When Market Illiquidity Generates Volume

Serge Darolles* Gaëlle Le Fol† Gulten Mero‡

January 21, 2011

Abstract

We develop a model of the daily return-volume relationship which incorporates information and liquidity shocks. First, we distinguish between two trading strategies, information-based and liquidity-based trading and suggest that their respective impacts on returns and volume should be modeled differently. Second, we integrate the microstructure setting of Grossman and Miller (1988) with the information flow perspective of Tauchen and Pitts (1983) and derive a modified MDH model with two latent factors related to information and liquidity. Our model explains how the liquidity frictions can increase the daily traded volume, in the presence of liquidity arbitragers. Finally, we propose a stock-specific liquidity measure using daily return and volume observations of FTSE100 stocks.

JEL classification: C51, C52, G12

Key words: Volatility-volume relationship, mixture of distribution hypothesis, liquidity shocks, information-based trading, liquidity arbitrage, GMM tests.

*LYXOR AM and CREST-INSEE, France, serge.darolles@ensae.fr, Tel. +33 (0)1 58 98 13 29.

†University of Paris-Dauphine, DRM Finance (CEREG) and CREST-INSEE, France, gaelle.lefol@ensae.fr, Tel. +33 (0)1 44 05 46 03.

‡University of Rennes 1, CREST-INSEE and CREM-CNRS, France, gulten.mero@ensae.fr, Tel. +33 (0)1 41 17 77 21.

1 Introduction

In this article, we develop a model of the daily return-volume relationship which takes into account both information and liquidity shocks. To do so, we reconcile the information flow perspective of the mixture of distribution hypothesis (MDH) with the microstructure setting of Grossman and Miller (1988) which captures market liquidity. We develop a modified MDH model with two latent factors related to information and liquidity shocks.

Several empirical studies [see Ying (1966), Crouch (1970), Clark (1973), Copeland (1976), Copeland (1977), Epps and Epps (1976), Westerfield (1977), Rogalski (1978), Tauchen and Pitts (1983), Harris (1982), Harris (1986) and Harris (1987)] of both futures and equity markets find a positive association between price variability[1] and the contemporaneous trading volume[2] at the daily frequency. The usual theoretical explanation of this positive volume-return volatility relation comes from microstructure models which analyze how information is disseminated into prices, and how market prices convey information. Thus, several models predict a positive return volatility-volume relation that depends on the rate of information flow and the interaction between specialists, informed and liquidity traders [Kyle (1985), Glosten and Milgrom (1985), Easley and O'Hara (1987), Diamond and Verrechia (1987), Admati and Pfleiderer (1988), Foster and Viswanathan (1990), Foster and Viswanathan (1993) and Easley et al. (1996)], the market size [Gallant et al. (1992)] or the existence of a short sales constraint [Diamond and Verrechia (1987)].

The mixture of distribution hypothesis (MDH) models attempt to explore the microstructure framework in which information asymmetries and liquidity needs motivate trade in response to information arrivals. The MDH, pioneered by Clark (1973) and extended by Harris (1982), Tauchen and Pitts (1983) and Andersen (1996) among others, provides an explanation of the positive correlation between volume and the squared value of price change at a daily frequency. For example, Clark (1973) model assumes that events important to the

[1]As measured by either the square price change or the price change per se.
[2]See Karpoff (1987) for a detailed review of the literature.

pricing of a security occur at a random rate through time. It appears that price data are generated by a conditional normal stochastic process with a changing variance parameter that can be proxied by volume whose distribution is assumed to be lognormal. Clark (1973) shows that the lognormal-normal mixture outperforms several members of stable family. Using the same assumption, Harris (1982), Harris (1986), Harris (1987) and Tauchen and Pitts (1983) show that the joint distribution of daily price changes and volume can also be modeled by a mixture of bivariate normal distributions. They assume that both variables (the daily price change and daily volume) are conditioned by the rate of information which is random and serially uncorrelated. Assuming a lognormal distribution for the mixing variable, the model can be estimated by maximum likelihood [see Tauchen and Pitts (1983) for further discussion]. As pointed out by Harris (1982), Harris (1986) and Harris (1987), the MDH can explain the fat tailed probability distribution of the daily price change, and the positive correlation between return volatility and volume. The standard MDH models assume that information inflow drives the positive volatility-volume relationship.

If earlier tests find evidence supportive of the MDH model [Clark (1973), Epps and Epps (1976), Tauchen and Pitts (1983), Harris (1982), Harris (1986) and Harris (1987)], later studies are less favorable [Heimstra and Jones (1994), Lamoureux and Lastrapes (1994), Richardson and Smith (1994), and Andersen (1996)]. Different authors propose various extensions of the standard MDH model in order to improve its explicative power. Lamoureux and Lastrapes (1994) extension assumes that the information-arrival rate is serially correlated[3]. Andersen (1996) develops a modified MDH model that includes a conditional Poisson distribution for the trading process and a volume component that is not information sensitive. His tests suggest that the modified version significantly outperforms the standard MDH, which assumes that both returns and volume are normally distributed.

[3]However their model fails to explain GARCH persistence in return variance. Their finding is consistent with the results of Richardson and Smith (1994), who used the generalized method of moments (GMM) to test the mixture model but did not account for time dependencies in the data. Thus, the evidence against the model isolates the inability of the model to jointly accommodate the dynamic properties of squared returns and volume.

Previous MDH tests are performed under the assumption that markets are perfectly liquid and the impact of liquidity frictions on the volatility-volume relation is disregarded. However, several studies show that liquidity shocks are priced by the market [see Amihud (2002), and Acharya and Pedersen (2005) among others] and that they impact both returns and traded volume [see Chordia et al. (2001), Chordia et al. (2000), and Darolles and Fol (2005)]. In particular, as discussed by Darolles and Fol (2005), some large investors, such as Hedge Funds, play the role of liquidity arbitragers by tracking price pressures due to liquidity frictions and entering the market in order to provide immediacy and to cash the liquidity premium. Their intervention tends to correct price imperfections due to liquidity shocks and thus lowers the intra-day return volatility. Once the prices are back to their fully revealing information level, the arbitrage traders will liquidate their positions in order to benefit from the price reversals. As a consequence, the volume they trade adds to the volume that would prevail in the absence of liquidity frictions.

It follows that the observed daily traded volume of a particular stock is the result of both information-based trading and liquidity arbitragers. Thus, understanding and decomposing the traded volume can give some insights concerning the market liquidity. In particular, the raw traded volume is commonly used in the literature as a proxy for liquidity risk or market quality [Gallant et al. (1992), Domowitz and Wang (1994), Gourieroux and Fol (1998)]. However, more recent studies are less favorable to the idea that the raw traded volume is an efficient measure of liquidity. For example, Borgy et al. (2010) point out that price-impact based indicators are more accurate than raw traded volume in order to identify liquidity problems in the currency exchange (FX) market. In this paper, we suggest that total daily volume can be misleading since, in the presence of liquidity arbitragers, liquidity frictions can be associated with higher volume.

These observations motivate us to extend the standard MDH model framework by incorporating liquidity effects on daily stock returns and traded volume. To do so, we focus on the theoretical framework of Grossman and Miller (1988) who develop a market microstructure model that captures the essence of market liquidity. They consider two types of market participants. The first one trades in response to information shocks and can be assimilated to the active traders of Tauchen and Pitts (1983). The second type of traders enters the market to exploit the presence of the liquidity events and will be called *liquidity arbitragers*. A liquidity event is represented by a temporary order imbalance due to trade asynchronization among the active traders. In the presence of a liquidity event, trades occur at two dates. At time 1, the liquidity arbitragers observe price imperfections due to the order imbalance among the active traders and enter the market to provide immediacy. At time 2, they liquidate their positions as other active traders arrive to the market with opposite order imbalances[4]. The Grossman and Miller (1988) model implies that the volume traded by liquidity arbitragers at date 2 increases the aggregated traded volume.

Using the implications of the Grossman and Miller (1988) model at an aggregated level across times 1 and 2, we include an additional latent mixing variable L in the model of Tauchen and Pitts (1983) to take into account the liquidity shocks which are supposed to arrive randomly within the trading day. Our modified MDH model with two latent variables – called the MDHL model – permits us to decompose the trading volume into two components driven respectively by information and liquidity. Moreover, following Richardson and Smith (1994), we propose a direct test of the modified MDH model. Indeed, the model imposes restrictions on the joint moments of price changes and volume as a function of only a few parameters. It is then possible to form overidentifying restrictions. These restrictions can be tested using the generalized method of moments (GMM) procedure of Hansen (1982). Based on FTSE100 stock daily return and volume time series ranging from January 2005 to July 2007, we show that the MDHL model with two latent factors outperforms the standard MDH, suggesting that the inclusion of a latent liquidity variable may reconcile previous divergent results found in the literature.

[4]Note that, Grossman and Miller (1988) assume that if all the active traders were present at time 1, there would be no order imbalance and no benefit for liquidity arbitragers to enter the market.

The contribution[5] of this paper is threefold. First, it distinguishes between two trading strategies, information-based trading and liquidity arbitrage, and suggests that their respective impacts on returns and traded volume should be modeled differently. The former is incorporated into the daily price changes and traded volume and drives the positive volatility-volume relationship. The latter impacts the intraday price variations and volumes but does not affect the daily price changes, while increasing the daily traded volume. Although previous literature distinguishes between active traders and liquidity providers [see for example Grossman and Miller (1988)], we are the first to use the arbitrage trading impact on individual stock returns and volume in order to decompose the total traded volume into two components due to information and liquidity shocks. To do so, we blend the Grossman and Miller (1988) microstructure framework into the Tauchen and Pitts (1983) standard MDH and develop a two-latent factor model accounting for information inflow and liquidity frictions. Second, we use a structural model, the MDHL model herein proposed, to exploit the volume-volatility relation in order to extract a stock-specific liquidity measure μ_v^{la} using daily data. Finally, the MDHL model confirms previous studies by implying a positive volatility-volume relation driven by the common dependence of the observables on the information flow. However, in our framework, this positive correlation does not depend on the total traded volume but is function of the volume component due to information-based trading μ_v^{at} after controlling for the impact of liquidity shocks μ_v^{la}. The standard MDH model appears to be a special case of the MDHL model in the absence of liquidity frictions.

The paper is organized as follows. In Section 2, we briefly present the standard MDH model based on the Tauchen and Pitts (1983) framework. In Section 3, we develop our model. We first summarize the Grossman and Miller (1988) microstructure framework, and then discuss its implications concerning aggregated data. Finally, we develop a modified MDH model accounting for both information and liquidity shocks. In Section 4, we present the GMM tests and discuss the empirical results. Section 5 concludes the paper.

[5]Table 7 given in Appendix G compares our paper's contributions to those of previous results in the literature.

2 The standard MDH model (Tauchen and Pitts (1983))

This section provides a brief summary of the standard MDH model based on the theoretical framework of Tauchen and Pitts (1983), henceforth TP[6]. The model considers a simple economy with only one risky asset and J active traders. J is fixed over time. Each trading day, the market experiences a series of different Walrasian within-day equilibria; the information inflow triggers market progression from one equilibrium to the next[7]. No assumptions are made concerning liquidity problems since, in the TP economy, assets are deemed perfectly liquid.

The authors first assume that the number of within-day equilibria I_t is random since the number of new pieces of information hitting the market varies significantly across the trading days. Using, in addition, a variance-component model for the trader's reservation price increments, TP demonstrate that the intraday price change and the traded volume, denoted respectively by ΔP_i and V_i $(i = 1, ..., I_t)$ are normally distributed:

$$\Delta P_i \sim N(0, \sigma_p^2), \quad V_i \sim N(\mu_v, \sigma_v^2), \tag{2.1}$$

where price increment variance σ_p^2 as well as volume mean and variance parameters denoted respectively by μ_v and σ_v^2 are given in Appendix A.

In order to illustrate the TP model's world mechanism, we consider a simple example given in Figure 1. Let I_t be the number of intra-day equilibria of the t-th trading day and P_{t-1} be the closing price of the previous trading day. To show how intra-day price varies in response to the inflow of new information, we assume that only three pieces of information arrive in the course of day t, I_1, I_2 and I_3. Should I_1 be perceived as good news, the trader's expected value for the risky asset will increase resulting in a new equilibrium price $P_1 > P_{t-1}$; in this case the price increment due to the arrival of I_1, ΔP_1, is positive. I_2 being seen as

[6]A more detailed presentation of the TP model is provided in Appendix A.

[7]According to TP, "the intervals between successive equilibria are not necessarily of the same length; since buy/sell orders are executed sequentially, many actual transactions at the exchange can comprise what we think of as a single market clearing or transaction".

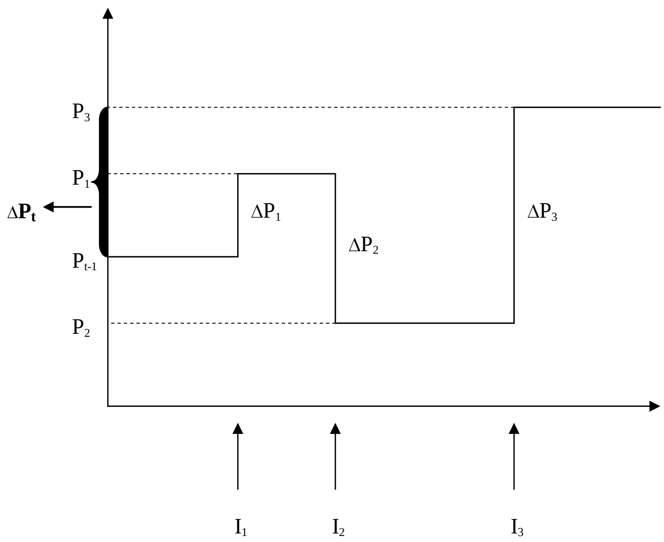

Figure 1: Day t price change as a function of intra-day price variations due to information shocks.

bad news, the next price increment ΔP_2 is negative. Lastly, I_3, which turns out to be good news, initiates the movement to the third intra-day equilibrium and ΔP_3 is positive. At the end of day t, we observe the daily price increment $\Delta P_t = P_3 - P_{t-1}$. The daily price change is the sum of intra-day price increments due to the arrival of the new information. More generally, summing the within-day price changes and trading volumes, we obtain the day-t price change ΔP_t and traded volume V_t:

$$\Delta P_t = \sum_{i=1}^{I_t} \Delta P_i, \qquad V_t = \sum_{i=1}^{I_t} V_i. \tag{2.2}$$

Both ΔP_t and V_t appears to be mixtures of independent normals with the same mixing variable I_t. Conditional on I_t, the daily price change ΔP_t is $N(0, \sigma_p^2 I_t)$ and the daily volume is $N(\mu_v I_t, \sigma_v^2 I_t)$, which yields the bivariate normal mixture:

$$\Delta P_t = \sigma_p \sqrt{I_t} Z_{1t}, \tag{2.3}$$

$$V_t = \mu_v I_t + \sigma_v \sqrt{I_t} Z_{2t}, \tag{2.4}$$

where Z_{1t} and Z_{2t} are i.i.d. standard normal variables and mutually independent. At the end of the day t, all the incoming information is incorporated into the price change ΔP_t and traded volume V_t. From (2.3)-(2.4), it follows that the contemporaneous relation between ΔP_t^2 and V_t is:

$$Cov(\Delta P_t^2, V_t) \quad = \quad \sigma_p^2 \mu_v Var[I_t] > 0. \tag{2.5}$$

Following TP, volume and prices can only change through the information arrival process. The TP framework is appealing as it defines an interesting factorial structure that we aim at extending to incorporate a liquidity shock arrival process.

3 Our theoretical framework

Based on the theoretical analysis of Grossman and Miller (1988), henceforth GM, we modify the standard MDH model by incorporating the effect of market liquidity on volatility-volume relationship. In our framework, we use "liquidity arbitragers" to refer to a particular family of liquidity providers who adopt a strategic behavior in order to take advantage of price distortions due to liquidity frictions. The aim of liquidity arbitragers is to cash the liquidity premium by offsetting their positions once prices revealing the information are established. In practice, many financial actors, such as proprietary trading desks or Hedge Funds, may play the role of liquidity arbitragers. In Subsection 1, we discuss the implications of the GM framework concerning total price changes and traded volumes, i.e. price changes and volumes related to information and liquidity shocks. In Subsection 2, we develop our modified MDH model accounting for both information and liquidity shocks.

3.1 Modeling the impact of liquidity shocks on total returns and volume

This paragraph adapts the GM model[8] to our economy and discusses how it can be extended to model the impact of intraday liquidity shocks on price changes and traded volumes. GM focus on a market in which liquidity is modeled as being determined by the demand and supply of immediacy. They consider two types of traders: the outside customers who trade in response to information inflow, and the market makers who trade in response to liquidity shocks. In our framework, the outside customers are called active traders as in TP and the market makers of GM correspond to our liquidity arbitragers.

The GM model focuses on a single risky asset and considers only three dates (1, 2, and 3). Dates 1 and 2 are trading dates, while date 3 is introduced only as a terminal condition; the liquidation value of the risky asset at date 3 is denoted by $\tilde{P}_3$. Information concerning $\tilde{P}_3$ is assumed to arrive before trading at period 1 and before trading at period 2. Let J be the number of all the potential active traders in the market. The active trader j ($j = 1, ..., J$) at time 1 has an endowment of size z_j in the security, which is inappropriate given the trade-off between his risk preferences and information at that date. At period 1, some liquidity frictions may arise because of asynchronization of order flows. This will result in a temporary order imbalance; if all the active participants were present in the market at date 1, the order imbalance would vanish and the net trading demand would be zero at the current price.

Generally speaking, it is important to distinguish between:

(i) the aggregated endowment shock across active participants and across periods 1 and 2, by definition equal to zero. If all the active traders were simultaneously present in the market at date 1, there would be no liquidity event and the equilibrium price would reveal all the available information about the future liquidation value of the asset;

[8]The GM model is presented in details in Appendix B. Here, we only report some important results helping to understand trader motivations, as well as the implications of the GM model concerning total price changes and total traded volumes, i.e., price changes and traded volumes related to both information and liquidity shocks.

