Liquidity and Market Crashes

Jennifer Huang
McCombs School of Business, University of Texas

Jiang Wang
MIT Sloan School of Management, CCFR, and NBER

In this paper, we develop an equilibrium model for stock market liquidity and its impact on asset prices when constant market presence is costly. We show that even when agents' trading needs are perfectly matched, costly market presence prevents them from synchronizing their trades and hence gives rise to endogenous order imbalances and the need for liquidity. Moreover, the endogenous liquidity need, when it occurs, is characterized by excessive selling of significant magnitudes. Such liquidity-driven selling leads to market crashes in the absence of any aggregate shocks. Finally, we show that illiquidity in the market leads to high expected returns, negative and asymmetric return serial correlation, and a positive relation between trading volume and future returns. We also propose new measures of liquidity based on its asymmetric impact on prices and demonstrate a negative relation between these measures and expected stock returns. (*JEL* D53, G12)

1. Introduction

Market crashes refer to large, sudden drops in asset prices in the absence of big news on the fundamentals, such as future payoffs. Crashes exhibit several distinct features: They are one-sided—market surges are less likely; they are typically accompanied by large selling pressures in the market; and while the drop in prices occurs quickly, the recovery is slow. The extant literature provides no clear consensus on what causes a crash. The lack of liquidity,

Part of this work was completed during Jiang Wang's visit at the Federal Reserve Bank of New York. The authors thank Tobias Adrian, Franklin Allen, Joel Hasbrouck, Nobu Kiyotaki, Pete Kyle, Arzu Ozoguz, Lubos Pastor, Lasse Pedersen, Jacob Sagi, Matthew Spiegel (the editor), Suresh Sundaresan, Sheridan Titman, Andrey Ukhov, Dimitri Vayanos, S. "Vish" Viswanathan, and participants at 2005 Adam Smith Asset Pricing Conference, 2005 China International Conference in Finance, 2005 Duke–UNC Asset Pricing Conference, 2005 Financial Management Association annual meeting, 2005 Utah Winter Finance Conference, 16th Financial Economics and Accounting Conference at UNC, 2006 Far Eastern Meeting of the Econometric Society, 2006 Bank of Canada and Norges Bank Workshop on the Microstructure of Foreign Exchange and Equity Markets, 2007 WFA annual meeting, seminars at Baruch College, Columbia, the Federal Reserve Bank of New York, HEC, Hong Kong University of Science and Technology, INSEAD, Stanford University, University of Michigan, University of Texas at Austin, University of Texas at Dallas, University of Washington at Seattle, and University of Wisconsin at Milwaukee for comments and suggestions. Support from Morgan Stanley (Equity Market Microstructure Grant, 2006) (Huang and Wang) and NSFC (Grant Number 70440420490) (Wang) is gratefully acknowledged.
Address correspondence to Jennifer Huang, McCombs School of Business, the University of Texas at Austin; telephone: (512) 232-9375; fax: (512) 471-5073. E-mail: jennifer.huang@mccombs.utexas.edu.

doi:10.1093/rfs/hhn086 Advance Access publication October 4, 2008

however, is always identified as its symptom and is blamed for exacerbating its consequences.[1]

This view is supported by increasing evidence that despite the profitable buying opportunities after a crash—at least as perceived by some observers—new capital flows in only after long lags. For example, following the 1987 stock market crash, a large number of companies announced repurchases of their own shares, reflecting the belief that their stocks were undervalued; however, these announcements were spread over many months and took even longer to be implemented. Similarly, following the Long-Term Capital Management (LTCM) episode in 1998, the substantial capital outflows from hedge funds operating in the same markets as LTCM (e.g., fixed income arbitrage and global macro strategies) only started to reverse several quarters later, despite the opportunities in these markets.[2] This evidence suggests that capital movements are costly. The costs range from informational costs to institutional rigidities (see Merton 1987, among others). When abnormal trading pressure hits, only a limited supply of liquidity is available to accommodate the trades, and hence prices have to shift significantly (see, for example, Shleifer 1986; and Grossman and Miller 1988).

This perspective focuses on the lack of liquidity supply, especially during market crises. But it does not explain what gives rise to the initial need for liquidity, why it is usually in the form of excessive selling, and why it is of large magnitudes. In this paper, we show that the same cost that hinders the *ex post* supply of liquidity also generates the need for liquidity in the first place. Despite the symmetric nature intrinsic to market participants' idiosyncratic trading needs, the aggregate need for liquidity, when it arises, is asymmetric (usually on the selling side) and of large size. With limited supply of liquidity in the market, these sudden surges of endogenous liquidity needs lead to large price drops, as in market crashes.

We start with a model that captures two important aspects of liquidity, the need to trade and the cost of trading. Trading needs arise from idiosyncratic shocks to agents' wealth, which the agents want to unload in the market by adjusting their asset holdings. By definition, idiosyncratic shocks sum to zero at the aggregate level. As a result, agents' trading needs are always symmetric and perfectly matched—that is, for each potential seller there is a potential buyer with offsetting trading needs. If market presence is costless, all potential buyers and sellers will be in the market at all times. Their trades will be perfectly synchronized and matched, and there will be no need for liquidity.

[1] For example, the report by the Committee on the Global Financial System (CGFS 1999) provides an overview of the "deterioration in liquidity and elevation of risk spreads" in many international financial markets in autumn 1998.

[2] Gammill and Marsh (1988) and Netter and Mitchell (1989) analyze the share repurchases after the 1987 crash. Mitchell, Pedersen, and Pulvino (2007) document the inefficiency in the convertible bond market, and Tremont (2006) reports the capital flows into and out of different hedge fund strategies around the collapse of LTCM in 1998.

In this case, the market-clearing price always reflects the fundamental value of the asset, such as asset payoffs and investor preferences, and idiosyncratic shocks generate trading but have no impact on prices.

In contrast, when market presence is costly, the need for liquidity arises endogenously and idiosyncratic shocks can affect prices. Costly market presence has two important effects. First, it prevents potential traders from being in the market constantly. They will enter the market only when they are far away from their desired positions and the expected gains from trading outweigh the cost. Infrequent trading implies that traders who are hit by idiosyncratic risks will not always be able to unload them in the market, which makes them more risk averse. Second, potential traders with offsetting trading needs perceive different gains from trading. In particular, the gains from trading for potential sellers are always larger than the gains from trading for potential buyers. The reason is that, as idiosyncratic shocks push them away from their optimal positions, traders become more risk averse and less willing to hold the asset. This increased risk aversion reduces their preferred asset holding, exacerbates the selling need for potential sellers, and dampens the buying demand for potential buyers. The asymmetry in their desire to trade leads to order imbalances in the form of excess supply, and the price has to decrease in response.

Moreover, the endogenous liquidity need is highly nonlinear in the idiosyncratic shocks that drive agents' trading needs. When the magnitude of idiosyncratic shocks is moderate, gains from trading are relatively small. As a result, all traders will stay out of the market and there is no need for liquidity. Only when the idiosyncratic shocks are sufficiently large do gains from trading exceed the participation cost for some potential traders. They enter the market with large trading needs and more on the selling side. Thus, when the order imbalance and the need for liquidity occur, they are large in magnitude, causing the price to drop discretely in the absence of any aggregate shocks. Such market behavior—namely, infrequent but large price drops accompanied by large selling pressure absent big news on the fundamentals—clearly resembles the features of market crashes.

This mechanism for crashes, driven solely by liquidity, differs from those proposed in the literature that rely on the presence of information asymmetry among investors about the fundamentals.[3] Our analysis shows that purely idiosyncratic and non-fundamental shocks can cause market crashes if capital flow is costly. Moreover, information-based models for crashes have two undesirable features from an empirical perspective: Both crashes and surges are possible and a crash reflects a permanent shift in the price instead of a transitory price change.[4] In contrast, our liquidity-based explanation for crashes predicts one-sided and transitory price movements—that is, it is less likely to

---

[3] For example, Grossman (1988), Gennotte and Leland (1990), and Romer (1993) consider models with information asymmetry in incomplete markets.

[4] The symmetry simply comes from the fact that when information moves prices, it can be either positive or negative. The permanent nature of the price change follows from the fact that the change reflects additional

see surges, and the crash represents a deviation from fundamentals that will eventually recover.

The impact of liquidity also leads to testable implications on the behavior of prices, returns, and trading volume. First, crashes caused by endogenous liquidity needs lead to extra volatility unrelated to changes in fundamentals. They also give rise to negative skewness and fat tails in the return distribution. Second, since the price impact of liquidity is transitory, it leads to return reversals (i.e., negative serial correlation in returns). More importantly, the negative and discrete nature of endogenous liquidity needs implies that return reversals are more prominent for negative returns than for positive returns. Third, in our model trading volume is positively related to liquidity needs, and thus it is negatively correlated with the contemporaneous return but positively correlated with the future return. Consequently, higher volume predicts higher future returns. Fourth, the asymmetric nature of the liquidity impact further implies that low returns accompanied by high volume exhibit stronger reversals than high returns. Fifth, since lower returns and higher volume are indicative of aggregate liquidity demand, they are also accompanied by higher asset volatility. In addition, given that the level of liquidity varies across markets, our analysis also implies that the liquidity effects on return and volume described above are stronger in less-liquid markets.

Furthermore, we show that an asset with lower liquidity has a lower price and a higher average return. In our model, the level of liquidity is negatively related to several observable variables such as the average volume and the price impact measures of Campbell, Grossman, and Wang (1993). Thus, our model provides an explanation for the positive relation between the average stock return and these variables, which have been documented in several empirical studies (see, for example, Brennan, Chordia, and Subrahmanyam 1998; and Pastor and Stambaugh 2003). In addition, several studies find that various trading cost measures are at best noisy proxies of liquidity in explaining returns (see, for example, Hasbrouck 2006; and Spiegel and Wang 2007). Based on the asymmetric nature of liquidity's price impact, we propose more direct measures of liquidity, such as the asymmetry in the return serial correlation between high and low returns or between returns accompanied by high and low volume. The model predicts a positive link between expected returns and these liquidity measures.

In studying the impact of liquidity, much of the attention is focused on the supply of liquidity, taking the liquidity demand as given.[5] For example,

information about the fundamentals. Models with short sale or borrowing constraints, such as Hong and Stein (2003), Yuan (2005), and Bai, Chang, and Wang (2006), can generate negative skewness in returns. But the skewness arises from the asymmetric distribution of small price changes, not discrete price drops.

[5] For example, in the market microstructure literature, which has liquidity as a central focus, the need for liquidity, as described by the order-flow process, is often taken as given. Amihud and Mendelson (1980) and Ho and Stoll (1981) examine how dealers' inventory costs affect the supply of liquidity, and Glosten and Milgrom (1985) and Kyle (1985) consider the additional effect of information asymmetry. Admati and Pfleiderer (1988) and Spiegel and Subrahmanyam (1995), however, do allow the order-flow process to be influenced by equilibrium.

Grossman and Miller (1988) consider how participation costs limit market makers' supply of liquidity and reduce price volatility, taking as given the nonsynchronization in trades. Pagano (1989) and Allen and Gale (1994) consider the *ex ante* participation decisions of agents with different future liquidity needs. They show that the *ex ante* optimal level of participation can be inadequate *ex post* when the realized liquidity need is very large, causing additional volatility in prices.[6]

We extend the existing literature on liquidity by modeling how the need for liquidity arises endogenously and how it behaves. Our analysis shows that it is the participation costs that generate the nonsynchronization in trades and hence the need for liquidity in the first place. We capture the dynamic aspect of liquidity by allowing traders to make their participation decisions after observing their trading needs. The endogenously derived liquidity needs exhibit distinctive properties—in particular, one-sided and fat-tailed—which allow us to show that liquidity needs can lead to market crashes in the absence of fundamental news.

Furthermore, in our model, liquidity needs arise purely from idiosyncratic shocks, which would have no pricing implication in the absence of the liquidity effect. Most of the existing models rely on aggregate shifts in demand.[7] The presence of aggregate shocks makes market crashes and surges equally likely, as the shocks can be either positive or negative. Moreover, it blurs the distinction between the effects of liquidity and risk (and/or preferences). In these models, liquidity merely plays the role of exacerbating the impact of exogenous aggregate shocks. In our model, it is the idiosyncratic shock that generates endogenous selling demand at the aggregate level.

Our model is closely related to the model of Lo, Mamaysky, and Wang (2004), which is in a continuous-time stationary setting. They show that gains from trading are in general asymmetric between traders with offsetting shocks when trading is costly. In order to focus on the impact of trading cost on price levels, they avoid potential order imbalances by allocating the cost endogenously among buyers and sellers so that their orders are always synchronized. As we show in this paper, it is the order imbalances that lead to liquidity needs and the instability in asset prices.

This paper proceeds as follows. Section 2 describes the basic model. Section 3 solves for the intertemporal equilibrium of the economy. In Section 4, we examine how the endogenous need for liquidity affects asset prices, and in particular causes market crashes. In Section 5, we explore in more detail the testable implications of our model on the impact of liquidity on the behavior of returns and volume. Section 6 concludes. The Appendix contains the proofs.

[6] Participation costs can also take the form of capital or position constraints. For example, Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2008) consider how binding margins and collateral restrictions in down markets limit the supply of liquidity. Acharya and Viswanathan (2008) and He and Krishnamurthy (2008) further consider how capital constraints of market participants arise endogenously as a response to agency costs.

[7] See Campbell, Grossman, and Wang (1993), Campbell and Kyle (1993), and Allen and Gale (1994).

2. The Model

We construct a parsimonious model that captures two important factors in analyzing liquidity, the need to trade and the cost of participating in the market. We use a discrete-time, infinite-horizon setting.

2.1 Economy

2.1.1 Asset market. A stock is traded in a competitive asset market. It yields a risky dividend D_t at time t, where $t = 0, 1, 2, \ldots$ Dividends are i.i.d. normally distributed with a mean of $\bar{D}$ and volatility of σ_D. Let P_t denote the ex-dividend stock price at time t. In addition, there is a short-term riskless bond, which yields a constant interest rate of $r > 0$ per period.

2.1.2 Agents. At $t = 0, 1, 2, \ldots$, a set of agents are born who live for one period. Agents born at t are referred to as generation t. They are born with initial wealth W_t, which they invest in the stock and the bond. They sell all their assets for consumption at time $t + 1$.

Each generation consists of two types of agents who face different endowments and trading costs. As described below, agents' heterogeneity in endowments gives rise to their trading needs in our model. The first type of agents, denoted by m, are "market makers." They have no inherent trading needs, but are present in the market at all times, ready to trade with others. The second type of agents are "traders," who have trading needs. Traders are split between two equal subgroups with different trading needs, denoted by a and b, respectively. The population weights of the market makers and the traders are μ and 2ν, respectively.

