Journal of Banking & Finance 50 (2015) 92–105

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Measuring the liquidity part of volume ☆

Serge Darolles [a], Gaëlle Le Fol [a], Gulten Mero [b,*]

[a] Université Paris-Dauphine, DRM Finance and CREST, France
[b] Université de Cergy-Pontoise and THEMA (UMR 8184, CNRS), 33 Boulevard du Port, 95011 Cergy-Pontoise Cedex, France

ARTICLE INFO

Article history:
Received 24 November 2013
Accepted 14 September 2014
Available online 13 October 2014

JEL classifications:
C51
C52
G12

Keywords:
Volatility–volume relationship
Mixture of distribution hypothesis
Liquidity shocks
Information-based trading
Liquidity arbitrage
GMM tests

ABSTRACT

Based on the concept that the presence of liquidity frictions can increase the daily traded volume, we develop an extended version of the mixture of distribution hypothesis model (MDH) along the lines of Tauchen and Pitts (1983) to measure the liquidity portion of volume. Our approach relies on a structural definition of liquidity frictions arising from the theoretical framework of Grossman and Miller (1988), which explains how liquidity shocks affect the way in which information is incorporated into daily trading characteristics. In addition, we propose an econometric setup exploiting the volatility–volume relationship to filter the liquidity portion of volume and infer the presence of liquidity frictions using daily data. Finally, based on FTSE 100 stocks, we show that the extended MDH model proposed here outperforms that of Andersen (1996) and that the liquidity frictions are priced in the cross-section of stock returns.

1. Introduction

The use of total traded volume as a proxy for liquidity is well documented in the literature [see Gallant et al. (1992), Domowitz and Wang (1994), and Darolles and Fol (2014) among others]. However, recent studies support the idea that stocks with a high traded volume are not necessarily the most liquid ones. Indeed, the total traded volume can increase in response to both information and liquidity shocks. For example, Borgy et al. (2010) note that price-impact-based indicators are more accurate than raw traded volume for identifying liquidity problems in the currency exchange (FX) market. The flash market crash of May 6, 2010 is a good illustration of the simultaneous effects of information and liquidity shocks on traded volume. Recall that bad news concerning the European debt crisis resulted in sell-side pressures on U.S.-based product prices, thus increasing market volatility. In addition, buy-side liquidity in E-Mini S&P 500 futures contracts and S&P 500 exchange traded funds dropped sharply. In response to this situation, a large fund initiated a sell program for a substantial number of E-Mini contracts. This trading program was calibrated to target an order execution rate of 9% of the previous minute trading volume. This resulted in an extraordinarily high trading volume and further increased the feeding rate of automated executed orders for the considered fund, implying a liquidity crisis for both the E-Mini and individual stock markets. Consequently, during this day, at 2:42 pm, the DJIA Index had plunged by about 300 points, and by 600 more points at 2:47 pm (which resulted to an abnormal intradaily return of almost −9%) before recovering a few minutes after; by 3:07 pm, the market had regained almost 600 points. At the end of the traded day, the DIJA Index lost only 3% of its value which reflected bad news related to worries about the debt crisis in Greece. In contrast, the daily traded volume was more than twice as high as the average volume of the 30 past trading days. Three important lessons should be drawn from this event. First, using the total traded volume as a liquidity indicator can be misleading, especially in periods of significant volatility. In fact the total traded volume can increase in response to both information and liquidity shocks; observing an important traded volume does not necessarily mean that the market is liquid. It is thus important to be able to separate information from liquidity components of volume. Second, the volatility–volume relation, rather than volume alone,

☆ We gratefully acknowledge financial supports from the chair of the QUANTVALLEY/Risk Foundation: Quantitative Management Initiative, as well as from the project ECONOM & RISK (ANR 2010 blanc 1804 03).

* Corresponding author. Tel. +33 (0)6 30 01 33 70.
E-mail addresses: serge.darolles@dauphine.fr (S. Darolles), gaelle.le_fol@dauphine.fr (G.L. Fol), gultenmero@gmail.com (G. Mero).

http://dx.doi.org/10.1016/j.jbankfin.2014.09.007
0378-4266/

should be exploited to distinguish between information-based and liquidity-based trading and used to build more efficient trading algorithms. Third, understanding the determinants of total traded volume is not a trivial exercise; we must be able to model the respective impacts of information and liquidity shocks on trading characteristics.