(ii) the aggregated endowment shock across active traders willing to trade at date 1:

$$z = \sum_{j=1}^{J_1} z_j \neq 0, \tag{3.1}$$

where $J_1 < J$ is the number of active traders being present in the market at date 1. In this case, z represents a temporary order imbalance caused by trade asynchronization.

Liquidity arbitragers, who continuously observe the market, provide immediacy at date 1 by taking trading positions that they hold until date 2. At date 2, they liquidate their positions as other active traders arrive with the opposite order imbalance.

Assuming exponential preferences for both types of traders, GM use backward induction to obtain the optimal excess demands for active traders as well as liquidity arbitragers at both dates. Then, given the market clearing conditions, the equilibrium price at period 1 denoted by P_1 is:

$$P_1 = E_1\tilde{P}_3 - \frac{z\alpha Var_1(E_2\tilde{P}_3)}{1+M}, \tag{3.2}$$

where α represents trader preferences which are assumed to be identical for all market participants, M is the number of liquidity arbitragers and $Var_1(E_2\tilde{P}_3)$ represents the risk from the point of view of period 1 that $P_2 = E_2\tilde{P}_3$ is not known. From equation (3.2), the equilibrium price at date 1 will deviate from the price revealing the information $E_1\tilde{P}_3$ and the equilibrium aggregate excess demand for active traders at date 1 is:

$$Q_1^{at} = -\frac{M}{1+M}z. \tag{3.3}$$

In the same way, the equilibrium excess demand per liquidity arbitrager at date 1 is:

$$Q_1^{la} = \frac{z}{1+M}, \tag{3.4}$$

Since the GM world assumes that liquidity arbitragers face a participation cost $c > 0$, their number M will be finite, which implies a limited capacity in providing immediacy and a deviation of P_1 from its fully revealing information level $E_1\tilde{P}_3$. As discussed by Brunnermeier and Pedersen (2009), funding liquidity constraints can also explain why the liquidity is not fully provided.

Generally speaking, the GM framework focuses on the consequences of an order imbalance on the intraday patterns of price change and transaction volume. At this stage, the model shows that in the presence of liquidity frictions and exogenous transaction costs:
(i) the traded volume at date 1 is lower than it would have been if there were no order imbalance $|Q_1^{at}| < |z|$.
(ii) the transaction price at date 1 deviates from its revealing information level ($P_1 \neq E_1\tilde{P}_3$).

However, from the assumptions that the order imbalance sums to zero across periods 1 and 2, and that the liquidity arbitragers offset their positions at date 2, it follows that the traded volume across dates 1 and 2 is higher than it would have been in the absence of liquidity frictions if the condition $M \geq 1$ is satisfied[9].

A. Order imbalances and the price change

As discussed by TP, the trading day can be considered as a set of successive equilibria and the movement from one equilibrium to the next is driven by the arrival of new information. Let us consider a trading day consisting of only 2 information arrivals[10]. Let δ_i be an indicator variable such as $\delta_i = 1$ ($i = 1, 2$) in the presence of order imbalances and $\delta_i = 0$ ($i = 1, 2$) otherwise. Here, we take $\delta_1 = 1$ and $\delta_2 = 0$, hence the trading day reduces to a 3-date process in the sense of GM. In other words, we assume that trade asynchronization occurring just after the arrival of the first piece of information ($\delta_1 = 1$) results in a 3-date

[9] In other words, the order imbalance faced by active traders who exchange at date 1 is offset thanks to immediacy provided by liquidity arbitragers who will liquidate their positions at date 2 and thus increase the traded volume.

[10] In the next subsection we will generalize this simple example by allowing for multiple information arrivals within a given trading day.

GM-process with the second piece of information arriving before trading at date 2 and $\tilde{P}_3$ being the liquidation value of the asset at the end of the trading day.

Note that, z_1 can be expressed as a function of δ_1:

$$z_1 = \sum_{j=1}^{J_1} z_{1j} + (1 - \delta_1) \sum_{j=J_1+1}^{J} z_{1j} = -\delta_1 \sum_{j=J_1+1}^{J} z_{1j}, \tag{3.5}$$

which implies that: (i) for $\delta_1 = 0$, $z_1 = \sum_{j=1}^{J} z_{1j}$ equals zero by definition, and (ii) for $\delta_1 = 1$ equation (3.5) is equivalent to equation (3.1).

Let $P_0 = E_0\tilde{P}_3$ be the price prevailing at the beginning of the trading day and $E_0\tilde{P}_3 = E_0E_1\tilde{P}_3$ be the expectation concerning $\tilde{P}_3$ before the arrival of new information to the market. From equation (3.2), the total price change at date 1 ($i = 1$), $\Delta P_1 = P_1 - P_0$, is:

$$\Delta P_1 = (E_1\tilde{P}_3 - E_0\tilde{P}_3) - \frac{z_1 \alpha Var_1(E_2\tilde{P}_3)}{1+M} = \Delta P_1^{'} + \Delta P_1^{''}(z_1), \tag{3.6}$$

where z_1 is the order imbalance occurring at date 1, $\Delta P_1^{'} = E_1\tilde{P}_3 - E_0\tilde{P}_3$ is the price change due to information hitting the market at date 1 and $\Delta P_1^{''}(z_1) = -\frac{z_1 \alpha Var_1(E_2\tilde{P}_3)}{1+M}$ is the price change due to order imbalance at date 1. In the same way, the total price change at date 2, $\Delta P_2 = P_2 - P_1$, can be written as:

$$\Delta P_2 = (E_2\tilde{P}_3 - E_1\tilde{P}_3) + \frac{z_1 \alpha Var_1(E_2\tilde{P}_3)}{1+M} = \Delta P_2^{'} - \Delta P_1^{''}(z_1), \tag{3.7}$$

where $\Delta P_2^{'} = E_2\tilde{P}_3 - E_1\tilde{P}_3$ represents the price change due to information arrival at date 2 and $-\Delta P_1^{''}(z_1)$ represents the price adjustment as new active traders arrive at date 2 with opposite order imbalance.

From (3.6)-(3.7), the total price change across periods 1 and 2 is equal to:

$$\Delta P_1^{'} + \Delta P_2^{'} = E_2\tilde{P}_3 - E_0\tilde{P}_0, \tag{3.8}$$

i.e. the impact of the order imbalance on the total price change vanishes; price variation due to liquidity shocks and price adjustments offset each other and the aggregated price change is only due to information flow.

Figure 2 illustrates how intraday price increments behave in response to both information flow and liquidity shocks in the simple example considered here. Suppose that the two successive pieces of information reaching the market, denoted respectively by I_1 and I_2, are perceived as good news. The intraday price behavior in the absence of liquidity shocks is visually described by the dashed lines and corresponds exactly to Figure 1. Trade asynchronization occurring just after the arrival of I_1, results in a 3-date GM-process with I_2 arriving before trading at date 2 and date 3 being a terminal condition; the liquidation value of the risky asset is $\tilde{P}_3$.

Let $\Delta P_1^{'} = E_1\tilde{P}_3 - P_0$ be the price increment due to I_1, and $\Delta P_1^{''}(z_1)$ be the price variation due to the liquidity friction at date 1. As for I_1, the active trader expectations concerning $\tilde{P}_3$ will rise, resulting in a positive $\Delta P_1^{'}$. The active traders face sell-side liquidity shortage due to trade asynchronization at date 1 and the asset price increases more than if there were no liquidity problems, resulting in a positive $\Delta P_1^{''}(z_1)$.

In particular, the liquidity arbitragers observing the exchange enter the market to provide the missing liquidity. At date-1-equilibrium, they sell the stock at $P_1 = P_0 + \Delta P_1^{'} + \Delta P_1^{''}(z_1)$, where $\delta_1 = 1$. At date 2, the liquidity arbitragers enter the market to buy the stock at $P_2 = P_2^{'} = E_2\tilde{P}_3$ (the price revealing the information at date 2), as new active traders arrive with the opposite order imbalance. The date-2-equilibrium price can be written: $P_2 = P_0 + \Delta P_1^{'} + \Delta P_1^{''}(z_1) + \Delta P_2^{'} - \Delta P_1^{''}(z_1)$, where $-\Delta P_1^{''}(z_1)$ is the price adjustment as new active traders arrive at date 2 to offset the order imbalance. Since price distortion due to liquidity event at date 1 and price variation due to liquidity adjustment cancel out, the price returns to its fully revealing information level which corresponds also to the price that would prevail in the absence of trade asynchronization at date 1 (as shown by the dashed lines). It follows that the intraday price distortion due to liquidity shocks does not impact

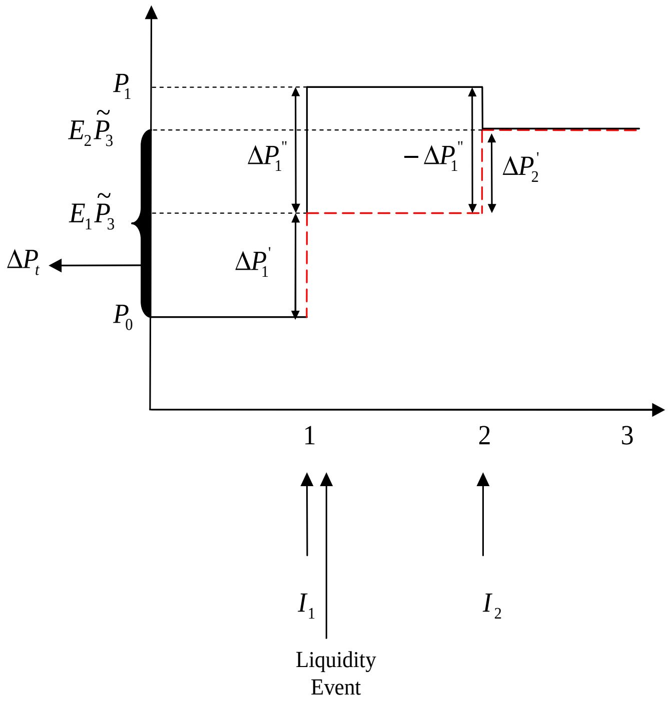

Figure 2: Day t price change as a function of intra-day price fluctuations due to information and liquidity shocks.

the total price change and equation (3.8) is always satisfied.

B. Order imbalances and the traded volume

We focus on a simple trading day with two pieces of information and a unique order imbalance occurring after the first information arrival and discuss the impact of liquidity frictions on the intraday traded volume. Let V_1 $(i = 1)$ be the total traded volume due to the first information arrival and the liquidity friction; it is the sum of the volume V_1' due to information flow and the volume V_1'' due to the intervention of liquidity arbitragers. As discussed by GM, when a new piece of information hits the market, active traders, who revise their expectations concerning the future liquidation value of the asset, are willing to rebalance their positions in order to share risk through the market. Let z_{1j} be the quantity that trader j $(j = 1, ..., J)$ is willing to trade in response to the first information arrival of the trading day. We obtain:

$$V_1^{'} = \frac{1}{2}\sum_{j=1}^{J} \mid z_{1j} \mid . \tag{3.9}$$

$V_1^{'}$ corresponds to the traded volume due to the first piece of information ($i = 1$) in the TP model. If all the active traders arrive simultaneously in the market at date 1, $V_1^{'}$ represents the total traded volume due to the first piece of information and equals V_1.

Any liquidity event occurring at date 1, creates a temporary order imbalance. Since the order imbalance sums to zero across periods 1 and 2, the liquidity arbitragers offset their positions at date 2 as other active traders arrive with the opposite order imbalance. It follows that the total traded volume V_1 is higher than $V_1^{'}$. The difference $V_1^{''}$ is the amount of immediacy provided by liquidity arbitragers at equilibrium at date 1, MQ_1^{la} ($MQ_1^{la} = -Q_1^{at}$). From equation (3.4) or (3.3), $V_1^{''}$ is given by:

$$V_1^{''} = MQ_1^{la} = Q_1^{at} = \frac{M}{1+M} \mid z_1 \mid . \tag{3.10}$$

Generally speaking, the total traded volume V_1 due to the first information arrival can be written as:

$$V_1 = V_1^{'} + V_1^{''}(z_1), \tag{3.11}$$

where the $V_1^{''}(z_1)$ is used to denote the dependence of $V_1^{''}$ on z_1 as expressed in equation (3.10). It follows that:

(i) In the absence of liquidity frictions ($\delta_1 = 0$) $V_1^{''} = 0$ and we obtain the TP model which states that the total traded volume is completely explained by information inflow: $V_1 = V_1^{'}$.

(ii) The occurrence of liquidity events ($\delta_1 = 1$) increases the total traded volume related to the first piece of information: $V_1 > V_1^{'}$. This is due to the intervention of the liquidity arbitragers in the market.

In conclusion, the liquidity shocks can increase the total traded volume[11] but have no impact on the total price change. This result motivates us to blend the GM model's implications into the standard framework of TP in order to model the impact of liquidity events on the daily price change and traded volume and thus separate the observed traded volume into two parts due to information and liquidity.

3.2 A modified MDH model with information and liquidity shocks

This section develops the modified MDH model which takes into account both information and liquidity shocks. Indeed, as discussed above, some liquidity frictions may arise from trade asynchronization, even if the number of active traders J is large. The equilibrium price then differs significantly from the price revealing the information, which motivates the liquidity arbitragers to enter the market, provide immediacy and cash the liquidity premium.

We focus on a simple economy with a risk-free asset and a single risky security having a liquidation value $\tilde{P}_T$ at the end of the trading day. The risk-free rate is normalized to zero. To generalize the simple example of the previous subsection (based on the GM analysis) at a daily frequency by allowing for multiple information arrivals within the trading day, we consider each 3-date-process as a 2-trading-date (or 2-equilibria) process, henceforth GM process, and report the terminal condition at the end of the trading day. There are only two kinds of traders in the market: the active traders who trade in response to new information, and the liquidity arbitragers who trade in response to liquidity frictions. The number of each category of traders − respectively J and M − is nonrandom and fixed over time. We then assume that, within the day, the market passes through a sequence of distinct equilibria in the sense of TP. The movement from one equilibrium to the next is initiated by the arrival of new information to the market. Given the new information, the active traders decide to rebalance their positions in order to share risk through the market. Let z_{ij} be the endowment

[11] At this stage of the analysis, for simplicity purpose, we make abstraction of the impact of the second piece of information on the traded volume. Since in this simple example $\delta_2 = 0$, the volume due to the second information arrival corresponds to that of TP for $i = 2$; This amount will be added to V_1 when considering the total traded volume across periods 1 and 2.

shock of trader j ($j = 1, ..., J$) given the ith piece of information ($i = 1, ..., I_t$). As in TP, I_t is assumed to be random since the number of pieces of information reaching the market each day varies significantly. If all the active traders are present in the market, the aggregated endowment shock across traders is zero and the ith equilibrium price equals its fully revealing information level.

However, if a liquidity event occurs, the aggregated endowment shock across the active traders being present in the market ($J_1 < J$) represents the order imbalance: $z_i = \sum_{j=1}^{J_1} z_{ij} \neq 0$. Liquidity arbitragers who observe this market imperfection enter the market in order to provide immediacy and the trade is generated from a GM process. Date 1 of the GM-process coincides with the ith piece of information and the equilibrium price at this date deviates from the price revealing the information[12]. In order to denote the appartenance to the ith within-day equilibrium, we index by i all the intraday variables of interest, such as price changes, excess demands of traders, as well as traded volumes. In the previous subsection, we introduced an indicator variable, δ_i, such as $\delta_i = 1$ in the presence of liquidity frictions and $\delta_i = 0$ otherwise. Then, equation (3.5) can be generalized as follows:

$$z_i = \sum_{j=1}^{J_1} z_{ij} + (1 - \delta_i) \sum_{j=J_1+1}^{J} z_{ij} = -\delta_i \sum_{j=J_1+1}^{J} z_{ij}, \quad i = 1, ..., I_t. \tag{3.12}$$

In addition, we assume that $z_i \sim N(0, \sigma_z^2)$ when a liquidity event occurs ($\delta_i = 1$) and $z_i = 0$ otherwise ($\delta_i = 0$).

Let us consider a GM-process debuting at the ith intraday equilibrium and comprising of two successive information arrivals: the ith and the $(i + 1)$th pieces of information which

[12]If there were some noise (non-informed) traders at date 1 who trade in response to liquidity needs, it would be possible for the arbitrage participants to liquidate their positions before the arrival of the next piece of information by trading with the noise traders at date 1. Trade between strategic and non-strategic traders would take place at a disadvantageous price for the noise traders who would bear, in that case, the liquidity premium perceived by the liquidity arbitragers. Since we do not allow for the presence of noise traders in our model, the liquidity arbitragers have to wait from period 1 to period 2 to trade as new active traders arrive with the opposite order imbalance. For this reason, the arbitragers face the risk that a new piece of information arrives at date 2 causing the date-2-equilibrium price to change towards a disadvantageous direction for them.

arrive respectively before trading at date 1 and before trading at date 2. We can now generalize (3.6) and (3.7) by getting the corresponding price changes at each date ΔP_i (date 1) and ΔP_{i+1} (date 2):

$$\begin{aligned} \Delta P_i &= \Delta P_i^{'} + \Delta P_i^{''}(z_i), & (3.13) \\ \Delta P_{i+1} &= \Delta P_{i+1}^{'} - \Delta P_i^{''}(z_i). & (3.14) \end{aligned}$$

In these equations, $\Delta P_i^{'} = E_i\tilde{P}_T - E_{i-1}\tilde{P}_T$ and $\Delta P_{i+1}^{'} = E_{i+1}\tilde{P}_T - E_i\tilde{P}_T$ represent price changes due to information inflow, $\Delta P_i^{''}(z_i) = -\frac{z_i \alpha Var_i(E_{(i+1)}\tilde{P}_T)}{1+M}$ represents price distortion due to the liquidity event occurring at date 1 and coinciding with the ith information arrival, while $-\Delta P_i^{''}(z_i)$ represents the price adjustment as other active traders arrive at date 2 with the opposite order imbalance. The number of GM-processes within a trading day t, denoted by L_t, corresponds to the number of liquidity events and is given by $L_t = \sum_{i=1}^{I_t} \delta_i$. L_t and I_t are assumed to be conditionally independent $Cov(L_t|I_t, I_t) = 0$. Moreover, we suppose that a liquidity event may occur at any intraday equilibrium of the trading day t except the last one, which yields $\delta_{I_t} = 0$. This assumption is necessary for generalizing the GM 3-period world at a daily frequency; for instance, for $I_t = 2$, i.e. 2 pieces of information and only one liquidity event, we obtain the GM model as a particular case of our modified model.