The per capita supply of the stock is $\bar{\theta}$, which is positive (i.e., $\bar{\theta} > 0$). In addition, each agent i of generation t receives a nontraded payoff N^i_{t+1} at the end of his lifespan, given by

$$N^i_{t+1} = \lambda^i Z\, n_{t+1}, \quad i = m, a, b, \tag{1}$$

where Z and n_{t+1} are mutually independent, normal random variables with a mean of zero and a volatility of σ_Z and σ_n, respectively, and λ^i is a binomial random variable drawn independently for each agent within his group, where[8]

$$\lambda^m = 0, \quad \lambda^a = -\lambda^b = \begin{cases} 1, & \text{with probability } \lambda \\ 0, & \text{with probability } 1 - \lambda. \end{cases} \tag{2}$$

Thus, market makers receive no nontraded payoff, while a fraction λ of traders within each trader group receives nontraded payoffs. Since $\lambda^a = -\lambda^b$, the two

[8] Since our analysis focuses only on generation t, we omit the time subscript for brevity whenever there is little room for confusion. For example, λ^i and Z have no time subscript.

groups of traders receive perfectly offsetting nontraded payoffs. By construction, we have

$$\sum_{i=a,b,m} N^i_{t+1} = 0. \tag{3}$$

The nontraded payoff is assumed to be correlated with the stock dividend D_{t+1}. In particular, we let $n_{t+1} = D_{t+1} - \bar{D}$.[9]

In the absence of risks from nontraded payoffs, all agents are identical and there is no need to trade among them. However, in the presence of nontraded risks, traders who receive them want to trade in order to share these risks. In particular, given the correlation between the nontraded payoff and the stock payoff, they want to adjust their stock positions in order to hedge their nontraded risks. Thus, traders' idiosyncratic risk exposures give rise to their inherent trading needs.

Since the nontraded risks sum to zero as in (3), the traders' underlying trading needs are perfectly matched. If all traders are present in the market at all times, a seller is always matched with a buyer and there is perfect synchronization in their trades. If, however, only some traders are present at a given time, trades may not be always synchronized and the need for liquidity may arise.

For tractability, we assume that all agents have a utility function of constant absolute-risk aversion over their terminal wealth. The utility function for generation-t agents is

$$\mathrm{E}\left[-e^{-\alpha W^i_{t+1}}\right], \qquad i = a, b, m, \tag{4}$$

where W^i_{t+1} denotes agent i's terminal wealth.

Given agents' nontraded payoff and utility function, we need the following condition to guarantee that their expected utility is always well defined (i.e., finite):

$$\frac{1}{2}\,\alpha^2\sigma_D^2\sigma_Z^2 < 1. \tag{5}$$

2.1.3 Participation costs. All agents can trade in the market at no cost at the beginning and the end of their lifespan. That is, agents of generation t can trade in the market at t and $t+1$ without cost. In addition, market makers can also trade at no cost at any time between t and $t+1$. The traders, however, face a fixed cost $c \geq 0$ if they want to trade between t and $t+1$.

2.1.4 Time line. We now describe in detail the timing of events and actions. At t, agents of generation t are born. They purchase shares of the stock from

[9] Our modeling of the heterogeneity in endowments is similar to that of Glosten (1989) and Wang (1994) and is meant to capture the need to trade for risk sharing. We only need the correlation between n_{t+1} and D_{t+1} to be nonzero. The qualitative nature of our results is independent of the sign and the magnitude of the correlation. To fix ideas, we set it to one.

the old generation and construct their optimal portfolio θ_t^i, $i = a, b, m$. Market equilibrium at t determines P_t.

After t, traders learn if they will be exposed to any idiosyncratic risks (i.e., their draws of λ^i). Those subject to such risks ($\lambda^i \neq 0$) also observe a signal S about the potential magnitude of the risk, Z, that is

$$S = Z + u, \tag{6}$$

where u is the noise in the signal, normally distributed with a mean of zero and a variance of $\sigma_u^2 > 0$. For future convenience, we denote by X the expectation of Z conditional on signal S and σ_z^2 the conditional variance. We then have

$$X \equiv \mathrm{E}[Z|S] = \frac{\sigma_Z^2}{\sigma_Z^2 + \sigma_u^2}\, S, \quad \sigma_z^2 \equiv \mathrm{Var}[Z|S] = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_Z^2}\, \sigma_Z^2. \tag{7}$$

Under normality, X is a sufficient statistic for signal S. Thus, we will use X to denote these traders' information about the magnitude of their idiosyncratic risks.

In addition, agents also receive a signal S_D about the next-period dividend payment

$$S_D = D_{t+1} + e, \tag{8}$$

where e is the signal noise with a mean of zero and a variance of σ_e^2. For convenience, we set $\sigma_e = \sigma_D$ so that half of the uncertainty about D_{t+1} is resolved at $t + 1/2$.[10]

After learning about their idiosyncratic risks, traders face the choice of staying out of the market (until their terminal date) or paying a cost c to enter the market. Those who choose to enter will then trade among themselves as well as with market makers. To fix ideas, we assume that signal X and entry decisions occur at $t + 1/2$, and that trading occurs right after.

A trader's choice to enter the market depends on his draw of λ^i and the signal X on the magnitude of the idiosyncratic risk if $\lambda^i \neq 0$. Let η^i be the discrete choice variable of trader i ($i = a, b$) for whether to enter the market, where $\eta^i = 1$ denotes entry and $\eta^i = 0$ denotes no entry. Among group i traders ($i = a, b$) who receive idiosyncratic shocks (i.e., $\lambda^i \neq 0$), we use $\omega^{i,L}$ to denote the fraction of traders who choose to enter the market. Similarly, $\omega^{i,NL}$ denotes the fraction of traders without idiosyncratic shocks who choose to enter. We also use $\theta_{t+1/2}^i(\eta^i)$ to denote the number of stock shares agent i ($i = m, a, b$) holds after trading at date $t + 1/2$. Of course, $\theta_{t+1/2}^i(\eta^i = 0) = \theta_t^i$. Summarizing the description above, Figure 1 illustrates the time line of the economy.

[10] The signal S_D is not essential for the model. It is introduced so that the fundamental risk of the stock is i.i.d. in the two subperiods, making it easier to draw empirical predictions.

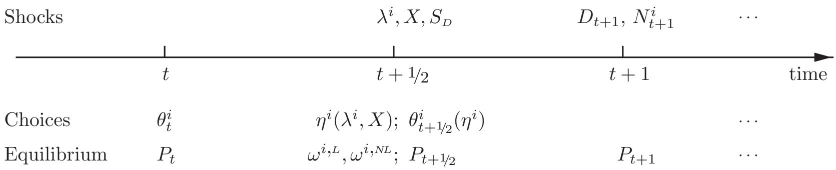

Figure 1
The time line of the economy

For agent i, his terminal financial wealth, denoted by V^i_{t+1}, is

$$
\begin{aligned}
V^i_{t+1} = R_F^2 W_t - R_F \eta^i c^i + \theta^i_t R_F \big(P_{t+1/2} - R_F P_t\big) \\
+ \theta^i_{t+1/2}(\eta^i)\big(D_{t+1} + P_{t+1} - R_F P_{t+1/2}\big),
\end{aligned} \tag{9}
$$

where $R_F = (1+r)^{1/2}$ is the gross interest rate for each half-period, $c^i = c$ for $i = a, b$, and $c^i = 0$ for $i = m$. His total wealth at date $t+1$ is then given by

$$
W^i_{t+1} = V^i_{t+1} + N^i_{t+1}, \tag{10}
$$

where N^i_{t+1} is the income from the nontraded asset in (1).

2.2 Discussions and simplifications

In this subsection, we provide additional discussions on several aspects of the model. A key ingredient of our model is the cost to participate in the market. The cost is intended to capture frictions that prevent either the full participation of all potential players in a market or the instant capital flow to a market. Information costs and institutional rigidities are abundant. Gathering and processing information, devising trading strategies and their support systems in response to new information, raising capital, and making changes in business practice to implement these strategies all involve costs and time. After an extensive discussion on the importance of these costs, Merton (1987) observes, “On the time scale of trading opportunities, the capital stock of dealers, market makers and traders is essentially fixed. Entry into the dealer business is neither costless nor instantaneous.”[11] While direct measurements of participation costs are difficult, there is increasing evidence demonstrating their significance (see, for example, Coval and Stafford 2007; Gabaix, Krishnamurthy, and Vigneron 2007; and Mitchell, Pedersen, and Pulvino 2007).

Our model also makes an important technical assumption that, at the time of participation decisions, traders only partially learn about their future idiosyncratic risks, i.e., they receive a noisy signal S about Z. If Z is fully known at the time of the participation decision, a single trader can remove all future nontraded risks, and the model becomes essentially static. By assuming a partial observation of Z, we capture the intertemporal effect that a trader, even when

[11] See also Brennan (1975), Hirshleifer (1988), Leland and Rubinstein (1988), Chatterjee and Corbae (1992), and Vissing-Jorgensen (2002), among others, for more discussion of participation costs in financial markets.

he chooses to enter the market now, still expects to bear some idiosyncratic risk since he may not be in the market in the future. Such an expectation influences his current participation decision. As we will see in the next section, this remaining uncertainty leads to asymmetric participation decisions for traders with matching trading needs. Thus, this result arises from the intertemporal nature of the model. In a fully intertemporal setting, Lo, Mamaysky, and Wang (2004) show that when participation costs force traders to trade infrequently, they always expect to bear some idiosyncratic risks and the asymmetry in their trading is a general outcome. Our setup provides a simple way to capture the same effect.

As long as it occurs after the participation decision, the exact timing of the full revelation of Z is not critical. For simplicity, we assume that by the time of trading (right after $t + 1/2$) all traders who receive idiosyncratic risks also observe the realization of Z.[12]

2.3 Equilibrium with costless participation

Before solving for the equilibrium, we describe the special case of participation costs being zero for all agents. This case serves as a benchmark when we examine the impact of participation costs on liquidity and stock prices. At zero cost, all traders and market makers will be in the market at all times.

At any time t, we define the conditional mean and variance of the stock's future payoff, discounted at the risk-free rate r, as

$$F_t \equiv \mathrm{E}_t\left[\sum_{s>t} \frac{1}{(1+r)^{s-t}} D_s\right], \quad \sigma_t^2 \equiv \mathrm{Var}_t\left[\sum_{s>t} \frac{1}{(1+r)^{s-t}} D_s\right], \tag{11}$$

where $\mathrm{E}_t[\cdot]$ and $\mathrm{Var}_t[\cdot]$ denote the expectation and variance conditional on the information at time t. The equilibrium price and agents' equilibrium stock holdings are given by

$$\begin{aligned} P_t &= F_t - \alpha\sigma_t^2\bar{\theta}, & \theta_t^i &= \bar{\theta} \\ P_{t+1/2} &= F_{t+1/2} - \alpha\sigma_{t+1/2}^2\bar{\theta}, & \theta_{t+1/2}^i &= \bar{\theta} - \lambda^i Z, \end{aligned} \tag{12}$$

where $t = 0, 1, 2, \ldots,$ and $i = a, b, m$.

In this case, the stock price P_t is determined by the stock's expected future dividends F_t, the dividend risk σ_t^2, and the aggregate (per capita) risk exposure $\bar{\theta}$. We call these the "fundamentals." Prices do not depend on the idiosyncratic risk Z. For traders exposed to nontraded risks, their stock holdings equal the per capita endowment $\bar{\theta}$ plus an additional component $\lambda^i Z$, which reflects the traders' hedging demand to offset the exposure to the nontraded risk. It

[12] We also solve the model under the assumption that Z is revealed at $t + 1$. The equilibrium price and participation decisions are qualitatively the same, except that there is an extra risk premium for the unhedged risk even under full participation. Since our focus is on the price difference between the full and the partial participation equilibrium, we choose the current setup to have a simpler full participation benchmark.

is important to note that because these traders' underlying trading needs are perfectly matched ($\lambda^a = -\lambda^b$), so are their trades when they are all in the market. In this case, the market is perfectly liquid in the sense that order flows have no price impact. There is no need for liquidity and market makers perform no role (their holdings stay at $\theta^m = \bar{\theta}$).

3. Equilibrium

We now solve for the equilibrium with costly participation as follows. First, taking the stock price at $t+1$, agents' initial stock holdings, and participation decisions as given, we solve for the stock market equilibrium at $t+1/2$. Second, we solve for individual agents' participation decisions and the participation equilibrium, given the market equilibrium at $t+1/2$ and the agents' initial stock holdings at t. Finally, we solve for the market equilibrium at time t, and use the condition $P_{t+1} = P_t$ to obtain the full stationary equilibrium of the economy.

In the first two steps (Sections 3.1 to 3.3), we assume that traders who receive no idiosyncratic shocks ($\lambda^i = 0$) stay out of the market until the end of their horizon—that is, $\omega^{i,NL} = 0, i = a, b$. Thus, we consider only those traders who receive shocks and solve for their participation decisions, the participation equilibrium, and the market equilibrium at $t+1/2$. In these subsections, unless stated otherwise, traders refer only to those with $\lambda^i \neq 0$, and $\omega^a \equiv \omega^{a,L}$ and $\omega^b \equiv \omega^{b,L}$ refer to fractions of the traders that choose to participate. In the last step (Section 3.4), we include all traders and confirm that, indeed, in equilibrium those who receive no idiosyncratic shocks choose not to participate in the market.

3.1 Market equilibrium at $t+1/2$

At $t+1/2$, we take agents' initial stock holdings and their participation decisions as given and solve for the market equilibrium. Let $\theta \equiv (\theta_t^a, \theta_t^b, \theta_t^m)$ denote agents' stock holdings at t and $\omega \equiv (\omega^a, \omega^b)$ denote the participation decision. Together with the idiosyncratic shock Z, $\{\theta, \omega\}$ defines the state of the economy at $t+1/2$. Two variables are of particular importance in describing the market condition

$$\hat{\theta} \equiv \frac{\mu\theta^m + \lambda\,\nu(\omega^a\theta^a + \omega^b\theta^b)}{\mu + \lambda\,\nu(\omega^a + \omega^b)}, \qquad \delta \equiv \frac{\lambda\,\nu}{\mu + \lambda\,\nu(\omega^a + \omega^b)}(\omega^a - \omega^b), \quad (13)$$

where $\hat{\theta}$ gives the per capita stock supply in the market (brought in by participating agents) and δ measures the difference in participation between the two trader groups. Since the participation equilibrium at t depends on the information X about the nontraded risk, ω^a and ω^b, and thus $\hat{\theta}$ and δ are all functions of X.