This example leads to a natural question that determines the scope of this paper: how can we separate information from liquidity shock impact on the daily traded volume and thus measure the liquidity portion of volume? To answer this question, we develop an extended version of the mixture of distribution hypothesis (MDH) of Tauchen and Pitts (1983) that accounts for the presence of liquidity frictions. The standard MDH framework assumes that the market is perfectly liquid and that only information affects price changes and traded volumes. However, recent liquidity events suggest that we cannot address information shock impact on trading characteristics without considering liquidity frictions.

Distinguishing between the informative and non-informative portion of volume is not new in the literature. Andersen (1996) proposes an extended MDH version to separate the informed from the noise (or liquidity) trading components of volume. However, Andersen's definition of the noise trading component of volume is *ad hoc* and does not rely on any structural definition of liquidity trading.[1] In addition, Andersen's MDH version is difficult to calibrate, and its empirical validity is bounded when the model is tested using larger samples of data.[2] In contrast, the extended MDH version developed here is based on the theoretical framework of Grossman and Miller (1988), henceforth GM, which admits the presence of liquidity frictions within the trading day and is structured to explain how they impact the way in which information is incorporated into prices and volumes. According to GM, liquidity is determined by the demand for and the supply of immediacy. Two types of market participants are considered. The first type, active traders, trades in response to information shocks. The second type of trader acts as a market maker. They provide immediacy when the market faces liquidity frictions and liquidate their positions once prices return to their fundamental levels to obtain the liquidity premium. As discussed by Hendershott and Riordan (2013), in practice, algo traders who monitor the market, are not trading for information purposes so they can definitely be considered as market makers. Liquidity frictions result from temporary order imbalances due to trade asynchronization among the information-based traders. These order imbalances are resorbed by the market within the trading day. In fact, market markers manage their inventory position so as to get back home flat at the end of the day.[3] The GM framework implies that the liquidity frictions occurring at the intradaily frequency do not impact the daily price change. However, they increase the daily traded volume because the volume traded by market makers to liquidate their positions adds to the volume that would prevail in the absence of liquidity frictions.

In this paper, we first exploit the theoretical definition of liquidity frictions in the sense of GM allows us to put enough structure in the standard MDH model of Tauchen and Pitts (1983) to capture the impact of liquidity frictions on daily trading characteristics. Recall that the MDH models represent reduced econometric forms of microstructure models, thus facilitating the estimation of the information flow impact on the relation between price change and volume. The standard MDH model provides an explanation of the positive correlation between volume and squared price change at the daily frequency as well as other stylized facts, such as the fat tailed probability distributions of the daily time series of returns and volumes [see Harris (1982), Harris (1986), Harris (1987), Tauchen and Pitts (1983), among others]. The basic idea behind the MDH is that the joint distribution of daily price changes and volume can be modeled by a mixture of bivariate normal distributions conditioned by a unique latent variable, the information flow, which is supposed to be random. However, the standard MDH assumes that the market is perfectly liquid, disregarding the presence of liquidity frictions. From this perspective, we reconsider the standard MDH of Tauchen and Pitts (1983) by incorporating a second latent variable to capture the effect of liquidity frictions on daily traded volume. Our version of the MDH with two latent variables, called the MDHL model, allows us to exploit the volatility–volume relation to decompose the daily traded volume for a given stock into two components due to information and liquidity shocks.

Second, because our econometric specification can be tested empirically, we can filter the liquidity portion of total traded volume and infer the presence of liquidity frictions using the daily time series of returns and volumes. In particular, the model imposes restrictions on the joint moments of price change and volume as a function of only a few parameters. It is thus possible to use the generalized method of moments (GMM) procedure of Hansen (1982) to estimate model parameters and to test the model's global validity by forming overidentifying restrictions.[4] We provide a stock-specific liquidity indicator using daily return and volume observations.

The contribution of this paper is twofold. First, relying on the theoretical microstructure framework of GM, we extend the information-based standard MDH model to account for the impact of liquidity frictions on daily trading characteristics. The MDHL model suggests that the volatility–volume relation has two determinants, information flow and liquidity frictions and that their respective impacts on returns and volumes should be modeled differently. The former is incorporated into the daily price changes and traded volumes. The latter impacts intradaily price variations and volumes; it does not affect daily price changes but increases daily traded volumes. Second, the empirical estimation of the MDHL model allows us to exploit the volatility–volume relation in order to separate the total traded volume into two parts due to information and liquidity shocks. We propose a measure of the liquidity part of volume, thus providing a more accurate proxy for market liquidity for individual stocks. In particular, this information could be useful for practitioners, such as high frequency traders or fund managers, who want to hedge against liquidity risk or to track liquidity frictions to construct liquidity arbitrage strategies.