Let V_i be the cumulated traded volume across periods 1 and 2 due to the ith piece of information and the liquidity event occurring at the ith equilibrium:

$$V_i = V_i^{'} + V_i^{''}(z_i), \qquad (3.15)$$

where $V_i^{'} = \frac{1}{2}\sum_{j=1}^{J} \mid z_{ij} \mid$ is the traded volume due to the ith information arrival to the market [see equation (3.9)], and $V_i^{''}(z_i) =\mid MQ_i^{la}(z_i) \mid = \frac{M}{1+M} \mid -\delta_i \sum_{j=J_1+1}^{J} z_{ij} \mid$ is the traded volume due to the intervention of the liquidity arbitragers, as measured by the amount of active traders that is completed by liquidity arbitragers at date 1 [see equation (3.10)]. If all the active traders were present in the market after the ith information arrival ($\delta_i = 0$),

$V_i = V_i^{'}$ would correspond to the traded volume implied by the standard MDH of TP at the ith equilibrium of the trading day.

Generally speaking, ex-post, the GM-process includes the TP model as a particular case in the absence of the liquidity events. The rapprochement can be done in two ways:
(i) a GM-process with new information hitting the market at date 2, corresponds to two successive TP equilibria in the absence of liquidity frictions[13].
(ii) in the absence of new information at date 2 and with no liquidity frictions, the GM-process would be assimilated to a unique TP (ith) equilibrium.

Let now consider the traded volume V_i resulting from the ith information arrival and the liquidity event occurring at the ith equilibrium, as given in equation (3.15). As discussed above, the volume component $V_i^{'}$ due to information is the same as that considered by TP. As shown in Appendix A, TP demonstrate that the total traded volume V_i^{TP} is due to the ith piece of information and is given by:

$$V_i^{TP} = \frac{\alpha}{2} \sum_{j=1}^{J} \mid \psi_{ij} - \bar{\psi}_i \mid, \tag{3.16}$$

where α is a constant, ψ_{ij} is drawn from a normal distribution with mean zero and variance σ_ψ^2 and $\bar{\psi}_i = \frac{1}{J} \sum_{j=1}^{J} \psi_{ij}$. Since in the TP world the market is deemed perfectly liquid, V_i^{TP} corresponds to $V_i^{'}$ of our model. TP show that, for large J, V_i^{TP} is approximately normally distributed with first two moments:

$$\mu_v^{at} \equiv E[V_i^{'}] = \left(\frac{\alpha}{2}\right) \sigma_\psi \sqrt{\frac{2}{\pi}} \left(\sqrt{\frac{J-1}{J}}\right) J, \tag{3.17}$$

$$(\sigma_v^{at})^2 \equiv Var[\Delta V_i^{'}] = \left(\frac{\alpha}{2}\right)^2 \sigma_\psi^2 \left(1 - \frac{2}{\pi}\right) J + o(J). \tag{3.18}$$

[13] In order to facilitate the comparison between our model with liquidity frictions and the standard MDH of TP and without any loss of generality, the active traders who were present at the ith equilibrium are also allowed to trade as the $(i+1)$th piece of information hits the market. This has no impact on the $(i+1)$th equilibrium price, but forces the traded volume due to information to equal the cumulated amount that would prevail in the absence of liquidity frictions across equilibria i and $(i+1)$.

Since $V_i^{TP} = V_i^{'}$, from equation (3.9) generalized to $i = 1, ..., I_t$ and equation (3.16), we obtain:

$$\frac{1}{2}\sum_{j=1}^{J} \mid z_{ij} \mid \;\; = \;\; \frac{\alpha}{2}\sum_{j=1}^{J} \mid \psi_{ij} - \bar{\psi}_i \mid . \tag{3.19}$$

We set $z_{ij} = \alpha(\psi_{ij} - \bar{\psi}_i)$. Then, combining equations (3.1) generalized to $i = 1, ..., I_t$ and (3.19) yields:

$$z_i = \sum_{j=1}^{J_1} z_{ij} = \alpha \sum_{j=1}^{J_1} (\psi_{ij} - \bar{\psi}_i). \tag{3.20}$$

It follows that the variance of the order imbalance σ_z^2 is a function of σ_ψ^2:

$$\sigma_z^2 \;\; = \;\; \alpha^2 J_1^2 \left(\frac{J - J_1}{J J_1} \right) \sigma_\psi^2. \tag{3.21}$$

When a liquidity event occurs at the ith intraday equilibrium, only J_1 out of J active traders participate at the exchange. As discussed by TP, the ith equilibrium price change is the average of the reservation price increments of active traders being present at the market. Let ΔP_{ij}^* be the reservation price of trader j ($j = 1, ..., J$) at the ith equilibrium ($i = 1, ..., I_t$). Following TP, $\Delta P_{ij}^* = \phi_i + \psi_{ij}$ with $\phi_i \sim N(0, \sigma_\phi^2)$ and independent of ψ_{ij}. Then, from (3.13), we obtain:

$$\begin{aligned}
\Delta P_i^{''} &= \Delta P_i - \Delta P_i^{'}, \\
\Delta P_i^{''} &= \frac{1}{J_1}\sum_{j=1}^{J_1} \Delta P_{ij}^* - \frac{1}{J}\sum_{j=1}^{J} \Delta P_{ij}^*, \\
\Delta P_i^{''} &= \frac{1}{J_1}\sum_{j=1}^{J_1} (\phi_i + \psi_{ij}) - \frac{1}{J}\sum_{j=1}^{J} (\phi_i + \psi_{ij}), \\
\Delta P_i^{''} &= \frac{1}{J_1}\sum_{j=1}^{J_1} \psi_{ij} - \frac{1}{J}\sum_{j=1}^{J} \psi_{ij}.
\end{aligned} \tag{3.22}$$

It follows that $\Delta P_i^{''}$ is a normally distributed variable with mean zero and variance:

$$Var(\Delta P_i^{''}) = \sigma_\psi^2 \left(\frac{J - J_1}{J J_1} \right). \quad (3.23)$$

Replacing $\frac{1}{J}\sum_{j=1}^{J} \psi_{ij}$ by $\bar{\psi}_i$ and rearranging the terms of the last equation of (3.22) yields:

$$\Delta P_i^{''} = \frac{1}{J_1} \sum_{j=1}^{J_1} (\psi_{ij} - \bar{\psi}_i). \quad (3.24)$$

Then, from equations (3.20) and (3.24), it follows that:

$$z_i = \alpha J_1 \Delta P_i^{''}. \quad (3.25)$$

Replacing (3.25) into (3.10) generalized to $i = 1, ..., I_t$, we can show that the traded volume due to order imbalance z_i is a function of $\Delta P_i^{''}$:

$$V_i^{''} = a \mid \Delta P_i^{''} \mid, \quad (3.26)$$

where $a = \alpha \frac{M}{1+M} J_1$. Thus, $V_i^{''}$ is the absolute value (multiplied by a) of a normally distributed variable $\Delta P_i^{''}$ with mean zero and variance given in equation (3.23). The first two moments of $V_i^{''}$ denoted respectively by μ_v^{la} and $(\sigma_v^{la})^2$ are:

$$\mu_v^{la} \equiv E[V_i^{''}] = a\sigma_\psi \sqrt{\frac{2}{\pi}} \left(\sqrt{\frac{J - J_1}{J J_1}} \right), \quad (3.27)$$

$$(\sigma_v^{la})^2 \equiv Var[\Delta V_i^{''}] = a^2 \sigma_\psi^2 \left(1 - \frac{2}{\pi}\right) \left(\frac{J - J_1}{J J_1} \right). \quad (3.28)$$

From (3.21) and (3.27) as well as the relation $a = \alpha \frac{M}{1+M} J_1$, we have: $\mu_v^{la} = \sigma_z \sqrt{\frac{2}{\pi}} \left(\frac{M}{1+M}\right)$ when a liquidity event occurs and $\mu_v^{la} = 0$ otherwise. This means that μ_v^{la} can be explained by the combined effect of the occurrence of order imbalance z_i and the intervention of liquidity arbitragers. The average traded volume due to liquidity frictions μ_v^{la} is an increasing function of σ_z and M. The absence of liquidity events yields $\mu_v^{la} = 0$; this result follows independently from $\mu_v^{la} = \sigma_z \sqrt{\frac{2}{\pi}} \left(\frac{M}{1+M}\right)$ when $z = 0$ and from equation (3.22) when $J_1 = J$.

Generally speaking, summing the within-day price changes due to information $\Delta P_i^{'}$ and the price imperfections due to lacks of liquidity $\Delta P_i^{''}$, as well as the liquidity adjustments $-\Delta P_i^{''}$, yields the day-t price change ΔP_t:

$$\Delta P_t \;=\; \sum_{i=1}^{I_t} \Delta P_i^{'} + \sum_{i=1}^{I_t} \delta_i \Delta P_i^{''} - \sum_{i=1}^{I_t} \delta_i \Delta P_i^{''}. \tag{3.29}$$

Therefore, the intraday liquidity events do not impact the daily price change. As in TP, the daily price increment is normally distributed with mean zero and variance σ_p^2, which yields:

$$\Delta P_t \;=\; \sum_{i=1}^{I_t} \Delta P_i^{'}, \tag{3.30}$$

where $\Delta P_i^{'} \sim N(0, \sigma_p^2)$. Consequently, only the information flow impact is integrated in the daily price change. In our model, by definition, there is no order imbalance at the last trading date of the day (i.e., date $I_t - 1$). Thus, the closing price of the day reveals the information available up to that date: $P_{I_t} = E_{I_t} \tilde{P}_T$. Since the liquidation value of the asset $\tilde{P}_T$ is revealed at the end of the trading day, the closing price $E_{I_t} \tilde{P}_T$ converges[14] to the liquidation value of the asset P_T.

However, as discussed above, the volume traded by liquidity arbitragers adds to the volume that would be traded in the absence of liquidity imperfections. Summing the within-day traded volume motivated by information flow $V_i^{'}$ and the traded volume due to liquidity shocks $V_i^{''}$, we obtain the day-t traded volume:

$$V_t \;=\; \sum_{i=1}^{I_t} V_i^{'} + \sum_{i=1}^{I_t} \delta_i V_i^{''}. \tag{3.31}$$

[14]Indeed, during the trading day, thanks to the arrival of new information to the market, the equilibrium price converges to the liquidation value of the asset that will prevail at the end of the trading day. Here, even if this convergence is blurred by the presence of liquidity frictions at intraday frequency, it is successfully achieved at the end of the trading day since no liquidity friction occurs at the last equilibrium.

In this equation, $V_i^{'} \sim N(\mu_v^{at}, (\sigma_v^{at})^2)$, $V_i^{''} \sim$ Half-$N(\mu_v^{la}, (\sigma_v^{la})^2)$ and $\sum_{i=1}^{I_t} \delta_i V_i^{''} = \sum_{l=1}^{L_t} V_l^{''}$ where $l = 1, ..., L_t$ is a subsequence of $i = 1, ..., I_t$ such as $\delta_i = 1$.

In this paper, we assume that I_t and L_t are conditionally independent $Cov[(L_t|I_t), I_t] = 0$ which implies that $Cov(f(L_t|I_t), g(I_t)) = 0$, where $f(L_t|I_t)$ and $g(I_t)$ can be any function of $L_t|I_t$ and I_t, respectively. We consider that the indicator variable δ_i is independently drawn from a Bernoulli distribution with parameter[15] p. Then, L_t's first two unconditional[16] moments are respectively: $E(L_t) = pE(I_t)$ and $Var(L_t) = p(1-p)E(I_t) + p^2 Var(I_t)$. The unconditional covariance between I_t and L_t is given by:

$$Cov(I_t, L_t) = pVar(I_t). \tag{3.32}$$

From equations (3.30) and (3.31), we obtain a mixture of distribution model with two latent variables, I_t and L_t. Note that, conditional on I_t and L_t, $V_i^{'}$ and $V_i^{''}$ are independent. In addition, $(\sigma_v^{la})^2$ given in (3.28) can be considered as $o(JJ_1)$ when added to $(\sigma_v^{at})^2$ given in (3.18). It follows that, conditional on I_t and L_t, the daily volume V_t can be considered as $N(\mu_v^{at} I_t + \mu_v^{la} L_t, (\sigma_v^{at})^2 I_t)$ without any loss of generality[17]. Henceforth, for notation simplicity, we replace $(\sigma_v^{at})^2$ by σ_v^2. The bivariate normal mixture can then be written:

$$\Delta P_t = \sigma_p \sqrt{I_t} Z_{1t}, \tag{3.33}$$

$$V_t = \mu_v^{at} I_t + \mu_v^{la} L_t + \sigma_v \sqrt{I_t} Z_{2t}, \tag{3.34}$$

where $Cov(\Delta P_t, V_t \mid I_t, L_t) = 0$, and Z_{1t} and Z_{2t} are mutually independent standard normal variables (and independent of I_t and L_t). Conditional on I_t, the daily price change is normally distributed: $\Delta P_t \sim N(0, \sigma_p^2 I_t)$. Our model implies that the information flow impacts both

[15]This means that for each $i = 1, ..., I_t$, δ_i takes value 1 with success probability p and value zero with failure probability $(1-p)$: $\delta_i \sim B(p)$. Its first two moments are $E(\delta_i) = p$ and $Var(\delta_i) = p(1-p)$.

[16]Since $L_t = \sum_{i=1}^{I_t} \delta_i$, conditional on I_t, L_t has a binomial distribution with parameters I_t and p: $L_t|I_t \sim B(I_t, p)$. Let $E(I_t)$ and $Var(I_t)$ be the unconditional mean and variance of I_t.

[17]For large L_t, the sum of L_t absolute values of normally distributed variables $\sum_{l=1}^{L_t} V_l^{''} = \sum_{i=1}^{I_t} V_i^{''}(z_i)$ can be approximated by a normal distribution.

the daily price change and the traded volume, while only the daily volume is affected by the random liquidity shocks.

Note that the standard MDH of TP as well as the GM model are implied by (3.33)-(3.34) as particular cases:
(i) When $z_i = 0$, $\mu_v^{la} = 0$ and the system (3.33)-(3.34) reduces to the standard MDH of TP;
(ii) If the trading day consists of only two trading periods and a terminal condition date as in Figure 2, the bivariate mixture given in (3.33)-(3.34) reduces to the standard GM model.

From equations (3.32) and (3.33)-(3.34), the unconditional contemporaneous relation between ΔP_t^2 and V_t is:

$$\begin{aligned} Cov(\Delta P_t^2, V_t) &= \sigma_p^2(\mu_v^{at} + p\mu_v^{la})Var(I_t), \\ &= \sigma_p^2 \mu_v Var(I_t). \end{aligned} \tag{3.35}$$

The volatility-volume covariance predicted by our model is positive as is that of TP given in (2.5). However, while in the TP world the average total volume μ_v is due to information, in our model the average total volume is decomposed into two parts, μ_v^{at} and $p\mu_v^{la}$, due to information and liquidity shocks, respectively: $\mu_v = \mu_v^{at} + p\mu_v^{la}$. Since TP do not account for liquidity shocks, the standard MDH model may overestimate the average volume related to information inflow: $\mu_v \geq \mu_v^{at}$.

The model given in (3.34) is called the modified MDH model with liquidity (henceforth MDHL model), and forms the basis of our empirical work. The particularity of this model is that it takes into account both information and liquidity shocks. Based on the MDHL model, we can exploit the volume-volatility correlation in order to decompose the traded volume for a given stock into two components and thus separate information from the liquidity trading impact on the observed daily volume.

4 Empirical application

4.1 The data

Our sample consists in all FTSE100 stocks listed on 10 July, 2007. The daily returns R_t are measured by the daily (log) price change. We consider the period from 4 January 2005 to 26 June 2007, i.e. 636 observation dates. We exclude stocks with missing observations ending up with 93 stocks. Daily returns and transaction volumes are extracted from Bloomberg databases. Following Bialkowski et al. (2008), we retain the turnover ratio as a measure for volume which controls for dependency between the traded volume and the float. The latter represents the difference between annual common shares outstanding and closely held shares for any given fiscal year. Common and closely held shares are extracted from Factset databases. Let q_{kt} be the number of shares traded for asset k, $k = 1, ..., K$ on day t, $t = 1, ..., T$, and N_{kt} the float for asset k on day t. The individual stock turnover for asset k on day t is $V_{kt} = \frac{q_{kt}}{N_{kt}}$.

	Returns				Volume			
	Average	Dispersion	Min	Max	Average	Dispersion	Min	Max
Mean	**0,0007**	0,0005	-0,0005	0,0024	**0,0087**	0,0052	0,0018	0,0405
Volatility	**0,0137**	0,0031	0,0074	0,0263	**0,0065**	0,0062	0,0011	0,0545
Skewness	**0,2853**	0,9271	-4,0840	3,1510	**3,4636**	1,7526	1,0041	9,8661
Kurtosis	**9,9205**	9,8313	3,2134	61,3788	**28,4178**	26,5025	4,8613	133,8895
$(\text{Return})^2$ with Volume Correlation	-	-	-	-	**0,42**	0,14	0,17	0,75

Table 1: Summary statistics for return and turnover across securities.

