Why is capital slow moving? Liquidity hysteresis and the dynamics of limited arbitrage.

James Dow[*], Jungsuk Han[†] and Francesco Sangiorgi[‡]

February 1, 2018

Abstract

Will arbitrage capital flow into a market experiencing a liquidity shock, mitigating the adverse effect of the shock on liquidity? Using a stochastic dynamic model of equilibrium pricing with privately informed capital-constrained arbitrageurs, we show that arbitrage capital may actually flow out of the illiquid market. When some arbitrage capital flows out, the remaining capital in the market becomes trapped because it becomes too illiquid for arbitrageurs to want to close out their positions. This mechanism creates endogenous liquidity regimes under which temporary shocks can trigger flight-to-liquidity resulting in "liquidity hysteresis" which is a persistent shift in market liquidity.

[*]London Business School, Regent's Park, London, United Kingdom, NW14SA, E-mail: jdow@london.edu.

[†]Stockholm School of Economics, Drottninggatan 98, Stockholm, Sweden, SE 11160, E-mail: jungsuk.han@hhs.se

[‡]Frankfurt School of Finance and Management, Adickesallee 32-34, 60322, Frankfurt am Main, Germany, E-mail: F.Sangiorgi@fs.de

1 Introduction

Traditional finance theory derives rational prices for assets based on the arbitrage mechanism. Arbitrage (including trading on private information) pushes market prices towards fundamental value.This pricing mechanism may break down when arbitrageurs are capital-constrained for various reasons. An extensive literature studies the limits to arbitrage and hence, how prices may diverge from fundamental value due to constrained arbitrage capital (e.g., Allen and Gale (1994), Shleifer and Vishny (1997), Gromb and Vayanos (2002), and Brunnermeier and Pedersen (2009)). In principle, any surplus capital should flow to exploit those arbitrage opportunities, but it sometimes happens too slowly (e.g., Duffie (2010)). In other words, mispricing may still persist even with plenty of capital around because capital does not flow to the right markets. But what stops capital from flowing to correct mispricing? It seems paradoxical that capital does not quickly flow to correct prices. So endogenizing the rate of flow of arbitrage capital is a priority for research. This paper seeks to address this need.

First, we distinguish between informed and uninformed capital (much of the existing literature studies models with full information). We assume that uninformed capital is plentiful but informed capital is limited. Even so, isn't there enough informed capital in the economy to flow to where it is needed?

We find that the resolution of this paradox lies in the difference between stock and flow of arbitrage capital; a stock of arbitrage capital does not necessarily translate into a capital flow to arbitrage opportunities. For a capital-constrained trader to invest in a new position, he or she must close out some existing positions. Unless the new position is more profitable, the trader would stick to the existing positions until those asset prices revert closer to fundamental value because of public information or because of trades by subsequent privately informed traders. This means that arbitrage capital plays a dual role; the wedge of mispricing not only decides the profitability of new investment but also decides the speed at which engaged arbitrage capital is released (thus deciding the availability of arbitrage capital).

The dual role of arbitrage capital has several important implications. First, markets may be inefficient because arbitrage capital is "trapped" and efficiency may change over time as trapped capital is released. Second, there can be "liquidity hysteresis" in the form of a long-lasting shift in efficiency as a response to temporary changes in market liquidity. Arbitrage capital indeed does not immediately flow to seemingly profitable arbitrage opportunities, but only do that slowly with a delay. Third, flight-to-liquidity arises when a market suffers liquidity hysteresis. Arbitrage capital will flow to more liquid investment opportunities as illiquidity reinforces itself.

To formalize these ideas, we study a dynamic model of arbitrage with two markets where arbitrageurs freely move between the two, but are capital-constrained. One market is populated with short maturity assets (henceforth "liquidity market"), and the other market is populated with long maturity assets (henceforth, "illiquid market"). Arbitrageurs collect private information on assets, and then trade those assets for speculative gains. In equilibrium, the two markets should offer the same expected speculative profits – otherwise arbitrageurs will move across to the one with higher profits. This means that the illiquid market should have a higher mispricing wedge than that of the liquid market (to compensate for the opportunity cost of longer maturity of investment); price efficiency of the illiquid market should be lower than that of the liquid market. Lower price efficiency in the illiquid market in turn implies that more capital is trapped because it becomes too illiquid for arbitrageurs to want to close out their positions.

The overall efficiency of the markets is determined by how much arbitrage capital is active as opposed to trapped. In other words, efficiency depends on the pool of active capital as a state variable. This matters because while the total stock of arbitrage capital may be large, the stock of active capital may be much smaller. The efficiency of a market may change over time as trapped capital is released from other markets. Furthermore, there is a delayed response in efficiency to changes such as shocks to liquidity trading.

The active (as opposed to trapped) capital, being the state variable of the economy, creates a feedback channel between liquidity and active capital. As more active arbitrage capital flow to the illiquid market, those who are trapped in the market become active again more quickly, and this in turn creates a larger capital flow to the illiquid market by increasing the overall size of active capital in the economy. This virtuous cycle leads to a high information steady state where arbitrage capital is redeployed at a faster rate (thus giving rise to higher liquidity). On the other hand, a vicious cycle may arise, thereby leading to a low information steady state in which arbitrage capital flows to the liquid market leaving locked-in investment in the illiquid market being trapped for a long time.

The feedback channel between active capital and liquidity leads to multiple steady state equilibria in our model; there is a threshold of active capital that separates domains of attraction for liquidity. In addition to comparing properties of these equilibria and impulse responses to unanticipated liquidity shocks, we illustrate our model's implications for liquidity and capital flow dynamics by studying shock responses to a Markov stationary system where (either good or bad) liquidity shocks randomly hit the illiquid market. With a small adverse shock to the illiquid market, market liquidity recovers on its own thanks to a virtuous cycle of liquidity. As more trapped arbitrageurs become active again, they quickly replenish market liquidity. On the other hand, a large adverse shock can trigger a vicious cycle of illiquidity with flight-to-liquidity where arbitrage capital flows to the liquid market; more and more arbitrageurs choose to invest in the liquid market over time because they expect further deterioration of low future liquidity in the illiquid market. This leads to an illiquidity regime where there is a persistent overall lack of liquidity in the market. We call this "liquidity hysteresis" because a shock to the system moves the equilibrium to a different path even after the shock is removed.

We illustrate how the market can move in and out of this illiquidity regime with numerical simulations of the stochastic equilibrium of the model. A sequence of temporary bad shocks to the illiquid market can trigger a flight-to-liquidity resulting in the illiquidity regime, from which the market can recover only after a sequence of shocks in the opposite direction. Thus, the market features persistent (endogenous) liquidity regimes even when (exogenous) liquidity trading is at its normal level most of the time. These results provide a theoretical explanation of slow-moving capital regarding why capital moves slowly, how fast (or slowly) it moves, and to which directions it moves. They further provide interesting policy implications.

The paper is organized as follows. In Section 2, we discuss related literature. In Section 3, we describe the basic model. In Section 4, we solve equilibrium of the model. In Section 5, we discuss empirical and regulatory implications of our model. In Section 6, we conclude.

2 Literature Review

There is a growing literature which explains fluctuations in market liquidity using on the idea of slow-moving capital. Those papers extend the limits-to-arbitrage argument suggesting that slow-moving capital could be the source of prolonged illiquidity even when there is enough capital in the economy. For example, Mitchell, Pedersen, and Pulvino (2007) show convertible bonds traded at prices well below the arbitrage price (relative to the stock and a straight bond) during an extended period when the convertible bond hedge funds (that normally arbitrage these assets) were short of capital, and multi-strategy hedge funds (that opportunistically redeploy capital to wherever returns are high) were slow to enter the market. Duffie (2010) suggest that institutional impediments such as search frictions, taxes, regulations, and market segmentation can slow down capital flow. Our paper contributes to the discussion of slow-moving capital by combining the argument of limits to arbitrage with a dynamic mechanism of capital flow.

Traditional literature in limits to arbitrage shows that the level of arbitrage capital is the state variable that determines market liquidity: cash-in-the-market-pricing (e.g., Allen and Gale (1994)), investors' fund flow (e.g., Shleifer and Vishny (1997)), leverage constraint (e.g., Gromb and Vayanos (2002)), margin constraint (e.g., Brunnermeier and Pedersen (2009)). In these papers, the capital available for arbitrage in the asset market is limited. This is due to frictions, but the frictions that limit the availability of capital in the market are often not explicitly modeled. One of our objectives is to model those frictions. But, in real life, capital is often engaged in other investment activities. So, even though they may not be constrained ex-ante, they are effectively (or ex-post) constrained due to existing investment. Our paper differs from the existing models in that we endogenize such engagement in a dynamic setting. So, we extend the existing argument on limits to arbitrage by showing that the pool of active arbitrage capital (rather than the pool of arbitrage capital itself) is the key state variable determining price efficiency and liquidity.

Our paper is closely related to papers studying the dynamics of arbitrage or intermediary capital movement across multiple markets (or multiple arbitrage opportunities). In most of models in this line of literature, (marginal) expected returns of new investment should be equalized across markets in equilibrium. Consequently, mispricing wedge for long duration assets becomes higher than that of short duration assets as discussed in Shleifer and Vishny (1990). Furthermore, a shock in one market tends to create a spillover effect across other markets where shocks are transferred to other markets through the channel of wealth effects (e.g., Kyle and Xiong (2001)) or collateral constraints (e.g., Gromb and Vayanos (2017) and Gromb and Vayanos (Forthcoming)). The literature also studies the adjustment process of capital after the arrival of a shock. Duffie, Garleanu, and Pedersen (2005) and Duffie, Garleanu, and Pedersen (2007) find a gradual process of recovery after a shock to investors' preference in search-based models. In Duffie and Strulovici (2012) where financial intermediaries trade off the cost against the benefit of intermediation, the speed of capital flow is governed by the imbalance of capital as well as the level of intermediation competition across markets. In Gromb and

Vayanos (Forthcoming), there is a phase with an immediate increase in the spread where arbitrageurs decrease their positions (thus, causing contagion effect), and then it is followed by a recovery phase. Our paper has several important differences from this line of literature. In this line of literature, unlike our paper, there is full information. Prices diverge from fundamentals because of exogenous demand shocks. Price is determined by supply and demand (unlike our model where price is equal to expected value conditional on public information), but the amount of arbitrage capital in the economy is too small to make prices equal fundamentals. By contrast in our paper arbitrage capital is the capital belonging to informed traders; if their information becomes public, the price incorporates it. Because informed capital is often already engaged in investment, the level of active informed capital becomes the state variable which governs market liquidity.[1] We find a temporary shock can leave a long-lasting (or even permanent) impact to market liquidity if active informed capital is impaired to the point under which it cannot recover on its own. We contribute to this line of literature by providing alternative explanations of flight-to-liquidity and slow-moving capital in light of liquidity hysteresis.