The following proposition solves the market equilibrium at $t+1/2$.

Proposition 1. *Let P_{t+1} be the equilibrium price at time $t+1$. Given the market condition $\hat{\theta}$ and δ, the equilibrium stock price at $t+1/2$ is*

$$P_{t+1/2} = R_F^{-1}\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - \tfrac{1}{2}\alpha\sigma_D^2\hat{\theta} - \tfrac{1}{2}\alpha\sigma_D^2\delta Z\right) \tag{14}$$

and the equilibrium stock holding of participating agent i is

$$\theta^i_{t+1/2} = \hat{\theta} + \delta Z - \lambda^i Z, \qquad i = a, b, m. \tag{15}$$

When $\delta = 0$, the participation of the two groups of traders is symmetric. The participating agents' holdings are equal to the per capita holding $\hat{\theta}$ minus the hedging demand $\lambda^i Z$. Since $\lambda^a = -\lambda^b$, there is a perfect match between the buy and sell orders among traders, and the equilibrium price is not affected by the idiosyncratic shock Z. This situation is reminiscent of the benchmark case when participation is costless.

When $\delta \neq 0$, the participation of the two groups of traders is asymmetric. The quantity δZ measures the excess exposure (per capita) to the nontraded risk due to the asymmetric participation of traders. In this case, the optimal holding in (15) has an extra term δZ for all participating agents since they equally share this additional source of risk. The idiosyncratic shock Z now affects the equilibrium price. Thus, in our model, even though traders face offsetting shocks, asymmetry in their participation can give rise to a mismatch in their trades and cause the price to change in response to these shocks.

Here, we have taken traders' participation and the resulting δ and $\hat{\theta}$ as given. In the following subsections, we show that when individual participation decisions are made endogenously, asymmetric participation occurs as an equilibrium outcome.

3.2 Optimal participation decision

Given the market equilibrium at $t+1/2$ and the signal X for future idiosyncratic shocks, we now solve the optimal participation policy of an individual trader, taking as given the participation decision of others. In the next subsection, we find the competitive equilibrium for traders' participation decisions.

For trader i, let J_P and J_{NP} denote his utility from participation and no participation, respectively. In general, trader i's utility depends on his initial stock holding θ^i, his exposure to the nontraded risk given by λ^i and X, and the market condition given by $\hat{\theta}$ and δ. His net gain from participation can be defined as the certainty equivalence gain in wealth:

$$g(\theta^i; \lambda^i, X; \hat{\theta}, \delta) = -\frac{1}{\alpha} \ln \frac{J_P(\theta^i; \lambda^i, X; \hat{\theta}, \delta)}{J_{NP}(\theta^i; \lambda^i, X; \hat{\theta}, \delta)}. \tag{16}$$

The minus sign on the right-hand side adjusts for the fact that $J_P(\cdot)$ and $J_{NP}(\cdot)$ are negative. The following proposition describes the optimal participation policy for an individual trader.

Proposition 2. *For trader i with initial stock holding θ^i, idiosyncratic shock $\lambda^i \neq 0$ and X, and market condition $\hat{\theta}$ and δ, his net gain from participation is*[13]

$$g(\theta^i; \lambda^i, X; \hat{\theta}, \delta) = g_1(\theta^i; \lambda^i, X; \hat{\theta}, \delta) + g_2(\lambda^i; \delta) - R_F c^i, \tag{17}$$

where

$$\begin{aligned} g_1(\cdot) &= \frac{\alpha\sigma_D^2(1-k\,\lambda^i\delta)^2}{4(1-k)[1-k+k\,(1-\lambda^i\delta)^2]}\,(\theta^i - \hat{\theta}^i)^2, \\ g_2(\cdot) &= \frac{1}{2\alpha}\ln[1+(1-\lambda^i\delta)^2 k/(1-k)], \end{aligned} \tag{18}$$

and

$$\hat{\theta}^i \equiv \frac{1-k}{1-k\lambda^i\delta}\,\hat{\theta} - \frac{1-\lambda^i\delta}{1-k\lambda^i\delta}\,\lambda^i X, \quad k \equiv \tfrac{1}{2}\,\alpha^2\sigma_D^2\sigma_z^2. \tag{19}$$

The trader chooses to participate if and only if $g(\cdot) > 0$.

The first term of the gain, $g_1(\cdot)$, represents the expected gain from trading given the current signal X on nontraded risks. This term depends on trader i's initial holding θ^i, the per capita stock supply of all participating agents $\hat{\theta}$, and the expected idiosyncratic risk, $\lambda^i X$. The second term, $g_2(\cdot)$, captures the expected gain from trading to offset future shocks to nontraded risks. This term depends on the market condition δ and the quantity k, which depends on σ_z in (7) and captures the variation in future trading needs. The last term, $-R_F c^i$, simply reflects the cost of participation.

The gain is always positive when the participation cost is small, i.e., when $c \leq R_F^{-1} g_2(\cdot)$. Trader i always participates in this case, independent of X. The more interesting case is when $c > R_F^{-1} g_2(\cdot)$ and trader i chooses to participate only if the expected gain $g_1(\cdot)$ from trading against his current expected exposure is sufficiently large. Note that $g_1(\cdot)$ is zero when his current holding θ^i is equal to $\hat{\theta}^i$. Thus, we can interpret $\hat{\theta}^i$ as trader i's desired stock holding after observing his idiosyncratic risk. In this case, a trader chooses to participate when his holding θ^i is sufficiently far away from the desired position $\hat{\theta}^i$.

Gains from participation depend on traders' initial stock holding θ^i. When λ is small, we expect that in equilibrium θ^i ($i = a, b$) and θ^m are not too far apart and both are close to $\bar{\theta}$, the per capita supply of stock. For the discussion to follow we assume that this is the case, in particular, where agents' initial holdings satisfy the following condition:

$$|\theta^i - \theta^m| \leq \min\left\{\frac{\mu\,\sigma_z}{\mu + \lambda\nu},\ k\,\theta^m\right\}, \quad i = a, b. \tag{20}$$

[13] The gain from participation for those with $\lambda^i = 0$ is different and is given in the Appendix.

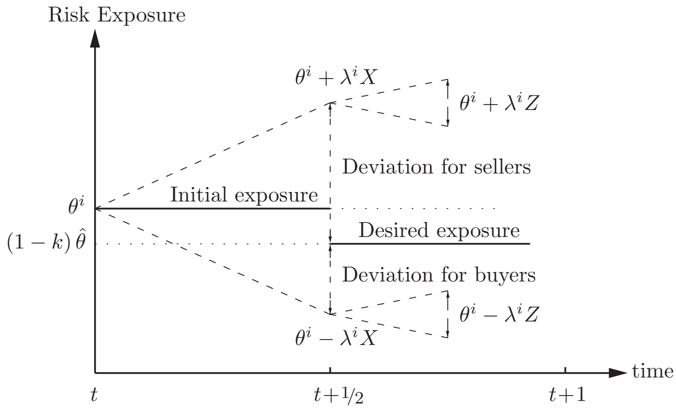

Figure 2
Traders' desired risk exposure before and after idiosyncratic shocks

We verify later that this condition is indeed satisfied in equilibrium (see Theorem 1).

From the expressions in Proposition 2, it is obvious that the gains from trading are not symmetric between the two trader groups (with $\lambda^a \neq \lambda^b$). To understand the intuition, we consider the simple situation in which the market participation rate is symmetric ($\delta = 0$) and show that the gains from trading are not symmetric even in this case. Note that when $\delta = 0$, $g_2(\cdot)$ and c^i are identical for both trader groups, and $g_1(\cdot)$ reduces to the following:

$$g_1(\theta^i; \lambda^i, X; \hat{\theta}, 0) = \frac{\alpha\sigma_D^2}{4(1-k)} [\theta^i - (1-k)\hat{\theta} + \lambda^i X]^2, \quad i = a, b. \qquad (21)$$

Under (13) and (20), we have $\theta^i > (1-k)\hat{\theta}$. Thus, the trading gain is always higher for the group with $\lambda^i X > 0$ (potential sellers) than for the group with $\lambda^i X < 0$ (potential buyers).

Figure 2 illustrates the asymmetric trading gains between the buyers and the sellers. We plot the case in which $\lambda^i X > 0$. The solid lines correspond to traders' desired stock holding before and after their idiosyncratic shocks. A trader i starts with an initial holding θ^i, which is optimal before receiving any idiosyncratic shock. After learning that he will receive a shock ($\lambda^i \neq 0$), the trader's preferred stock exposure changes to $(1-k)\,\hat{\theta}$, which is only $(1-k)$ share of the per capita stock supply in the market, and clearly lower than his initial holding θ^i. This change in the desired risk exposure is independent of the actual sign or the magnitude of the shock X. Thus, conditional on the idiosyncratic shock, potential sellers who have received additional positive exposure via the nontraded risk (i.e., $\lambda^i X > 0$) are further away from their optimal holding than potential buyers are. As a result, the gains from trading are higher for the potential sellers.

The reason traders prefer a lower risk exposure upon receiving the idiosyncratic shock is that the cost of participation prevents the trader from trading in the market at all times. As a result, the trader expects to bear some of the idiosyncratic risk Z. This extra risk effectively reduces his risk tolerance and lowers his desired stock exposure relative to market makers, who face no cost and can always trade.[14] The percentage reduction in the trader's desired position, captured by k, is proportional to the level of the remaining uncertainty in his idiosyncratic risk exposure.

In summary, the main intuition behind the asymmetric trading gains is as follows. Since traders choose their initial holdings before they learn whether or not they will receive idiosyncratic shocks, they rationally choose a high initial holding if they expect a low probability of ever receiving a shock. However, once they are hit with shocks, their initial holding level becomes too high given the possibility of bearing some unhedged risk. Irrespective of the sign of his idiosyncratic shock, a trader prefers to decrease his stock exposure. Obviously, potential sellers who have received additional positive exposure are further away from the desired holding level than are potential buyers. As a result, sellers enjoy larger gains from trading.[15]

3.3 Participation equilibrium

Intuitively, the asymmetry in gains from trading will lead to asymmetric participation between the traders. In particular, since potential sellers always have higher gains from trading than potential buyers in our setting, we further expect that sellers are more likely to participate in the market than buyers. We confirm this intuition by considering the participation equilibrium.

In order to solve for the equilibrium ω^a and ω^b, we substitute the expression of $\hat{\theta}$ and δ in (13) into the definition of $g(\cdot)$ and define a function of participation gain for group-a and group-b traders, respectively, as

$$g^a(\omega^a, \omega^b) \equiv g(\theta^a; \lambda^a, X; \hat{\theta}, \delta), \qquad g^b(\omega^a, \omega^b) \equiv g(\theta^b; \lambda^b, X; \hat{\theta}, \delta). \quad (22)$$

The following proposition describes the participation equilibrium.

Proposition 3. *When agents' initial stock holdings satisfy (20), there exists a unique participation equilibrium. Let*

$$\hat{s}^a = \begin{cases} 0, & \textit{if } g^a(0,0) \le 0 \\ 1, & \textit{if } g^a(1,0) \ge 0 \\ s^a, & \textit{otherwise} \end{cases} \quad \textit{and} \quad \hat{s}^b = \begin{cases} 0, & \textit{if } g^b(1,0) \le 0 \\ 1, & \textit{if } g^b(1,1) \ge 0 \\ s^b, & \textit{otherwise,} \end{cases}$$

[14] The result that traders become effectively more risk averse with unhedged idiosyncratic risks is clearly preference dependent. Kimball (1993) shows that it is true for "standard risk aversion," which is defined as a class of utility function that exhibits both DARA and decreasing absolute prudence.

[15] In a setting similar to ours, Lo, Mamaysky, and Wang (2004) show that even in continuous time the gain from trading is asymmetric around the optimal holding due to the fact that traders only trade infrequently.

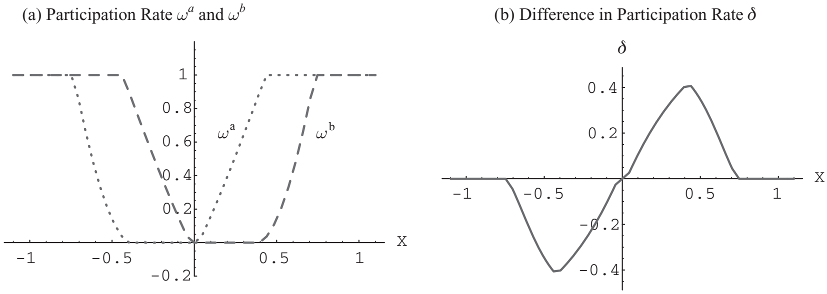

Figure 3
Equilibrium participation
The figure plots the equilibrium participation rate for the two trader groups for different values of idiosyncratic shock X. Panel (a) reports the equilibrium fraction of group i traders who choose to participate, where the dotted and the dashed lines refer to group a and b traders, respectively. Panel (b) reports the difference in participation decisions, $\delta = \lambda\nu(\omega^a - \omega^b)/[\mu + \lambda\nu(\omega^a + \omega^b)]$. Other parameters are set at the following values: $\bar{\theta} = 1, \alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 1$, $\nu = 5$, and $\lambda = 0.15$.

where s^a and s^b are the solutions to $g^a(s^a, 0) = 0$ and $g^b(1, s^b) = 0$, respectively. For $X > 0$, the equilibrium is fully specified as follows:

A. For $g^a(1, \hat{s}^b) \geq 0$, $\omega^a = 1$ and $\omega^b = \hat{s}^b$.
B. For $g^a(1, \hat{s}^b) < 0$ and $g^b(\hat{s}^a, 0) \leq 0$, $\omega^a = \hat{s}^a$ and $\omega^b = 0$.
C. Otherwise, $\omega^a, \omega^b \in (0, 1)$ and satisfy both $g^a(\omega^a, \omega^b) = 0$ and $g^b(\omega^a, \omega^b) = 0$.

Moreover, $\omega^a \geq \omega^b$. For $X < 0$, the equilibrium is given by exchanging subscripts a and b.

When $X > 0$, group-a traders are potential sellers and group-b traders are potential buyers. Cases A and B thus describe two polar cases, either all potential sellers participate (Case A) or no buyers do (Case B). Case A corresponds to the situation in which trading gains for sellers are overwhelming so that they all enter the market, irrespective of what buyers do. The presence of a large number of sellers increases the trading gain for buyers. Thus, in this case some buyers may also choose to participate. Case B corresponds to the situation in which all sellers will not participate but, independent of what they do, the net trading gains for buyers remain negative. In this case, some sellers choose to participate while buyers do not. Case C corresponds to the intermediate case when we have a partial interior solution. In this case, participation of each group depends on the degree of participation of the other group.

