, in the context of our Example above. As soon as that happens, repo holders of an SR who had taken on the leverage level of .99 would start a Run, taking the immediate trading price as the maximal liquidation value of the SR they had funded, as in the model of He and Xiong (2009). If sufficiently many SRs, with similar leverage levels as well as absence of inside liquidity (m^*) to finance such withdrawals, then try to sell ALL of their assets immediately, that would lower market prices, ultimately to a level equalling ($\delta\eta\rho$), at which point the secondary market will collapse, and remain so into period $t = 2$, when asymmetric information about qualities of offered assets takes hold. The reason is, of course, that the liquidity available from LR agents for buying these assets equals at most - because some had bought the asset earlier in immediate trading from less exuberant SRs - the level required to support a price level of P_d, for a volume/measure $(1 - q\eta)$ units of assets to be sold in Delayed trading. It would thus not suffice even to support the price level of $P_e(\lambda^*)$ in the Run.

Note that, when one introduces the possibility of endogenous leverage as above, one must specify what one means by the aggregate availability of funds in (or of) the SR Sector. One specification is that this amount - normalized to unity above - represents the maximum amount of equity and debt funds this sector can raise in the aggregate, with any overall mix of debt and equity that it chooses. The other interpretation, as noted by Rajan (2010) and others, is one that arises from "global imbalances" leading to much savings being available in economies without sophisticated financial markets, which clamours for (even ostensibly) "safe" investment products, thereby enhancing the debt funding capacity of institutions such as the SRs in our model, which are then constrained - given a pre-specified level of equity capital - only via some acceptable maximum (possibly regulated) leverage ratio, vis-a-vis aggregate funding capacity. In our interpretation, this distinction may be of importance, in terms of regulatory policy.

Two major regulatory policy interventions that are natural to consider in our setting are minimum Capital, or equivalently maximum Leverage, ratio restrictions, as well as minimum asset price Guarantees, possibly coupled with restrictions on Liquidity ratios m^* ex ante. Let us first consider the former. One may set a maximum leverage ratio on investments in the risky technology to be $P_e(\lambda^*) = q\eta\delta\rho + (1 - q\eta)P_d(\lambda^*)$, in which λ^* represents, as before, the switch point above which SR agents would prefer the Delayed over the Early trading equilibrium. Without much detailed knowledge of the LR agents' opportunity cost function for providing liquidity to the asset market, or $F(I)$, this would not be an easy policy to implement: doing so based on the lowest λ that satisfies our necessary condition in Lemma 3 above, will result in a too generous leverage ratio which may result in Runs as above arising. Even if a regulator has the informational capacity to calculate λ^*, there are still two potential difficulties. For $\lambda > \lambda^*$, $P_d(\lambda)$ in Delayed trading equilibrium with $m^* = 0$ would be decreasing in λ, as with a set of Early trading equilibrium with $m^* = 0$, for $\lambda > \lambda_c > \lambda^*$, as described in Proposition 1 and its Corollary above. On the other hand, if the regulatory maximum leverage ratio is set at the level of say $P_e(\lambda_u)$, that may be overly restrictive in terms of decreasing the pledgeable value arising from securitizing assets, via lowering d in essence. In our Example above, that would mean decreasing the maximum allowed leverage ratio from say .815 to $[1 - .6(1.2(1 - .6)] = .70$, which is not required.

In any event, a maximal leverage ratio will not prevent the planned trading strategies of at least a subset of SR agents – all assigning probabilities $\lambda > \lambda^*$ to the aggregate adverse shock not occurring – being delayed trading coupled with no inside liquidity holding (setting $m = 0$), since there would not be enough external liquidity to absorb so much of risky assets in Immediate or early equilibrium trading. As we have noted above, such delayed trading - in our model, as opposed to that of Bolton et al (2010) - would lead to part of the feasible gains from trade between SR and LR agents being unrealized, decreasing (in our numerical simulations) payoffs aggregated over them.