3 Setup

We consider an infinite horizon discrete time economy with a continuum of long-lived agents. All agents have risk neutral preferences with a discount factor of β. There exists a risk-free asset in the economy whose return is equal to $r_f = 1/\beta - 1$.

There is a continuum of financial securities which are claims to a single random liquidation value. There are two classes of securities that differ in their maturity: (i) "liquid assets" which are short-lived, and (ii) "illiquid assets" which are long-lived. At this point in the paper, calling the assets "liquid" and "illiquid" is just convenient terminology, since liquidity is an equilibrium property of an asset and we have not yet characterized the equilibrium. We will show later that this terminology is justified and the "liquid" assets are indeed more liquid.

Illiquid assets are traded in market I, and liquid assets are traded in market L. It is important to note that they are not segmented markets because capital can freely

[1] Our model also differs from most of noisy rational expectations equilibrium models in that we study capital movement across multiple markets in a dynamic setup. For example, Dow and Gorton (1994) study a multiperiod model of limits to arbitrage where the cost of carry combined with arbitrageurs' short horizon break down the chain of arbitrage in a single asset market. Dow and Han (Forthcoming) study a static noisy rational expectations model with endogenous adverse selection in asset supply where the presence of informed capital facilitates movement of uninformed capital.

move between two markets without any friction. An asset in market L has a one period maturity; it pays its liquidation value in the subsequent period after issuance.[2] On the other hand, an asset in market I has a random maturity; it pays its liquidation value in every period with probability $q > 0$. At maturity, any asset i in market $h \in \{L, I\}$ pays V^i which is either high ($V^i = V_H^h$) or low ($V^i = V_L^h$) with equal probabilities.[3] We further assume that the present value of assets in both markets are identical.[4] Asset payoffs are independent across assets and time.

The assets are called "unrevealed" if their prices do not reflect true fundamental value (because payoffs are unknown), and "fully-revealed" if prices reflect true fundamental value (because payoffs are known). Payoffs can become known if the liquidation value is fully revealed by the trading process and asset prices. For simplicity, we assume that the mass of unrevealed assets is fixed to one unit in each market at any point of time. That is, new assets are issued to replace those which either realized payoffs, or become fully-revealed.[5]

There is a unit mass of capital-constrained "arbitrageurs" who trade to generate speculative profits. Each arbitrageur can produce private information about the payoff of one asset in each period. All arbitrageurs who investigate an asset can perfectly observe the value of its liquidation value. For mathematical tractability, we assume a simple form of capital constraint under which each arbitrageur can hold at most one unit of unrevealed assets at any point of time. That is, they can acquire new unrevealed positions upon liquidating their existing unrevealed positions.

There is a continuum of competitive risk-neutral market makers who set prices to clear the market as in the Kyle (1985) model. There are also noise traders who trade for exogenous reasons such as liquidity needs. In each period, arbitrageurs and noise

[2]For simplicity, we assume that one of the class of assets pays every period (liquid assets). It would be possible, but considerably more complex, to analyze the case with a payoff probability less than one for all assets.

[3]The assumption that payoffs are high or low with equal probabilities simplifies the analysis by making profits from long and short positions symmetric.

[4]This is simply for mathematical convenience. Under this assumption, we have

$$\frac{\beta q}{1-\beta(1-q)}(V_H^I - V_L^I) = \beta(V_H^L - V_L^L),$$

where the left- and right- hand side is the present value of difference in high and low payoff in market I and L, respectively.

[5]One can consider that there exists a unit mass of firms which issue new securities to invest in new projects whenever their existing projects pay liquidation value or become fully-revealed. If we assume, instead, that unrevealed assets that become fully-revealed are not immediately replaced, the model would require an additional state variable and would be considerably more complex to analyze.

traders submit market orders to the market makers. In period t, for each asset i in market $h \in \{L, I\}$, noise traders submit an aggregate order flow of z_t^i which follows an independent uniform distribution on $[-\bar{z}_t^h, \bar{z}_t^h]$. The magnitude of $\bar{z}_t^h$ captures the intensity of noise trading in market h in period t. We assume that $\bar{z}_t^L$ is equal to a constant $\bar{z}^L$ for any period t whereas $\bar{z}_t^I$ follows a Markov process with N states and corresponding values $Z_1 < Z_2 < ... < Z_N$. The transition matrix between states is given by

$$\Omega = \begin{bmatrix} \omega_{11} & \dots & \omega_{1N} \\ \vdots & \ddots & \vdots \\ \omega_{N1} & \dots & \omega_{NN} \end{bmatrix}$$

We further assume that there are enough noise trading activities in the market to prevent the price for every asset from being fully-revealing; the support of aggregate noise trading is strictly greater than that of arbitrageurs' aggregate order flow: $\bar{z}_t^I + \bar{z}_t^L > 1$ for all t. Finally, we assume that the process of noise trading intensity and the realizations of asset payoffs are mutually independent.

The timing of events in each period is as follows. At the beginning of the period, asset payoffs realize and they are distributed among claim holders. Next, new assets are issued. After these events, arbitrageurs collect private information on unrevealed assets, then submit orders to market makers. At the end of the period, market makers post asset prices and trades are finalized.

4 Equilibrium

4.1 Laws of Motion

In each period t, each arbitrageur is in one of two situations: "active" or "locked-in". An active arbitrageur does not have any existing position in unrevealed assets, thus, has available capital for new investment whereas a locked-in arbitrageur has already an existing position in unrevealed assets, thus, does not have available capital for new investment unless the existing position is liquidated.[6] We denote ξ_t to be the the mass of active arbitrageurs, and π_t to be the mass of locked-in arbitrageurs where $\xi_t + \pi_t = 1$. Each active arbitrageur chooses to hold a new position in either market I or L. δ_t denotes the portion of those choosing to trade assets in market I in period t (thus, $1 - \delta_t$

[6]Notice that those who hold assets with fully-revealing prices are active.

denotes the portion of those choosing to trade assets in market L).

We focus on a class of market-wise symmetric equilibria in which price efficiency is symmetric across all assets in the same market. In other words, the mass of informed trading is equally distributed across all assets in the same market. In that class of equilibria, the mass of informed traders who acquire information on each asset should be given by the total mass of investors who acquire information on market h assets divided by the total mass of market h assets. Then, we can show the following result for price efficiency in each market:

Lemma 1 *The probability of information revelation in market I is given by*

$$\lambda_t^I = \frac{\delta_t \xi_t}{\bar{z}_t^I}. \tag{1}$$

Likewise, the probability of information revelation in market L is given by

$$\lambda_t^L = \frac{(1-\delta_t)\xi_t}{\bar{z}_t^L}. \tag{2}$$

Proof. See Appendix. ■

Then, the laws of motion of the mass of each group of arbitrageurs are given by

$$\xi_{t+1} = (1-\delta_t)\xi_t + (\delta_t \xi_t + \pi_t)(q + (1-q)\lambda_t^I); \tag{3}$$

$$\pi_{t+1} = (\delta_t \xi_t + \pi_t)(1-q)(1-\lambda_t^I). \tag{4}$$

The first equation describes the evolution of active capital ξ_t. The right hand side of Eq. (3) is the sum of two terms. The first term is the mass of arbitrageurs invested in market L at time t; this mass becomes entirely active in $t+1$ as the L assets are short-lived. The second term is the mass of arbitrageurs invested in market I at time t (i.e., $\delta_t \xi_t$ new arbitrageurs from the current period and π_t arbitrageurs locked-in from the previous period) that become available for new investment in $t+1$. This happens either if the asset pays off or if the market price fully reveals the asset value (in which case the position becomes risk-free, thus relaxing the portfolio constraint). Overall, a fraction $q + (1-q)\lambda_t^I$ of the arbitrageurs invested in the I market at time t becomes free for new investment in $t+1$. Finally, we remark that equation Eq. (4) for the evolution of locked-in capital π is redundant given the fact that $\xi_t + \pi_t = 1$ for all t.

4.2 Arbitrageurs' Investment Decisions

We denote the vector of state variables $\theta_t = (\xi_t, \bar{z}_t^I)$, and $E_t[y] = E[y|\theta_t]$ to be the expectation of any random variable y conditional on θ_t. Given state θ_t, we denote $J_I(\theta_t), J_L(\theta_t)$ to be the value function of an arbitrageur who invests in a new position in market I and market L, respectively. Also, we denote $J_l(\theta_t)$ to be the value function of a locked-in arbitrageur in an asset in market I given θ_t. Because any active arbitrageur can choose between the two markets, the value function of an active arbitrageur given θ_t equals

$$J_f(\theta_t) = \max\left(J_I(\theta_t), J_L(\theta_t)\right). \tag{5}$$

Due to Lemma 1, asset prices are either fully-revealing or non-revealing. We denote P_L^h and P_H^h to be the fully-revealing price of an asset in market h whose value is revealed low and high, respectively. We also denote P_0^h to be the non-revealing price of an asset in market h. Using the symmetry of trading profits between long and short positions, we can obtain the following value functions given θ_t:

$$\begin{aligned}
J_I(\theta_t) &= -(\lambda_t^I P_H^I + (1-\lambda_t^I)P_0^I) + \beta\Big[qV_H^I + (1-q)\lambda_t^I P_H^I + (1-(1-\lambda_t^I)(1-q))E_t[J_f(\theta_{t+1})] \\
&\quad + (1-\lambda_t^I)(1-q)E_t[J_l(\theta_{t+1})]\Big], \\
J_L(\theta_t) &= -(\lambda_t^L P_H^L + (1-\lambda_t^L)P_0^L) + \beta\Big[V_H^L + E_t[J_f(\theta_{t+1})]\Big].
\end{aligned}$$