For each of the 93 stocks, we compute the empirical first moments (mean, volatility, skewness and kurtosis) of volume and returns as well as the correlation between squared returns and volume. The cross-security distribution of these statistics are summarized in

Table 1. The first row reports the average, the dispersion, the minimum, and the maximum of the means of returns and volume across the 93 stocks. The second row gives the same cross-section statistics (average, dispersion, minimum and maximum) of the volatilities of returns and volume, and so on for the skewness, kurtosis, and the correlation between squared returns and volume. We perform a Pearson test to check the significance of the correlation coefficients. These correlation coefficients are statistically significants for 92 over 93 stocks at the 95% confidence level. The statistics reported in the last row of Table 1 are computed using only the statistically significant correlations between squared returns and volume.

The implications of the MDH for the joint distribution of daily returns and volume, are examined in details by Clark (1973), Westerfield (1977), Tauchen and Pitts (1983), Harris (1986), Harris (1987) among others. They assume that both variables (the daily (log) price change and daily volume) are conditioned by a random and serially uncorrelated mixing variable represented by the information flow. They show that the MDH can explain why the sample distribution of daily returns is kurtotic relative to the normal distribution, why the distribution of the associated traded volume is positively skewed and kurtotic relative to the normal distribution and why squared returns are positively correlated with trading volume. The randomness of the mixing variable is crucial to the MDH analysis. If the mixing variable were constant, there would be no reason to observe the above empirical patterns, and the daily returns and volume should be mutually independent and normally distributed.

The results reported in Table 1 are then consistent with the MDH. The average and minimum statistics of the volume skewness and squared return correlation with volume are positive; and the average and minimum statistics of return and volume kurtosis are greater then 3, as predicted by the mixture model. Moreover, these cross-security statistics are larger than their corresponding constant mixing variable expected values[18].

Finally, we present in Figure 3 the scatter plots of returns and squared returns against turnover for two FTSE100 stocks: ANGLO AMERICAN (AAL LN) and AVIVA (AV LN).

[18]The expected value of the volume skewness and correlation coefficient is zero, and the expected value of return and volume kurtosis is 3 when the mixing variable is constant.

The upper (lower) graphs are pairwise scatter plots for AAL LN (AV LN) with return-turnover on the left, and volatility-turnover on the right. The graphs highlight the well-documented positive[19] relation between volatility and volume.

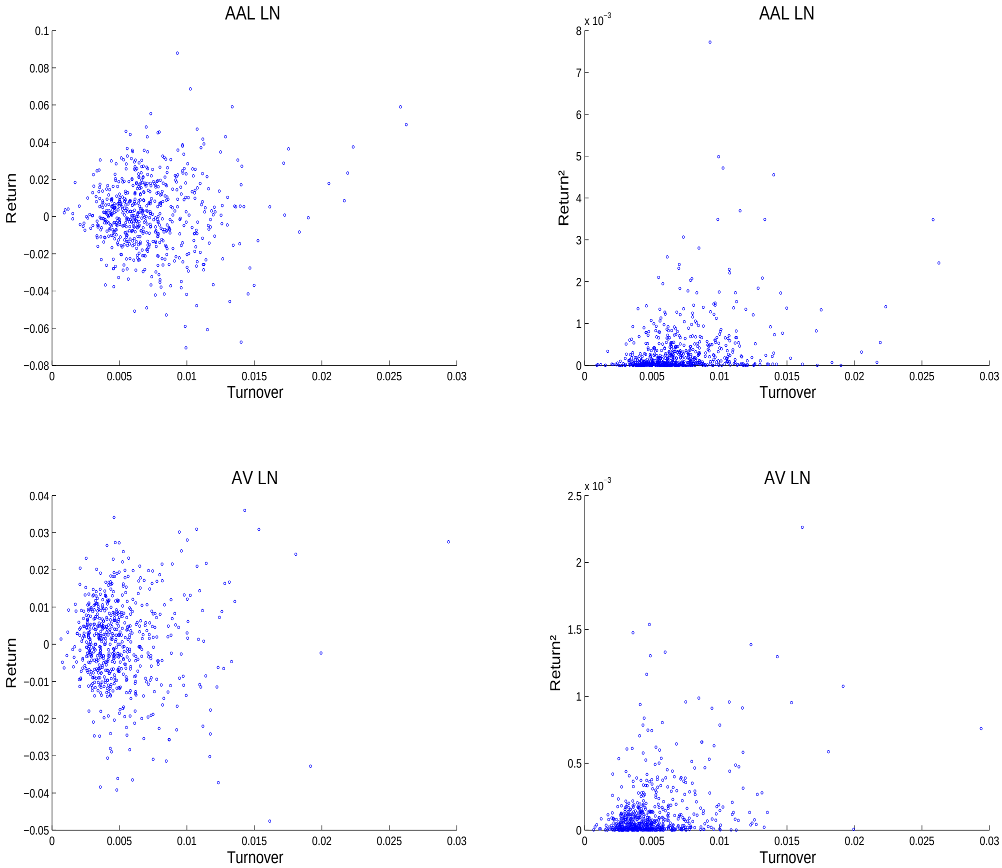

Figure 3: Scatter plots of returns and squared returns against turnover for two FTSE100 stocks: Anglo American (AAL LN) and AVIVA (AV LN).

[19]Clark (1973), Copeland (1976), Copeland (1977), Tauchen and Pitts (1983), Harris (1982), Harris (1986), Harris (1987), Epps and Epps (1976), and Westerfield (1977) among others show a positive correlation between the variability of price change and volume.

4.2 The MDHL test

4.2.1 Test methodology

Following Richardson and Smith (1994), we use the Generalized Method of Moments (GMM) of Hansen (1982) to test the validity of the MDHL model. Since our bivariate mixture with two latent variables imposes restrictions on the unconditional joint moments of the observables as a function of model parameters, it is possible to form overidentifying restrictions on the data. Optimization methods can then be used to estimate the coefficients and test the global validity of the model simultaneously.

Let $X_t = (R_t, V_t)$ be the vector of return and volume observations prevailing at day t for a given stock and $\theta = (\mu_v^{at}, \mu_v^{la}, \sigma_p^2, \sigma_v^2, m_{2I}, p)$ be the 6×1 vector of the MDHL model parameters. The first four coefficients are related to the observables and correspond to the mean and variance parameters of equations (3.33)-(3.34), m_{2I} is the second moment of the latent variable I_t and p is the Bernoulli distribution parameter which drives the distribution of the latent variable L_t.

If X_t is generated by the MDHL model, there is some true set of parameters θ_0 for which:

$$E[h_t(X_t, \theta_0)] = 0, \tag{4.1}$$

where h_t is a column vector of H unconditional moment conditions implied by our model. Since we do not observe the true expectation of h_t in practice, we define a vector $g_T(\theta)$ containing the sample averages corresponding to the elements of h_t. For large T, if X_t is generated by the MDHL model, $g_T(\theta_0)$ should be close to zero[20]:

$$g_T(\theta_0) \equiv \frac{1}{T} \sum_{t=1}^{T} h_t(X_t, \theta_0) \longrightarrow 0, \;\; when \;\; T \to \infty. \tag{4.2}$$

In order to derive the moment restrictions implied by the MDHL model, we focus on the

[20]The GMM procedure of Hansen (1982) is presented in details in the Appendix C.1.

first four moments of the return and volume time series and on some of their corresponding cross-moments such as the covariances between returns and either volume or squared volume.

In the previous section we assumed that, conditional on I_t, L_t is drawn from a binomial distribution with parameters I_t and p. It follows that the unconditional moments of L_t are functions of p and the unconditional moments of I_t. In addition, we need to choose a distribution function for the latent variable I_t. TP assume a lognormal distribution for the mixing variable I_t in order to ensure its positiveness. Lognormality has also been suggested by several authors, such as **?** as well as Foster and Viswanathan (1993). Richardson and Smith (1994) tested several distribution functions for the information inflow and conclude that the data reject the lognormal distribution less frequently than the other distribution candidates, such as inverted gamma and Poisson distributions. These results motivate us to retain a lognormal distribution for I_t.

As discussed by TP, the mathematical formulations of the latent factor models, such as the MDHL model, are invariant with respect to scalar transformations of the unobserved variables. It follows that, if a is any positive constant such as $I_t^* \equiv I_t/a$, the model:

$$R_t \sim N(0, [\sigma_p^2 a] I_t^* \mid I_t, L_t), \tag{4.3}$$

$$V_t \sim N([a\mu_v^{at}] I_t^* + [\mu_v^{la}] L_t, [a\sigma_v^2] I_t^* \mid I_t, L_t), \tag{4.4}$$

is empirically the same as the MDHL model given in (3.33)-(3.34). By setting $E[I_t^*] = 1$, we can identify the transformed parameters which are given by: $\mu_v^{at*} = \mu_v^{at} m_{1I}$, $\sigma_p^{*2} = \sigma_p^2 m_{1I}$, $\sigma_v^{*2} = \sigma_v^2 m_{1I}$, $m_{2I}^* = m_{2I}/m_{1I}^2$, $m_{3I}^* = m_{3I}/m_{1I}^3$ and $m_{4I}^* = m_{4I}/m_{1I}^4$. Henceforth, we will consider only these transformed parameters. However, for notation simplicity, we omit the "$*$" symbol.

The lognormality assumption for I_t implies the following moment restrictions [see Richardson and Smith (1994)]:

$$\begin{aligned} m_{3I} - m_{2I}^3 - 3m_{2I}^2 &= 0 \\ m_{4I} + 4(1+m_{2I})^3 + 3 - (1+m_{2I})^6 - 6(1+m_{2I}) &= 0 \end{aligned} \quad (4.5)$$

where $m_{iI}, (i = 2, 3, 4)$ is the i^{th} centered moment for the mixing variable I_t.

Given the scalar transformations of the parameters depending on I_t, as well as the distribution assumptions for I_t ($I_t \sim LogN(1, m_{2I})$) and L_t ($L_t \mid I_t \sim B(I_t, p)$), the sample moment vector $g_T(\theta)$ is given by:

$$g_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} \begin{pmatrix} (V_t - E(V_t)) & (1) \\ (R_t - E(R_t))^2 & (2) \\ (V_t - E(V_t))^2 & (3) \\ (R_t^2 - E(R_t^2))(V_t - E(V_t)) & (4) \\ (R_t^2 - E(R_t^2))(V_t^2 - E(V_t^2)) & (5) \\ (V_t - E(V_t))^3 & (6) \\ (R_t - E(R_t))^4 & (7) \\ (V_t - E(V_t))^4 & (8) \\ (R_t - E(R_t))^2(V_t - E(V_t))^2 & (9) \end{pmatrix}. \quad (4.6)$$

The functional forms of the sample moments (1)-(9) are given in Appendix $C.2$. We obtain a system of nine equations and only six parameters to be estimated which yields three overidentifying restriction to test[21].

4.2.2 Test results

We apply the GMM procedure described in the previous paragraph to the 93 stocks of our sample using the entire data history. To restrict the Bernoulli parameter p to evolve between

[21] When working with an overidentified system, the GMM chooses $\hat{\theta}_T$ as the value of θ that minimizes the quadratic form of $g_T(\theta)$ which requires the selection of a weighting matrix. For this purpose, we use the Newey and West (1987) methodology which is described in Appendix $C.1$

0 and 1, we use a logistic-transform with x being the unconstrained parameter. Tables 2 and 3 of Appendix D report the estimation results. The test statistics of Hansen (1982) allowing to assess the global validity of the MDHL model are given in column 9. With three overidentifying restrictions, they are asymptotically distributed as a χ^2_3. For 83% of the stocks, the test statistic values do not exceed their critical value of $7,82$. Consequently, we can not reject the MDHL model at the 95% level of significance.

Columns 2 to 5 in Tables 2 and 3 provide parameter estimates for returns and volume distributions, while columns 6 to 8 report estimated parameters related to the latent variables I_t and L_t distributions. Since we set $E(I_t) = 1$, the estimated μ_v^{at} can be interpreted as the time-series-average of the impact of information inflow on the daily traded volume. On the other hand, $p\mu_v^{la}$ can be interpreted as the time-series-average of the impact of liquidity shocks on the daily traded volume. In particular, $p\mu_v^{la}$ represents a stock-specific measure for liquidity which is determined by both the amplitude of trade asynchronization, as measured by μ_v^{la}, and its probability of occurrence p. The higher the trade asynchronization for a given stock the higher its frequency and the liquidity-arbitrage-based traded volume. This in turn results in a higher volume and thus a higher $p\mu_v^{la}$.

Since our model implies that information moves the market from one equilibrium to the next and liquidity shocks appear within some of these equilibria, we should expect to observe a statistically significant μ_v^{la} parameter only for stocks having also a significant μ_v^{at}. The results reported in Appendix D confirm our intuitions. The 43 stocks for which we obtain significant μ_v^{la} have also a μ_v^{at} parameter statistically different from zero. Note that, for these stocks, we also obtain statistically significant x parameters. Reported are in column 9 of Tables 2 and 3 the relative values of the average liquidity volume as measured by $p\mu_v^{la}$ divided by the sum of μ_v^{at} and[22] $p\mu_v^{la}$, henceforth *relative* $p\mu_v^{la}$. At this stage of the analysis, two additional remarks can be made:

(i) A significantly positive $p\mu_v^{la}$ suggests that the stock faces time-average intraday liquidity

[22]Note that, under our model specification the unconditional mean of the daily traded volume is: $E(V_t) = \mu_v^{at} + p\mu_v^{la}$.

frictions. This motivates the liquidity arbitragers to enter the market and thus increase the average traded volume. Since we do not observe liquidity shocks, we can infer their occurrence from liquidity arbitrage trading which directly impacts the volume. Our model helps identify the intraday impact of this type of market participants on the traded volume using daily data: 39 out of the 43 stocks with a significantly positive $p\mu_v^{la}$ are concerned by significant liquidity problems[23].

(ii) If $p\mu_v^{la}$ is not significant, our model comes down to that of Tauchen and Pitts (1983) which assumes that the total traded volume is a proxy of the information flow.

4.3 The MDHL-based liquidity measure

We use a structural model to separate the respective impacts of the two latent variables I_t and L_t on the average-raw-traded volume of individual stocks. The model is particularly attractive in practice since it provides a static, stock-specific liquidity measure $p\mu_v^{la}$ which helps identify the presence of intraday liquidity frictions using daily data. Based on the μ_v^{at}, μ_v^{la} and p parameters, we can distinguish stocks concerned by liquidity frictions for a given period (on average) from liquid equities whose average daily traded volume is driven only by information inflow. In addition, using the relative $p\mu_v^{la}$ reported in column 9 of Tables 2 and 3, stocks facing liquidity frictions can be ranked according to their respective degree of illiquidity, which is determined for any given stock by (i) the amplitude of trade asynchronization and (ii) its probability of occurrence. Thus, estimating μ_v^{la} and p separately provides additional insights concerning the liquidity profile of a given stock. The liquidity-based average volume for a particular period can be explained by frequent but small liquidity accidents, rare but large liquidity accidents, or simultaneously frequent and large liquidity accidents. For example, HAMMERSON PLC (stock 32), SEGRO PLC (stock 77), SCOTTISH & SOUTHERN ENERGY (stock 81) and XSTRATA PLC (stock 92), exhibiting the

[23]The 4 remaining stocks, CADBURY PLC (stock 18), MAN GROUP PLC (stock 27), SABMILLER PLC (stock 72) and UNILEVER PLC (stock 85), have negligible relative $p\mu_v^{la}$ characterized by both p and μ_v^{la} values evolving in the neighborhood of zero.

4 highest relative $p\mu_v^{la}$ of our sample, are characterized by both important μ_v^{la} parameters (being 2 to 3 times higher than the corresponding μ_v^{at}) and important probabilities of trade asynchronization p whose values fall in the sample's highest decile. On the other hand, some other equity assets, such as LONMIN PLC (stock 50) and MITCHELLS & BUTLERS PLC (stock 51) face liquidity shocks characterized by much higher amplitude of trade asynchronization than in the former case (of an order of 7 to 9 times higher then the corresponding μ_v^{at}) yet much lower p values.

Previous literature relates stock liquidity to total traded volume and suggests that illiquid equity assets have low traded volume or turnover[24]. Thus, the total traded volume appears to be a good proxy for liquidity. Moreover, using market capitalization as a proxy for stock liquidity is a common practice in financial markets where small stocks are assumed to face more liquidity problems than blue chip stocks. We now confront these 2 measures to the MDHL-based liquidity indicator $p\mu_v^{la}$.

Figures 4 and 5 focus on the 39 stocks of our sample presenting a significantly positive relative $p\mu_v^{la}$ and show the relative liquidity volume against the average raw daily volume[25] and the average market capitalization[26] over the estimation period, respectively. The first graph points out that there is no systematic relation between relative $p\mu_v^{la}$ and total traded volume. For example, the highest time-average-raw-volume stock, XSTRATA PLC (stock 92), presents a greater relative $p\mu_v^{la}$ than some others with lower MDHL-based liquidity measure, such as HSBC HOLDINGS (stock 34) and BP PLC (stock 15). More generally, within the groups of large traded volume and low traded volume stocks, there is an important dispersion of the illiquidity level. As a result, the total traded volume does not help discriminate stocks facing liquidity shocks according to their degree of illiquidity.

[24]See Datar et al. (1998), and Chordia et al. (2000) among others.
[25]The traded volume is measured by the turnover.
[26]The market capitalization is measured by the float.

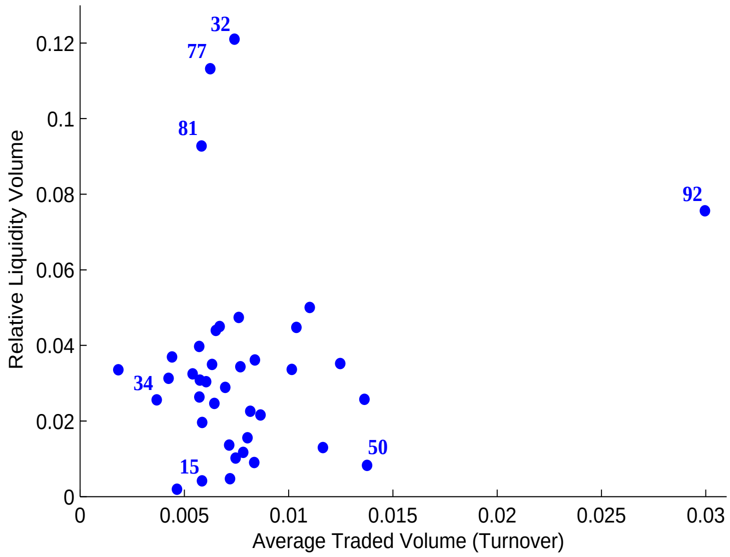

Figure 4: Relative liquidity volume versus average daily traded volume.