An alternative regulatory tool, analysed in the BSS paper, would be for the regulator – with access to fiscal or monetary powers - to provide a minimum price guarantee on the asset. The purpose of such a guarantee in our setting would be the opposite of what it is in the BSS setting, which is to support a delayed trading equilibrium when private liquidity provision by LR agents, in the face of the lemons discount in pricing given adverse selection, is insufficient to support it. Our price guarantee will apply to only Early trading, at a level $P_e(\lambda^*)$, or for Immediate trading at an unit price. For $\lambda < \lambda^*$, no SR agent would (strictly) prefer to sell to the regulator at these prices. Instead, they would invest and trade as in the BSS Early trading equilibrium, with $m^* > 0$, or implement an immediate trading equilibrium with a higher level of m^*. However, in these states, they would not be constrained by (overly restrictive, see above) leverage regulations. Indeed, now even when LR agents would believe that $\lambda > \lambda^*$, they would plan to sell their assets early, and thus the possibility of Runs of the sort we outlined above would not arise; LR agents would have no reason to make immediate trading offers that are more advantageous to SRs. However, SR agents would invest - given the simple linear-in-payoffs expected utility functions we have assumed - all of their funding capacity in the long-term asset, setting $m^* = 0$. As a result, some of the sales of assets originated by SR agents would indeed be to the regulator/government, since private external liquidity provide by LR agents would not suffice to support such a price for this volume of asset sales. If the regulator's marginal cost of providing such liquidity is no lower than that of LR agents at the hypothetical Early equilibrium for these levels of λ, the regulator may seek to couple such a price support with minimum liquidity ratios. A tight one would be at $m^*(\lambda^*)$ and a looser one would be at $m^*(\lambda_u)$ provided the latter is indeed non-zero, i.e., $\lambda_u < \lambda_c$ in Proposition 1 above. In setting the level of this minimum Liquidity ratio, optimal regulatory policy would thus need to trade off governmental (deadweight or opportunity) costs of providing asset price support, for excessive asset origination at levels of λ in the neighbourhood of λ^*, as against constraining such asset origination excessively when λ is closer to $\lambda_u >> \lambda^*$.

Maximum leverage ratios could, however, lead to additional advantageous effects via another, a "general equilibrium" channel, if one assumes that levels of equity capital available to SR institutions is limited, and hence leverage regulations would serve to reduce the overall funding capacity of such agents, relative to that of LR agents. In that case, conditions may be created such that we would have Early trading outcomes in which we would have $m^* = 0$ for lower levels of λ_c, possibly lower than λ^*, such that SR agents would thereby obtain an interior share of the surplus created by asset origination in Early trading equilibrium as well, for all λ in $(\lambda^*, \lambda_u]$. In that event, they would be less likely to be tempted to carry out trading in a Delayed equilibrium, thus eliminating the possibility of excessive asset and leverage creation, and Runs.

Finally, we should note another recent contribution, by Diamond and Rajan (2010), which also has the feature that banks – funded with short-maturity demandable liabilities – may trade their assets late rather than earlier, even if earlier trading would have preserved their solvency. In their model, such a phenomenon arises owing to decisions being made by (in the interest of) the levered equity holders of banks, who realize that – given the pattern of anticipated prices in the secondary market for their assets – they would be unable to fully repay their depositors even after asset sales in future, contingent on a "liquidity shock" leading a fraction of these to withdraw earlier. These choices give rise to the possibility of (avoidable) Runs in the high interim liquidty-demand state,They also discuss alternative regulatory interventions, including their impact on banks' chosen trading strategies.

Appendix: Proofs

Proof of Proposition 2. We first observe that if there exists a delayed equilibrium with $m^* = 0$ then the market clearing condition implies that $M_d = (1 - q\eta)P_d$. Substituting the expression for liquidity M_d into LR's first order condition (6) in delayed equilibrium we obtain that the price P_d in delayed equilibrium is given by $P_d = (1 - q\eta)x$, where x solves a nonlinear equation (15).

By comparing the payoffs from early and delayed trades we obtain the following condition guaranteeing that SRs prefer to trade late:

$$P_e \leq q\eta\rho\delta + (1 - q\eta)P_d, \tag{a.1}$$

where P_e denotes the early equilibrium price that the SRs expect to see if they switch to early trade. The expression on the right-hand side of (a.1) represents the expected gain from the delayed trade conditional on observing a bad shock at $t = 1$.