Note that, the standard MDH theory considers only informed agents with homogenous beliefs trading simultaneously after the arrival of each piece of information in a frictionless market. This implies that, at each point of time, the equilibrium prices reflect the fundamental value of the assets. However, this theory does not account for more complex interactions between traders having different trading motives, such as informed traders, noise traders and institutional arbitragers, nor does it account for the effects of sequential trading and the dispersion of beliefs among traders on the trading characteristics. In this context, three other theories building on the volatility–volume relation refine and complete the MDH framework by providing a better understanding of how information is incorporated in price changes and traded volume. The first one is represented by the sequential arrival of information

[1] The noise trading component of volume is simply (and arbitrarily) supposed to follow a time-invariant Poisson process in Andersen (1996).

[2] The empirical tests in Andersen (1996) are based on only five common stocks listed on the U.S. market. As discussed in Section 4 of this paper, Andersen's MDH version is rejected by the data for more than 50% of the common stocks listed on the FTSE 100 index.

[3] see Menkveld (2013), Fig. 2 and page 720.

[4] This procedure is initially used by Richardson and Smith (1994) to test the standard MDH model.

model initially developed by Copeland (1976) and Copeland (1977). It states that the fact that different groups of traders don't receive information in the same time is responsible for sequential trading, implying a lagged volatility–volume relation. The second theory focuses on the dispersion of beliefs among traders receiving the same information at the same time [see Shalen (1993) and Harris and Raviv (1993)]. According to this theory, the dispersion of beliefs tends to increase return volatility as well as the traded volume, thus affecting their contemporaneous relation. The third theory deals with the effect of noise trading activity on the intervention of institutional arbitragers to the market [see DeLong et al. (1990a) and DeLong et al. (1990b)]. In fact the former are responsible for price distortions from the fundamentals, which motivates the latter to enter the market in order to take advantage of asset mispricing or, alternatively, to reduce their risk related to mispricing. For a more detailed discussion of these theories, as well as test results of their empirical implications, see Chen and Daigler (2008).

From this point of view, our MDHL model can be considered as an alternative extension of the standard MDH literature, focusing on how the presence of liquidity frictions can modify the way information is incorporated in daily price changes and traded volume. In our framework, the occurrence of liquidity frictions is responsible for sequential trading as well as the intervention of liquidity arbitragers who enter the market to provide immediacy in order to cash the liquidity premium. Liquidity arbitragers respond to the increasing price volatility rather than being responsible for it and their intervention tends to increase the daily traded volume. According to our framework, the presence of liquidity frictions represents an alternative way to explain why the total traded volume increases in the absence of new information arrival, and why prices deviates from their fundamentals even in the presence of informed traders. Our MDHL model differs from other MDH model extensions trying to separate informative from non-informative trading activity, such as in Andersen (1996) or in Li and Wu (2006), in terms of the number of determinants of volatility–volume relation, as well as the random pattern of the liquidity friction process. In fact, Andersen (1996) and Li and Wu (2006), the volatility–volume relation is driven by only one latent variable, the information flow process, whereas the non-informative trading is captured by a constant. Alternatively, in our framework, the volatility–volume relation is driven by two latent variables, the information flow and the liquidity friction process, the liquidity par of volume being captured by the distribution parameters of the second latent variable.[5]

The empirical results provided in this paper show that our model is validated by the data for 83% of the FTSE 100 stocks. We also show that our model presents a similar level of global validity when compared to the standard MDH model which does not account for the liquidity part of volume. In addition, the MDHL model outperforms the modified MDH version of Andersen (1996) when tested on a large sample of data (i.e., the 100 stocks belonging to the FTSE 100 index). Finally, portfolio and factor analyses performed in this paper provide consistent evidence that the amplitude of liquidity frictions commands a premium in the cross-section of returns. We also find that the presence of liquidity frictions can increase the daily traded volume and that the positive effect of liquidity events on total traded volume is captured by the size of order imbalances.

The paper is organized as follows. In Section 1, we summarize the basic features of the Grossman and Miller (1988) model and discuss its implications concerning daily trading characteristics. Section 2 develops an empirically testable version of the revisited MDH, accounting for both information and liquidity shocks. The data used in this paper are described in Section 3. In Section 4, we challenge the MDHL model and we discuss our empirical results, as well as their contribution to the literature. In Section 6, we analyze the features of the liquidity portion of volume and perform some cross-sectional analysis to assess whether the liquidity frictions command a premium in the cross-section of returns. Finally, Section 7 concludes the paper.