Likewise, the value function of a locked-in arbitrageur given θ_t equals

$$J_l(\theta_t) = \max\left(J_E(\theta_t), J_S(\theta_t)\right),$$

where $J_E(\theta_t)$ is the value function from liquidating the position right away and $J_S(\theta_t)$ is the value function from holding on the position one more period:

$$\begin{aligned}
J_E(\theta_t) &= \lambda_t^I P_H^I + (1-\lambda_t^I)P_0^I + \beta E_t[J_f(\theta_{t+1})], \\
J_S(\theta_t) &= \beta\Big[qV_H^I + (1-q)\lambda_t^I P_H^I + (q + (1-q)\,\lambda_t^I)E_t[J_f(\theta_{t+1})] + (1-\lambda_t^I)(1-q)E_t[J_l(\theta_{t+1})]\Big].
\end{aligned}$$

4.3 Equilibrium Liquidity Dynamics

Equilibrium is defined in a standard manner:

Definition 1 *A stationary equilibrium is a collection of processes for asset prices $\{P_t^i\}$, the mass of active arbitrageurs $\{\xi_t\}$, and arbitrageurs' portfolio choice $\{x_t^j\}$ such that:*

1. *Each active arbitrageur j optimally chooses to acquire private information and hold a position $x_t^j \in [-1, 1]$ in a unrevealed asset to maximizes the present value of his trading profits.*

2. *Asset prices equal expected discounted liquidation values conditional on order flow information:*
$$P_t^{i,} = E_t\left[\beta^{\tau_i} V^i \middle| X_t^i\right] \qquad \text{for all asset } i \tag{6}$$
where τ_i is the the maturity of asset i and $X_t^i = \int x_t^j dj + z_t^i$ is the aggregate order flow for asset i.

3. *The law of motion of $\{\xi_t\}$ satisfies Eq. (3).*

An equilibrium is said to be interior in period t if $J_I(\theta_t) = J_L(\theta_t)$, so that active arbitrageurs are indifferent between investing in the L and I markets. In case of an interior equilibrium, we can show the following result:

Lemma 2 *In an interior equilibrium at t, locked-in arbitrageurs do not have incentives to liquidate early, $J_l(\theta_t) = J_S(\theta_t)$.*

Proof. See Appendix. ■

Intuitively, by liquidating early, an arbitrageur incurs the risk of losing the capital gain (or liquidating dividend) he could obtain for sure by holding on to the position. Hence, an arbitrageur whose capital is engaged finds it optimal not to close out the position (and therefore be inactive) until either the price fully reveals the asset value or the asset pays off.

An interior equilibrium has the following implications for cross sectional and dynamic properties of price informativeness and liquidity.

Lemma 3 *Whenever the equilibrium is interior in t and $t+1$, equilibrium price informativeness satisfies*

$$\lambda_t^L - \lambda_t^I = \beta\left(1-q\right)\left(1-\lambda_t^I\right)\left(1-E_t[\lambda_{t+1}^I]\right); \tag{7}$$

Proof. See Appendix. ∎

The left-hand side of Eq. (7) is the difference in probabilities of trading at fully-revealing price in market L over market I at time t. Equivalently, this difference measures how more likely an arbitrageur is to make a speculative profit when trading in market I compared to trading in market L at time t.

The right-hand side of Eq. (7) is the current probability $(1-q)(1-\lambda_t^I)$ of remaining locked in a trade in market I, weighted by the discount factor β, and multiplied by expected *future* illiquidity in market I, captured by the term $1 - E_t[\lambda_{t+1}^I]$.

The indifference condition Eq. (7) requires $\lambda_t^L > \lambda_t^I$. This is intuitive: by trading in market I, a speculator gives up the certainty of being able to re-trade in the next period; for arbitrageurs to be indifferent between the two markets the I market must compensate this opportunity cost with a higher probability of trading at non-revealing price in the current period. Therefore, a shock to expected future illiquidity in market I requires active arbitrageurs to flow out of market L in the current period for $\lambda_t^L - \lambda_t^I$ to increase so (see Eqs. (1)-(2)).

If equilibrium is interior for all t, recursively substituting Eq. (7) into itself yields

$$\lambda_t^L - \lambda_t^I = E_t\left[\sum_{\tau=1}^{\infty} \beta^\tau (1-q)^\tau \prod_{j=0}^{\tau-1} \left(1-\lambda_{t+j}^I\right)\left(1-\lambda_{t+\tau}^L\right)\right]. \tag{8}$$

Notice that $(1-q)^\tau \left(1-\lambda_t^I\right) ... \left(1-\lambda_{t+\tau-1}^I\right)$ is the probability that by investing in market I in the current period, an arbitrageur's capital is not available for a new trade in period $t+\tau$. Because $\left(1-\lambda_{t+\tau}^L\right)$ is the probability of realizing a speculative profit in $t+\tau$ in market L, the right hand side of Eq. (8) measures the expected loss in future speculative profits arising from capital being trapped in future periods. This is the opportunity cost of trading in market I.

4.4 Steady State Equilibrium

In this subsection, we derive steady state equilibrium under the assumption that noise trading intensity is fixed at a constant level, i.e., $\bar{z}_t^I = \bar{z}^I$ for all t for some constant $\bar{z}^I$. We denote ξ and π to be the steady-state-level mass of active and locked-in arbitrageurs, respectively, and also denote λ^L and λ^I to be the steady-state-level price informativeness in market L and I, respectively.

In steady state, the indifference condition in Eq. (7) can be expressed in terms of λ^L and λ^I as follows:

$$\lambda^L - \lambda^I = \beta(1-q)(1-\lambda^I)^2. \tag{9}$$

Eq. (9) reveals that price informativeness (or efficiency) plays a dual role. On the one hand, price informativeness determines the profitability of investment opportunities: higher λ^I (and also λ^L) decreases the probability of acquiring a new position at non-revealing prices. We term this the "first lambda" effect of price efficiency on speculative profits. On the other hand, price informativeness determines the maturity of investment opportunities in longed lived assets: higher λ^I increases the likelihood of closing out a position with profits earlier. We term this the "second lambda" effect of price efficiency on speculative profits.

Substituting Eqs. (1) and (2) into Eq. (9) yields the following steady state relationship between δ and ξ implied by arbitrageurs' indifference condition:

$$\frac{\bar{z}^L - (1-\delta)\xi}{\bar{z}^L} = \left(\frac{\bar{z}^I - \delta\xi}{\bar{z}^I}\right)\left[1 - \beta(1-q)\left(\frac{\bar{z}^I - \delta\xi}{\bar{z}^I}\right)\right]. \tag{IC}$$

For a fixed δ, a decrease in active arbitrage capital ξ decreases price efficiency in both markets. This has a (positive) first lambda effect on speculative profits in both markets but a (negative) second lambda effect in market I, which becomes relatively less attractive. Hence, δ must decrease to restore arbitrageurs' indifference condition across markets.[7] As arbitrage capital becomes scarce, more capital leaves the illiquid market to join the liquid market, thereby making the illiquid market even more illiquid.

We summarize these findings in the following lemma:

Lemma 4 *When $\frac{\bar{z}^I}{\bar{z}^L} + 1 \geq 2\beta(1-q)$, the IC curve implicitly defines δ as an increasing function of ξ.*

Proof. See Appendix. ■

An interior steady state equilibrium is found at the intersection of the (IC) curve and the following capital movement (CM) curve obtained from the law of motion for

[7] Lemma 4 provides the sufficient condition for the net benefit of trading in market I to decrease as δ increases, for a fixed value of ξ.

active arbitrage capital in Eq. (3) together with Eq. (1) for λ^I:

$$\xi = (1-\delta)\xi + (\delta\xi + 1 - \xi)\left(q + (1-q)\frac{\delta\xi}{\bar{z}^I}\right). \tag{CM}$$

Notice that an increase in the fraction of active arbitrageurs that invest in market I has two opposing effects. On the one hand, as δ increases, more arbitrageurs remain trapped in market I. This tends to reduce steady state value for active capital ξ. On the other hand, an increase in δ improves price efficiency in market I, which increases the rate at which arbitrage capital is released form this market. This feedback effect tends to increase ξ. Which effect dominates depends on the model parameters. The first effect dominates in the right hand panel of Figure 1 for δ is small, while the second effect dominates for δ large. Intuitively, increasing the rate at which trapped capital is released has a bigger effect when the mass of arbitrageurs that are invested in market I is larger.

We can show existence of the steady state equilibrium:

Proposition 1 *A steady state equilibrium exists, and there is either one or two stable (saddle point) equilibria. Equilibrium is always interior ($\delta \in (0,1)$) if $\beta\,(1-q)\,\bar{z}^L < 1$, whereas $\delta = 0$ is also an equilibrium if $\beta\,(1-q)\,\bar{z}^L \geq 1$.*

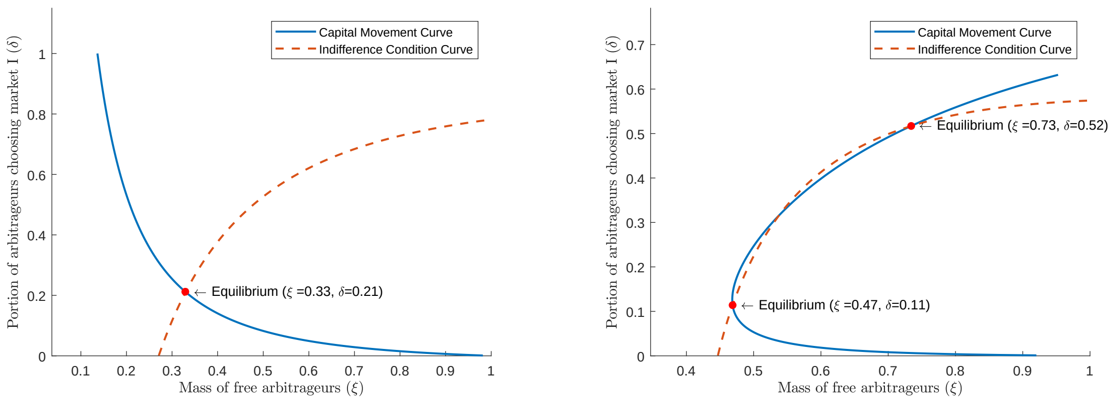
Proof. See Appendix. ■

Figure 1 illustrates the steady state equilibrium values for ξ and δ determined by the intersection of the IC and CM curves. Equilibrium is unique in panel (a), whereas there are three equilibria in panel (b), of which two are stable and one (for intermediate values of ξ and δ) is unstable. Figure 2 illustrates the region of noise trading intensity in the illiquid market where there is uniqueness or multiplicity.