Proposition 3 confirms that there are always more sellers entering the market than buyers in equilibrium, generating an excess sell order in the market and the need for liquidity. Market makers provide the necessary liquidity in equilibrium.

Figure 3 illustrates the equilibrium participation decisions as functions of the idiosyncratic shock X. Panel (a) reports the fraction ω^i of traders within group i who choose to participate. The dotted line plots ω^a and the dashed line plots ω^b. Panel (b) reports the difference in participation ratio between the two groups of traders, δ, defined in equation (13). Consistent with our earlier intuition, more sellers are participating than buyers as ω^a is always above ω^b when $X > 0$. In particular, when X is not too far from zero, $\omega^a > 0$ and $\omega^b = 0$, that is, no group-b traders choose to participate because the benefit from trading is too small, and only a fraction of group-a traders participates. This corresponds to Case B in Proposition 3. As X increases, the gains from trading increase for both groups and both ω^a and ω^b increase. In particular, for medium levels of X, ω^b becomes positive and ω^a reaches one. That is, the gain from trading dominates the cost for group-a traders and they all choose to participate. This corresponds to Case A in Proposition 3. When $X < 0$, group-a traders become potential buyers and group-b traders become potential sellers. The above results remain the same after we switch subscripts a and b. In fact, ω^b is simply the mirror image of ω^a around the vertical axis, reflecting the fact that traders a and b face opposite idiosyncratic shocks. Neither ω^a nor ω^b is symmetric around zero, consistent with the fact that a trader's gain from trading is asymmetric between positive and negative idiosyncratic shocks.

Panel (b) of Figure 3 shows that the normalized difference between ω^a and ω^b is always positive when $X > 0$, indicating that more group-a traders are participating. Since they are potential sellers when $X > 0$, the aggregate order imbalance is skewed toward sell orders. Similarly, when $X < 0$, δ is always negative, indicating more group-b traders are participating. Since group-b traders are potential sellers when $X < 0$, the order imbalance is again skewed toward sell orders.

3.4 Full equilibrium of the economy

We now solve the full equilibrium of the economy. We start by computing the value function for all agents at time t, including traders who receive no idiosyncratic risks. For trader $i = a, b$, his indirect utility function, J_P or J_{NP}, depends on his own λ^i and X, given his initial stock holding θ^i_t. For a trader with $\lambda^i \neq 0$, his unconditional value function becomes

$$J^L\big(\theta^i_t;\theta_t\big) = \mathrm{E}\big[\max\big\{J_P\big(\theta^i_t;\lambda^i, X;\hat{\theta},\delta\big),\; J_{NP}(\theta^i_t;\lambda^i, X;\hat{\theta},\delta)\big\}\;\big|\;\lambda^i \neq\big] \quad (23)$$

and for a trader with $\lambda^i = 0$, who does not observe X, his value function is

$$\begin{aligned} J^{NL}\big(\theta^i_t;\theta_t\big) = \max\big\{&\mathrm{E}\big[J_P\big(\theta^i_t;\lambda^i, X;\hat{\theta},\delta\big)\;\big|\;\lambda^i = 0\big],\\ &\mathrm{E}\big[J_{NP}\big(\theta^i_t;\lambda^i, X;\hat{\theta},\delta\big)\;\big|\;\lambda^i = 0\big]\big\}, \end{aligned} \quad (24)$$

where $\hat{\theta}$ and δ, defined in (13), depend on the equilibrium participation ratio ω^a and ω^b in Proposition 3 and thus are functions of X (and θ_t), and $\mathrm{E}[\cdot]$ denotes expectation over X. The *ex ante* utility of any trader before receiving any information on idiosyncratic shocks can then be defined as a weighted average of J^L and J^{NL}

$$J^i\left(\theta_t^i;\theta_t\right) = \lambda J^L\left(\theta_t^i;\theta_t\right) + (1-\lambda)J^{NL}\left(\theta_t^i;\theta_t\right), \quad i = a, b. \tag{25}$$

Finally, for market makers, the *ex ante* utility simply is

$$J^m\left(\theta_t^m;\theta_t\right) = \mathrm{E}\left[J_P\left(\theta_t^m;\lambda^m, X;\hat{\theta},\delta\right) \,\middle|\, \lambda^m = 0, c^i = 0\right]. \tag{26}$$

To solve for the full equilibrium of the economy, we first take P_{t+1} as given to derive the equilibrium price P_t and stock holding θ_t from the following market clearing condition:

$$\mu\,\theta_t^m + \nu\left(\theta_t^a + \theta_t^b\right) = (\mu + 2\nu)\,\bar{\theta}. \tag{27}$$

We then impose the stationarity condition

$$P_{t+1} = P_t \tag{28}$$

to derive the full equilibrium. In addition, we need to confirm that in equilibrium, traders receiving no idiosyncratic shocks optimally choose to stay out of the market, that is,

$$\mathrm{E}\left[J_P\left(\theta_t^i;\lambda^i, X;\hat{\theta},\delta\right) \,\middle|\, \lambda^i = 0\right] \le \mathrm{E}\left[J_{NP}\left(\theta_t^i;\lambda^i, X;\hat{\theta},\delta\right) \,\middle|\, \lambda^i = 0\right]. \tag{29}$$

The following proposition describes the condition that defines the equilibrium.

Proposition 4. *A stationary equilibrium of the economy is determined by the set of prices and holdings $\{P_t, \theta_t\}$ that satisfies conditions (20) and (29) and solves agents' optimality condition*

$$0 = \frac{\partial}{\partial\theta_t^i}\, J^i\left(\theta_t^i;\theta_t\right), \quad i = a, b, m, \tag{30}$$

the market clearing condition (27), and the stationarity condition (28).

Equation (30) is agents' first-order condition for optimal portfolio choice at t before they receive any idiosyncratic shocks.

We can solve the equilibrium explicitly when the probability of idiosyncratic shock λ is small, as shown in the Appendix, which leads to the following theorem:

Theorem 1. *When the probability of idiosyncratic shock λ is small, there exists a stationary equilibrium as described by Proposition 4.*

For arbitrary λ, we have to solve the equilibrium numerically.

4. Endogenous Liquidity Demand and Market Crashes

The equilibrium under costly participation shows two striking features. First, despite the fact that the two groups of traders have perfectly offsetting trading needs, their actual trades are not synchronized. The nonsynchronization in their trades gives rise to the need for liquidity in the market. A group of traders may bring their orders to the market while traders with offsetting trading needs are absent, creating an imbalance of orders. The stock price adjusts in response to the order imbalance to induce market makers to provide liquidity and to accommodate the orders. As a result, the price of the stock depends not only on the fundamentals (i.e., its expected future payoffs and total risks), but also on idiosyncratic shocks that market participants face. Second, despite the symmetry between shocks to potential buyers and sellers, the order imbalance observed in the market tends to be asymmetric and on average is dominated by sell orders. Thus, the endogenous liquidity need typically takes the form of excessive selling, which causes the price to tank. We now examine these results and their implications in more detail.

By construction, the equilibrium stock price is stationary over time at the beginning of each generation, $P_{t+1} = P_t = P$. And it fluctuates during the lifespan of each generation as a function of the idiosyncratic shocks. As (14) indicates, the intermediate price consists of two components: the risk-adjusted fundamental value $R_F^{-1}(\mathrm{E}_{t+1/2}[D_{t+1}] + P - \frac{1}{2}\alpha\sigma_D^2\,\hat{\theta})$ and the liquidity component

$$p \equiv -\tfrac{1}{2}\left(\alpha\sigma_D^2 \big/ R_F\right)\delta Z. \tag{31}$$

The fundamental component is equal to the expected future payoffs (dividend plus resale price) minus a risk premium. It reflects the stock's "fundamental value" since it gives the stock price when the idiosyncratic shock is zero. The liquidity component p, on the other hand, captures price deviations caused by market illiquidity. It is nonzero only when agents receive idiosyncratic shocks. Moreover, it is proportional to the per capita order imbalance, driven by the asymmetric participation between buyers and sellers. Since our purpose here is to understand the endogenous nature of order imbalances and its impact on asset prices, we focus our discussion on the liquidity component p.

From (31), p depends on the difference in market participation rate δ, which is a function of the signal X, and the realized idiosyncratic shock Z, which is equal to the signal X plus an update, $Z - X \sim N(0, \sigma_z^2)$. We can average out the update term and consider the expected liquidity component conditional on the signal X:

$$\hat{p} = \mathrm{E}[\,p \mid X] = -\tfrac{1}{2}\left(\alpha\sigma_D^2 \big/ R_F\right)\delta X. \tag{32}$$

Panel (a) of Figure 4 plots the conditional liquidity component $\hat{p}$ as a function of X. Recall from Section 2.3 that $\hat{p}$ is always zero in the absence of

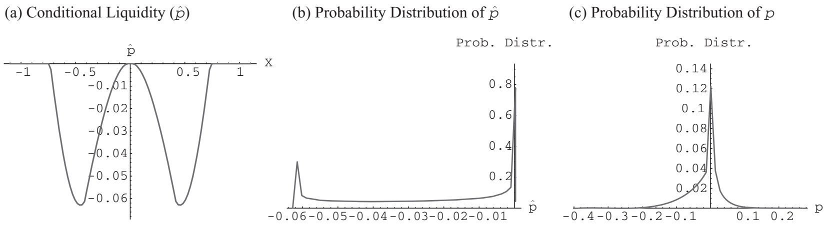

Figure 4
The liquidity component p in price and its expected value $\hat{p}$ conditional on signal X
Panel (a) plots the conditional liquidity component $\hat{p}$ as a function of the signal X. Panels (b) and (c) plot the probability density function of the conditional liquidity component $\hat{p}$ and the unconditional liquidity component p, respectively. The values at the point 0 correspond to the total probability mass at the point since the density function should be infinity at this point. Other parameters are set at the following values: $\bar{\theta} = 1$, $\alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 1$, $\nu = 5$, and $\lambda = 0.15$.

participation costs. Figure 4(a) shows that $\hat{p}$ is generally not zero and is always negative in the presence of costs. This result follows costs. This result follows directly from the asymmetric participation equilibrium obtained in Section 3. In particular, partial participation leads to nonsynchronized trades among traders and the need for liquidity. The stock price has to adjust to attract the market makers to provide the liquidity and accommodate the trades. In general, the stock price becomes dependent on the idiosyncratic shocks of individual traders and $\hat{p} \neq 0$. Moreover, potential sellers are more willing to enter the market to sell the stock. Thus, the average order imbalance, as captured by $-\delta X$, is always negative, which leads to a negative $\hat{p}$. The fact that $\hat{p}$ is independent of the sign of X indicates that, independent of the source of idiosyncratic shocks and its distribution among investors, costly participation always leads to an excess selling pressure in the market and a lower stock price. In summary, we have the following result.

Result 1. *Under costly participation, purely idiosyncratic trading needs can lead to aggregate demand for liquidity, which always takes the form of excess selling and causes asset prices to drop.*

The magnitude of $\hat{p}$ depends on the size of X. From Figure 4(a), we further observe that $\hat{p}$ is highly nonlinear in X. In particular, for small values of X, gains from trading are small for all traders and they do not enter the market. As a result, there is no need for liquidity and the price impact of liquidity is equal to zero. For very large values of X, gains from trading are sufficiently large for all traders and they all enter the market. There is no need for liquidity and $\hat{p}$ is also equal to zero. For intermediate ranges of X, the gains from trading are large enough for some traders to enter the market, but not for all traders to do so. It is in this case that trades are nonsynchronized and liquidity is needed in the market, which will in turn affect the stock price. As Figure 4(a) shows, the price impact of liquidity reaches the maximum for a certain magnitude of the idiosyncratic shock.

The result that the price impact of liquidity need is one-sided and highly nonlinear arises from the fact that liquidity needs are endogenous in our model. In most of the existing models of liquidity, such as that in Grossman and Miller (1988), liquidity needs are exogenously specified; consequently, the price impact is linear in the exogenous liquidity needs and symmetrically distributed. Our analysis shows that modeling liquidity needs endogenously is important for understanding their behavior and impact on prices.

The nonlinearity in the price impact of liquidity leads to another interesting result: large but infrequent price movements in the absence of any aggregate shocks. Figure 4(b) plots the probability distribution of $\hat{p}$. When participation is costless, there is no liquidity effect, and the distribution is simply a delta function at zero. When traders face costs to participate in the market, however, the stock price becomes dependent on the idiosyncratic shock X. Moreover, even though the underlying idiosyncratic shocks that drive the individual traders' trading needs are normally distributed, their price impact $\hat{p}$ as depicted in Figure 4(a) is always negative and highly nonlinear in X. In particular, its distribution peaks at a finite and negative value. Since such a price movement is caused by a large imbalance in trades, endogenously risen from idiosyncratic shocks, it represents a market crash driven purely by liquidity needs. We call it a "liquidity crash." We summarize this result as follows.

Result 2. *The impact of liquidity can lead to "liquidity crashes" in which large price drops occur in the absence of any shocks to the fundamentals.*

The above discussion focuses on $\hat{p}$, which gives the expected price impact of liquidity conditional on X, the signal on future idiosyncratic shocks. The actual liquidity component in price, as given in (31), depends on Z, the actual realization of idiosyncratic shock. Although the behavior of p is qualitatively similar to that of $\hat{p}$, its distribution is slightly different because of the additional shock. Figure 4(c) plots the unconditional distribution of p. Here, we observe that p exhibits negative skewness and fat tails. In the absence of liquidity effects, its distribution will simply be a delta function at zero. The total price of the stock also includes the news on the dividend, which is normally distributed. Thus, its distribution combines the distributions of dividend news and p, which also exhibits negative skewness and fat tails. Hence, we have the following result.

Result 3. *The impact of liquidity can significantly increase the downside risk and lead to negative skewness and fat tails in prices.*

5. Return and Volume

The impact of liquidity also leads to testable implications about the behavior of return and volume. In this section, we explore some of these implications.

The (excess) returns on the stock over the two relevant periods are given by

$$R_s = D_{t+s} + P_{t+s} - R_F P_{t+s-1/2}, \quad s = 1/2,\ 1. \tag{33}$$

The trading volume of the stock at $s = 1/2$ is given by

$$V_{1/2} \equiv \mu\,|\delta Z| + \lambda\nu \sum_{i=a,b} \omega^i |\delta Z - \lambda^i Z|. \tag{34}$$

5.1 Return and volume dynamics

First, let us examine the impact of liquidity on return dynamics. In contrast to shocks to fundamentals (i.e., cash flows), which cause permanent changes in prices, liquidity shocks give rise to transitory price changes. Consequently, the impact of liquidity generates negative serial correlation in returns. This is probably the most salient feature of liquidity's influence on prices, as emphasized in Ho and Stoll (1981), Grossman and Miller (1988), and Campbell, Grossman, and Wang (1993). Our analysis of liquidity leads to additional predictions. In particular, in our model it is the idiosyncratic shock to different agents that gives rise to the liquidity need, which leads to transitory price deviations from its fundamental value. Moreover, the impact of liquidity on prices is more likely to be negative. We thus have the following result.