2. The theoretical framework

Based on the theoretical framework of Grossman and Miller (1988), henceforth GM, we develop a version of the MDH to infer the presence of liquidity frictions from trading characteristics. The standard MDH framework assumes that the market is perfectly liquid; thus, only information affects price changes and traded volumes. However, the GM model accounts for the presence of liquidity frictions and is structured to explain how they impact the way in which information is incorporated into prices and volumes. We summarize the basic features of the GM model and impose sufficient structure on the standard MDH setting to explore the model's implications for how liquidity frictions are incorporated into daily trading characteristics.

The GM model focuses on a market with a single asset in which liquidity is modeled as determined by the demand for and the supply of immediacy. There are two distinct groups of traders: the outside customers (or active traders) who trade in response to information inflow and the market makers who trade in response to liquidity shocks. Only three dates are considered $(1, 2, \text{and } 3)$. Dates 1 and 2 are trading dates, and date 3 is introduced only as a terminal condition. The asset has a random liquidation value that is known at date 3; it is denoted by $\widetilde{P}_3$. Information concerning $\widetilde{P}_3$ is assumed to arrive before trading at period 1 and before trading at period 2. Let J be the number of all potential active traders in the market. The active trader j $(j = 1, \ldots, J)$ 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. By definition, if all of the active participants are present in the market at date 1, the net trading demand will be zero at the current price. However, at period 1, some liquidity frictions may arise because of the asynchronization of order flows occurring when only J_1 out of J active traders are present in the market. These frictions result in a temporary order imbalance.[6] Market makers, 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.[7]

In the presence of liquidity shocks, the equilibrium price at period $1, P_1$, deviates from the price revealing the information, i.e., $E_1\widetilde{P}_3$.[8] In addition, both, the equilibrium aggregate excess demand for active traders at date $1, Q_1^{at}$, and the equilibrium excess demand

[5] For a more detailed discussion see subsection 5.4.

[6] $z = \sum_{j=1}^{J_1} z_j \neq 0$.

[7] Some large investors, such as hedge fund algo traders, play the role of market makers by tracking price pressures due to liquidity frictions and entering the market to provide immediacy and to obtain the liquidity premium. Their intervention tends to correct price imperfections due to liquidity shocks and thus lowers the intradaily return volatility. Since they manage their inventory position so as to get back home flat at the end of the day, these large investors liquidate their positions to benefit from the intradaily price reversals [see Menkveld (2013)]. As a consequence, the volume that they trade adds to the volume that would prevail in the absence of liquidity frictions.

[8] More precisely, in the presence of liquidity frictions, the equilibrium price is $P_1 = E_1\widetilde{P}_3 - \frac{z\alpha Var_1(E_2\widetilde{P}_3)}{1+M}$, where α represents trader preferences, which are assumed to be identical for all market participants; M is the number of market makers, and $Var_1(E_2\widetilde{P}_3)$ represents the risk from the point of view of period 1 in which $P_2 = E_2\widetilde{P}_3$ is not known.

per market maker, Q_1^{la}, are different from zero.[9] Because GM assume that market makers face a participation cost $c > 0$, their number M will be finite, which implies a limited capacity to provide immediacy and a deviation of P_1 from $E_1\widetilde{P}_3$.[10]

Regarding the consequences of order imbalances on the intradaily patterns of trading characteristics, 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 ($|Q_1^{at}| < |z|$), and (ii) the transaction price at date 1 deviates from $E_1\widetilde{P}_3$. However, GM assume that the active trader's endowment shocks sum to zero across periods 1 and 2 and that the market makers 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 $M \geqslant 1$ is satisfied. In other words, the order imbalance faced by active traders who exchange at date 1 is offset because of the immediacy provided by the market makers, who will liquidate their positions at date 2 and thus increase the traded volume. This process demonstrates that the impact of liquidity frictions on price changes and traded volumes exhibits different properties when addressing aggregated data across periods 1 and 2. Note that the daily frequency can be considered a special case of data aggregation. In fact, as discussed by Tauchen and Pitts (1983), the daily price change and traded volume are obtained by aggregating their intradaily counterparts. It is thus intuitive to expect that the occurrence of intradaily order imbalances should not impact daily and intradaily data in the same way. Because we are interested in the impact of liquidity frictions – occurring at the intradaily level – on price changes and volume at the daily frequency, we focus on the implications of the GM microstructure framework concerning aggregated trading characteristics across dates 1 and 2, which can be formalized are as follows:[11]

(i) the impact of order imbalances on the total price change across dates 1 and 2 vanishes. In fact, price distortions due to liquidity shocks occurring at date 1 and price adjustments due to the arrival of new traders at date 2 with opposite order imbalances offset each other, and the total price change across dates 1 and 2 is only due to information flow.