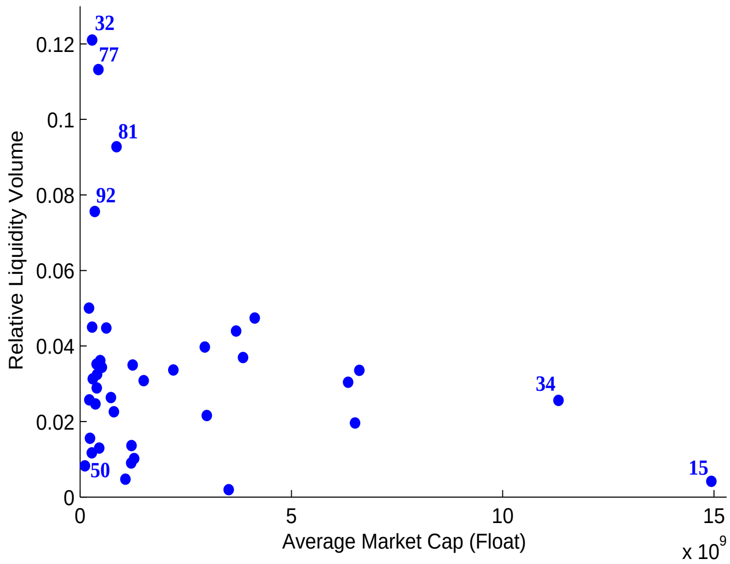

Figure 5: Relative liquidity volume versus average market cap measured by the float.

These results confirm the findings of Borgy et al. (2010) regarding the lack of the traded volume and the number of transactions to correctly measure market illiquidity. For example, a higher number of transactions may be due to a higher liquidity risk which induces market participants to split their trades, as well as to an increasing market liquidity due to a larger number of liquidity providers being present into the market. Similarly, in our framework, an increasing total traded volume for a given stock may be explained by a rise in information-based trading, or by an increase in liquidity trading activity due to the intervention of liquidity arbitragers who trade in response to liquidity frictions. This suggests that decomposing the total traded volume into two components due to information and liquidity shocks provides more precise indications on market liquidity.

Figure 5 shows that the biggest companies among the 39 stocks are also the most liquid ones. For large market capitalizations, there is indeed quite a strong negative relation between firm size and illiquidity level. However, within the group of small capitalizations, there is an important dispersion of $p\mu_v^{la}$ values. For example, some of the most illiquid firms, such as HAMMERSON PLC (stock 32), SEGRO PLC (stock 77), but also some of the less illiquid ones, such as LONMIN PLC (stock 50), belong to the lowest size deciles. These findings suggest that the market size is not a good proxy for liquidity shocks. In particular, considering small firms to be illiquid may be misleading since market size fails to discriminate small companies according to their illiquidity level.

Assessing the stock liquidity level through simultaneously total traded volume and market capitalization ends up to being quite disconcerting. Illustrating this point, XSTRATA PLC (stock 92) is considered as the less illiquid among the 39 firms according to the total traded volume criterion, but as one of the most illiquid ones as reported by the market capitalization indicator. Conversely, HSBS HOLDINGS (stock 34) and BP PLC (stock 15) seem to be highly illiquid when focusing on the total traded volume, while their (large) size ranks them among the less illiquid of the 39 equity assest considered here. These results highlight the relevance of such a structural liquidity measure as the $p\mu_v^{la}$, in order to obtain a better understanding of the market liquidity for a given stock. The $p\mu_v^{la}$ indicator provides additional insights on a firm's liquidity while reconciliating and explaining the results obtained using the total traded volume and the market capitalization criteria. In particular, for XSTRATA PLC (stock 92), HSBS HOLDINGS (stock 34) and BP PLC (stock 15), the MDHL-based liquidity measure reinforces the results provided by the size criterion at the expense of the total traded volume indicator.

4.4 Robustness checks

4.4.1 Global validity of MDHL relative to the standard MDH

We also estimate the standard MDH model using Richardson and Smith (1994) procedure[27]. The results are presented in Tables 4 and 5 in the Appendix E. The standard MDH model is accepted by the data for 89% of stocks versus 83% for the MDHL model. The slight under-performance of the MDHL model in terms of global validity can be explained by its higher degree of estimation complexity. Richardson and Smith (1994) estimate unbounded parameters while we restrict the values of the p to evolve between 0 and 1. On the other hand, Richardson and Smith (1994) modify the TP's price change equation by artificially introducing a mean parameter μ_p which allows them to obtain much simpler moment conditions than in the absence of μ_p. This is not the case in our framework; our model is directly derived from the standard MDH of TP without a mean parameter in the price variation equation.

The MDHL model has a two-dimensional structure, allowing separating information from liquidity shock impacts on the total traded volume. While providing a deeper comprehension of how the daily traded volume is built up, it enables us to obtain a similar level of global validity compared to the standard MDH model, which stands as its one-dimensional counterpart. When comparing the mean volume parameters obtained by the two models, we find that μ_v is approximately equal to the sum of μ_v^{at} and $p\mu_v^{la}$. For example, for ASSOCIATED BRITISH FOODS PLC (stock 2), we have $\mu_v = 0,00621$ and $(\mu_v^{at} + p\mu_v^{la}) = 0,00625$; for BARCLAYS PLC (stock 8), we obtain $\mu_v = 0,00592$ and $(\mu_v^{at} + p\mu_v^{la}) = 0,00589$. These results are intuitive and show that the MDHL model succeeds in decomposing the average

[27] To estimate the standard MDH model we use the implied unconditional means, variances, skewness, and corresponding cross-moments of the observable variables, R_t and V_t. With 9 moment conditions and only 6 parameters to be estimated, there are 3 overidentifying restrictions to be tested. For more details, see Richardson and Smith (1994).

traded volume into information-based and liquidity-based components.

4.4.2 Parameter stability

As discussed previously, $p\mu_v^{la}$ is, by construction, a static liquidity measure which quantifies, on average over a given test period, the daily volume driven by liquidity frictions. To assess its dynamics over time, we perform subperiod analysis for a set of 10 stocks of our sample, 8 of them representing different illiquidity levels as measured by relative $p\mu_v^{la}$, the other 2 being deemed perfectly liquid. For these 10 candidates, we split the data history into two subperiods of 318 observations extending from 4 January 2005 to 4 April 2006 and from 5 April 2006 to 10 July 2007, respectively. Our goal is to assess the stability of the stock liquidity profile over the 2 time intervals. In the presence of time-varying $p\mu_v^{la}$, we should observe an increase of the illiquidity level in the second subperiod since stock markets were impacted in 2007 by significant liquidity shocks in connection with the subprime crisis.

Table 6 in Appendix F gives the MDHL-estimated parameters for both subperiods as well as the overall time interval. Global validity of the MDHL model is confirmed for both subperiods; the χ_3^2 values do not exceed their critical value of $7,82$. This suggests that, for the selected stocks, the MDHL model is a plausible explanation of the bivariate distribution of stock returns and traded volume; its global validity is not sample-dependent. Moreover, the information-based volume parameter μ_v^{at} estimated using the overall time period for a given stock is included in the interval delimited by the μ_v^{at} values obtained using the two distinct subperiods. In particular, the overall-period-information-based measure is approximately equal to the mean of the two subperiod ones. The slight deviations may be due to the different lengths of the data history $-$ and thus different amplitudes of the standard errors. The same is true for the MDHL liquidity measure[28].

The subperiod analysis provides additional insights concerning the stability of the liq-

[28] In this case the differences between the full period $p\mu_v^{la}$ and the mean between the two subperiod $p\mu_v^{la}$ are larger than in the former case since our liquidity measure simultaneously depends on two estimated parameters (p and μ_v^{la}).

uidity measure $p\mu_v^{la}$ proposed in this paper. As for the 2 liquid stocks considered here, CAPITA GROUP PLC (stock 22) remains liquid over time, while NATIONAL GRID PLC (stock 54) is affected by significant liquidity shocks during the second time interval. In this case, working with the entire sample history hides the presence of liquidity frictions related to a particular subperiod. Within the group of firms impacted by liquidity shocks, we can distinguish two types of stocks: those having constant $p\mu_v^{la}$ over time, such as DIAGEO PLC (stock 24) and ICAP PLC (stock 35), and those exhibiting substantial variations in the liquidity measure, as for KELLN SOLAR (stock 44) and SEGRO PLC (stock 77). In the latter case, variations in the absolute illiquidity level are due to significant changes in the amplitude of liquidity-based volume as well as to the probability of order imbalance, reflecting a time-varying liquidity profile. KELLN SOLAR (stock 44) illustrates this point with large but infrequent liquidity shocks for subperiod 1 ($\mu_v^{la} = 0,0176$, $p = 0,008$) and lower but more probable liquidity frictions for subperiod 2 ($\mu_v^{la} = 0,0126$, $p = 0,038$). On the other hand, the time variation of the SEGRO PLC (stock 77) liquidity profile can be explained by an substantial growth of the order imbalance frequency which varies from $0,02$ in the first time interval to $0,11$ in the second one.

Generally speaking, subperiod 2 is characterized by an increasing stock illiquidity level as compared to the first time interval. Even some of the firms with $p\mu_v^{la} = 0$ during subperiod 1, such as SAGE GROUP PLC (stock 76), REED ELSEVIER PLC (stock 65) and NATIONAL GRID PLC (stock 54), turn out to face significant liquidity frictions during the second time interval. Such results are intuitive and reflect important liquidity shocks which affected financial markets during the summer of 2007. Our model enriches the analysis by providing a more acute explanation of the impact of liquidity shocks on trading volume. It enables us to characterize illiquid firms according to the amplitude of the liquidity shocks and its probability of occurrence, allowing traders to adapt their strategies accordingly.

To summarize these results, the global validity of the MDHL model seems to be time-invariant. However, the parameter stability varies from one stock to the other. Such a static illiquidity measure $p\mu_v^{la}$ can be directly applied to assets whose liquidity-based volume does not vary significatively over time. In this case, we can get a better understanding of firm liquidity and decompose the total traded volume into information-based and liquidity-based components. On the other hand, the subperiod analysis highlights an important drawback of our liquidity measure related to its failure to capture the time-dynamics of the stock illiquidity profile. This remark leads to a natural extension of our framework consisting in building a time-varying liquidity measure. This point will be discussed in the next section.

5 Concluding remarks

In this article, we first distinguish between two trading strategies, information-based and liquidity-based trading. We suggest that their respective impacts on returns and traded volume should be modeled differently. The former is incorporated into the daily price changes and the traded volume. The latter impacts the intraday price variations and volumes but do not affect the daily price changes, while increasing the daily traded volume. Second, we focus on the contemporaneous volatility-volume relationship and blend the microstructure setting of Grossman and Miller (1988) into the Tauchen and Pitts (1983) framework in order to develop an modified MDH model with two latent factors related to information arrivals and liquidity frictions. Our model provides a theoretical explanation of how the liquidity accidents increase the daily traded volume, in the presence of arbitrage participants. Third, the MDHL model gives a better comprehension of how the daily traded volume is built up; We show how to exploit the volatility-volume relation in order to separate information from liquidity impacts on the observed daily volume. In other words, the increase of volume due to liquidity arbitragers helps inferring the presence of liquidity frictions corresponding to order imbalances driven by asynchronization of order flows among active participants. In particular, our model exploits the time-series dimension of individual assets to provide an average (over time), stock-specific measure for liquidity shocks using daily data. This helps distinguish, for a given period, liquid (presenting not significant $p\mu_v^{la}$) from less liquid stocks (presenting significant $p\mu_v^{la}$). In addition, estimating p and μ_v^{la} separately provides a better comprehension of the stock liquidity profile determined by the amplitude of the order imbalances and the probability of their occurrence. This may be useful in order to build stock-picking strategies at a high trading frequency.

Our MDHL liquidity-based indicator is similar to that of Getmansky et al. (2004) who provide a static measure of the illiquidity affecting hedge fund returns. The authors systematically analyze various sources of the observed autocorrelation in hedge fund returns, such as time-varying expected returns, time-varying leverage, fee structures of hedge funds, as well as illiquidity and smoothed returns. They conclude that illiquid investments which drive "marking to model" returns and performance smoothing are the most plausible cause of the time-persistence of hedge fund returns. It follows that, serial correlation of fund returns may be a good proxy for illiquidity. Time-series of reported hedge fund returns can then be used to estimate the serial correlation of individual funds, which helps separate liquid from illiquid hedge funds for a given period.

Finally, our liquidity indicator presents two main limitations. First, it is a static indicator and as such it fails to capture the time-varying dynamics of liquidity frictions. The second limitation concerns the impossibility to build a common (market-wide) liquidity factor using stock-specific $p\mu_v^{la}$ parameters. Several recent studies are based on the commonality and time-varying properties of liquidity risk. Patton and Li (2009) extend Getmansky et al. (2004) analysis by allowing for serial correlation parameters to vary over time. They propose a model for time-varying hedge fund liquidity, building on the connection between liquidity and autocorrelation. In their empirical application over 600 individual hedge funds, they find strong evidence of time-varying liquidity for all hedge fund styles. They also provide a dynamic time-dependent proxy of liquidity for individual hedge funds. Nagel (2009) uses the profitability of contrarian strategies as a proxy for returns which compensate liquidity supplying activity. Using the cross-section of stock returns at each point in time, the author extracts a time-varying, market-wide liquidity indicator. The advantage of such an indicator is that it provides information on how market liquidity evolves over time and what determines its evolution. For example, Nagel (2009) finds that the liquidity indicator co-moves closely with the level of the VIX.

Therefore, it would be interesting to expand our stock-specific approach to first extract time-varying latent liquidity factors for individual stocks. For this purpose, the MDHL model developed in this paper can be extended to allow for serial dependence in L_t. Several studies show that liquidity shocks are not isolated events in time but rather seem to be time-persistent[29]. This suggests that serial correlation in L_t may explain the persistence of the traded volume. Signal extraction methods can then be used to filter the latent variable L_t for individual assets and thus to provide a time-varying, stock-specific liquidity indicator. Finally, factor decomposition analysis can be applied to the panel of individual liquidity indicators in order to build market-wide liquidity factors and thus to separate, for a given stock, common from specific liquidity components. This point is out of the scope of this paper and is part of current research.

[29] See, for example, Acharya and Pedersen (2005).

References

Acharya, V. V. and Pedersen, L. H. (2005). Asset Pricing with Liquidity Risk. *Journal of Financial Economics*, 77:375–410.

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

Amihud, Y. (2002). Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. *Journal of Financial Markets*, 5:31–56.

Andersen, T. (1996). Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. *The Journal of Finance*, 51(1):169–204.

Bialkowski, J., Darolles, S., and Fol, G. L. (2008). Improving VWAP Strategies: A Dynamic Volume Approach. *Journal of Banking and Finance*, 32:1709–1722.

Borgy, V., Idier, J., and Fol, G. L. (2010). Liquidity Problems in the FX Liquid Market: Ask for the "BIL". *Working paper, Banque de France.*

Brunnermeier, M. K. and Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *The Review of Financial Studies*, 22(6):2201–2238.

Chordia, T., Roll, R., and Subrahmanyam, A. (2001). Market Liquidity and Trading Activity.

Chordia, T., Subrahmanyam, A., and Anshuman, V. R. (2000). Trading Activity and Expected Stock Returns. *Journal of Financial Economics*, 59:3–32.

Clark, P. K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica*, 41:135–155.

Copeland, T. E. (1976). A Model of Asset Trading under the Assumption of Sequential Information Arrivals. *The Journal of Finance*, 31:1149–1168.

Copeland, T. E. (1977). A Probability Model of Asset Trading. *Journal of Financial and Quantitative Analysis*, 12:563–579.

Crouch, R. L. (1970). The Volume of Transactions and Price Changes on the New York Stock Exchange. *Financial Analysts Journal*, 26:104–109.

Darolles, S. and Fol, G. L. (2005). Trading Volume and Arbitrage. *Working Paper, CREST-INSEE.*

Datar, V. T., Naik, N. Y., and Radcliffe, R. (1998). Liquidity and Stock Returns: An Alternative Test. *Journal of Financial Markets*, 1:203–219.

Diamond, D. W. and Verrechia, R. E. (1987). Constraints on Short-Selling and Asset Price Adjustment to Private Information. *Journal of Financial Economics*, 18:277–311.

Domowitz, I. and Wang, J. (1994). Auctions as Algorith: Computerized Trade Execution and Price Discovery. *Journal of Economic Dynamics and Control*, 18:29–60.

Easley, D. and O'Hara, M. (1987). Price, Trade Size and Information in Security Markets. *Journal of Financial Economics*, 19:69–90.

Easley, D., O'Hara, M., and Paperman, J. B. (1996). Liquidity, Information and Infrequently Traded Stocks. *Journal of Finance*, 51:1405–1436.

Epps, T. W. and Epps, M. L. (1976). The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture of Distribution Hypothesis. *Econometrica*, 44:305–321.

Foster, D. F. and Viswanathan, S. (1990). A Theory of Intraday Variations in Volumes, Variances and Trading Costs in Securities Markets. *The Review of Financial Studies*, 3:593–694.

Foster, D. F. and Viswanathan, S. (1993). Variation in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models. *The Journal of Finance*, 48(1):187–211.

Gallant, R. A., Rossi, P. E., and Tauchen, G. (1992). Stock Prices and Volume. *The Review of Financial Studies*, 5(2):199–242.

Getmansky, M., Lo, A. W., and Makarov, I. (2004). An Econometric Analysis of Serial Correlation and Illiquidity in Hedge-Fund Returns. *Journal of Financial Economics*, 74:529–609.