Result 4. *The impact of liquidity leads to return reversals. Moreover, negative returns exhibit stronger reversals than positive returns.*

Given a set of parameters, we can simulate the returns and trading volumes from the model and compute their statistics. Table 1 reports several of these statistics for the benchmark parameters. The first three columns of Panel A report the unconditional mean of volume and returns and the fourth column reports the value of return serial correlation, denoted by $\rho \equiv \mathrm{Corr}[\mathrm{R}_{1/2}, \mathrm{R}_1]$. Clearly, $\rho < 0$, confirming the first part of Result 4. Panel B reports these statistics conditioning on the current return being above and below its average, respectively. The third and fourth columns of Panel B confirm that the expected future return (i.e., $\mathrm{E}[R_1]$) is higher and the return autocorrelation is more negative if the current return is below average.

The fifth column of Table 1 reports the excess volatility, which is defined as the volatility of the liquidity component p in (31) and captures the volatility of the price in excess of the fundamental volatility. Since lower returns and higher volume are indicative of liquidity demand, not surprisingly, we find that the excess volatility is higher in these cases. We summarize the result as follows.

Result 5. *Volatility is higher during negative return and high volume periods.*

Table 1
Return and volume dynamics

Conditioning information	$E[V_{1/2}]$	$E[R_{1/2}]$	$E[R_1]$	ρ	σ_p
		A. Unconditional			
Unconditional	0.079	0.349	0.375	−0.023	0.045
		B. Conditioning on current return			
$R_{1/2} > E[R_{1/2}]$	0.076	0.588	0.370	−0.012	0.043
$R_{1/2} < E[R_{1/2}]$	0.081	0.110	0.381	−0.016	0.047
		C. Conditioning on current volume			
$V_{1/2} > E[V_{1/2}]$	0.164	0.323	0.402	−0.035	0.056
$V_{1/2} < E[V_{1/2}]$	0.031	0.364	0.360	−0.009	0.028

Parameters take the following values: $\bar{\theta} = 1$, $\alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 0.2$, $\nu = 5$, and $\lambda = 0.02$.

While the asymmetric relation between return and volatility is well documented in the literature, our model offers a liquidity-based explanation in addition to the existing alternatives such as leverage and time-varying volatility.[16]

Next, we consider the joint behavior of returns and volume. In our setting, both trading and the need for liquidity are generated by agents' idiosyncratic shocks. Thus, they are closely related. In particular, higher trading volume also implies that the need for liquidity or the order imbalance is more likely to be high, which is associated with lower current returns but higher future returns. As shown in the second column of Panel C, Table 1, higher current volume is on average accompanied by lower returns. Given return reversals, it also implies higher future returns. The third column of Panel C confirms this result, while the fourth column shows that the return reversal is also stronger for higher volume. We summarize these findings as follows.

Result 6. *Higher volume implies higher future returns. Also, returns accompanied by higher volume exhibit stronger reversals.*

The negative relation between volume and return serial correlation is studied in Campbell, Grossman, and Wang (1993). The positive relation between volume and future returns, a rather surprising result, has been empirically documented by Gervais, Kaniel, and Mingelgrin (2001). Our analysis provides a liquidity-based explanation.

Given the negative correlation between contemporaneous return and volume, both driven by agents' endogenous liquidity needs and predictive of future returns, a natural question is whether volume provides additional information about the magnitude of liquidity needs and consequently future returns.

16 For a levered firm, a negative realized return reduces its value and further increases its financial leverage, which makes the firm's equity riskier and increases its volatility. See, for example, Black (1976), Christie (1982), and Schwert (1989). When volatility is time varying, an anticipated increase in volatility raises the required return on equity, which causes an immediate stock price decline and a lower realized return. See, for example, French, Schwert, and Stambaugh (1987), and Campbell and Hentschel (1992).

Table 2
Expected future return conditional on current returns and volume

Sorted by current returns, $R_{1/2}$	Sorted by current volume, $V_{1/2}$			
	Low	Medium	High	High − Low
Low (25%)	0.351	0.383	0.418	0.067
Medium (50%)	0.351	0.376	0.402	0.051
High (25%)	0.350	0.368	0.388	0.038

Parameters take the following values: $\bar{\theta} = 1$, $\alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 0.2$, $\nu = 5$, and $\lambda = 0.02$.

In Table 2, we report the expected future returns for each subgroup after double-sorting the simulated data into four-by-four subgroups based on current returns $R_{1/2}$ and volume $V_{1/2}$. For conciseness we combine the middle two quartiles in the table. In the last column, we report the difference in the average future return for the high- and low-volume groups. Clearly, higher volume still predicts higher future returns even after controlling for the current return. Moreover, the predictive power of volume is stronger (the difference between the high and low groups is larger) for lower current returns. These findings are consistent with the empirical results of Avramov, Chordia, and Goyal (2006).

A simple way to describe the double-conditioning results in Table 2 is to consider the following forecasting equation, proposed by Campbell, Grossman, and Wang (1993):

$$R_{t+1} = a + bR_{t+1/2} + cV_{t+1/2}R_{t+1/2} + \varepsilon_{t+1}. \tag{35}$$

If we use the whole sample to estimate (35), from Result 6, c is negative, which is confirmed by the more negative ρ for higher volume in the last column of Panel C, Table 1. However, suppose that we split the sample into two according to the current return, below or above its mean, and then run the regression for each subsample. Let c_- and c_+ denote the corresponding coefficients for the volume-return interaction term. Comparing the first and last rows of Table 2, we observe that c_- is much more negative than c_+.[17] This result again arises from the fact that the transitory impact of liquidity on prices is typically negative with large volume, an important feature of our model. Thus, we have the following result.

Result 7. *Negative returns accompanied by high volume exhibit stronger reversals than positive returns.*

5.2 The cross-section of returns and volume

The impact of liquidity on asset prices clearly depends on the level of liquidity available in the market. To the extent that the level of liquidity varies across

[17] Indeed, because $R_{1/2}$ is negative for the low-return group and positive for the high-return group, the coefficient c_- should be significantly negative while c_+ should be mildly positive.

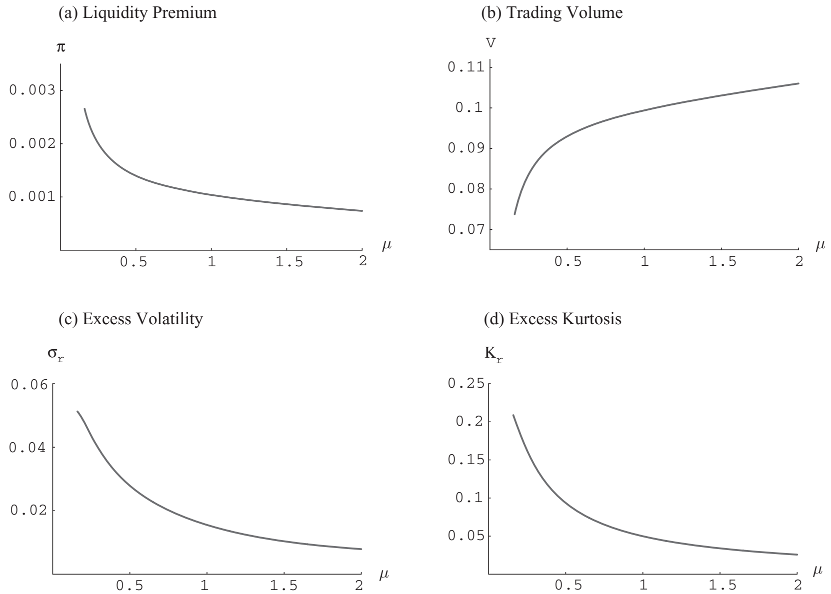

Figure 5
Liquidity premium, trading volume, excess volatility, excess kurtosis, and the level of liquidity μ
Panels (a), (b), (c), and (d) plot the liquidity premium, average trading volume, excess return volatility, and excess kurtosis (which equals zero for normal distribution) as a function of μ, respectively. Other parameters are set at the following values: $\bar{\theta} = 1$, $\alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 1$, $\nu = 5$, and $\lambda = 0.15$.

different markets, its influence also changes. In our model, the level of liquidity in the market is captured by μ, the population of market makers. Thus, in this subsection we analyze how the behavior of prices depends on μ. The resulting implications provide a theoretical motivation to examine the differences in price behavior across markets with different levels of liquidity.

We first consider how the level of liquidity influences expected stock returns, which is defined as $\mathrm{E}[R] = \mathrm{E}[(R_{1/2} + R_1)/2]$. In the absence of participation costs, the market is perfectly liquid and the expected stock return is determined only by the risk premium, which we denote by R^0. From Section 2.3, we have $R^0 = \frac{1}{2}\alpha\sigma_D^2\bar{\theta}$. In the presence of participation costs, we define the liquidity premium as

$$\pi = \mathrm{E}[R] - R^0. \tag{36}$$

In all discussions, R^0 is held constant and we use the terms liquidity premium and expected return interchangeably.

In Figure 5(a), we plot the liquidity premium π for different values of μ, the level of liquidity in the market. Clearly, the liquidity premium decreases with μ. Figure 5(b) plots the average trading volume against μ, which shows that with more market makers present, liquidity is higher and so is trading volume.

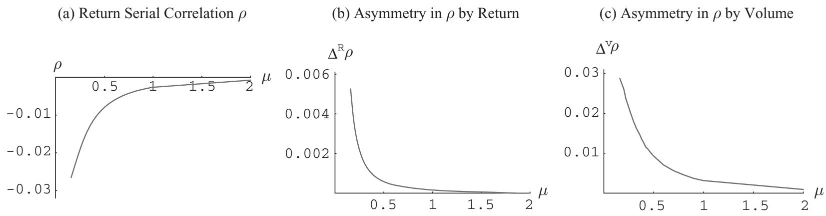

Figure 6
Dependence of unconditional and conditional return serial correlation on the level of liquidity μ
Panel (a) reports the unconditional return serial correlation as a function of μ. Panel (b) reports $\Delta^R\rho$ in (37), the difference in return serial correlation between positive and negative returns. Panel (c) reports $\Delta^V\rho$ in (38), the difference in return serial correlation between high trading volume and low trading volume states. Other parameters are set at the following values: $\bar{\theta} = 1$, $\alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 1$, $\nu = 5$, and $\lambda = 0.15$.

As discussed in Section 4, in our model, non-fundamental shocks, idiosyncratic to individual agents, give rise to the endogenous liquidity need and its impact on price deviations. This liquidity effect leads to more volatile and fat-tailed prices. We next examine how liquidity influences the behavior of returns. In Figures 5(c) and (d), we report the excess volatility and the excess kurtosis for different values of μ. We see that returns are more volatile and exhibit fatter tails in less liquid markets (i.e., with lower μ). We summarize the above results as follows.

Result 8. *Lower levels of liquidity lead to higher expected returns, higher excess volatility, and fatter tails in the return distribution, but lower average trading volume.*

Next, we examine how the return and volume dynamics vary with liquidity. As discussed earlier, a general feature of the price impact of liquidity is negative serial correlation in returns. Figure 6(a) plots the return serial correlation for different values of μ. It clearly shows that for markets with higher levels of liquidity, there is less negative serial correlation in returns. An important implication of our model is that return serial correlation is asymmetric, stronger for negative returns than for positive returns. We define the asymmetry in return serial correlation as follows:

$$\Delta^R\rho \equiv \mathrm{Corr}\big[\mathrm{R}_{1/2}, \mathrm{R}_1 \,\big|\, \mathrm{R}_{1/2} \geq \mathrm{E}\big[\mathrm{R}_{1/2}\big]\big] - \mathrm{Corr}\big[\mathrm{R}_{1/2}, \mathrm{R}_1 \,\big|\, \mathrm{R}_{1/2} < \mathrm{E}\big[\mathrm{R}_{1/2}\big]\big]. \tag{37}$$

Result 4 states that $\Delta^R\rho$ is positive. Figure 6(b) further shows that the magnitude of $\Delta^R\rho$ decreases with μ. In other words, the return serial correlation becomes less symmetric, and in particular, stronger for negative returns in less-liquid markets.

Result 6 states that the negative return serial correlation is stronger when trading volume is higher. We define the difference in the return serial correlation

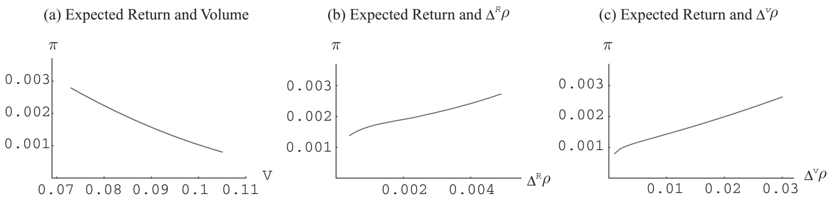

Figure 7
The expected return and different liquidity measures
Panels (a), (b), and (c) plot the liquidity premium (which in turn determines the expected return) against trading volume $E[V]$, asymmetry in return serial correlation $\Delta^R\rho$, and the volume sensitivity of return serial correlation $\Delta^V\rho$, respectively. Other parameters are set at the following values: $\bar{\theta} = 1$, $\alpha = 4$, $r = 0.05$, $\bar{D} = 0.36$, $c = 0.09$, $\sigma_D = 0.42$, $\sigma_z = 0.7$, $\sigma_u = 0.7$, $\mu = 1$, $\nu = 5$, and $\lambda = 0.15$.

between low-volume and high-volume states as follows:

$$\Delta^V\rho \equiv \mathrm{Corr}\left[R_{1/2}, R_1 \,\middle|\, V_{1/2} < E\left[V_{1/2}\right]\right] - \mathrm{Corr}\left[R_{1/2}, R_1 \,\middle|\, V_{1/2} \geq E\left[V_{1/2}\right]\right]. \tag{38}$$

Then, $\Delta^V\rho$ should be positive. Figure 6(c) confirms the result and further shows that $\Delta^V\rho$ decreases with μ. For markets with higher liquidity (i.e., high μ values), return serial correlation diminishes and $\Delta^V\rho$ decreases to zero. We thus have the following result:

Result 9. *Both the asymmetry in return serial correlation, as measured by $\Delta^R\rho$, and the volume sensitivity of return serial correlation, as measured by $\Delta^V\rho$, are negatively related to the level of liquidity in the market.*

5.3 Liquidity and expected returns

Our above analysis, when applied to a cross-section of stocks, implies a negative relation between a stock's average return and the level of its liquidity. However, the level of liquidity μ is not directly observable. Many empirical studies employ different proxies as measures of liquidity and find that higher liquidity is generally associated with lower expected stock returns. For example, Brennan, Chordia, and Subrahmanyam (1998) use volume as one of the liquidity proxies and find that it has a negative correlation with average stock returns.