(ii) the occurrence of liquidity events increases the total traded volume across dates 1 and 2; this is due to the intervention of market makers, who provide immediacy at date 1 and liquidate their positions at date 2 as other active traders arrive with the opposite order imbalance. Only in the absence of liquidity frictions is the total traded volume completely explained by information flow.

These implications show that the presence of liquidity frictions, which occur at the intradaily frequency and are reabsorbed by the market within the trading day modify the way in which information is incorporated into intradaily and daily trading characteristics. In particular, liquidity frictions increase the daily traded volume while leaving the daily price variations unchanged. It follows that the observed daily traded volume of a particular stock results from both information-based and liquidity arbitrage trading. Thus, understanding and decomposing the traded volume can provide some insights concerning market liquidity. In particular, the raw traded volume is used in the literature as a proxy for market liquidity [Gallant et al. (1992), Domowitz and Wang (1994)]. It is thus commonly stated that the higher the traded volume, the higher the market liquidity. In contrast, the implications of GM's framework show that the total daily volume can be misleading because the presence of liquidity frictions can be associated with a higher volume.

These observations motivate us to develop a new MDH version that allows us to infer the presence of liquidity frictions from observed trading characteristics. In fact, the standard MDH model of Tauchen and Pitts (1983), henceforth TP, explains the positive daily volatility–volume relation using the dependence of both variables on the same random latent factor, the information flow. Based on the implications of GM's model concerning aggregated trading characteristics, we reconsider the standard MDH by incorporating a second latent variable to capture the effect of liquidity frictions on daily traded volume. Because this new MDH version with liquidity frictions can be tested empirically, we can filter the liquidity portion of the total traded volume, which allows us to infer the presence of liquidity frictions using the daily time series of returns and volume.

3. The empirical specification

This section develops an empirical, testable version of the MDH that accounts for the presence of liquidity frictions in the sense of GM.

We focus on a simple economy with a risk-free asset and a single risky security with liquidation value $\widetilde{P}$ at the end of the trading day. The risk-free rate is normalized to zero. There are two types of traders in the market: active traders who react to information flow and market makers who trade in response to liquidity frictions. We assume that within the trading day, the market passes through a sequence of distinct equilibria. The movement from one equilibrium to the next is initiated by the arrival of new information to the market. Given a new piece of information, the active traders decide to rebalance their positions to share risk through the market. The number of pieces of information that hit the market during the trading day t, I_t, should be random. Because information generates trades, I_t simultaneously impacts intradaily price increments and transaction volumes. Note that daily price changes, ΔP_t, and traded volumes, V_t, are obtained by summing their intradaily counterparts, which implies that they are both impacted by I_t.

In addition, we allow for the presence of liquidity frictions, result from the order imbalances that occur at the intradaily frequency and are resorbed by the market within the trading day.[12] As discussed in the previous section, the presence of liquidity shocks increases the daily traded volume while leaving the daily price increments unchanged. We assume that the number of liquidity events occurring within the trading day t, L_t, is random. In fact, as discussed in the previous section, if a liquidity event occurs, the aggregated endowment shock across the active traders present in the market represents the order imbalance. Market makers who observe this market imperfection enter the market to provide immediacy just

[9] The equilibrium prices and the trader equilibrium demands are obtained using backward induction to obtain the optimal excess demands for active traders as well as market makers (at both trading dates) under the assumption of exponential preferences for both types of traders. Then, the solution of the market clearing conditions at each date yields the equilibrium price. Finally, replacing the equilibrium price in the optimal excess demand equations yields the equilibrium excess demands for both types of traders at each date.

[10] As discussed by Brunnermeier and Pedersen (2009), funding liquidity constraints can also explain why liquidity is not fully provided.

[11] The implications of GM framework regarding daily returns and volume can be easily derived from their microstructure model. We have detailed these implications in a separate appendix, which is not reported here but can be available upon request.