4.5 Shock Response in Steady State Equilibrium

Now, we turn to analysis of the response of the system to an unanticipated temporary shock to noise trading intensity in market I, whereby $\bar{z}^I$ deviates from its normal level to a higher value only in one period.

To shed light on the response to this temporary liquidity shock, we rearrange the



(a) Unique steady state equilibrium

(b) Multiple steady state equilibria

Figure 1: **Steady State Equilibrium.** Parameter values for the unique steady state equilibrium in panel (a): $q = .05, \bar{z}^I = 1.5, \bar{z}^L = .3, \beta = .95$. Parameter values for the multiple steady state equilibria in panel (b): $q = .01, \bar{z}^I = .65, \bar{z}^L = .475, \beta = .95$

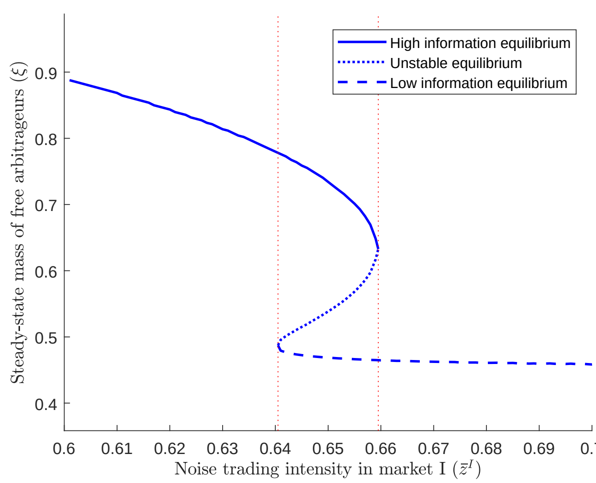
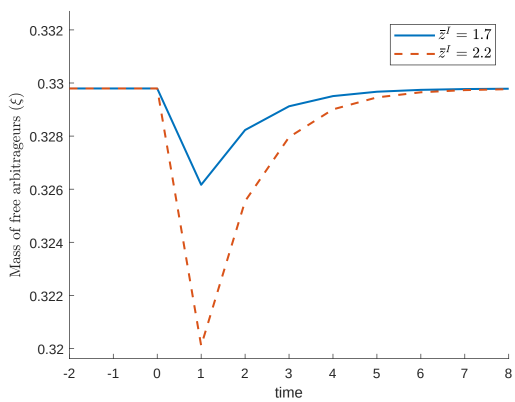


Figure 2: **Steady State Equilibria With Respect to Various Values of $\bar{z}^I$.** Parameter values: $q = .01, \bar{z}^L = .475, \beta = .95$

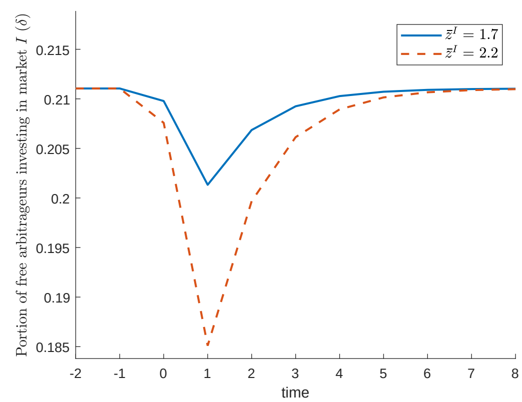
indifference condition in Eq. (7) as follows:

$$1 - \lambda_0^L = \left(1 - \lambda_0^I\right)\left(1 - \beta\left(1 - q\right)\left(1 - E_0[\lambda_1^I]\right)\right) \tag{10}$$

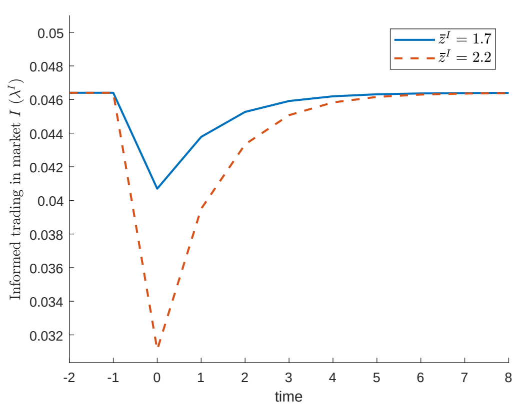
Consider the effect of the shock to both sides of Eq. (10) in case δ_0 does not react to the shock. By Eq. (2), the left hand side of Eq. (10) is unaffected, while the right hand side is affected via two channels. First, λ_0^I would drop (see Eq. (1)), making investment in market L more attractive (first lambda effect). But, lower λ_0^I implies that ξ_1 would also drop because current locked-in capital is released at a lower rate (see Eq. (3)). This decreases λ_1^I and implies that market I is more illiquid in $t = 1$ (second lambda effect). Arbitrageurs that consider investing in market I at $t = 0$ must trade off the larger probability of trading at a non-revealing price in the current period with the longer expected duration of the investment and therefore the larger opportunity cost of being inactive in future periods. When this second effect is sufficiently strong, a larger fraction of active arbitrageurs flows into market L and away from market I.



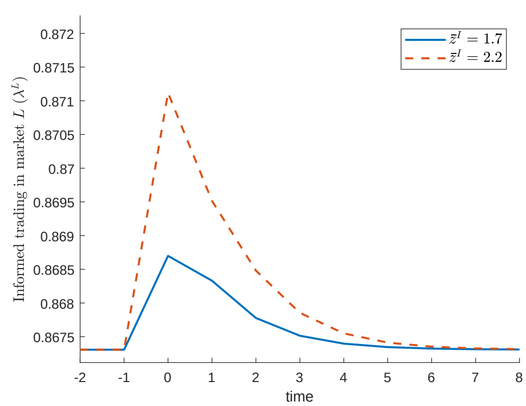
(a) Free arbitrageurs



(b) Investment in illiquid market



(c) Price informativeness in illiquid market



(d) Price informativeness in liquid market

Figure 3: **Transitional Dynamics for a Temporary Shock under a Unique Steady State Equilibrium.** Parameter values: $q = .05, \bar{z}^I = 1.5, \bar{z}^L = .3, \beta = .95$, a shock of $\bar{z}^I = 1.7$ (solid line) and $\bar{z}^I = 2.2$ (dotted line) is given at $t = 0$

Figures 3 shows the impulse response to shocks to $\bar{z}^I$ when the parameters of the model are as in panel (a) of Figure 1. Arbitrageurs react to the shock by flowing out of market I as they anticipate lower liquidity and larger opportunity cost of being locked in this market going forward. The resulting reduction in δ_t continues up to the point where the large probability of realizing a speculative gain in market I compensates for the larger opportunity cost of remaining locked in. The initial shock at $t = 0$ leads to a drop in price informativeness in market I and therefore a decrease in ξ_t from its steady state value starting from $t = 1$. Meanwhile, price informativeness increases in market L as arbitrageurs flow into this market, exacerbating the differential in price informativeness across markets. Then the shock is absorbed and the economy converges to its (unique) steady state. A larger value for the initial shock (dashed line) increases the magnitude of the response but not its qualitative features.

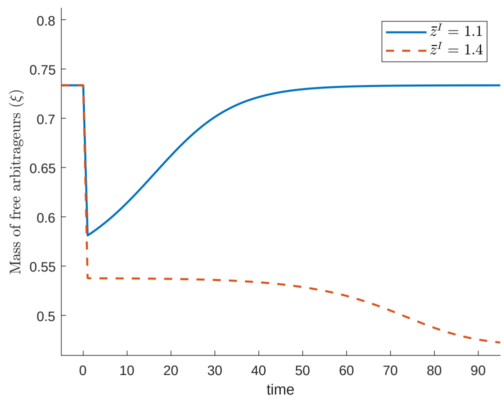
Figure 4 shows the impulse response to shocks to $\bar{z}^I$ when the parameters of the model are as in panel (b) of Figure 1 and the economy is in the high liquidity steady state ($\xi = 0.73$, $\delta = 0.52$). The response to a small shock is qualitatively the same as in Figure 3.[8] By contrast, the response to a larger shock has different dynamics. Instead of reverting back to the initial steady state value, the outflow of arbitrageurs persists as the economy transitions to the stable interior equilibrium shown in Figure 1 that features low values for δ and ξ ($\xi = 0.46$, $\delta = 0.11$). This temporary shock has permanent effects on price informativeness in both markets.

Figure 5 illustrates the dynamics of the mass of active arbitrageurs after an initial shock that pushes ξ away from its steady state value. ξ converges back to its initial steady state value after small shocks. However, there exists a threshold level of ξ such that the economy transitions to a different steady state equilibrium when ξ crosses this threshold. This is the value of ξ corresponding to the the unstable equilibrium in panel (b) of Figure 1 (the middle intersection between the IC and CM curves).

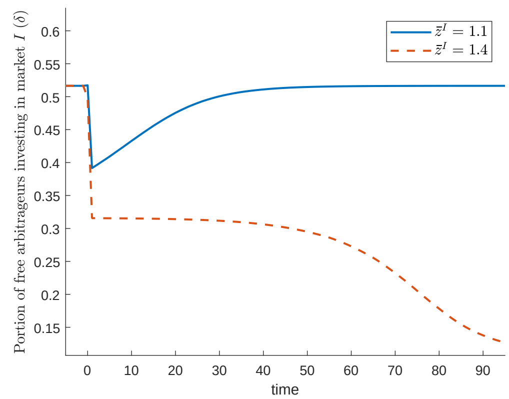
4.6 Stochastic Equilibrium

Here we consider shock response in the stochastic model. In contrast to the analysis in the previous subsection, arbitrageurs anticipate the possibility that a change in noise trading in market I might occur on the equilibrium path. Figures 6 and 7 are obtained from the numerical solution of the stochastic model assuming that $\bar{z}_t^I$ is a Markov process as outlined in Section 3.