Moreover, if there exists a delayed equilibrium with $m^* = 0$ then P_d should satisfy the following two inequalities:

$$P_d \geq \eta\rho\delta, \qquad P_d \leq \rho\delta. \tag{a.2}$$

Indeed, if the first inequality in (a.2) is violated only lemons are traded at time $t = 2$, which is not consistent with having a non-trivial delayed equilibrium. The second inequality in (a.2) guarantees that the SRs receiving good news at $t = 2$ do not trade the assets (as discussed in Section II). Substituting $P_d = (1 - q\eta)x$ into the inequalities (a.1) and (a.2) and rewriting them as inequalities on x we obtain the following inequality:

$$\max\left\{\eta\rho\delta, \frac{P_e(\lambda) - q\eta\rho\delta}{1 - q\eta}\right\} \leq x(\lambda) \leq \rho\delta. \tag{a.3}$$

The inequality (a.3) imposes restrictions on $\{\lambda, \delta\}$ in equilibrium. Resolving the inequality (a.3) with respect to δ we obtain an equivalent inequality:

$$\delta_*(\lambda) \leq \delta \leq \delta^*(\lambda), \tag{a.4}$$

where δ_* and δ^* are given in (14).

So far, the inequality (a.4) has been derived as a necessary condition for the existence of the delayed equilibrium. However, we observe that this inequality is equivalent to inequality (a.3) which also gives a sufficient condition for the existence of a delayed equilibrium. Indeed, x solves a nonlinear equation (15) and $P_d = x$ defines the price in the delayed equilibrium since all the equilibrium conditions are satisfied. In particular, from (a.3) it follows that inequalities (a.1) and (a.2) are satisfied, and hence under the price P_d the SRs prefer to trade late. Moreover, noting that $M_d = (1 - q\eta)P_d$ we rewrite the non-linear equation (15) as follows:

$$\frac{dF(I = K - M_d)}{dI} = \lambda + (1 - \lambda)\frac{\eta\rho(1 - q)}{(1 - q\eta)P_d}, \tag{a.5}$$

which gives the FOC for LRs. Therefore, M_d defined as $M_d = (1-q\theta)P_d$ indeed gives the optimal liquidity level chosen by LRs that anticipate the delayed equilibrium. This completes the proof that inequality (a.4) defines both necessary and sufficient condition for the existence of a delayed equilibrium.

The uniqueness of the delayed equilibrium follows from the properties of production function $F(\cdot)$. Since $F(\cdot)$ is an increasing and concave function the left-hand side of the equation (15) for x is a monotonically increasing function of x on the interval $(0, K)$ and goes to infinity as $x \to K$. On the other hand, the right-hand side of (15) is a monotonically decreasing function x which becomes infinite when $x \to 0$. Therefore, there exists the unique solution of equation (15) defining the price P_d.

We now demonstrate that the inequality (13) presents an equivalent way of rewriting the necessary and sufficient condition for the existence of the delayed equilibrium. From the inequality (a.4) on δ we observe that the necessary and sufficient condition for the existence of equilibrium is given by the inequality $\delta^*(\lambda) \geq \delta_*(\lambda)$ which implies that there exists at least one equilibrium pair $\{\lambda, \delta\}$ satisfying inequality (a.4). By comparing $\delta^*(\lambda)$ and $\delta_*(\lambda)$ in (14) we obtain that the inequality $\delta^*(\lambda) \geq \delta_*(\lambda)$ is equivalent to the following inequality:

$$\frac{x}{\eta\rho} \geq \frac{P_e - (1-q\eta)x}{q\eta\rho}.$$

Resolving the above inequality with respect to x we obtain that the necessary and sufficient condition for the existence of the delayed equilibrium is given by $x \geq P_{\min}$, where $P_{\min}$ is defined in Proposition 2. Thus, we obtain an exogenous lower bound $P_{\min}$ on the delayed equilibrium price. Since $F'(K-(1-q\eta)x)$ is an increasing function of x the inequality $x \geq P_{\min}$ is equivalent to the inequality (13) in Proposition 2. Therefore, the inequality (13) gives a necessary and sufficient condition for the existence of delayed equilibrium.

We also note that the necessary condition for the existence of a delayed trading equilibrium (10), derived in Lemma 2, is implied by the inequality (13). To demonstrate this, we first observe that $F'(K-(1-q\eta)x) > 1$ by assumption. Therefore, the left-hand side of the equation (15) for x exceeds unity. Consequently, from (15) we obtain the following upper bound on price P_d:

$$P_d \leq \frac{(1-q)\eta\rho}{1-q\eta}. \tag{a.6}$$

Noting that the early equilibrium price satisfies inequality $P_e \geq (1-\lambda\rho)/(1-\lambda)$ allows us to rewrite the inequality $P_d \geq P_{\min}$ as follows:

$$P_d \geq \frac{P_e}{1+q(1-\eta)} \geq \frac{1-\lambda\rho}{1-\lambda}\frac{1}{1+q(1-\eta)}. \tag{a.7}$$

The inequalities (a.6) and (a.7) imply that:

$$\frac{1-\lambda\rho}{1-\lambda}\frac{1}{1+q(1-\eta)} \leq \frac{(1-q)\eta\rho}{1-q\eta}.$$

After simple algebraic manipulations it can easily be demonstrated that the above inequality is tantamount to the necessary condition (10).