In our model, both average return and volume are related to the market liquidity level as an equilibrium outcome. As a result, they are closely related to each other. In particular, as Figure 5(b) shows, higher volume indicates higher liquidity, which requires a lower liquidity premium. This negative relation between volume and liquidity premium is confirmed in Figure 7(a).

Besides volume, other proxies of liquidity have also been used to explain the cross-section of stock returns. But the results are less definitive (see, for example, Hasbrouck 2006). Since most empirical proxies of liquidity are also related to price volatility, Spiegel and Wang (2007) further examine the marginal contribution of these liquidity measures and idiosyncratic return volatility in explaining returns (also see Kieschnick, Cook, and Arugaslan 2007). They find that the impact of idiosyncratic risk is stronger and often consumes the explanatory power of various cost-based liquidity proxies.

Our analysis suggests new empirical measures of liquidity. In particular, Result 9 implies that both $\Delta^R\rho$ and $\Delta^V\rho$ provide measures of liquidity. They capture a unique aspect of liquidity—namely, its general downward pressure on prices, accompanied by high volume and strong return reversals. Using the above measures as proxies for the level of liquidity in the market, we expect a positive relation between them and the liquidity premium π. Figures 7(b) and (c) plot the liquidity premium π for different values of $\Delta^R\rho$ and $\Delta^V\rho$, respectively, and clearly show a positive relation between π and these two measures.

6. Conclusion

In this paper, we show that frictions such as costs to market presence can induce nonsynchronization in agents' trades even when their trading needs are perfectly matched. Each trader, when arriving at the market, faces only a partial demand/supply of the asset. The mismatch in the timing and the size of trades creates temporary order imbalances and a need for liquidity, causing asset prices to deviate from the fundamentals. Purely idiosyncratic shocks can affect prices, introducing additional price volatility. Moreover, the price deviations tend to be highly skewed and of large size. In particular, the shortage of liquidity always causes the price to decrease, and when this happens, the price tends to drop significantly, resembling a crash due to a sudden surge in liquidity needs. We also show that the impact of liquidity leads to interesting implications on return and volume behavior, consistent with existing empirical findings.

A few additional comments are in order. First, our analysis takes as given the population weight of market makers, which determines the amount of liquidity they can provide and the equilibrium impact of liquidity needs. As Huang and Wang (2008) show, the population weight of market makers can be endogenized. In particular, they assume that all agents can pay either a low cost *ex ante* to become a market maker or a high cost *ex post* when trading needs arise. They show that typically only a small fraction of agents will choose to become market makers. In light of their analysis, we can interpret the relative population weight of market makers and traders as an equilibrium outcome. Second, in our model the idiosyncratic shocks are transitory. Thus, when a liquidity crash occurs, the stock price tanks but eventually recovers. The possibility of such a price pattern might seem puzzling since it seems to leave profitable opportunities. However, this is not so given the costs. With a small probability for such an event to happen, it is profitable for only a small number of market makers to enter the market *ex ante* even if the cost for becoming a market maker is rather small. For others, the significant cost to jump in on the spot prevents them from taking advantage of the opportunities. Finally, in our setting, the cost to jump into the market on the spot does impose an upper bound on the potential impact of liquidity on prices. But, this is true only in the absence of aggregate shocks as we have assumed in the model. In the presence of aggregate shocks, the potential impact of endogenous liquidity needs on prices becomes unbounded.

Appendix

Proof of Proposition 1. Given P_{t+1}, participating agent i maximizes his expected utility over his terminal wealth W^i_{t+1} given in (10), which is obtained by integrating over the distribution of D_{t+1} given $\theta^i_{t+1/2}$

$$\max_{\theta^i_{t+1/2}} -e^{-\alpha\left[R_F^2 W_t - R_F c^i + R_F\theta^i_t(P_{t+1/2}-R_F P_t)+\theta^i_{t+1/2}(\mathrm{E}_{t+1/2}[D_{t+1}]+P_{t+1}-R_F P_{t+1/2})-\frac{1}{4}\alpha\sigma_D^2\left(\theta^i_{t+1/2}+\lambda^i Z\right)^2\right]}. \tag{A1}$$

His optimal holding is calculated by solving the first-order condition with respect to $\theta^i_{t+1/2}$,

$$\theta^i_{t+1/2} = \frac{2}{\alpha\sigma_D^2}\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F P_{t+1/2}\right) - \lambda^i Z, \quad i = a, b, m. \tag{A2}$$

The market-clearing condition is given by

$$\mu\theta^m_{t+1/2} + \nu\sum_{i=a,b}\left[\lambda\,\omega^{i,L}\theta^{i,L}_{t+1/2} + (1-\lambda)\omega^{i,NL}\theta^{i,NL}_{t+1/2}\right] = \mu\theta^m_t + \nu\sum_{i=a,b}[\lambda\omega^{i,L} + (1-\lambda)\omega^{i,NL}]\theta^i_t. \tag{A3}$$

Its solution yields the equilibrium price $P_{t+1/2}$. Plugging in the definition of δ and $\hat{\theta}$ in (13) yields the expression of $P_{t+1/2}$ in the proposition. The optimal holding in the proposition is obtained by substituting the equilibrium price $P_{t+1/2}$ back into (A2).

Proof of Proposition 2. For trader i, his (indirect) utility if he chooses to participate or not, denoted by J_P and J_{NP}, respectively, is given by

$$J_P(\theta^i;\lambda^i, X;\hat{\theta},\delta) = \mathrm{E}\left[\max_{\theta^i_{t+1/2}} \mathrm{E}_{t+1/2}\left[-e^{-\alpha W^i_{t+1}}\right] \,\middle|\, \lambda^i, X;\ \eta^i = 1\right] \tag{A4a}$$

$$J_{NP}(\theta^i;\lambda^i, X;\hat{\theta},\delta) = \mathrm{E}\left[-e^{-\alpha W^i_{t+1}} \mid \lambda^i, X;\ \eta^i = 0,\ \theta^i_{t+1/2} = \theta^i\right]. \tag{A4b}$$

Substituting the equilibrium $P_{t+1/2}$ and $\theta^i_{t+1/2}$ in Proposition 1 into (A1) and integrating over the distribution of D_{t+1} and Z conditional on X, we obtain J_P for the participating traders,

$$J_P(\cdot) = -\frac{1}{\sqrt{1-k+k(1-\lambda^i\delta)^2}}\, e^{-\alpha\left[R_F^2 W_t - R_F c^i + \theta^i_t\left(\mathrm{E}_{t+1/2}[D_{t+1}]+P_{t+1}-R_F^2 P_t\right)-\frac{\alpha\sigma_D^2}{4(1-k)}\left(\theta^i_t+\lambda^i X\right)^2+g_1(\cdot)\right]}, \tag{A5}$$

where $g_1(\cdot)$ and k are defined in (18) and (19). Next, we calculate the value function for non-participating traders J_{NP} in (A4b) by integrating over D_{t+1} and Z conditional on X,

$$J_{NP}(\cdot) = -\frac{1}{\sqrt{1-k}}\, e^{-\alpha\left[R_F^2 W_t + \theta^i_t\left(\mathrm{E}_{t+1/2}[D_{t+1}]+P_{t+1}-R_F^2 P_t\right)-\frac{\alpha\sigma_D^2}{4(1-k)}(\theta^i_t+\lambda^i X)^2\right]}. \tag{A6}$$

Finally, we substitute J_P and J_{NP} into (16) to derive the gains from participation. Obviously, trader i chooses to participate in the market if and only if $g(\cdot) > 0$.

We also calculate the value function for traders with $\lambda^i = 0$. Conditional on the signal X, the utility if they choose to participate is

$$J_P(\lambda^i = 0) = -\frac{1}{\sqrt{1+k\,\delta^2}} e^{-\alpha\left[R_F^2 W_t - R_F c^i + \theta_t^i\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t\right) - \frac{1}{4}\alpha\sigma_D^2 {\theta_t^i}^2 + \frac{1}{4}\frac{\alpha\sigma_D^2}{(1+k\,\delta^2)}(\theta_t^i - \hat{\theta} - \delta X)^2\right]}. \tag{A7}$$

If they choose not to participate, the utility is

$$J_{NP}(\lambda^i = 0) = -e^{-\alpha\left[R_F^2 W_t + \theta_t^i\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t\right) - \frac{1}{4}\alpha\sigma_D^2 {\theta_t^i}^2\right]}. \tag{A8}$$

Applying the definition in (16), the gains from participation for traders with $\lambda^i = 0$ is

$$g(\theta^i; \lambda^i = 0, X; \hat{\theta}, \delta) = \frac{\alpha\sigma_D^2}{4(1+\delta^2 k)}\left(\theta^i - \hat{\theta} - \delta X\right)^2 + \frac{1}{2\alpha}\ln(1+\delta^2 k) - R_F c^i. \tag{A9}$$

Proof of Proposition 3.

Lemma 1. *When traders' initial stock holdings satisfy (20), the gain from participation $g^a(\omega^a, \omega^b)$ for group-a traders decreases with ω^a and increases with ω^b, while the opposite is true for group-b traders' gain $g^b(\omega^a, \omega^b)$.*

The proof of Lemma 1 is as follows. Given the definition of $g(\cdot)$ in (17), we compute its partial derivative with respect to ω^a and ω^b. Define $\delta^i \equiv \lambda^i \delta$ and $d^i \equiv 1-k+k(1-\delta^i)^2$. Following (22), let $g^i \equiv g(\theta_t^i; \lambda^i, X; \hat{\theta}, \delta)$. Then

$$\frac{\partial g^i}{\partial \omega^j} = \left(\frac{\partial g_1}{\partial \delta^i} + \frac{\partial g_2}{\partial \delta^i}\right)\frac{\partial \delta^i}{\partial \omega^j} + \frac{\partial g_1}{\partial \hat{\theta}}\frac{\partial \hat{\theta}}{\partial \omega^j}, \quad j = a, b,$$

where

$$\begin{aligned}
\frac{\partial g_1}{\partial \delta^i} &= \frac{\alpha\sigma_D^2(1-k\delta^i)}{2(d^i)^2}\left[k\,\delta^i\theta_t^i + k(1-\delta^i)\hat{\theta} + \lambda^i X\right]\left(\theta_t^i - \frac{1-k}{1-k\delta^i}\,\hat{\theta} + \frac{1-\delta^i}{1-k\delta^i}\,\lambda^i X\right)\\
\frac{\partial g_2}{\partial \delta^i} &= -\frac{(1-\delta^i)k}{\alpha d^i}, \quad \frac{\partial \delta^i}{\partial \omega^j} = \lambda^i\,(\lambda^j - \delta)\hat{\lambda}, \quad \hat{\lambda} \equiv \frac{\lambda\nu}{\mu + \lambda\nu(\omega^a + \omega^b)}\\
\frac{\partial g_1}{\partial \hat{\theta}} &= -\frac{\alpha\sigma_D^2(1-k\delta^i)}{2d^i}\left(\theta_t^i - \frac{1-k}{1-k\delta^i}\,\hat{\theta} + \frac{1-\delta^i}{1-k\delta^i}\,\lambda^i X\right), \quad \frac{\partial \hat{\theta}}{\partial \omega^j} = \hat{\lambda}\left(\theta_t^i - \hat{\theta}\right).
\end{aligned}$$

We now consider the cases of $j = i$ and $j \neq i$ separately. When $j = i$, then $\lambda^i\lambda^j = 1$ and

$$\frac{\partial g^i}{\partial \omega^i} = -\frac{\alpha\sigma_D^2\hat{\lambda}(1-\delta^i)^2}{2(d^i)^2}\left[\lambda^i X + k\,\hat{\theta} + \frac{1-k\,\delta^i}{1-\delta^i}\left(\theta_t^i - \hat{\theta}\right)\right]^2 - \frac{k\hat{\lambda}(1-\delta^i)^2}{\alpha d^i}.$$

From (13), δ increases in ω^a and decreases in ω^b. Hence, $\delta \in [-\bar{\delta}, \bar{\delta}]$, where $\bar{\delta} = \frac{\lambda\nu}{\mu+\lambda\nu} < 1$. Since $\delta^i = \lambda^i\delta \in [-\bar{\delta}, \bar{\delta}]$, we have $\partial g^i/\partial\omega^i < 0$.

When $j \neq i$, then $\lambda^i \lambda^j = -1$ and

$$\frac{\partial g^i}{\partial \omega^j} = \frac{\alpha \sigma_D^2 \hat{\lambda} (1-\delta^2)}{2(d^i)^2} \left[\lambda^i X + k \hat{\theta} + \frac{(1+k)\,\delta^i - 2k\,\delta^2}{1-\delta^2} \left(\theta_t^i - \hat{\theta}\right) \right]^2 + \frac{\alpha \sigma_D^2 \hat{\lambda}}{2(1-\delta^2)} \left[\frac{k\,(1-\delta^2)^2}{\frac{1}{2}\alpha^2 \sigma_D^2 d^i} - \left(\theta_t^i - \hat{\theta}\right)^2 \right]. \tag{A10}$$

Since $\delta^2 \in [0, \bar{\delta}^2]$, we have $(1-\delta^i)^2 \in [0, 1+\bar{\delta}^2]$. From (5), $k = \frac{1}{2}\alpha^2 \sigma_D^2 \sigma_z^2 \in [0, 1]$. Thus,

$$\frac{k(1-\delta^2)^2}{\frac{1}{2}\alpha^2 \sigma_D^2 d^i} \geq \frac{\sigma_z^2 (1-\bar{\delta}^2)^2}{1-k+k(1+\bar{\delta})^2} > \sigma_z^2 (1-\bar{\delta})^2.$$

However, $\hat{\theta}$ in (13) is a weighted average of θ_t^i and θ_t^m with weights in-between zero and one. We have

$$\left(\theta_t^i - \hat{\theta}\right)^2 \leq \left(\theta_t^i - \theta_t^m\right)^2 < \left(\frac{\mu\,\sigma_z}{\mu + \lambda \nu}\right)^2 = \sigma_z^2 (1-\bar{\delta})^2,$$

where the second inequality is due to condition (20). Thus, $\partial g^i / \partial \omega^j > 0$ for $j \neq i$, proving the lemma.