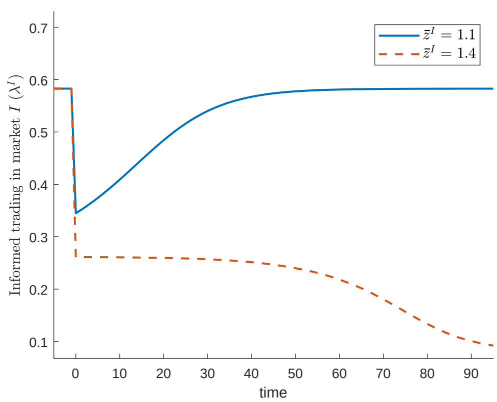
[8]The only qualitative difference is that λ^I decreases after the shock; this is because the reduction in active capital ξ more than offsets the flow of active arbitrageurs in market L.



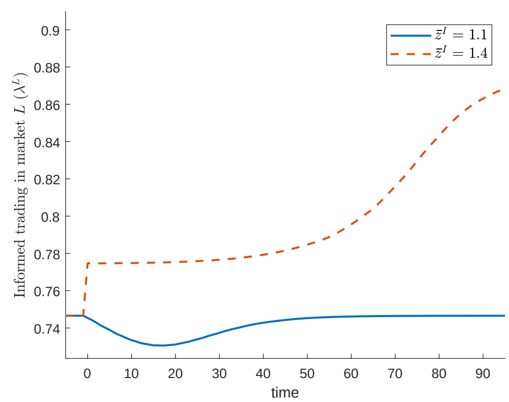
(a) Free arbitrageurs



(b) Investment in illiquid market



(c) Price informativeness in illiquid market



(d) Price informativeness in liquid market

Figure 4: **Transitional Dynamics for a Temporary Shock under a Multiple Steady State Equilibrium.** Parameter values: $q = .01, \bar{z}^I = .65, \bar{z}^L = .475, \beta = .95$, a shock of $\bar{z}^I = 1.1$ (solid line) and $\bar{z}^I = 1.4$ (dotted line) is given at $t = 0$

Figure 6 reproduces the qualitative features of Figure 4 while arbitrageurs anticipate shocks to occur on the equilibrium path. Here we assume that $\bar{z}^I_t$ can take two states: a normal and persistent state, and a higher and transitory "shock" state. Figure 6 shows responses to different shock sizes when the shock lasts only one period. Liquidity has convergent dynamics after a small shock, whereas a larger shock has persistent effects: the shock triggers a flight-to-liquidity and the system transitions to a different regime. In this regime, liquidity in market I remains low and a large fraction of capital remains locked in, reducing active capital on a permanent basis.

In equilibrium, the market can move in and out the illiquidity regime. We illustrate this in Figure 7, which shows a simulation of the stochastic model when there is a

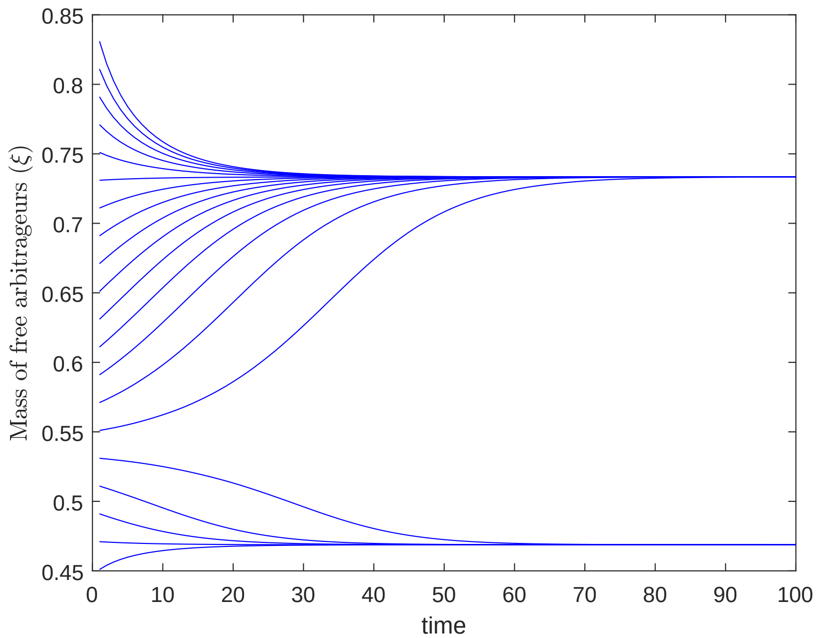
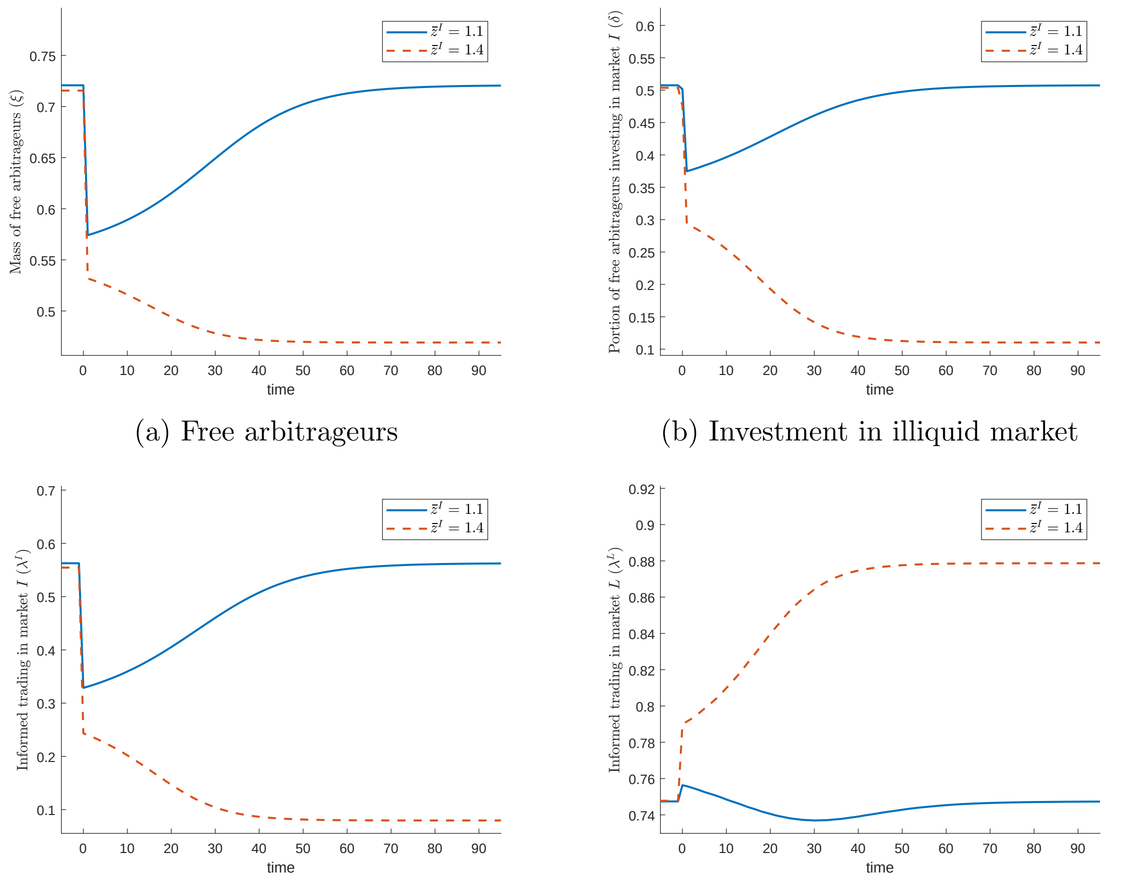


Figure 5: **Evolutionary Paths of the Mass of Free Arbitrageurs under Various Initial Values of ξ.** parameter values in the initial equilibrium: $q = .01, \bar{z}^I = .65, \bar{z}^L = .475, \beta = .95$

"normal" and persistent level for noise trading intensity in market I but with small probability noise trading intensity can jump up or down from its normal level, and these shocks are not persistent. The occurrence of temporary shocks does not have persistent effects in the first portion of the simulation. It is only when bad shocks occur for several consecutive periods (around $t = 220$ in the figure) that there is a sustained flight-to-liquidity and the economy enters a different regime. In the figure, the initial high liquidity regime for market I (white area) is followed by a low liquidity regime for market I (shaded area) after the occurrence of a sequence of bad shocks in this market. The economy is therefore trapped in this regime for many periods even though noise trading is in its normal state most of the time during these periods. It takes a sequence of good shocks in market I for the economy to exit this illiquidity regime revert back to the high liquidity regime for market I. Along the transition, capital flows to market I and improves liquidity in this market; as a result locked in capital is released at a faster rate, further increasing liquidity and price informativeness.



(a) Free arbitrageurs

(b) Investment in illiquid market

(c) Price informativeness in illiquid market

(d) Price informativeness in liquid market

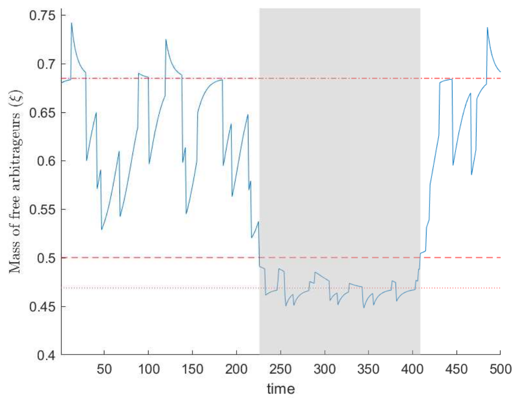
Figure 6: **Transitional Dynamics for a Temporary (Stochastic) Shock under a Multiple Steady State Equilibrium.** Parameter values: $q = .01, \bar{z}^I = .65$ (normal level), $\bar{z}^L = .475, \beta = .95$, a shock of $\bar{z}^I = 1.1$ (solid line) and $\bar{z}^I = 1.4$ (dotted line) is given at $t = 0$, the transition probability is given by $\omega_{11} = .95, \omega_{12} = .05, \omega_{21} = .95, \omega_{22} = .05$ where state 1 is the state with a normal level of $\bar{z}^I$, and state 2 is the state with a high level of $\bar{z}^I$

5 Discussion

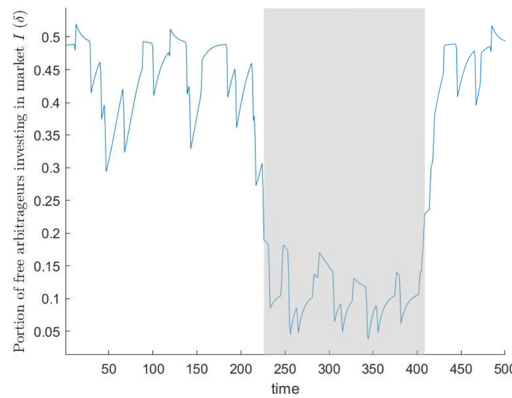
In this section, we discuss empirical implications of our model.