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

Gourieroux, C. and Fol, G. L. (1998). Effets des Modèles de Négociation sur les Echanges. *Revue Economique*, 49:795–808.

Grossman, S. J. and Miller, M. H. (1988). Liquidity and Market Structure. *The Journal of Finance*, 43(3):617–633.

Hansen, L. P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50:1029–1054.

Harris, L. (1982). A Theoretical and Empirical Analysis of the Distribution of Speculative Prices and of the Relation between Absolute Price Change and Volume. *Unpublished Ph.D. Dissertation, University of Chicago.*

Harris, L. (1986). Cross-Security Tests of the Mixture of Distribution Hypothesis. *Journal of Financial and Quantitative Analysis*, 21:39–46.

Harris, L. (1987). Transaction Date Tests of the Mixture of Distribution Hypothesis. *Journal of Financial and Quantitative Analysis*, 22:127–141.

Heimstra, C. and Jones, J. (1994). Testing for Linear and Nonlinear Granger Causality in the Stock Price - Volume Relation. *Journal of Finance*, 49:1639–1664.

Karpoff, J. (1987). The Relation Between Price Changes and Trading Volume: A Survey. *Journal of Financial and Quantitative Analysis*, 22:109–126.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, 53:1315–1336.

Lamoureux, C. and Lastrapes, W. (1994). Endogenous Trading Volume and Momentum in Stock-Return Volatility. *Journal of Business and Economic Statistics*, 12:253–260.

Nagel, S. (2009). Evaporating Liquidity. *Working Paper NBER*.

Newey, W. K. and West, K. D. (1987). Hypothesis Testing with Efficient Method of Moments Estimation. *International Economic Review*, 28:777–787.

Richardson, M. and Smith, T. (1994). A Direct Test of the Mixture of Distributions Hypothesis: Measuring The Daily Flow of Information. *The Journal of Financial and Quantitative Analysis*, 29(1):101–116.

Rogalski, R. J. (1978). The Dependence of Prices and Volume. *Review of Economics and Statistics*, 60:268–274.

Tauchen, G. and Pitts, M. (1983). The Price Variability-Volume Relationship on Speculative Markets. *Econometrica*, 51(2):485–505.

Westerfield, R. (1977). The Distribution of Common Stock Price Changes: An Application of Transaction Time Subordinated Stochastic Models. *The Journal of Financial and Quantitative Analysis*, 12(5):743–765.

Ying, C. C. (1966). Stock Market Prices and Volumes of Sales. *Econometrica*, 34:676–685.

Appendices

A The MDH model of Tauchen and Pitts (1983)

The economy of Tauchen and Pitts (1983), henceforth TP, comprises a single risky asset and J active participants who trade in response to information arrival to the market. Each trading day consists of a series of intraday sucessive equilibria initiated by information shocks. The number of intraday equilibria I_t is random which drives the variability of price changes and traded volume.

Let Q_{ij} be the quantity that the trader j ($j = 1, ..., J$) is willing to trade at the ith intra-day equilibrium ($i = 1, ..., I_t$). Q_{ij} is then given by the linear relation:

$$Q_{ij} = a[P^*_{ij} - P_i], \quad (j = 1, 2, ..., J), \tag{A.1}$$

where $a > 0$ is a constant, P^*_{ij} is the reservation price of trader j at the intra-day equilibrium i and P_i is the current market price[30]. The reservation price heterogeneity among traders comes from different expectation about the future liquidation value $\tilde{P}$, as well as different needs to transfer the risk through the market. The ith peace of information hitting the market will result in a price increment ΔP_i and a corresponding traded volume V_i.

In particular, the market clearing condition $\sum_{j=1}^{J} Q_{ij} = 0$ and equation (A.1) yield the ith equilibrium price:

$$P_i = \frac{1}{J}\sum_{j=1}^{J} P^*_{ij}. \tag{A.2}$$

[30]Note that transaction costs are not considered in equation (A.1); the model assumes that the traders differ only in their reservation prices.

From the market clearing condition and equation (A.1), it follows that:

$$\Delta P_i = \frac{1}{J}\sum_{j=1}^{J} \Delta P_{ij}^*, \tag{A.3}$$

and

$$V_i \equiv \frac{1}{2}\sum_{j=1}^{J} |Q_{ij} - Q_{i-1,j}| = \frac{\alpha}{2}\sum_{j=1}^{J} |\Delta P_{ij}^* - \Delta P_i|, \tag{A.4}$$

where ΔP_{ij}^* is the increment of the jth trader reservation price.

TP make some additional assumptions concerning the distribution of trader's reservation price increments in order to obtain testable implications of the model. They assume a variance-component model:

$$\Delta P_{ij}^* = \phi_i + \psi_{ij}, \tag{A.5}$$
$$with \quad \phi_i \sim N(0, \sigma_\phi^2), \;\; \psi_{ij} \sim N(0, \sigma_\psi^2),$$

where ϕ and ψ are mutually independent both across traders and through time. Note that, ϕ_i is common to all traders and represents common variations of equilibrium price in response to new information. ψ_{ij} is supposed to be the trader-specific component of price increment related to trader subjectif interpretation of new information. The higher the absolute value of ϕ_i relative to ψ_{ij}, the higher the signal-to-noise ratio concerning information inflow. Using equations (A.3)-(A.5), ΔP_i and V_i can be written as:

$$\Delta P_i = \phi_i + \bar{\psi}_i, \tag{A.6}$$
$$V_i = \frac{\alpha}{2}\sum_{j=1}^{J} \mid \psi_{ij} - \bar{\psi}_i \mid, \tag{A.7}$$

where $\bar{\psi}_i = \frac{1}{J}\sum_{j=1}^{J} \psi_{ij}$.

From normality assumption for ϕ_i and ψ_{ij} as well as equations (A.6)-(A.7), it follows show that: (i) Intraday price change ΔP_i is normally distributed: $\Delta P_i \sim N(\mu_p, \sigma_p^2)$; ($ii$) Intraday traded volume V_i is approximately normally distributed for large J: $V_i \sim N(\mu_v, \sigma_v^2)$; ($iii$) ΔP_i and V_i are stochastically independent and their first two moments are[31]:

$$\begin{aligned}
\mu_p &\equiv E[\Delta P_i] = 0, \\
\sigma_p^2 &\equiv Var[\Delta P_i] = \sigma_\phi^2 + \frac{\sigma_\psi^2}{J}, \\
\mu_v &\equiv E[V_i] = \left(\frac{\alpha}{2}\right)\sigma_\psi^2\sqrt{\frac{2}{\pi}}\left(\sqrt{\frac{J-1}{J}}\right)J \\
\sigma_v^2 &\equiv Var[\Delta V_i] = \left(\frac{\alpha}{2}\right)^2\sigma_\psi^2\left(1-\frac{2}{\pi}\right)J + o(J).
\end{aligned} \tag{A.8}$$

Daily price change ΔP_t and trading volume V_t are obtained by summing their within-day counterparts ΔP_i and V_i:

$$\Delta P_t = \sum_{i=1}^{I_t} \Delta P_i, \quad \Delta P_i \sim N(0, \sigma_p^2), \tag{A.9}$$

$$V_t = \sum_{i=1}^{I_t} V_i, \quad V_i \sim N(\mu_v, \sigma_v^2). \tag{A.10}$$

Both ΔP_t and V_t are mixtures of independent normals with the same mixing variable I_t. Conditional on I_t, the bivariate normal mixture is:

$$\begin{aligned}
\Delta P_t &= \sigma_p\sqrt{I_t}Z_{1t}, \\
V_t &= \mu_v I_t + \sigma_v\sqrt{I_t}Z_{2t},
\end{aligned} \tag{A.11}$$

where Z_{1t} and Z_{2t} are i.i.d. standard normal variables and mutually independent. At the end of the day t, all the incoming information is incorporated into the price change ΔP_t and traded volume V_t.

[31]Point (i) is trivial. For more details and proofs of (ii) and (iii), see TP (1983), page 490-91.

Using a lognormal distribution for I_t and the maximum likelihood method, TP show that the standard MDH model captures the positive relationship between price change variance and volume on the 90-day T-bills futures market. Richardson and Smith (1994) extend TP work by introducing a mean parameter for daily price change and use GMM tests to validate the model. In this paper, we use Richardson and Smith (1994) version when estimating the standard MDH model for robustness checks (see section 4).

B The Grossman and Miller (1988) model

Grossman and Miller (1988), henceforth GM, consider a simple world with only three dates. Date 1 and date 2 are trading dates, while date 3 is used only as a terminal condition. There are only two assets in the GM economy: (i) a risky asset whose liquidation value at date 3 is $\tilde{P}_3$ and (ii) a risk-free asset whose return is normalized to zero. GM consider two types of traders, the outside customers who trade in response to information inflow, and the market makers who trade in response to liquidity shocks. In our framework, the outside customers are called active traders as in TP. Moreover, the market makers of GM correspond exactly to our liquidity arbitragers: they provide liquidity when it is needed in order to cash the liquidity premium.

Information concerning $\tilde{P}_3$ is assumed to arrive before trade at period 1 and before trade at period 2. Let J be the number of all the potential active traders in the market. The active trader j $(j = 1, ..., J)$ at time 1 has an endowment of size z_j in the security which is unsuitable given the trade-off between his risk preferences and information at that date. At period 1, some liquidity frictions arise because of asynchronization of time of trade among the active traders. This will result in a temporary order imbalance of magnitude z given by:

$$z = \sum_{j=1}^{J'} z_j \neq 0, \quad J' < J, \tag{B.1}$$

where $J^{'}$ is the number of active traders being present in the market at date 1. If all the active participants were present in the market at date 1, the order imbalance would vanish and the net trading demand would be zero:

$$\sum_{j=1}^{J} z_j = 0. \tag{B.2}$$

In the GM world, a liquidity event occurs at date 1 which motivates the liquidity arbitragers to enter the market in order to provide immediacy and thus compensate for the order disproportion; they liquidate their positions at date 2 as other active traders arrive with the opposite order imbalance. At date 2, the remaining active participants arrive with the opposite aggregated endowment shock which, by definition, cancels out the time-1-order imbalance. This assumption is crucial to discerning the advantages for the active traders arriving at date 1 to postpone their trades to date 2.

Let B_s be the cash-position of the active trader j at date s ($s = 1, 2$) and $\bar{Q}_s$ be the quantity of the risky asset he holds after trading at time s:

$$\bar{Q}_s = Q_s + z_j, \tag{B.3}$$

where Q_s is trader's excess demand. Using exponential preferences:

$$U(W_3) = -e^{-\alpha W_3}, \tag{B.4}$$

and backward induction, we can obtain the optimal excess demand at period s ($s = 1, 2$) by maximizing the expected utility of terminal wealth W_3:

$$E_s U(W_3) = E_s(-e^{-\alpha W_3}), \tag{B.5}$$

under (i) the normality assumption concerning $\tilde{P}_1$, $\tilde{P}_2$ as well as $\tilde{P}_3$, and (ii) the following budget constraints:

$$W_3 = B_2 + \bar{Q}_3 \tilde{P}_3, \tag{B.6}$$

$$\tilde{P}_2 \bar{Q}_2 + B_2 = W_2 = B_1 + \tilde{P}_2 \bar{Q}_1, \tag{B.7}$$

$$\tilde{P}_1 \bar{Q}_1 + B_1 = W_1 = \tilde{P}_1 z_j + W_0, \tag{B.8}$$

where W_0 represents other wealth possessed by the active participant before trade at date 1.

In particular, date-2-participants of the GM world consist of: (i) the active traders who arrived in the market at date 1; (ii) the active traders arriving at date 2 with opposite order imbalance, as well as (iii) the liquidity arbitragers willing to liquidate the positions taken at date 1. At date 2, the maximization program for active trader j belonging to the first group can be written as:

$$\max_{Q_2} E_2 U(W_2 - P_2 z_j + (\tilde{P}_3 - P_2) Q_2 + \tilde{P}_3 z_j), \tag{B.9}$$

where $W_3 = W_2 - P_2 z_j + (\tilde{P}_3 - P_2) Q_2 + \tilde{P}_3 z_j$ is deduced by equations (B.3) and (B.6)-(B.8). Solving for Q_2 yields the optimal excess demand denoted by Q_2^{at}:

$$Q_2^{at} = \frac{E_2 \tilde{P}_3 - P_2}{\alpha Var_2 \tilde{P}_3} - z_j, \tag{B.10}$$

where mean and variance operators reflect the information available at date 2. Assuming that active traders differ only concerning z_j and from linearity between Q_2^{at} and z_j, Q_2^{at} corresponds to the aggregated optimal excess demand across active traders when z_j is replaced by z in equation (B.10):

$$Q_2^{at} = \frac{E_2 \tilde{P}_3 - P_2}{\alpha Var_2 \tilde{P}_3} - z. \tag{B.11}$$

In the same way, the aggregated optimal excess demand of active traders arriving at date

2 with opposite order imbalance is given by:

$$\frac{E_2\tilde{P}_3 - P_2}{\alpha Var_2\tilde{P}_3} + z. \tag{B.12}$$

Assuming that there are M liquidity arbitragers in the market having the same preferences as the active traders except that for them the endowment shock is zero, their total optimal excess demand at date 2 is given by:

$$MQ_2^{la} = M\frac{E_2\tilde{P}_3 - P_2}{\alpha Var_2\tilde{P}_3}, \tag{B.13}$$

where Q_2^{la} is the optimal excess demand per liquidity arbitrager.

Given the excess demand functions (B.11), (B.12) and (B.13), the market clearing condition at date 2 can be written as:

$$\frac{E_2\tilde{P}_3 - P_2}{\alpha Var_2\tilde{P}_3} - z + \frac{E_2\tilde{P}_3 - P_2}{\alpha Var_2\tilde{P}_3} + z + M\frac{E_2\tilde{P}_3 - P_2}{\alpha Var_2\tilde{P}_3} = 0, \tag{B.14}$$

which implies that:

$$P_2 = E_2\tilde{P}_3. \tag{B.15}$$

It follows that at the equilibrium we get:

$$Q_2^{at} = -z, \tag{B.16}$$

$$Q_2^{la} = 0. \tag{B.17}$$

At date 1, the active participants who are willing to trade maximize the expected utility depending on date-1 information. Note that the risk at date 1 comes from the fact that new information may arrive at date 2 causing $P_2 = E_2\tilde{P}_3$ to be different from $E_1\tilde{P}_3$. From (B.5), (B.6)-(B.8), (B.15) and (B.16), we get:

$$\max_{Q_1} E_1 U(W_0 + Q_1(E_2\tilde{P}_3 - P_1) + zE_2\tilde{P}_3), \tag{B.18}$$

which yields the optimal aggregated excess demand Q_1^{at} as given by:

$$Q_1^{at} = \frac{E_1\tilde{P}_3 - P_1}{\alpha Var_1(E_2\tilde{P}_3)} - z, \tag{B.19}$$

where $E_1E_2\tilde{P}_3 = E_1\tilde{P}_3$, as implied by the law of iterated expectations.

Liquidity arbitragers, who continually observe the market, provide immediacy at date 1 by taking trading positions that they hold until date 2. In the same way as for active participants, the optimal excess demand per liquidity arbitrager is given by:

$$Q_1^{la} = \frac{E_1\tilde{P}_3 - P_1}{\alpha Var_1(E_2\tilde{P}_3)}. \tag{B.20}$$

The market clearing condition at period 1 gives:

$$\frac{E_1\tilde{P}_3 - P_1}{\alpha Var_1(E_2\tilde{P}_3)} - z + M\frac{E_1\tilde{P}_3 - P_1}{\alpha Var_1(E_2\tilde{P}_3)} = 0, \tag{B.21}$$

which yields the equilibrium price at time 1, P_1:

$$P_1 = E_1\tilde{P}_3 - \frac{z\alpha Var_1(E_2\tilde{P}_3)}{1+M}. \tag{B.22}$$

From (B.19), (B.20) and (B.22) we get the equilibrium excess demands for both time-1 market participants:

$$Q_1^{at} = -\frac{M}{1+M}z, \tag{B.23}$$

and

$$Q_1^{la} = \frac{z}{1+M}. \tag{B.24}$$

Note that P_1 depends on the magnitude and the sign of the order imbalance. When $z = 0$, P_1 equals the price revealing the information $E_1\tilde{P}_3$. For $z \neq 0$, the equilibrium price depends on the number of liquidity providers present in the market at date 1. The higher the number of liquidity arbitragers, the lower the order imbalance impact on the equilibrium price P_1. Let $\tilde{R}_2 = \tilde{P}_2/P_1 - 1$ be the excess return earned by arbitragers at date 2. From (B.22) it follows that:

$$E_1\tilde{R}_2 = \frac{P_1 z}{1+M}\alpha Var_1(\tilde{R}_2). \tag{B.25}$$

Note that, if either $z \to 0$ or $M \to \infty$, $E_1\tilde{R}_2 = 0$. This means that the combined effect of the order asynchronization and the finite number of liquidity arbitragers results in departures of $E_1\tilde{R}_2$ from zero.

Finally, GM assume that liquidity arbitragers face an exogenous cost of maintaining a market presence – denoted by c – and that the order imbalance z is not known when this cost is paid out. Supposing that z behaves as a centered normally distributed variable which is independent from information shocks, the expected utility for a given liquidity arbitrager is[32]:

$$EU(W_3) = EU(W_0 - c + (\tilde{P}_2 - P_1)Q_1^{la}). \tag{B.26}$$

It follows that, arbitragers will be motivated to enter the market until the transaction costs

[32]From (B.17), it follows that the profit between period 2 and period 3, $(\tilde{P}_3 - P_2)Q_2^{la}$ vanishes.

offset the expected profits between dates 1 and 2:

$$EU(W_0 - c + (\tilde{P}_2 - P_1)Q_1^{la}) = EU(W_0). \tag{B.27}$$

Using (B.20) and exponential utility for (B.27), it can be shown that M is a decreasing function of c[33]. When $c > 0$ the number of arbitragers is finite. This result is critical to understanding the benefits, for liquidity arbitragers, of providing immediacy at date 1.