Finally, we prove the inequality (16) that gives an upper bound on the size of the interval for the discount δ that supports the delayed trading equilibrium. In particular, from the expressions for the upper and lower bounds for δ in (14) we obtain:

$$\delta^* - \delta_* \leq \frac{x}{\eta\rho} - \frac{x}{\rho} = \frac{x}{\rho}\frac{1-\eta}{\eta}. \tag{a.8}$$

Combining the inequality (a.8) with the inequality (a.6) we obtain:

$$\begin{aligned} \delta^* - \delta_* &\leq \frac{(1-q)(1-\eta)}{1-q\eta} \\ &= \frac{(1-q)(1-\eta)}{1-q+q(1-\eta)\eta} = \frac{1}{\frac{1}{1-\eta} + \frac{q}{1-q}} \\ &\leq \min\{1-\eta, \frac{1-q}{q}\}. \end{aligned}$$

The above inequality demonstrates that the equilibrium existence region shrinks as $\eta \to 1$ or $q \to 1$.

Q.E.D.

References

Acharya, V., H.S. Shin, and T. Yorulmazer (2009), "Crisis Resolution and Bank Liquidity", NBER Working Paper 15567, forthcoming in *Review of Financial Studies.*

Adrian, T. and H.S. Shin (2009a), "The Shadow Banking System: Implications for Financial Regulation", Working Paper, Federal Reserve Bank of New York.

Adrian, T. and H.S. Shin (2009b), "Liquidity, Leverage, and Monetary Policy", *American Economic Review* 89, *Papers and Proceedings.*

Allen, F. and D. Gale (1995), "A Welfare Comparison of Intermediaries and Financial Markets in Germany and the US", *European Economic Review* 39, 179-209.

Bolton, P., T. Santos, and J. Scheinkman (2010), "Outside and Inside Liquidity" forthcoming in *Quarterly Journal of Economics.*

Bolton, P., T. Santos, and J. Scheinkman (2009), "Market and Public Liquidity", *American Economic Review* 89, *Papers and Proceedings*, 59499.

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

Diamond, D.W. and R. Rajan (2000), "A Model of Bank Capital", *Journal of Finance* 55, 2431-2465.

Diamond, D.W. and R. Rajan (2010), "Fear of Fire Sales, Illiquidity Seeking, and the Credit Freeze", Chicago Booth School of Business

Chari, V.V, A. Shourideh and A. Zetlin-Jones (2010), "Adverse Selection, Reputation and Sudden Collapses in Secondary Loan Markets", NBER Working Paper 16080.

Cifuentes, R., G. Ferrucci, and H.S. Shin (2005), "Liquidity Risk and Contagion", *Journal of the European Economic Association* 3, 555-566.

Downing, C., D. Jaffee and N. Wallace (2009), "Is the Market for Mortgage-Backed Securities a Market for Lemons?", *Review of Financial Studies* 22, 2457-2494.

Geanakoplos, J. (2010), "The Leverage Cycle", *NBER Macroeconomics Annual.*

Gorton, G. and A. Metrick (2009), "Runs on Repos", NBER Working Paper.

Gorton, G. and A. Metrick (2010), "Haircuts", NBER Working Paper.

He, J. and W. Xiong (2009), "Dynamic Debt Runs," NBER Working Paper.

Hellwig, M. (1998), "Banks, Markets, and the Allocation of Risks in an Economy", *Journal Institutional and Theoretical Economics* 154, 328-345.

Hellwig, M. (2008), "Systemic Risk in the Financial Sector: an Analysis of the Subprime-Mortgage Financial Crisis", Max Planck Institute for Research, Bonn.

Rajan, R.G (2010), "Fault Lines", Princeton University Press.

Shin, H.S. (2009), "Securitisation and Financial Stability", *Economic Journal*, 118, 309-332.

Stein, J. (2010), "Monetary Policy as Financial Stability Regulation", Harvard mimeo.