5.1 First Lambda vs. Second Lambda Effects

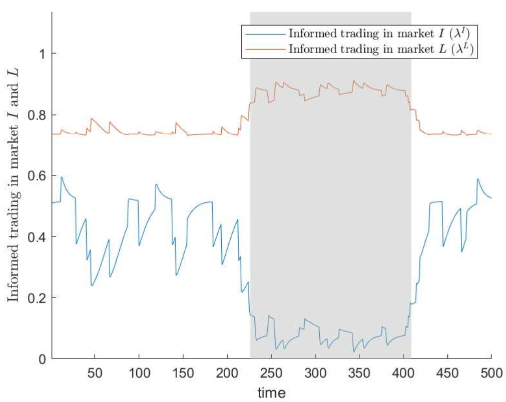
In our model, price efficiency (lambda) plays a dual role in long-lived assets. It determines mispricing wedge which determines the profitability of investment opportunities



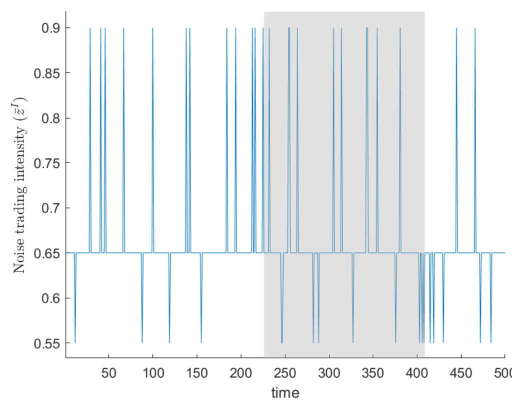
(a) Free arbitrageurs



(b) Investment in illiquid market



(c) Price informativeness



(d) noise trading intensity in illiquid market

Figure 7: **Simulation.** Parameter values: $q = .01, \bar{z}^I = .65$ (normal level), $\bar{z}^L = .475, \beta = .95$, there are three possible values of $\bar{z}^I \in \{.55, .65, .9\}$. Transition probabilities are given by $\omega_{11} = .03, \omega_{12} = .97, \omega_{12} = 0, \omega_{21} = .04, \omega_{22} = .92, \omega_{23} = .04, \omega_{31} = 0, \omega_{32} = .92, \omega_{33} = .08$ where states 1,2 and 3 correspond to low, normal and high level of $\bar{z}^I$, respectively.

but also determines the maturity of new investment which determines its liquidity. Upon the arrival of a liquidity shock, two opposing effects arise together. We call the effect of increased profitability the first lambda effect, and the effect of decreased liquidity the second lambda effect. The first lambda effect is closer to the traditional interpretation of Kyle's lambda. In that context, lambda measures price impact of trade and is often interpreted as a measure of illiquidity from the point of view of uninformed traders. On the other hand, the second lambda determines the speed at which arbitrageurs can close out their positions at a profit because subsequent trades make prices efficient. Hence, second lambda is a measure of liquidity from the point of view of informed arbitrageurs in our model.

Our model implies that the mass of active capital is the key state variable that determines the market liquidity. However, active capital is difficult to observe. On the other hand, price efficiency is often measured using various empirical measures. Our theory predicts that those empirical measures of price efficiency can be used as a proxy for the actual state variable–the mass of active capital. In the following subsections, we discuss how this idea can be applied to some empirical implications.

5.2 Liquidity Crises

There are several well-known episodes of liquidity crisis such as the 1987 stock market crash, the 1998 Long-Term Capital Management crisis, and the subprime mortgage crisis of 2007-2009. These episodes are often characterized by a delayed recovery of liquidity in the aftermath (e.g., Mitchell, Pedersen, and Pulvino (2007); Coval and Stafford (2007)). Existing literature often explains those liquidity crises as a result of shock amplifications which impair capital itself.[9] In our model, a liquidity crisis can happen even in the absence of any reduction in arbitrage capital itself – what matters is a reduction in *active* arbitrage capital.

Our simulations illustrate that all it takes to create a full-blown liquidity crisis is merely a transient shock which causes engaged capital to get redeployed more slowly. While market liquidity recovers rather quickly after a small shock, a sizable shock (or a sequence of small shocks) can trigger a change in regime and have long-lasting impact. At the core of this argument lies the multiplicity of steady state equilibria; a sufficiently large shock can disturb the system enough to put the state variable (active capital) in another path.[10] This mechanism allows us to give a distinct prediction that equilibrium may be shifted toward low liquidity as a result of shocks. In the case of stochastic shocks, it takes a long time to have a series of enough good shocks to push the state variable in the upward trajectory. This prediction matches empirical observations of long periods of illiquidity in the market.

[9] For example, capital becomes increasingly less available through the channel of tightened collateral (e.g., Gromb and Vayanos (2002)) or margin constraints (e.g., Brunnermeier and Pedersen (2009)).

[10] Even in cases where the system has only one regime (or one steady state), a shock that initially reduces liquidity may lead to further drops in liquidity and long delays before liquidity is re-established. But, it can take arbitrarily long time to recover in case of multiple regimes.

5.3 Flight-to-Liquidity

Using our model, we show that active arbitrageurs may optimally choose to invest in the liquid market upon the arrival of liquidity shocks. There is indeed ample evidence about flight-to-liquidity in various markets: investors tend to prefer liquid assets during bad times. For example, Beber, Brandt, and Kavajecz (2007) find that capital flow in case of the Euro-area government bonds is mostly determined by liquidity rather than credit quality. Acharya, Amihud, and Bharath (2013) also document that there are two liquidity regimes for corporate bonds. In one regime, liquidity shocks have mostly insignificant effects on bond prices whereas in another regime, liquidity shocks produce significant effects. In particular, they find empirical evidence of flight-to-liquidity: prices of investment-grade bonds rise while prices of speculative-grade bonds fall. Ben-Rephael (2017) also find that mutual funds reduce their holdings of illiquid stocks during bad times.

In our model, capital tends to flow out of a market if this market's future liquidity is expected to deteriorate. Because lower future liquidity means longer maturity of new investment, mispricing wedge should become larger to compensate arbitrageurs with lower price efficiency in return for longer maturity. Furthermore, we also show the conditions under which capital flows in or out of a market hit by a liquidity shock.

One of the key observations in our model is well summarized by Eq. (8). It states that the cross-sectional difference in liquidity across markets predict future illiquidity. That is, when there is a larger cross-sectional difference in liquidity, we expect a period of low liquidity in the subsequent periods. This prediction is consistent with flight-to-liquidity episodes in which there is a divergence of liquidity across markets, and this is followed by a period of overall low liquidity. There is some empirical support on this hypothesis. Cao, Chen, Liang, and Lo (2013) find that hedge fund managers can time market liquidity based on their forecasts of future market liquidity conditions. Furthermore, Cao, Liang, Lo, and Petrasek (2017) find that hedge funds contribute to price efficiency by investing in relatively more mispriced stocks, but those stocks tend to experience large decline in price efficiency during liquidity crises.

5.4 Reaching-for-Yield

As an opposite situation to flight-to-liquidity, traders sometimes seek more risk by investing in illiquid assets. That is, traders tend to reach for yield during good times ample with liquidity (e.g., Becker and Ivashina (2015)). Our model can also contribute to the discussion by suggesting an alternative mechanism of reaching for yield. Our theory suggests that as the market starts having more capital, there would a reinforcement effect in which more locked-in capital is further released. This will raise price efficiency and shorten maturities of investment in illiquid assets. Consequently, capital starts flowing into more illiquid asset classes as active capital expands. This can reduce mispricing wedge greatly by transferring to high liquidity equilibrium. While lower mispricing wedge is good for price efficiency, it puts pressure on financial institutions to reach for higher yields. We interpret this situation as reaching-for-yield because arbitrageurs invest more in riskier assets in that situation.

6 Conclusion

We study a dynamic stationary model of informed trading with two markets. The model features endogenous liquidity regimes where temporary shocks to noise trading can trigger a shift of the regime. We show that upon the arrival of a shock arbitrage capital may actually flow out of the illiquid market and only come back later. With some arbitrage capital flowing out, the remaining capital in the market becomes trapped because it is too illiquid for arbitrageurs to want to close out their positions. This in turn deepens illiquidity in a self-reinforcing manner, thereby creating liquidity hysteresis where illiquidity persists even when the initial cause is removed.

In our model, arbitrage capital plays a dual role; the wedge of mispricing not only decide the profitability of new investment but also decides the speed at which engaged arbitrage capital is released (thus deciding the availability of arbitrage capital). The dual role of arbitrage capital implies that efficiency depends on the pool of active capital as a state variable. Furthermore, it creates a feedback channel between active capital and liquidity which leads to multiple steady state equilibria where there is a threshold of active capital that separates domains of attraction for liquidity. Therefore, a large adverse shock can trigger a vicious cycle of illiquidity with flight-to-liquidity where arbitrage capital flows to the market with short-lived assets.

In case of a stochastic model with anticipated shocks, the market can move in and out of endogenous regimes. We show that it may take quite a long time to come back to a normal liquidity regime from an illiquidity regime because it requires a sequence of good shocks strong enough to push the mass of active capital toward the path of a normal liquidity regime. Our results shed light on why capital moves slowly, how fast (or slowly) it moves, and to which directions it moves. The results further provide interesting implications on liquidity crises, flight-to-liquidity, and cross-section of liquidity.