Generally speaking, the GM framework focuses on the consequences of an order imbalance on the intraday patterns of price change and transaction volume. At this stage, the model shows that in the presence of liquidity frictions and exogenous transaction costs:

(i) The traded volume at date 1 is lower than it would have been if there were no order imbalance[34].

(ii) The transaction price at date 1 deviates from its revealing information level ($P_1 \neq E_1\tilde{P}_3$).

However, from the assumptions that the order imbalance sums to zero across periods 1 and 2, and that the liquidity arbitragers offset their positions at date 2, it follows that the traded volume across dates 1 and 2 is higher than it would have been in the absence of liquidity frictions if the condition $M \geq 1$ is verified[35]. This reasoning motivates us to extend the GM framework in order to model the impact of liquidity frictions on total price changes and total traded volume.

[33] For a detailed demonstration see Grossman and Miller (1988).

[34] From (B.23) it follows that a finite M implies $|Q_1^{at}| < |z|$.

[35] In other words, the order imbalance faced by outside customers who exchange at date 1 is offset thanks to immediacy provided by market makers who will liquidate their positions at date 2 and thus increase the traded volume.

C The GMM estimations

C.1 The GMM procedure of Hansen (1982) with Newey and West (1987) weighting matrix

Let $X_t = (R_t, V_t)$ be the vector of return and volume observations prevailing at day t for a given stock and θ be the $(N_p \times 1)$ vector of the MDHL model parameters: $\theta = (\mu_v^{at}, \mu_v^{la}, \sigma_p^2, \sigma_v^2, m_{2I}, p)$. If X_t is generated by the MDHL model, there is some true set of parameters θ_0 for which:

$$E[h_t(X_t, \theta_0)] = 0, \tag{C.1}$$

where h_t is a column vector of N_h unconditional moment conditions implied by our model. Since we do not observe the true expectation of h_t, we define a vector $g_T(\theta)$ containing the sample averages corresponding to the elements of h_t. We index the vector of sample moments by T to indicate its dependence on the sample size. For large T, if X_t is generated by the MDHL model, $g_T(\theta_0)$ should be close to zero:

$$g_T(\theta_0) \equiv \frac{1}{T}\sum_{t=1}^{T} h_t(X_t, \theta_0) \longrightarrow 0, \;\; when \;\; T \to \infty. \tag{C.2}$$

In this paper, we work with an overidentified system, i.e., $N_h > N_p$, which allows us to estimate θ and test the global validity of the MDHL model simultaneously. In this case, the GMM chooses $\hat{\theta}_T$ as the value of θ that minimizes the quadratic form of $g_T(\theta)$ given by:

$$Q_T(\theta) \equiv g_T(\theta)' W_T g_T(\theta), \tag{C.3}$$

where W_T is an (N_h x N_h) symmetric-positive-definite-weighting matrix. Since the problem is nonlinear, this minimization is performed numerically. The first order condition is:

$$D_T(\hat{\theta}_T)'W_T g_T(\hat{\theta}_T) = 0, \tag{C.4}$$

where $D_T(\hat{\theta})$ is the sample approximation of the true partial derivative matrix and is given by:

$$D_T(\hat{\theta}_T) = \partial g_T(\hat{\theta}_T)/\partial\hat{\theta}'_T. \tag{C.5}$$

The asymptotic distribution of the coefficient estimate is:

$$\sqrt{T}(\hat{\theta}_T - \theta) \sim^{asy} N(0, V), \tag{C.6}$$

where V is its asymptotic covariance matrix. An important point of the GMM analysis is to pick a weighting matrix W_T that minimizes V and hence deliver an asymptotically efficient estimator. In this article, we use the Newey and West (1987) methodology to estimate the optimal weighting matrix denoted by S_T. The Newey-West estimator accounts for serial correlation and heteroskedasticity among the terms of the matrix h_t and is given by:

$$S_T(q, \hat{\theta}_T) = \Gamma_{0,T}(\hat{\theta}_T) + \sum_{j=1}^{q} \left(\frac{q-j}{q}\right) \left(\Gamma_{j,T}(\hat{\theta}_T) + \Gamma'_{j,T}(\hat{\theta}_T)\right), \tag{C.7}$$

where q is the number of autocovariances one wishes to include in the computation and $\Gamma_{j,T}(\hat{\theta}_T)$ is the sample autocovariance matrix of h_t as given by:

$$\Gamma_{j,T}(\hat{\theta}_T) \equiv T^{-1} \sum_{t=j+1}^{T} h_t(\hat{\theta}_T) h_t(\hat{\theta}_T)'. \tag{C.8}$$

Finally, Hansen (1982) provides an overidentifying test statistic $J_T(\hat{\theta})$ as follows:

$$J_T(\hat{\theta}) \equiv T g_T(\hat{\theta}_T) S_T^{-1}(q, \hat{\theta}_T) g_T(\hat{\theta}_T) \sim^{asy} \chi^2_{N_h - N_p},$$

which allows us to test the global validity of the model.

C.2 Sample moment conditions for the MDHL model

The sample moment conditions in equation (4.6) are given as follows:

$$
\begin{aligned}
(V_t - E(V_t)) &= 0, && \text{(C.9)}\\
(R_t - E(R_t))^2 &= \sigma_p^2, && \text{(C.10)}\\
(V_t - E(V_t))^2 &= (\mu_v^{at})^2 m_{2I} + (\mu_v^{la})^2[p(1-p) + p^2 m_{2I}] && \text{(C.11)}\\
&\quad + 2\mu_v^{at}\mu_v^{la} p m_{2I} + \sigma_v^2,\\
(R_t^2 - E(R_t^2))(V_t - E(V_t)) &= \sigma_p^2(\mu_v^{at} + p\mu_v^{la}) m_{2I}, && \text{(C.12)}\\
(R_t^2 - E(R_t^2))(V_t^2 - E(V_t^2)) &= (\mu_v^{at})^2\sigma_p^2(m_{3I} + 2m_{2I}) && \text{(C.13)}\\
&\quad + (\mu_v^{la})^2\sigma_p^2[p^2 m_{3I} + p(1-p)m_{2I}]\\
&\quad + 2\mu_v^{at}\mu_v^{la}\sigma_p^2 p(m_{3I} + 2m_{2I}) + \sigma_v^2\sigma_p^2 m_{2I},\\
(V_t - E(V_t))^3 &= 3\mu_v^{at}\sigma_v^2 m_{2I} + (\mu_v^{at})^3 m_{3I} && \text{(C.14)}\\
&\quad + (\mu_v^{la})^3[p^3 m_{3I} + 3p^2(1-p)m_{2I} + p(1-3p+2p^2)]\\
&\quad + 3\mu_v^{la}\sigma_v^2 p m_{2I} + 3(\mu_v^{at})^2\mu_v^{la} p m_{3I}\\
&\quad + 3\mu_v^{at}(\mu_v^{la})^2[p^2 m_{3I} + p(1-p)m_{2I}],\\
(R_t - E(R_t))^4 &= 3\sigma_p^4(m_{2I} + 1), && \text{(C.15)}\\
(V_t - E(V_t))^4 &= (\mu_v^{la})^4[p^4 m_{4I} + (6p^3 - 6p^4)m_{3I}] && \text{(C.16)}\\
&\quad + (\mu_v^{la})^4[(4p^2 - 6p^3 + 2p^4)m_{2I} + (p - 7p^2 + 12p^3 - 6p^4)]\\
&\quad + 6(\mu_v^{at})^2(\mu_v^{la})^2[p^2 m_{4I} + p(1-p)(m_{3I} + m_{2I})]\\
&\quad + 6(\mu_v^{at})^2\sigma_v^2(m_{3I} + m_{2I}) + 4(\mu_v^{at})^3\mu_v^{la} p m_{4I}\\
&\quad + 4\mu_v^{at}(\mu_v^{la})^3[p^3 m_{4I} + 3(p^2 - p^3)m_{3I} + (p - p^3)m_{2I}]\\
&\quad + 6(\mu_v^{la})^2\sigma_v^2[p^2 m_{3I} + p m_{2I} + p(1-p)] + (\mu_v^{at})^4 m_{4I}\\
&\quad + 12\mu_v^{at}\mu_v^{la}\sigma_v^2 p(m_{3I} + m_{2I}) + 3\sigma_v^4(m_{2I} + 1),\\
(R_t - E(R_t))^2(V_t - E(V_t))^2 &= (\mu_v^{at})^2\sigma_p^2(m_{3I} + m_{2I}) + \sigma_v^2\sigma_p^2(m_{2I} + 1) && \text{(C.17)}\\
&\quad + (\mu_v^{la})^2\sigma_p^2[p^2 m_{3I} + p m_{2I} + p(1-p)] + 2\mu_v^{at}\mu_v^{la}\sigma_p^2 p(m_{3I} + m_{2I}).
\end{aligned}
$$

Equations (C.9) to (C.17) correspond to sample moment conditions (1) to (9) in (4.6). The third and fourth central moments of I_t, m_{3I} and m_{4I}, are functions of its respective second central moment m_{2I} as given in equation (4.5). Note that, the central moments of L_t being functions of p and the central moments of I_t, need not to be estimated. Finally, expectation operators of the observables are also functions of θ:

$$\begin{aligned} E(R_t) &= 0, \\ E(V_t) &= \mu_v^{at} + p\mu_v^{la}, \\ E(R_t^2) &= \sigma_p^2, \\ E(V_t^2) &= \sigma_v^2 + 2\mu_v^{at}\mu_v^{la}p(m_{2I}+1) \\ &\quad +(\mu_v^{at})^2(m_{2I}+1) + (\mu_v^{la})^2(p+p^2m_{2I}). \end{aligned} \tag{C.18}$$

D GMM estimation results for MDHL model

ID	μ_v^{at}	μ_v^{la}	σ_p^2	σ_v^2	m_{2I}	x	p	χ_3^2	$p\mu_v^{la}$ (%)
1	0,007040**	0,014930*	0,000325**	0,00000000	0,173**	6,09**	0,002	2,78	0,47
2	0,006095**	0,024615**	0,000075**	0,00000000	0,327**	5,07**	0,006	2,95	2,46
3	0,006440**	0,050036**	0,000133**	0,00000000	0,487**	5,56**	0,004	3,23	2,89
4	0,016193**	0,035676	0,000377**	0,00000000	0,239**	6,89	0,001	7,81	-
5	0,004827**	0,007781	0,000127**	0,00000000	0,207**	4,02	0,018	3,68	-
6	0,005632**	0,000000	0,000110**	0,00000000	0,192**	4,11**	0,016	10,57	-
7	0,008299**	0,080822**	0,000180**	0,00000183	0,234**	6,09**	0,002	3,48	2,16
8	0,005706**	0,012923**	0,000137**	0,00000000	0,256**	4,27**	0,014	2,25	3,04
9	0,005497**	0,024871**	0,000105**	0,00000143	0,228**	4,95**	0,007	0,28	3,08
10	0,013645**	0,000005	0,000217**	0,00000313	0,271**	4,47**	0,011	15,93	-
11	0,010347**	0,030288**	0,000235**	0,00001010**	0,212**	4,00**	0,018	1,76	5,01
12	0,004579**	0,023868**	0,000202**	0,00000029	0,194**	7,89**	0,000	2,21	0,20
13	0,007147**	0,020656*	0,000160**	0,00000302	0,223**	4,38**	0,012	5,39	3,44
14	0,005174**	0,017671	0,000348**	0,00000000	0,211**	6,95*	0,001	7,43	-
15	0,005821**	0,019152**	0,000129**	0,00000000	0,157**	6,66**	0,001	3,24	0,42
16	0,011817**	0,067002**	0,000102**	0,00000000	0,332**	5,94**	0,003	8,30	-
17	0,005713**	0,000000	0,000121**	0,00000000	0,219**	8,36**	0,000	8,95	-
18	0,006297**	0,000000**	0,000102**	0,00000000	0,316**	15,41**	0,000	1,11	0,00
19	0,002458**	0,016717	0,000159**	0,00000000	0,322**	9,52**	0,000	4,56	-
20	0,006078**	0,009091**	0,000155**	0,00000000	0,242**	3,45**	0,031	1,97	4,40
21	0,009133**	0,021376	0,000146**	0,00000000	0,367**	4,52	0,011	6,23	-
22	0,006114	0,010793	0,000120**	0,00000000	0,258**	2,87	0,054	3,83	-
23	0,012486**	0,024083	0,000197**	0,00000000	0,364**	5,72	0,003	6,03	-
24	0,005370**	0,010659**	0,000063**	0,00000072	0,130**	3,85**	0,021	4,84	3,97
25	0,005098**	0,019464**	0,000119**	0,00000000	0,352**	4,73**	0,009	2,19	3,25
26	0,011474	0,014244	0,000141**	0,00000200	0,259	3,98	0,018	6,87	-
27	0,010676**	0,000000**	0,000246**	0,00000000	0,310**	15,53**	0,000	6,49	0,00
28	0,009013**	0,025501	0,000136**	0,00000000	0,347**	7,09	0,001	1,14	-
29	0,007865**	0,064835	0,000161**	0,00000000	0,486**	6,72**	0,001	2,70	-
30	0,004305	0,024855	0,000104	0,00000000	0,212	6,16	0,002	2,57	-
31	0,004146**	0,008845**	0,000101**	0,00000000	0,223**	4,00**	0,018	5,11	3,70
32	0,006383**	0,016126**	0,000179**	0,00000000	0,389**	2,85**	0,054	2,52	12,10
33	0,007054**	0,000000	0,000182**	0,00000000	0,442**	6,65**	0,001	2,92	-
34	0,003395**	0,006673**	0,000045**	0,00000067	0,120**	4,30**	0,013	7,55	2,56
35	0,007926**	0,056664**	0,000264**	0,00000945**	0,373**	5,24**	0,005	2,57	3,62
36	0,008806**	0,000000	0,000188**	0,00000000	0,320**	5,05**	0,006	12,55	-
37	0,009513**	0,000000	0,000147**	0,00000000	0,388**	7,18**	0,001	4,89	-
38	0,007438	0,026125	0,000122	0,00000415	0,219	4,72	0,009	3,81	-
39	0,005334**	0,019771**	0,000087**	0,00000000	0,237**	4,91**	0,007	4,89	2,63
40	0,008118**	0,006594	0,000192**	0,00000000	0,231**	2,96	0,049	13,08	-
41	0,011131**	0,000008**	0,000176**	0,00000000	0,315**	13,45**	0,000	7,95	-
42	0,009672	0,000000	0,000284**	0,00000000	0,332**	0,71**	0,328	6,00	-
43	0,006575**	0,000000	0,000146**	0,00000000	0,299**	9,94**	0,000	5,24	-
44	0,006229**	0,013975**	0,000120**	0,00000026	0,258**	3,84**	0,021	1,42	4,50
45	0,009898**	0,045066**	0,000162**	0,00000000	0,364**	4,87**	0,008	4,38	3,36
46	0,005729**	0,013238**	0,000139**	0,00000000	0,236**	4,28**	0,014	9,71	-

"*" and "**" indicate significance at 90% and 95% levels of confidence respectively.

Table 2: MDHL model estimated parameters (1)

ID	μ_v^{at}	μ_v^{la}	σ_p^2	σ_v^2	m_{2I}	x	p	χ_3^2	$p\mu_v^{la}$ (%)
47	0,005670**	0,039195**	0,000166**	0,00000029	0,299**	5,84**	0,003	0,30	1,96
48	0,005226**	0,008135	0,000119	0,00000124	0,270**	3,34	0,034	4,17	-
49	0,005967**	0,007442	0,000088	0,00000000	0,236**	3,12	0,042	4,70	-
50	0,012743**	0,097959**	0,000364**	0,00000000	0,337**	6,83**	0,001	3,81	0,82
51	0,011330**	0,072937**	0,000175**	0,00000000	0,480**	5,17**	0,006	2,92	3,52
52	0,008719**	0,000000	0,000124**	0,00000000	0,413**	4,01**	0,018	11,47	-
53	0,009159**	0,000000	0,000143**	0,00000000	0,468**	6,73**	0,001	8,04	-
54	0,004448**	0,007292	0,000095**	0,00000000	0,225**	3,19	0,040	4,86	-
55	0,008263**	0,004292	0,000131**	0,00000000	0,335**	1,75**	0,148	6,85	-
56	0,012171	0,003088	0,000128**	0,00000000	0,285**	-0,28	0,570	9,09	-
57	0,007277**	0,042687**	0,000219**	0,00000070	0,401**	4,76**	0,008	1,30	4,74
58	0,006668**	0,011587	0,000187	0,00000000	0,218**	5,12	0,006	6,28	-
59	0,007672**	0,022692**	0,000256**	0,00000000	0,276**	5,43**	0,004	1,32	1,17
60	0,007863**	0,011006	0,000113**	0,00000000	0,269**	3,22	0,039	1,32	-
61	0,013258**	0,115264**	0,000175**	0,00000000	0,397**	5,79**	0,003	3,01	2,57
62	0,005446**	0,007457	0,000102	0,00000000	0,237**	4,38	0,012	7,01	-
63	0,004559**	0,030197**	0,000079**	0,00000000	0,141**	6,33**	0,002	9,95	-
64	0,001733**	0,005458**	0,000120**	0,00000000	0,244**	4,50**	0,011	3,31	3,36
65	0,006773**	0,018750**	0,000099	0,00000000	0,236**	5,20**	0,005	5,49	1,37
66	0,007098**	0,015138	0,000104**	0,00000000	0,335**	3,71**	0,024	4,43	-
67	0,005627	0,002416	0,000291**	0,00000000	0,142**	0,18	0,454	13,08	-
68	0,010011**	0,000003	0,000194**	0,00000165	0,254**	4,10**	0,016	2,10	-
69	0,008553**	0,014322	0,000177**	0,00000300	0,365**	8,82**	0,000	6,84	-
70	0,007432	0,024748	0,000164	0,00000000	0,143**	2,13	0,106	5,55	-
71	0,010442**	0,082827	0,000166**	0,00000000	0,343	6,78**	0,001	4,33	-
72	0,003652**	0,000001**	0,000157**	0,00000000	0,264**	7,96**	0,000	4,91	0,00
73	0,011692**	0,000001	0,000099**	0,00000000	0,544**	7,09**	0,001	4,04	-
74	0,007525**	0,043429**	0,000121	0,00000319	0,341**	5,61**	0,004	1,97	2,26
75	0,004046**	0,011888**	0,000218	0,00000000	0,307**	4,48**	0,011	1,63	3,13
76	0,007274**	0,037411**	0,000176	0,00000000	0,316**	6,39**	0,002	2,79	1,02
77	0,005522**	0,010229**	0,000152**	0,00000000	0,278**	2,60**	0,069	1,18	11,32
78	0,010804**	0,086171**	0,000193**	0,00000000	0,357**	6,41**	0,002	5,62	1,30
79	0,008293**	0,015345	0,000125**	0,00000000	0,285**	3,65*	0,025	3,68	-
80	0,008059**	0,036491**	0,000120**	0,00000000	0,256	6,37**	0,002	9,28	-
81	0,005287**	0,008437**	0,000108**	0,00000150	0,157**	2,68**	0,064	2,86	9,28
82	0,006074**	0,032030**	0,000182**	0,00000000	0,391**	4,97**	0,007	1,23	3,50
83	0,007775**	0,047332**	0,000119**	0,00000382	0,228**	5,95**	0,003	1,57	1,56
84	0,005163**	0,015509**	0,000091**	0,00000000	0,227**	5,96**	0,003	10,72	-
85	0,002348**	0,000005**	0,000094**	0,00000012	0,334**	14,48**	0,000	3,53	-
86	0,007168**	0,000002	0,000076**	0,00000083	0,201**	5,17**	0,006	2,19	-
87	0,018066**	0,013050	0,000513**	0,00000024	0,398**	1,97	0,123	7,75	-
88	0,004953**	0,000001**	0,000158**	0,00000000	0,228**	15,31**	0,000	9,16	-
89	0,063894	0,000030	0,000229**	0,00170417	0,347**	0,03**	0,493	20,42	-
90	0,008200**	0,032734**	0,000118**	0,00000048	0,206**	6,08**	0,002	1,65	0,09
91	0,008935**	0,000000	0,000137**	0,00000000	0,491**	7,40**	0,001	4,37	-
92	0,026785**	0,091897**	0,000438**	0,00014244**	0,144**	3,71**	0,024	2,36	7,56
93	0,009267**	0,046187**	0,000148**	0,00000000	0,412	4,66**	0,009	4,29	4,48

"*" and "**" indicate significance at 90% and 95% levels of confidence respectively.