Lemma 2. *Given (20), under symmetric participation, sellers always enjoy larger gains from trading than buyers. That is, when $X > 0$, $g^a(\omega, \omega) \geq g^b(\omega, \omega)$, $\forall\, \omega \in [0, 1]$.*

The proof of Lemma 2 is as follows. When $\omega^a = \omega^b$, $\delta = 0$ and $g(\cdot)$ in (17) reduce to

$$g\left(\theta_t^i; \lambda^i, X; \hat{\theta}, 0\right) = \frac{\alpha \sigma_D^2}{4(1-k)} \left[\theta_t^i - (1-k)\,\hat{\theta} + \lambda^i X\right]^2 - \frac{1}{2\alpha} \ln(1-k) - R_F c^i.$$

Hence,

$$g^a(\omega^a, \omega^b) - g^b(\omega^a, \omega^b) = \frac{\alpha \sigma_D^2}{2(1-k)} \left[\theta_t^i - (1-k)\,\hat{\theta}\right] (\lambda^a - \lambda^b) X.$$

When $X > 0$, group-a traders are sellers and group-b are buyers, and $\lambda^a X \geq 0 \geq \lambda^b X$. Since $\hat{\theta}$ is the weighted average of θ_t^i and θ_t^m, $\theta_t^i - (1-k)\,\hat{\theta} \geq \theta_t^i - (1-k)\,\theta_t^m > 0$, where the last inequality comes from (20). Hence, $g^a(\omega^a, \omega^b) \geq g^b(\omega^a, \omega^b)$ when $\omega^a = \omega^b$, proving the lemma.

Now we prove Proposition 3. First, from Lemma 1, we know that $g^a(\omega^a, 0)$ is a monotonically decreasing function of ω^a. If $g^a(0, 0) > 0 > g^a(1, 0)$, then there exists an $s^a \in (0, 1)$ that solves $g^a(s^a, 0) = 0$. Similarly, $g^b(1, \omega^b)$ is monotonically decreasing in ω^b and $g^b(1, 0) > 0 > g^b(1, 1)$ guarantees that the solution $s^b \in (0, 1)$. Hence, $\hat{s}^a,\ \hat{s}^b \in [0, 1]$.

The three cases in the proposition are clearly exhaustive. Case A has three subcases depending on the value of $\hat{s}^b$. First, if $\hat{s}^b = 0$, then $g^b(1, 0) \leq 0 \leq g^a(1, 0)$. From Lemma 1, $g^b(1, \omega) \leq g^b(1, 0)$ and $g^a(1, 0) \leq g^a(1, \omega)$ for any ω. Hence, $\omega^a = 1$ and $\omega^b = 0$ in equilibrium. Second, if $\hat{s}^b = 1$, then $g^b(1, 1) \geq 0$. Lemma 2 implies $g^a(1, 1) \geq 0$ as well. Hence, all traders participate and $\omega^a = \omega^b = 1$. Third, if $\hat{s}^b = s^b \in (0,\ 1)$, then $g^b(1, \hat{s}^b) = 0$. The condition $g^a(1, \hat{s}^b) \geq 0$ confirms that sellers enjoy positive gains in this case. Hence, at equilibrium participation $\omega^a = 1$ and $\omega^b = \hat{s}^b$, trader a enjoys a positive gain and trader b is indifferent between participating or not.

In Case B, there are only two subcases depending on the value of $\hat{s}^a$. Note that $\hat{s}^a = 1$ is not feasible under the condition $g^a(1, \hat{s}^b) < 0$, since $g^a(1, 0) \leq g^a(1, \hat{s}^b) < 0$ according to Lemma 1, while $\hat{s}^a = 1$ requires $g^a(1, 0) \geq 0$. The first subcase is $\hat{s}^a = 0$. Then $g^a(0, 0) \leq 0$. Since

$g^b(0,0) < g^a(0,0)$ by Lemma 2, $\omega^a = \omega^b = 0$ is the only solution. The second subcase is $\hat{s}^a = s^a \in (0,1)$, where $\hat{s}^a$ solves $g^a(\hat{s}^a, 0) = 0$. At $\hat{s}^a$, trader a is indifferent between participating or not. The condition $g^b(\hat{s}^a, 0) \leq 0$ confirms that trader b does not want to participate when $\omega^a = \hat{s}^a$. Hence, $\omega^a = \hat{s}^a$ and $\omega^b = 0$ in equilibrium.

In Case C, the condition is that $g^a(1, \hat{s}^b) < 0 < g^b(\hat{s}^a, 0)$. Similar to Case B, the condition $g^a(1, \hat{s}^b) < 0$ still rules out the possibility that $\hat{s}^a = 1$. In addition, $g^b(\hat{s}^a, 0) > 0$ rules out the possibility that $\hat{s}^b = 0$, since $0 < g^b(\hat{s}^a, 0) < g^b(1, 0)$ by Lemma 1. Similarly, we can rule out $\hat{s}^a = 0$ and $\hat{s}^b = 1$. Hence, the condition in Case C reduces to $g^a(1, s^b) < 0 < g^b(s^a, 0)$. Note that $g^a(s^a, 0) = 0$ and $g^b(s^a, 0) > 0$ implies that in equilibrium, $\omega^b > 0$. To prove this, assume by contradiction that $\omega^b = 0$. Then at the optimal $\omega^a = s^a$, trader a is indifferent while trader b can gain from participating. Thus, $\omega^b = 0$ cannot be the equilibrium. Similarly, $g^a(1, s^b) < 0$ and $g^b(1, s^b) = 0$ implies $\omega^a < 1$ in equilibrium. Lemma 2 guarantees that both ω^a and ω^b are interior solutions. Both traders need to be indifferent between participating or not, that is, $g^a(\omega^a, \omega^b) = 0$ and $g^b(\omega^a, \omega^b) = 0$. The monotonicity of the g^a and g^b functions ensures the existence of a solution in this case. Finally, to prove that $\omega^a \geq \omega^b$, we assume by contradiction that $\omega^a < \omega^b$. Then $0 = g^a(\omega^a, \omega^b) > g^a(\omega^a, \omega^a) > g^b(\omega^a, \omega^a) > g^b(\omega^a, \omega^b) = 0$, yielding a contradiction. Note that the first inequality is because $g^a(\omega^a, \omega^b)$ increases in ω^b, and the last is because $g^b(\omega^a, \omega^b)$ decreases in ω^b. The middle inequality is because of Lemma 2.

Proof of Proposition 4. We substitute in the participation and market equilibrium from Propositions 1 and 3 and integrate over X to derive the unconditional value function J^L and J^{NL} in (23) and (24). Hence, the *ex ante* value function $J^i(\cdot)$ in (25) is well defined for all traders. Moreover, the utility conditional on X for the market maker is the same as $J_P(\lambda^i = 0)$ in (A7) except that his initial holding is θ_t^m and his cost is $c^m = 0$. Thus,

$$J_P^m(X) = -\frac{1}{\sqrt{1+k\delta^2}}\, e^{-\alpha\left[R_F^2 W_t + \theta_t^m\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t\right) - \frac{1}{4}\alpha\sigma_D^2 \theta_t^{m\,2} + \frac{\alpha\sigma_D^2}{4(1+k\delta^2)}\left(\theta_t^m - \hat{\theta} - \delta X\right)^2\right]}. \tag{A11}$$

Integrating over X then yields the *ex ante* utility $J^m(\cdot) = \mathrm{E}[J_P^m(X)]$.

With the market clearing condition (27), stationarity condition (28), and three first-order conditions in (30), we have five equations and five unknowns $(\theta_t^a, \theta_t^b, \theta_t^m, P_t, P_{t+1})$. A solution to the system gives a full equilibrium of the economy.

Proof of Theorem 1. Proposition 4 describes conditions for an equilibrium. The *ex ante* symmetry between the two groups of traders implies that $J^a = J^b$ and $\theta_t^a = \theta_t^b$. For simplicity, we use index i to denote traders a or b. Substituting in the stationarity condition ($P_{t+1} = P_t$) directly, we are left with three variables $\{P_t, \theta_t^i, \theta_t^m\}$ and three equilibrium conditions: two first-order conditions (30) for agents i and m, respectively, and one market-clearing condition (27).

When $\lambda = 0$, traders face no idiosyncratic shocks. Clearly, traders never participate whenever $c^i > 0$, that is, $\omega^i = 0\ \forall\ i = a, b$. The equilibrium prices are determined by market makers as representative agents, and are identical to those derived in (12). Equilibrium holdings of the stock for all agents are always equal to the per capita supply $\bar{\theta}$.

For small λ, we expand the solution to equilibrium in λ to the first order

$$P_t = \bar{P} + P_\lambda \lambda + o(\lambda), \tag{A12a}$$

$$\theta_t^i = \bar{\theta} + \theta_\lambda^i \lambda + o(\lambda), \tag{A12b}$$

$$\theta_t^m = \bar{\theta} + \theta_\lambda^m \lambda + o(\lambda), \tag{A12c}$$

where $\bar{P}$ is equal to P_t in (12) and $o(\lambda)$ denotes higher-order terms of λ. We then solve the equilibrium up to the first order of λ.

Given (A12), condition (20) is satisfied for small λ since both $\mu\,\sigma_z/(\mu+\lambda\nu)$ and $k\,\theta^m$ in condition (20) are of order $O(1)$ while $\theta_t^i-\theta_t^m=\lambda(\theta_\lambda^i-\theta_\lambda^m)$ is of order $O(\lambda)$, where $O(\cdot)$ denotes terms of the same order. Hence, Proposition 3 holds. And the trading gain in (17) can be simplified to

$$g^i(\cdot)=-\frac{1}{2\alpha}\ln(1-k)+\frac{\alpha\sigma_D^2}{4(1-k)}(k\bar{\theta}+\lambda^i X)^2+O(\lambda).$$

Thus, trader i participates iff $g^i(\cdot)>0$, which occurs iff $X>X_+^i$ or $X<X_-^i$, where

$$X_\pm^i=-\lambda^i\,k\,\bar{\theta}\pm h+O(\lambda),$$

and

$$h\equiv\begin{cases}\dfrac{1}{\alpha\sigma_D}\sqrt{2(1-k)[2\alpha c^i R_F+\ln(1-k)]}, & \text{if }\ 2\alpha c^i R_F+\ln(1-k)\geq 0\\ 0, & \text{if }\ 2\alpha c^i R_F+\ln(1-k)\geq 0.\end{cases}$$

Since δ and $\hat{\theta}$ depend on ω^i only through term $\lambda\omega^i$, we can ignore all $O(\lambda)$ terms for the calculation of ω^i. The equilibrium participation in Proposition 3 can be simplified to

$$\begin{cases}\omega^a=\omega^b=1, & \delta=0, & \text{if } X\leq -k\,\bar{\theta}-h\\ \omega^a=0,\ \omega^b=1, & \delta=-\bar{\delta}, & \text{if } -k\,\bar{\theta}-h<X\leq-|k\,\bar{\theta}-h|\\ \omega^a=\omega^b=1, & \delta=0, & \text{if } -|k\,\bar{\theta}-h|<X\leq|k\,\bar{\theta}-h| \ \text{ and }\ k\,\bar{\theta}>h\\ \omega^a=\omega^b=0, & \delta=0, & \text{if } -|k\,\bar{\theta}-h|<X\leq|k\,\bar{\theta}-h| \ \text{ and }\ k\,\bar{\theta}<h\\ \omega^a=1,\ \omega^b=0, & \delta=\bar{\delta}, & \text{if } |k\,\bar{\theta}-h|<X<k\,\bar{\theta}+h\\ \omega^a=\omega^b=1, & \delta=0, & \text{if } X\geq k\,\bar{\theta}+h.\end{cases}\tag{A13}$$

Since $\bar{\delta}$ is of order $O(\lambda)$, so is δ. The following equation linearizes δ and $\hat{\theta}$:

$$\delta\big(\theta_t^i,\theta_t^m,X\big)=\delta_\lambda(X)\lambda\tag{A14a}$$

$$\hat{\theta}\big(\theta_t^i,\theta_t^m,X\big)=\bar{\theta}+\frac{\mu\theta_\lambda^m+\lambda\,\nu(\omega^a+\omega^b)\theta_\lambda^i}{\mu+\lambda\,\nu(\omega^a+\omega^b)}\,\lambda+o(\lambda)=\bar{\theta}+\theta_\lambda^m\lambda+o(\lambda).\tag{A14b}$$

Using (A12) and (A14) and the definition of $\bar{P}$ in (12), the first-order condition for market makers can be written as

$$\begin{aligned}0&=\frac{\partial J^m}{\partial\theta_t^m}\\&=\mathrm{E}\left[-\alpha J_P^m(X)\left(\mathrm{E}_{t+1/2}[D_{t+1}]+P_{t+1}-R_F^2P_t-\tfrac{1}{2}\alpha\sigma_D^2\theta_t^m-\frac{\alpha\sigma_D^2}{2(1+k\delta^2)}(\hat{\theta}-\theta_t^m+\delta X)\right)\right]\\&=\mathrm{E}\big[-\alpha J_0\,e^{\alpha r\bar{\theta}P_\lambda\,\lambda+o(\lambda)}\big[-(r\,P_\lambda+\tfrac{1}{2}\alpha\sigma_D^2\theta_\lambda^m+\tfrac{1}{2}\alpha\sigma_D^2\delta_\lambda X)\lambda+o(\lambda)\big]\big]\\&=\mathrm{E}\big[\alpha J_0\big(r\,P_\lambda+\tfrac{1}{2}\alpha\sigma_D^2\theta_\lambda^m+\tfrac{1}{2}\alpha\sigma_D^2\delta_\lambda X\big)\big]\lambda+o(\lambda),\end{aligned}$$

where $J_P^m(X)$ is defined in (A11), and $J_0\equiv-e^{-\alpha(R_F^2W_t+\frac{1}{4}\alpha\sigma_D^2\bar{\theta}^2)}$. Plugging in (A13), we have

$$r\,P_\lambda+\tfrac{1}{2}\alpha\sigma_D^2\,\theta_\lambda^m+c_1=0,\tag{A15}$$

where

$$c_1\equiv\frac{\nu}{\mu}\alpha\sigma_D^2\sigma_x\sqrt{\frac{1}{2\pi}}\big(e^{-h_1^2}-e^{-h_2^2}\big),\quad h_1\equiv\frac{k\,\bar{\theta}-h}{\sqrt{2}\,\sigma_x},\quad h_2\equiv\frac{k\,\bar{\theta}+h}{\sqrt{2}\,\sigma_x},$$

and $\sigma_x^2=\frac{\sigma_z^4}{\sigma_z^2+\sigma_u^2}$. Since $h_1\leq h_2$, we know that $c_1\geq 0$.