Appendix

Proof of Lemma 1: Let $X^i_{a,t}$ be the aggregate order flow of arbitrageurs for asset i. Suppose that there are μ^i_t mass of arbitrageurs investing in asset i. Because arbitrageurs are risk-neutral and informed, their aggregate order flow is given by $X^i_{a,t} = \mu^i_t$ if $V^i = V_H$, and $X^i_{a,t} = -\mu^i_t$ otherwise. The market makers observe the aggregate order flow $X^i_t = X^i_{a,t} + z_t$. Bayes' theorem implies that the posterior belief of the market makers is given by

$$\hat{p}^i_t(X^i_t) = \frac{pf^i_X(X^i_t|V^H)}{pf^i_X(X^i_t|V^H) + (1-p)f^i_X(X^i_t|V^L)}, \tag{11}$$

where $p = \frac{1}{2}$ is the prior belief and $f^i_X(\cdot|V^H)$ and $f^i_X(\cdot|V^L)$ are the distribution of X^i_t given V^H and V^L, respectively.

Recall that z^i_t follows a uniform distribution on the interval $[-\bar{z}^i, \bar{z}^i]$ in each period t where $\bar{z}^i = \bar{z}^I$ if asset i is an illiquid asset, and $\bar{z}^i = \bar{z}^L$ if asset i is a liquid asset. Therefore, X^i_t follows a uniform distribution on the interval $[\mu^i_t - \bar{z}^i, \mu^i_t + \bar{z}^i]$ if $V^i = V_H$, and $u[-\mu^i_t - \bar{z}^i, -\mu^i_t + \bar{z}^i]$ otherwise. Therefore, Eq. (11) implies

$$\hat{p}^i_t(X^i_t) \in \begin{cases} 0 & \text{if } X^i_t < -\bar{z}^i \\ p & \text{if } -\bar{z}^i \leq X^i_t \leq \bar{z}^i \\ 1 & \text{if } X^i_t > \bar{z}^i \end{cases}$$

Therefore, the probability of revealing the true value of V^i is given by

$$\lambda^i_t = \sum_{V^i \in \{V^H, V^L\}} \frac{1}{2}[Pr(X^i_t < -\bar{z}^i|V^i) + Pr(X^i_t > -\bar{z}^i|V^i)] = \frac{\mu^i_t}{2\bar{z}^i} + \frac{\mu^i_t}{2\bar{z}^i} = \frac{\mu^i_t}{\bar{z}^i}.$$

In a symmetric equilibrium, the future lambdas are equalized across assets in each market. In that case, arbitrageurs will want to invest in an asset with the lowest λ^i_t, thus, λ^i_t should be equalized across assets in each market in equilibrium. That is, $\mu^i_t = \delta_t \xi_t$ for market I, and $\mu^i_t = (1-\delta_t)\xi_t$ for market L.

Proof of Lemma 2: First, we show that, in an interior equilibrium, an arbitrageur who holds an existing position do not have incentives to liquidate it until the capital gain realizes. We can write $J_S(\theta_t)$ as

$$J_S(\theta_t) = J_I(\theta_t) + \lambda^I_t P^I_H + (1-\lambda^I_t)P^I_0.$$

In an interior equilibrium at t, $J_I(\theta_t) = J_L(\theta_t)$ and therefore

$$J_S(\theta_t) = J_L(\theta_t) + \lambda_t^I P_H^I + (1-\lambda_t^I)P_0^I = \beta V_H^L + E_t[J_f(\theta_{t+1})].$$

Hence,

$$J_E(\theta_t) = -[\beta V_H^L - (\lambda_t^L P_H + (1-\lambda_t^L)P_0)] + J_S(\theta_t) < J_S(\theta_t).$$

Proof of Lemma 3: We start rewriting $J_I(\theta_t)$ as

$$\begin{aligned}
J_I(\theta_t) &\overset{(i)}{=} \left(P_H^I - P_0^I\right)(1-\lambda_t^I) + \beta[(1-\lambda_t^I)(1-q)\left(E_t[J_l(\theta_{t+1})] - P_H^I - E_t[J_f(\theta_{t+1})]\right)] \\
&\quad +\beta E_t[J_f(\theta_{t+1})]; \\
J_I(\theta_t) &\overset{(ii)}{=} \left(P_H^I - P_0^I\right)(1-\lambda_t^I) + \beta[(1-\lambda_t^I)(1-q)E_t[J_I(\theta_{t+1}) - \left(P_H^I - P_0^I\right)\left(1-\lambda_{t+1}^I\right) - J_f(\theta_{t+1})] \\
&\quad +\beta E_t[J_f(\theta_{t+1})]; \\
J_I(\theta_t) &\overset{(iii)}{=} \left(P_H^I - P_0^I\right)(1-\lambda_t^I)[1-\beta(1-q)\left(1-E_t\left[\lambda_{t+1}^I\right]\right)] + \beta E_t[J_f(\theta_{t+1})],
\end{aligned}$$

where (i) uses $qV_H^I = P_H^I\left(R_f - (1-q)\right)$ and (ii) uses Lemma 2 and therefore $J_l(\theta_{t+1}) = J_I(\theta_{t+1}) + \lambda_{t+1}^I P_H^I + \left(1-\lambda_{t+1}^I\right)P_0^I$ and (iii) assumes interior equilibrium at $t+1$ and therefore $J_f(\theta_{t+1}) = J_I(\theta_{t+1})$. Similarly, we can use $\beta V_H^L = P_H^L$ to write

$$J_L(\theta_t) = \left(P_H^L - P_0^L\right)(1-\lambda_t^L) + \beta E_t[J_f(\theta_{t+1})]$$

Then, if the equilibrium is interior at t,

$$\begin{aligned}
J_L(\theta_t) &= J_I(\theta_t) \\
\left(P_H^L - P_0^L\right)(1-\lambda_t^L) &= \left(P_H^I - P_0^I\right)(1-\lambda_t^I)[1-\beta(1-q)\left(1-E_t\left[\lambda_{t+1}^I\right]\right)]
\end{aligned}$$

Because $\left(P_H^L - P_0^L\right) = \left(P_H^I - P_0^I\right)$, rearranging the above equation gives Eq. (7). $\square$

Proof of Lemma 4: Write the IC curve as $F(\delta,\xi) = 0$, where

$$F(\delta,\xi) = \frac{\bar{z}^L - (1-\delta)\xi}{\bar{z}^L} - \left(\frac{\bar{z}^I - \delta\xi}{\bar{z}^I}\right)\left[1-\beta(1-q)\left(\frac{\bar{z}^I - \delta\xi}{\bar{z}^I}\right)\right].$$

We wish to show that $\frac{\partial F(\delta,\xi)}{\partial\delta} > 0$ and $\frac{\partial F(\delta,\xi)}{\partial\xi} < 0$. We have:

$$\frac{\partial F(\delta,\xi)}{\partial\xi} = \frac{(1-\delta)}{\bar{z}^L} + \frac{\delta}{\bar{z}^I} - 2\frac{\delta}{\bar{z}^I}\beta(1-q)\left(\frac{\bar{z}^I - \delta\xi}{\bar{z}^I}\right) = \frac{1}{\xi}\left(\lambda^I - \lambda^L - 2\lambda^I\beta(1-q)\left(1-\lambda^I\right)\right).$$

Because $F(\delta, \xi) = 0$ requires $\lambda^I < \lambda^L$, then $\frac{\partial F(\delta,\xi)}{\partial \xi} < 0$. Furthermore,

$$\frac{\partial F(\delta, \xi)}{\partial \delta} = \frac{\xi}{\bar{z}^L} + \frac{\xi}{\bar{z}^I} - 2\frac{\xi}{\bar{z}^I}\beta(1-q)\left(\frac{\bar{z}^I - \delta\xi}{\bar{z}^I}\right) = \frac{\xi}{\bar{z}^I}\left(\frac{\bar{z}^I}{\bar{z}^L} + 1 - 2\beta(1-q)\left(1-\lambda^I\right)\right).$$

Clearly, $\frac{\partial F(\delta,\xi)}{\partial \delta} > 0$ if $\frac{\bar{z}^I}{\bar{z}^L} + 1 - 2\beta(1-q) \geq 0$. $\square$

Proof of Proposition 1: Suppose that Eq. (IC) is not satisfied. Then, it is one of the two cases: either everyone chooses market I or everyone chooses market L. In the former case, $\delta = 1$ and therefore $\lambda^L = 0$ and $\lambda^I \in (0, 1]$. However, we can show that there is no such equilibrium that satisfies Eq. (IC) because, for all $\lambda^I \in (0, 1]$

$$1 > \left(1 - \lambda^I\right)\left(1 - \beta\left(1 - \lambda^I\right)(1 - q)\right),$$

which implies $J_L(\xi) > J_I(\xi)$. In the latter case, we have $\delta = 0$ and therefore $\xi = 1$, $\lambda^I = 0$ and $\lambda^L = \min\{1, \frac{1}{\bar{z}^L}\}$. Hence, $\delta = 0$ is an equilibrium if $J_L(1)|_{\lambda^L = \min\{1, \frac{1}{\bar{z}^L}\}} \geq J_I(1)|_{\lambda^I=0}$ which is equivalent to

$$1 - \min\{1, \frac{1}{\bar{z}^L}\} \geq 1 - \beta(1-q) \Leftrightarrow \beta(1-q)\,\bar{z}^L \geq 1. \tag{12}$$

Next, we let $\beta(1-q)\,\bar{z}^L < 1$, for which there is no corner equilibrium, and proceed to show that there exist either one or three interior equilibria. We define $\hat{\xi}_t = \delta_t \xi_t$ as the net mass of arbitrageurs who are investing in the illiquid market at time t. Likewise, we define $\hat{\delta}_t = \delta_t \xi_t + \pi_t$ as the total mass of investors who are investing in the illiquid market at time t. Instead of the original problem stated in terms of of δ_t and ξ_t, we can solve an equivalent problem in terms of $\hat{\delta}_t$ and $\hat{\xi}_t$. Using the definition of $\hat{\xi}$ and $\hat{\delta}$, we find

$$\xi = \hat{\xi} + 1 - \hat{\delta}, \quad \delta = \frac{\hat{\xi}}{\hat{\xi} + 1 - \hat{\delta}}, \quad \lambda^I = \frac{\hat{\xi}}{\bar{z}^I}, \quad \lambda^L = \frac{1 - \hat{\delta}}{\bar{z}^L}. \tag{13}$$

Using Eq. (13), the CM equation in the text can be expressed as

$$\hat{\delta} = \frac{\hat{\xi}}{q + (1-q)\frac{\hat{\xi}}{\bar{z}^I}}. \tag{14}$$

Using Eq. (13) and Eq. (14), the IC equation in Eq. (IC) can be expressed as

$$Q(\hat{\xi}) = 0,$$

where Q is a third degree polynomial:

$$Q(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3,$$

with coefficients

$$\begin{aligned}
a_0 &= q\left(\bar{z}^I\right)^3\left(1-(1-q)\,\bar{z}^L\beta\right) \\
a_1 &= -\left(\bar{z}^I\right)^2\left(\bar{z}^I + q\bar{z}^L - (1-q)\left(1+(3q-1)\bar{z}^L\beta\right)\right) \\
a_2 &= -\bar{z}^I\bar{z}^L(1-q)(1+(3q-2)\beta) \\
a_3 &= -(1-q)^2\bar{z}^L\beta
\end{aligned}$$

Notice that $\beta(1-q)\,\bar{z}^L < 1$ implies $a_0 > 0$ and therefore $Q(0) > 0$. Furthermore, it is tedious but straightforward to verify that for $\bar{z}^I + \bar{z}^L > 1$ we have $Q(1) < 0$, which implies that Q has either one or three real roots in the $(0,1)$ interval. Each of these roots is an interior steady state equilibrium in which $\delta, \xi \in (0,1)$.