Table 3: MDHL model estimated parameters (2)

E GMM estimation to test the standard MDH model using Richardson and Smith (1994) procedure

ID	μ_p	μ_v	σ_p^2	σ_v^2	m_{2I}	m_{3I}	χ_3^2
1	0,001576**	0,006974**	0,000341**	0,000000	0,164821**	0,076466**	4,43
2	0,000293	0,006213**	0,000082**	0,000007	0,537526**	0,771725**	4,04
3	0,000459	0,006450	0,000124	0,000003	0,615128	1,182884	9,18
4	0,002793**	0,016191**	0,000348**	0,000015	0,168119**	0,086267**	7,89
5	0,000593	0,004944**	0,000117**	0,000000	0,251352**	0,200954*	6,16
6	0,000648	0,005961**	0,000141**	0,000013**	0,623407**	0,916119**	1,89
7	0,000992**	0,008348**	0,000198**	0,000008	0,370009**	0,002626	2,39
8	0,000304	0,005915**	0,000136**	0,000000	0,335995**	0,431232**	2,61
9	0,001127**	0,005421**	0,000104**	0,000001	0,273959**	0,000798	4,37
10	0,000924*	0,014019**	0,000277**	0,000043**	0,506739**	0,544113**	0,85
11	0,001128**	0,010572**	0,000247**	0,000020**	0,209113**	0,504891**	4,44
12	0,001227**	0,004564**	0,000203**	0,000000	0,201264**	0,081835	2,78
13	0,000790*	0,007445**	0,000185**	0,000008*	0,236387**	0,547731**	3,55
14	0,002101**	0,005181**	0,000354**	0,000002	0,218648**	0,422673**	7,01
15	0,000725	0,005650	0,000122	0,000000	0,142258	0,053039	8,15
16	0,000204	0,012450**	0,000116**	0,000020	0,626882**	1,485704**	2,77
17	0,000516	0,005783**	0,000136**	0,000004**	0,369925**	0,364957**	3,74
18	0,000308	0,006282**	0,000104**	0,000002	0,383053**	0,491836**	1,15
19	-0,000209	0,002463**	0,000163**	0,000001	0,462963**	0,457725**	2,70
20	0,000640	0,006393**	0,000152**	0,000002	0,261004**	0,290055**	2,57
21	0,000582	0,009301**	0,000138**	0,000003	0,358138*	0,402375**	7,50
22	0,000851**	0,006702**	0,000119**	0,000006	0,268992**	0,385132	3,77
23	0,001061*	0,012701**	0,000209**	0,000028	0,574036**	0,816102	7,33
24	0,000560**	0,005666**	0,000067**	0,000000	0,254055**	0,296119**	1,31
25	-0,000340	0,005180**	0,000103**	0,000018*	-0,213526	1,376168	4,97
26	0,000337	0,011794**	0,000171**	0,000062**	0,791139**	1,474336**	4,09
27	0,001466**	0,011202**	0,000271**	0,000015**	0,468301**	0,549208**	0,32
28	0,000753	0,008973**	0,000135**	0,000000	0,348388**	0,386472**	1,62
29	0,000160	0,008094**	0,000171**	0,000009	0,742010**	2,013111*	1,78
30	0,000213	0,004189**	0,000101**	0,000001	0,253895*	0,083571	4,78
31	0,000403	0,004333**	0,000110**	0,000001	0,338555**	0,400176**	1,81
32	0,000780	0,007169**	0,000187**	0,000011*	0,417406**	1,046547**	2,13
33	0,001097**	0,007001**	0,000178**	0,000018	0,360728*	10,347836	2,76
34	0,000092	0,003656**	0,000052**	0,000001	0,342485**	0,366668**	1,22
35	0,000013	0,008095	0,000271	0,000025	0,181488	0,632359	8,64
36	0,000806*	0,008736**	0,000207**	0,000013	0,505544**	0,636415*	5,52
37	0,000832*	0,009746**	0,000156**	0,000006	0,484987**	0,656846**	1,42
38	0,000898**	0,007722**	0,000143**	0,000002	0,315477**	0,334730**	3,73
39	0,000501	0,005558**	0,000091**	0,000001	0,392350**	0,757375**	3,18
40	0,001586**	0,008806**	0,000226**	0,000011**	0,448834**	0,447771**	0,68
41	-0,000652	0,010979	0,000175	0,000007	0,297440	0,761764	8,64
42	0,000931	0,009666**	0,000311**	0,000025	0,646960**	1,074764	1,94
43	0,000782*	0,006418**	0,000147**	0,000000	0,326012**	0,549908**	4,60
44	0,000667*	0,006578**	0,000119**	0,000005*	0,249873**	0,414520**	3,13
45	-0,001165**	0,009964**	0,000168**	0,000004	0,464029**	1,221262**	5,37

"*" and "**" indicate significance at 90% and 95% levels of confidence respectively.

Table 4: MDH model estimated parameters (1).

ID	μ_p	μ_v	σ_p^2	σ_v^2	m_{2I}	m_{3I}	χ_3^2
46	0,000308	0,006052**	0,000170**	0,000002	0,369355**	0,421853**	3,22
47	0,000359	0,005642**	0,000158**	0,000006	0,218251**	1,077850	2,74
48	0,000282	0,005360**	0,000126**	0,000002	0,304649**	0,321132**	6,57
49	0,000184	0,006046	0,000083	0,000001	0,235209	0,195631	9,36
50	0,002334**	0,013313**	0,000435**	0,000039*	0,612897**	1,071623**	1,61
51	0,001268**	0,012308**	0,000185**	0,000009	0,896280**	3,419218**	2,99
52	0,001181**	0,008821**	0,000129**	0,000015*	0,648002**	1,009654**	7,64
53	0,000696	0,009805**	0,000161**	0,000044**	1,140210**	3,849740**	7,01
54	0,000398	0,004967**	0,000107**	0,000003	0,266530**	0,480544**	3,24
55	0,000037	0,009627**	0,000139**	0,000010	0,518172**	0,794861*	3,86
56	0,000210	0,014942**	0,000170**	0,000108**	0,855629**	1,725279**	1,08
57	0,000429	0,007450**	0,000179**	0,000017**	0,429604**	2,473198**	12,86
58	0,000385**	0,006954**	0,000189**	0,000010**	0,491210**	0,718748**	6,44
59	0,000551	0,007730**	0,000253**	0,000004	0,229570**	0,237471**	2,13
60	0,000270	0,008247**	0,000112**	0,000001	0,327870**	0,370144**	2,49
61	0,000975*	0,013523**	0,000176**	0,000025	0,483996**	2,539710**	1,21
62	0,000763**	0,005644**	0,000105**	0,000001	0,310431**	0,257974**	6,28
63	0,000080	0,004747**	0,000093**	0,000004*	0,453089**	0,922374**	4,41
64	0,000421	0,001753**	0,000122**	0,000000	0,256621**	0,317692**	2,41
65	0,000399	0,007069**	0,000121**	0,000025**	0,807689**	1,646781**	1,44
66	0,000322	0,007495**	0,000113**	0,000006	0,532255**	0,735960**	3,05
67	0,002328	0,006563	0,000259	0,000004	0,081158	0,034710	13,13
68	0,001318*	0,010106**	0,000195**	0,000003	0,328610**	0,391733**	1,66
69	0,001144**	0,009141**	0,000197**	0,000016**	0,569769**	0,797812**	2,70
70	-0,000217**	0,011143**	0,000159**	0,000214**	0,015643	16,593616*	3,31
71	0,000831*	0,010479**	0,000168**	0,000007	0,539331**	1,714445	2,91
72	0,000684	0,003646**	0,000158**	0,000000	0,264416**	0,196433**	6,06
73	0,000830**	0,012073**	0,000104**	0,000082**	1,263259**	2,909492	2,95
74	0,000318	0,007707**	0,000123**	0,000005	0,415593**	0,990330**	2,07
75	0,000801	0,004110**	0,000216**	0,000001	0,329465**	0,565500**	3,21
76	0,000259	0,007227**	0,000179**	0,000003	0,358411**	0,352879	3,54
77	0,000297	0,005903**	0,000144**	0,000006	0,280831**	0,527107**	5,27
78	0,001124*	0,011166**	0,000217**	0,000021	0,626097**	1,111593	3,63
79	0,000378	0,008791**	0,000141**	0,000003	0,315523**	0,413985**	2,90
80	-0,000031	0,008275	0,000108	0,000022	0,003433	0,377380	10,43
81	0,000828**	0,005752**	0,000110**	0,000005**	0,165209**	0,225023**	1,37
82	0,000016	0,006422**	0,000177**	0,000012**	0,338610**	1,834843**	6,33
83	0,000415	0,007782**	0,000124**	0,000002	0,296313**	0,011812	3,64
84	0,000401	0,005160**	0,000101**	0,000004**	0,426593**	0,517322**	1,30
85	0,000494	0,002321**	0,000093**	0,000000	0,275580**	0,259745**	5,16
86	0,000398	0,007114**	0,000072**	0,000003	0,150148**	0,093268**	3,25
87	0,003429**	0,020237**	0,000541**	0,000001	0,418237**	0,420424**	4,32
88	0,000316	0,004931**	0,000184**	0,000011*	0,698633**	1,195637	2,12
89	0,000265	0,038437**	0,000218**	0,001783**	0,602895**	2,741132*	1,17
90	-0,000004	0,008261**	0,000115**	0,000006*	0,159434**	0,250851**	5,56
91	0,001053**	0,008996**	0,000137**	0,000025**	0,889680**	2,407741**	7,87
92	0,002545**	0,029124**	0,000449**	0,000298**	0,188494**	0,762121**	2,06
93	0,000493	0,010147**	0,000180**	0,000029**	1,005279**	3,151976**	1,73

"*" and "**" indicate significance at 90% and 95% levels of confidence respectively.

Table 5: MDH model estimated parameters (2).

F Subperiod analysis

Overall Period

ID	μ_v^{at}	μ_v^{la}	σ_p^2	σ_v^2	m_{2I}	x	p	χ_3^2	$p\mu_v^{la}$
22	0,006114	0,010793	0,000120**	0,00000000	0,258**	2,87	0,054	3,83	-
54	0,004448**	0,007292	0,000095**	0,00000000	0,225**	3,19	0,040	4,86	-
76	0,007274**	0,037411**	0,000176	0,00000000	0,316**	6,39**	0,002	2,79	0,000075
65	0,006773**	0,018750**	0,000099	0,00000000	0,236**	5,20**	0,005	5,49	0,000094
24	0,005370**	0,010659**	0,000063**	0,00000072	0,130**	3,85**	0,021	4,84	0,000222
44	0,006229**	0,013975**	0,000120**	0,00000027	0,258**	3,84**	0,021	1,42	0,000294
77	0,005522**	0,010229**	0,000152**	0,00000000	0,278**	2,60**	0,069	1,18	0,000705
31	0,004146**	0,008846**	0,000102**	0,00000000	0,223**	4,00**	0,018	5,11	0,000159
35	0,007926**	0,056664**	0,000264**	0,00000945**	0,373**	5,24**	0,005	2,57	0,000297
82	0,006074**	0,032010**	0,000182**	0,00000000	0,391**	4,97**	0,007	1,23	0,000220

Subperiod 1

ID	μ_v^{at}	μ_v^{la}	σ_p^2	σ_v^2	m_{2I}	x	p	χ_3^2	$p\mu_v^{la}$
22	0,006079**	0,012897	0,000120**	0,00000000	0,347**	2,927	0,051	1,40	-
54	0,004046**	0,004738*	0,000101**	0,000000000	0,215**	2,04	0,115	0,25	-
76	0,004747**	0,003933**	0,000129**	0,00000000	0,161**	0,80	0,309	3,20	-
65	0,006844**	0,009940	0,000082**	0,00000000	0,203**	3,95*	0,019	3,09	-
24	0,005668**	0,012285**	0,000057**	0,00000208**	0,082*	3,86**	0,021	3,57	0,000253
44	0,005301**	0,017626**	0,000104**	0,00000142	0,187**	4,87**	0,008	0,33	0,000134
77	0,004466**	0,011211**	0,000113**	0,00000163**	0,134**	3,89**	0,020	1,38	0,000224
31	0,004062**	0,010295**	0,000095**	0,00000000	0,210**	4,42**	0,012	2,02	0,000122
35	0,005758**	0,040619**	0,000217**	0,00000623	0,329**	5,01**	0,007	2,15	0,000269
82	0,005762**	0,034560**	0,000150**	0,00000000	0,369**	5,40**	0,004	3,98	0,000155

Subperiod 2

ID	μ_v^{at}	μ_v^{la}	σ_p^2	σ_v^2	m_{2I}	x	p	χ_3^2	$p\mu_v^{la}$
22	0,006560**	0,008112	0,0001223**	0,00000159	0,203**	3,80	0,022	2,57	-
54	0,004790**	0,014366**	0,0000852**	0,00000000	0,229**	4,28**	0,014	4,71	0,000196
76	0,008524**	0,033431**	0,000204**	0,00000046	0,257**	5,29**	0,005	2,86	0,000168
65	0,006520**	0,021978**	0,000116**	0,00000000	0,238**	5,36**	0,005	3,70	0,000103
24	0,004933**	0,007796**	0,000067**	0,00000000	0,142**	3,21**	0,039	3,48	0,000302
44	0,007154**	0,012640**	0,000134**	0,00000000	0,220**	3,22**	0,038	2,52	0,000485
77	0,006586**	0,009886**	0,000187**	0,00000000	0,224**	2,08**	0,111	0,27	0,001099
31	0,004208**	0,007810**	0,000106*	0,00000000	0,228**	3,65**	0,025	3,33	0,000198
35	0,009750**	0,029675*	0,000278**	0,00000718	0,272**	4,69**	0,009	3,08	0,000271
82	0,005623**	0,013323*	0,000170**	0,00000000	0,228**	3,62*	0,026	4,46	0,000346

"*" and "**" indicate significance at 90% and 95% levels of confidence respectively.

Table 6: MDHL model estimated parameters for subperiod analysis

G Summary results

	Data	MDH extension	Model validity	Contributions
Tauchen and Pitts (1983)	90-day T-bills futures market	–	Favorable	Explains $Cov(R_t^2, V_t) > 0$
Richardson and Smith (1994)	Dow Jones30 stocks	$E(R_t) \neq 0$	Less favorable	GMM test
Lamoureux and Lastrapes (1994)	10 NYSE stocks	$Cov(I_t, I_{t-1}) \neq 0$	Unfavorable	MDH explanation for GARCH effects?
Andersen (1996)	IBM common stocks	Non-informed part of volume	Unfavorable to standard MDH; Modified MDH does better.	Volume decomposition: informed versus uninformed part of volume with market maker
Roskelley (2001)	Dow Jones30 stocks	$Cov(I_t, I_{t-1}) \neq 0$	Unfavorable	Moment simplification
Li and Wu (2006)	Dow Jones30 stocks	Extend Andersen (1996): Non-informed part of return volatility	Rejection of Andersen (1996); Validation of their model.	Non-informed traders have negative impact on $Cov(R_t^2, V_t)$
MDHL model	FTSE 100 Stocks	Extend TP (1983): Information and Liquidity shocks	Favorable to standard MDH and to MDHL	Liquidity arbitragers are strategic agents and not noisy traders; Extends standard MDH by accounting for liquidity shocks; Volume decomposition; Proposes a new liquidity measure.

Table 7: Paper contributions compared to previous literature.