We now consider the first-order condition for trader i. First, we verify that traders with $\lambda^i = 0$ never participate. From (A9), the expected gain from participation is

$$g_{NL}\left(\theta_t^i;\theta\right) = \mathrm{E}[g(\theta^i;\lambda^i = 0, X;\hat{\theta},\delta)], \tag{A16}$$

where $\mathrm{E}[\cdot]$ is taken with respect to X. Given (A12) and (A14) and the fact that $\theta_t^i - \hat{\theta}$ and δ are both of the order $O(\lambda)$, we have

$$g_{NL}\left(\theta_t^i;\theta\right) = O(\lambda) - R_F c^i,$$

which is negative as long as c^i is finite and λ is small enough. Thus, $J^{NL} = \mathrm{E}[J_{NP}(\lambda^i = 0)]$. When $\lambda^i \neq 0$, a trader's first-order condition can be written as

$$\begin{aligned} 0 &= \lambda \frac{\partial J^L}{\partial \theta_t^i} + (1-\lambda)\frac{\partial J^{NL}}{\partial \theta_t^i} \\ &= \lambda \frac{\partial \mathrm{E}[J_{NP}]}{\partial \theta_t^i} + \lambda \frac{\partial \mathrm{E}\left[\mathbf{1}_{\{g(\cdot)>0\}}\,(J_P - J_{NP})\right]}{\partial \theta_t^i} + (1-\lambda)\,\frac{\partial J^{NL}}{\partial \theta_t^i}, \end{aligned} \tag{A17}$$

where J_P and J_{NP} are defined in (A5) and (A6), and $g(\cdot)$ is the trading gain in (17). The first term in (A17) can be simplified to

$$\begin{aligned} \lambda \frac{\partial \mathrm{E}[J_{NP}]}{\partial \theta_t^i} &= \lambda \frac{\alpha\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t - \frac{\alpha\sigma_D^2\theta_t^i}{2(1-k_z)}\right)}{\sqrt{1-k_z}}, \\ &\quad e^{-\alpha\left[R_F^2 W_t + \theta_t^i\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t\right) - \frac{1}{4}\frac{\alpha\sigma_D^2}{1-k_z}{\theta_t^i}^2\right]} \\ &= \lambda\, J_0\,\alpha\, c_2 + o(\lambda), \quad c_2 \equiv \frac{k_z\,\alpha\sigma_D^2\,\bar{\theta}}{2(1-k_z)^{3/2}}\; e^{\frac{1}{4}\frac{k_z}{1-k_z}\alpha^2\sigma_D^2\bar{\theta}^2}, \end{aligned}$$

where $k_z = \frac{1}{2}\alpha^2\sigma_D^2\sigma_z^2$ for σ_z in (7) captures the total uncertainty in idiosyncratic shocks. Since $J_P = J_{NP}$ when $g(\cdot) = 0$, for the second term in (A17), we have

$$\begin{aligned} \lambda\,\frac{\partial \mathrm{E}[\mathbf{1}_{\{g(\cdot)>0\}}\,(J_P - J_{NP})]}{\partial \theta_t^i} &= \lambda\,\mathrm{E}\left[\mathbf{1}_{\{g(\cdot)>0\}}\,\frac{\partial (J_P - J_{NP})}{\partial \theta_t^i}\right] \\ &= \mathrm{E}\Bigg[-\mathbf{1}_{\{g(\cdot)>0\}}\,J_0\,\frac{\alpha^2\sigma_D^2\,(k\,\bar{\theta} + X)}{2(1-k)^{3/2}} \\ &\qquad \times e^{\frac{\alpha^2\sigma_D^2}{4(1-k)}(k\,\bar{\theta}^2 + 2\bar{\theta}X + X^2)}\lambda + o(\lambda)\Bigg] \\ &= -J_0\,\alpha\,c_2 c_3\,\lambda + o(\lambda), \end{aligned}$$

where

$$\begin{aligned} c_3 &\equiv \sqrt{\frac{1-k_z}{2\pi(1-k)}}\,\frac{\sigma_x}{k_z\,\bar{\theta}}\left(e^{-h_3^2} - e^{-h_4^2}\right) - \frac{1}{2}\,(\mathrm{Erf}(h_3) + \mathrm{Erf}(h_4) - 2) \\ h_3 &\equiv \frac{h(1-k_z) - k_z(1-k)\bar{\theta}}{\sqrt{2(1-k)(1-k_z)}\,\sigma_x}, \quad h_4 \equiv \frac{h(1-k_z) + k_z(1-k)\bar{\theta}}{\sqrt{2(1-k)(1-k_z)}\,\sigma_x}. \end{aligned}$$

Note that $\frac{\partial c_3}{\partial h} = -\frac{h}{\sqrt{2\pi}\, k_z\, \sigma_x \bar{\theta}} (\frac{1-k_z}{1-k})^{3/2}(e^{-h_3^2} - e^{-h_4^2}) \le 0$ (since $h_3 \le h_4$.) Since $h \ge 0$, and $c_3 = 1$ when $h = 0$, we have $c_3 \le 1$. For the third term in (A17), we have

$$\begin{aligned}(1-\lambda)\frac{\partial J^{NL}}{\partial \theta_t^i} &= (1-\lambda)\,\alpha\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t - \tfrac{1}{2}\alpha\sigma_D^2\theta_t^i\right) \\ &\quad \times e^{-\alpha\left[R_F^2 W_t + \theta_t^i\left(\mathrm{E}_{t+1/2}[D_{t+1}] + P_{t+1} - R_F^2 P_t\right) - \frac{1}{4}\alpha\sigma_D^2 {\theta_t^i}^2\right]} \\ &= J_0\alpha\big(r P_\lambda + \tfrac{1}{2}\alpha\sigma_D^2\theta_\lambda^i\big)\lambda + o(\lambda).\end{aligned}$$

Hence, to the first order of λ, the first-order condition for trader i reduces to

$$r P_\lambda + \tfrac{1}{2}\alpha\sigma_D^2\,\theta_\lambda^i + c_2(1-c_3) = 0. \tag{A18}$$

Finally, the market clearing condition (27) reduces to

$$\mu\,\theta_\lambda^m + 2\nu\,\theta_\lambda^i = 0. \tag{A19}$$

Solving systems (A15), (A18), and (A19), we derive the linear stationary equilibrium

$$P_\lambda = -\frac{\mu\, c_1 + 2\,\nu\, c_2(1-c_3)}{r\,(\mu + 2\nu)} \tag{A20a}$$

$$\theta_\lambda^i = \frac{2\mu\,[c_1 - c_2(1-c_3)]}{\alpha\sigma_D^2\,(\mu + 2\nu)} \tag{A20b}$$

$$\theta_\lambda^m = -\frac{4\,\nu\,[c_1 - c_2(1-c_3)]}{\alpha\sigma_D^2\,(\mu + 2\nu)}. \tag{A20c}$$

Moreover, P_λ is always negative since $c_1 \ge 0$, $c_2 \ge 0$, and $c_3 \le 1$.

References

Acharya, V., and S. Viswanathan. 2008. Moral Hazard, Collateral, and Liquidity. Working Paper, London Business School and Duke University.

Admati, A. R., and P. Pfleiderer. 1988. A Theory of Intraday Patterns: Volume and Price Variability. *Review of Financial Studies* 1:3–40.

Allen, F., and D. Gale. 1994. Limited Market Participation and Volatility of Asset Prices. *American Economic Review* 84:933–55.

Amihud, Y., and H. Mendelson. 1980. Dealership Markets: Market Making with Inventory. *Journal of Financial Economics* 8:31–53.

Avramov, D., T. Chordia, and A. Goyal. 2006. Liquidity and Autocorrelations in Individual Stock Returns. *Journal of Finance* 61:2365–94.

Bai, Y., E. C. Chang, and J. Wang. 2006. Asset Prices Under Short-Sale Constraints. Working Paper, University of Hong Kong and Massachusetts Institute of Technology.

Black, F. 1976. Studies of Stock Price Volatility Changes. *Proceeding of the 1976 Meetings of the American Statistical Association, Business and Economical Statistics Section*: 177–81.

Brennan, M. J. 1975. The Optimal Number of Securities in a Risky Portfolio When There Are Fixed Costs of Transacting: Theory and Some Empirical Results. *Journal of Financial and Quantitative Analysis* 10:483–96.

Brennan, M. J., T. Chordia, and A. Subrahmanyam. 1998. Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Returns. *Journal of Financial Economics* 49:345–73.

Brunnermeier, M. K., and L. H. Pedersen. Forthcoming. Market Liquidity and Funding Liquidity. *Review of Financial Studies*.

Campbell, J. Y., S. J. Grossman, and J. Wang. 1993. Trading Volume and Serial Correlation in Stock Returns. *Quarterly Journal of Economics* 108:905–39.

Campbell, J. Y., and L. Hentschel. 1992. No News Is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics* 31:281–318.

Campbell, J. Y., and A. S. Kyle. 1993. Smart Money, Noise Trading, and Stock Price Behavior. *Review of Economic Studies* 60:1–34.

CGFS. 1999. *A Review of Financial Market Events in Autumn 1998.* Committee on the Global Financial System, Bank for International Settlements, Basel, Switzerland.

Chatterjee, S., and D. Corbae. 1992. Endogenous Market Participation and the General Equilibrium Value of Money. *Journal of Political Economy* 100:615–46.

Christie, A. A. 1982. The Stochastic Behavior of Common Stock Variances—Value, Leverage, and Interest Rate Effects. *Journal of Financial Economics* 10:407–32.

Coval, J., and E. Stafford. 2007. Asset Fire Sales (and Purchases) in Equity Markets. *Journal of Financial Economics* 86:479–512.

French, K. R., G. W. Schwert, and R. F. Stambaugh. 1987. Expected Stock Returns and Volatility. *Journal of Financial Economics* 19:3–29.

Gabaix, X., A. Krishnamurthy, and O. Vigneron. 2007. Limits of Arbitrage: Theory and Evidence from the Mortgage Backed Securities Market. *Journal of Finance* 62:557–95.

Gammill, J. F., and T. A. Marsh. 1988. Trading Activity and Price Behavior in the Stock and Stock Index Futures Markets in October 1987. *Journal of Economic Perspectives* 2:25–44.

Gennotte, G., and H. Leland. 1990. Market Liquidity, Hedging, and Crashes. *American Economic Review* 80:999–1021.

Gervais, S., R. Kaniel, and D. H. Mingelgrin. 2001. The High-Volume Return Premium. *Journal of Finance* 56:877–919.

Glosten, L. R. 1989. Insider Trading, Liquidity, and the Role of the Monopolist Specialist. *Journal of Business* 62:211–35.

Glosten, L. R., and P. Milgrom. 1985. Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. *Journal of Financial Economics* 14:71–100.

Gromb, D., and D. Vayanos. 2002. Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs. *Journal of Financial Economics* 66:361–407.

Grossman, S. J. 1988. An Analysis of the Implications for Stock and Futures Price Volatility of Program Trading and Dynamic Hedging Strategies. *Journal of Business* 61:275–98.

Grossman, S. J., and M. H. Miller. 1988. Liquidity and Market Structure. *Journal of Finance* 38:617–33.

Hasbrouck, J. 2006. Trading Costs and Returns for US Equities: Estimating Effective Costs from Daily Data. Working Paper, New York University.

He, Z., and A. Krishnamurthy. 2008. A Model of Capital and Crises. Working Paper, University of Chicago and Northwestern University.

Hirshleifer, D. 1988. Residual Risk, Trading Costs, and Commodity Futures Risk Premia. *Review of Financial Studies* 1:173–93.

Ho, T., and H. R. Stoll. 1981. Optimal Dealer Pricing under Transactions and Return Uncertainty. *Journal of Financial Economics* 9:47–73.

Hong, H., and J. C. Stein. 2003. Difference of Opinion, Short-Sales Constraints, and Market Crashes. *Review of Financial Studies* 16:487–525.

Huang, J., and J. Wang. Forthcoming. Market Liquidity, Asset Prices, and Welfare. *Journal of Financial Economics*.

Kieschnick, R., D. O. Cook, and O. Arugaslan. 2007. Is Liquidity a Priced Risk Factor? Working Paper, University of Texas at Dallas.

Kimball, M. S. 1993. Standard Risk Aversion. *Econometrica* 61:589–611.

Kyle, A. S. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53:1315–36.

Leland, H., and M. Rubinstein. 1988. Comments on the Market Crash: Six Months After. *Journal of Economic Perspectives* 2:45–50.

Lo, A., H. Mamaysky, and J. Wang. 2004. Asset Prices and Trading Volume under Fixed Transactions Costs. *Journal of Political Economy* 112:1054–90.

Merton, R. C. 1987. A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance* 42:483–510.

Mitchell, M., L. H. Pedersen, and T. Pulvino. 2007. Slow Moving Capital. *American Economic Review* 97:215–20.

Netter, J., and M. Mitchell. 1989. Stock Repurchase Announcements and Insider Transactions after the October 1987 Stock Market Crash. *Financial Management* 18:84–96.

Pagano, M. 1989. Endogenous Market Thinness and Stock Price Volatility. *Review of Economic Studies* 56:269–88.

Pastor, L., and R. F. Stambaugh. 2003. Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 111:642–85.

Romer, D. 1993. Rational Asset-Price Movements Without News. *American Economic Review* 83:1112–30.

Schwert. G. W. 1989. Why Does Stock Volatility Change Over Time? *Journal of Finance* 44:1115–54.

Shleifer, A. 1986. Do Demand Curves for Stocks Slope Down? *Journal of Finance* 41:579–90.

Spiegel, M., and A. Subrahmanyam. 1995. On Intraday Risk Premia. *Journal of Finance* 50:319–39.

Spiegel, M., and X. Wang. 2007. Cross-Sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk. Working Paper, Yale University and Pennsylvania State University.

Tremont, 2006. *Tremont Asset Flows Report*. Tremont Capital Management.

Vissing-Jorgensen, A. 2002. Limited Asset Market Participation and the Elasticity of Intertemporal Substitution. *Journal of Political Economy* 110:825–53.

Wang, J. 1994. A Model of Competitive Stock Trading Volume. *Journal of Political Economy* 102:127–68.

Yuan, K. 2005. Asymmetric Price Movements and Borrowing Constraints: A REE Model of Crisis, Contagion, and Confusion. *Journal of Finance* 60:379–411.