Next, we look for interior equilibria when $\beta(1-q)\,\bar{z}^L \geq 1$ and therefore $Q(0) \leq 0$. It is straightforward to verify that in this case we have $Q(1) < 0$. Because $a_3 < 0$, then Q has either two real roots or none in the $(0,1)$ interval. Each of these roots corresponds to an interior steady state equilibrium. Because $\delta = 0$ is an equilibrium for $\beta(1-q)\,\bar{z}^L \geq 1$, then there are either one (corner) or three (one corner and two interior) equilibria for when $\beta(1-q)\,\bar{z}^L \geq 1$.

For $\bar{z}^I + \bar{z}^L \leq 1$, it is immediate to verify that there exists a continuum of interior fully revealing equilibria in which $\hat{\xi} = \bar{z}^I$ and $\lambda^I = \lambda^L = 1$.

Now, we give a sketch of the proof of stability.

Lemma 5 *Consider the following dynamic system of x and y:*

$$x_{t+1} = f(x_t, y_t) \tag{15}$$

$$y_{t+1} = g(x_t, y_t) \tag{16}$$

where f and g are continuous and twice-differentiable. If there are three steady-state equilibria and two extreme steady state equilibria are stable, the middle steady state equilibrium is either a source (unstable) or a sink.

Proof. The steady state solution $(\bar{x}, \bar{y})$ solves

$$x = f(x, y) \tag{17}$$
$$y = g(x, y), \tag{18}$$

and it also has to satisfy

$$dx = f_x(\bar{x}, \bar{y})dx + f_y(\bar{x}, \bar{y})dy \tag{19}$$
$$dy = g_x(\bar{x}, \bar{y})dx + g_y(\bar{x}, \bar{y})dy \tag{20}$$

This implies

$$0 = [f_x(\bar{x}, \bar{y}) - 1]dx + f_y(\bar{x}, \bar{y})dy \tag{21}$$
$$0 = g_x(\bar{x}, \bar{y})dx + [g_y(\bar{x}, \bar{y}) - 1]dy \tag{22}$$

Or equivalently,

$$\left.\frac{dy}{dx}\right|_{x=f(x,y)} = -\frac{f_x(\bar{x}, \bar{y}) - 1}{f_y(\bar{x}, \bar{y})} \tag{23}$$
$$\left.\frac{dy}{dx}\right|_{y=g(x,y)} = -\frac{g_x(\bar{x}, \bar{y})}{g_y(\bar{x}, \bar{y}) - 1} \tag{24}$$

Now, suppose that there are three steady state equilibria. Because f and g are continuous, it has to be the case that the middle equilibrium has an opposite inequality on the slopes from the extreme ones. That is, if $\left.\frac{dy}{dx}\right|_{x=f(x,y)} > \left.\frac{dy}{dx}\right|_{y=g(x,y)}$ for the extreme equilibria, the middle one should have $\left.\frac{dy}{dx}\right|_{x=f(x,y)} < \left.\frac{dy}{dx}\right|_{y=g(x,y)}$, and vice versa. Equivalently, if $|D+1| < |T|$ for the extreme ones, $|D+1| > |T|$ for the middle one, and vice versa where

$$T = f_x(\bar{x}, \bar{y}) + g_y(\bar{x}, \bar{y}) \tag{25}$$
$$D = f_x(\bar{x}, \bar{y})g_y(\bar{x}, \bar{y}) - g_x(\bar{x}, \bar{y})f_y(\bar{x}, \bar{y}). \tag{26}$$

The linearized system of the original dynamic system around $(\bar{x}, \bar{y})$ is

$$dx_{t+1} = f_x(\bar{x}, \bar{y})dx_t + f_y(\bar{x}, \bar{y})dy_t \tag{27}$$

$$dy_{t+1} = g_x(\bar{x}, \bar{y})dx_t + g_y(\bar{x}, \bar{y})dy_t \tag{28}$$

But, it happens that the stability condition for the linearized system is given by $|D+1| < |T|$. Suppose the extreme equilibria are stable (i.e., $|D+1| < |T|$). Then it has to be the case that $|D+1| > |T|$ for the middle steady state. Then, the middle steady state is either sink if $|D| < 1$, or source if $|D| > 1$.□

Notice that we can express CM and IC curves as

$$\hat{\delta} = \frac{\hat{\xi}}{q + (1-q)\frac{\hat{\xi}}{\bar{z}^I}} \tag{29}$$

$$\frac{1-\hat{\delta}}{\bar{z}^L} - \frac{\hat{\xi}}{\bar{z}^I} = \beta(1-q)\left(1 - \frac{\hat{\xi}}{\bar{z}^I}\right)^2, \tag{30}$$

or equivalently,

$$h(\hat{\xi}) = \frac{\hat{\xi}}{q + (1-q)\frac{\hat{\xi}}{\bar{z}^I}} \tag{31}$$

$$l(\hat{\xi}) = 1 - \bar{z}^L\frac{\hat{\xi}}{\bar{z}^I} - \bar{z}^L\beta(1-q)\left(1 - \frac{\hat{\xi}}{\bar{z}^I}\right)^2. \tag{32}$$

Consider the case $\frac{1}{\bar{z}^L} > \beta(1-q)$, for which equilibrium is interior. Notice that $l(0) = \frac{1}{\bar{z}^L} - \beta(1-q) > h(0) = 0$. Thus, it has to be the case that $h'(\bar{\xi}) > l'(\bar{\xi})$ at the low information steady state $\bar{\xi}$. Then, Lemma 5 implies that $\bar{\xi}$ is a saddle point (because $|D+1| < |T|$) and also that the middle equilibrium is not a saddle point but the high information equilibrium is a saddle point. □

References

Acharya, V. V., Y. Amihud, and S. T. Bharath, 2013, "Liquidity risk of corporate bond returns: conditional approach," *Journal of Financial Economics*, 110(2), 358–386.

Allen, F., and D. Gale, 1994, "Limited Market Participation and Volatility of Asset Prices," *American Economic Review*, 84(4), 933–955.

Beber, A., M. W. Brandt, and K. A. Kavajecz, 2007, "Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market," *Review of Financial Studies*, 22(3), 925–957.

Becker, B., and V. Ivashina, 2015, "Reaching for Yield in the Bond Market," *Journal of Finance*, 70(5), 1863–1902.

Ben-Rephael, A., 2017, "Flight-to-liquidity, market uncertainty, and the actions of mutual fund investors," *Journal of Financial Intermediation*, 31, 30–44.

Brunnermeier, M. K., and L. H. Pedersen, 2009, "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22(6), 2201–2199.

Cao, C., Y. Chen, B. Liang, and A. W. Lo, 2013, "Can hedge funds time market liquidity?," *Journal of Financial Economics*, 109(2), 493–516.

Cao, C., B. Liang, A. W. Lo, and L. Petrasek, 2017, "Hedge fund holdings and stock market efficiency," *Review of Asset Pricing Studies*, (October), 1–61.

Coval, J., and E. Stafford, 2007, "Asset fire sales (and purchases) in equity markets," *Journal of Financial Economics*, 86(2), 479–512.

Dow, J., and G. Gorton, 1994, "Arbitrage Chains," *Journal of Finance*, 49(3), 819–849.

Dow, J., and J. Han, Forthcoming, "The Paradox of Financial Fire Sales: the Role of Arbitrage Capital in Determining Liquidityeness," *Journal of Finance*.

Duffie, D., 2010, "Asset Price Dynamics with Slow-Moving Capital," *Journal of Finance*, 65(4), 1237–1267.

Duffie, D., N. Garleanu, and L. H. Pedersen, 2005, "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

——— , 2007, "Valuation in Over-the-Counter Markets," *Review of Financial Studies*, 20(6), 1865–1900.

Duffie, D., and B. Strulovici, 2012, "Capital Mobility and Asset Pricing," *Econometrica*, 80(6), 2469–2509.

Gromb, D., and D. Vayanos, 2002, "Equilibrium and welfare in markets with financially constrained arbitrageurs," *Journal of Financial Economics*, 66(2-3), 361–407.

———, 2017, "Financially Constrained Arbitrage and Cross-Market Contagion," Working paper.

———, Forthcoming, "The Dynamics of Financially Constrained Arbitrage," *Journal of Finance*.

Kyle, A. S., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1335.

Kyle, A. S., and W. Xiong, 2001, "Contagion as a wealth effect," *The Journal of Finance*, 56(4).

Mitchell, M., L. H. Pedersen, and T. Pulvino, 2007, "Slow Moving Capital," *American Economic Review*, 97(2), 215–220.

Shleifer, A., and R. W. Vishny, 1990, "Equilibrium Short Horizons of Investors and Firms," *American Economic Review*, 80(2), 148–153.

Shleifer, A., and R. W. Vishny, 1997, "The Limits of Arbitrage," *Journal of Finance*, 52(1), 35–55.