[12] We assume that 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 to that date: $P_{I_t} = E_{I_t}\widetilde{P}$. In fact, because the liquidation value of the asset $\widetilde{P}$ is revealed at the end of the trading day, the closing price $E_{I_t}\widetilde{P}$ converges to the liquidation value. More precisely, during the trading day, because of 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. 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 because no liquidity friction occurs at the last equilibrium.

after the arrival of the ith piece of information, and the equilibrium price at this date deviates from $E_i\widetilde{P}_T$.[13] To denote membership in the ith within-day equilibrium, we index by i all of the intraday variables of interest, such as price changes, the excess demands of traders, and traded volumes.

Let ΔP_i denote the price increment due to the ith piece of information ($i = 1, \ldots, I_t$). Assuming that ΔP_i is normally distributed with mean zero and variance σ_p^2, it follows that the daily price change is a mixture of normal distributions with mixing variable I_t:

$$\Delta P_t = \sum_{i=1}^{I_t} \Delta P_i, \Delta P_i \sim N(0, \sigma_p^2). \tag{2.1}$$

This equation coincides with that of TP because the effect of liquidity friction on price changes vanishes when considering aggregated data.[14] Thus, conditional on $I_t, \Delta P_t$ is normally distributed:

$$\Delta P_t | I_t \sim N(0, \sigma_p^2 I_t). \tag{2.2}$$

However, the daily traded volume has information-based and liquidity-based components:

$$V_t = \sum_{i=1}^{I_t} V'_i + \sum_{i=1}^{I_t} \delta_i V''_i. \tag{2.3}$$

In this equation, V'_i is the volume component due to information during the ith intradaily equilibrium, V''_i represents the volume component due to the presence of order imbalances after the arrival of the ith piece of information, and δ_i is an indicator variable such that $\delta_i = 1$ if a liquidity event occurs at the ith equilibrium and $\delta_i = 0$ otherwise. It follows that $\sum_{i=1}^{I_t} \delta_i V''_i = \sum_{l=1}^{L_t} V''_l$, where $L_t = \sum_{i=1}^{I_t} \delta_i$ is the number of liquidity events within a trading day t and $l = 1, \ldots, L_t$ is a subsequence of $i = 1, \ldots, I_t$ such as $\delta_i = 1$.

We assume that the two volume components are normally distributed as follows:

$$V'_i \sim N(\mu_v^{at}, (\sigma_v^{at})^2), V''_i \sim N(\mu_v^{la}, (\sigma_v^{la})^2). \tag{2.4}$$

Moreover, the indicator variable δ_i is supposed to be drawn independently from a Bernoulli distribution with parameter p.[15] It follows that, 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. Then, L_t's first two unconditional 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

$$Cov(I_t, L_t) = pVar(I_t). \tag{2.5}$$

From Eqs. (2.1), (2.3) and (2.4), we obtain a bivariate mixture for the 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, because information flow generates trades, I_t is primarily responsible for volume variations. For this reason, we consider $(\sigma_v^{la})^2$ to be negligible compared to $(\sigma_v^{at})^2$. Thus, conditional on I_t and L_t, the daily volume V_t can be considered to be $N(\mu_v^{at} I_t + \mu_v^{la} L_t, (\sigma_v^{at})^2 I_t)$ without any loss of generality. Henceforth, for notation simplicity, we replace $(\sigma_v^{at})^2$ by σ_v^2. The bivariate normal mixture can then be written as follows:

$$\Delta P_t = \sigma_p \sqrt{I_t} Z_{1t}, \tag{2.6}$$

$$V_t = \mu_v^{at} I_t + \mu_v^{la} L_t + \sigma_v \sqrt{I_t} Z_{2t}, \tag{2.7}$$

where $Cov(\Delta P_t, V_t | I_t, L_t) = 0$, 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 the daily price change and the traded volume, whereas only the daily volume is affected by the random liquidity shocks. Note that the standard MDH of TP is implied by (2.6) and (2.7) as a particular case when $\mu_v^{la} = 0$.

From Eqs. (2.5) and (2.6) and (2.7), the unconditional contemporaneous relation between ΔP_t^2 and V_t is

$$Cov(\Delta P_t^2, V_t) = \sigma_p^2 (\mu_v^{at} + p\mu_v^{la}) Var(I_t). \tag{2.8}$$

The volatility–volume covariance predicted by our model is positive, as is that of TP as provided by

$$Cov(\Delta P_t^2, V_t) = \sigma_p^2 \mu_v Var[I_t].$$

Eq. (2.8) confirms that our version of MDH is consistent with the well-known stylized facts. However, whereas 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: $\mu_v = \mu_v^{at} + p\mu_v^{la}$. In the presence of liquidity shocks, the standard MDH model of TP overestimates the average volume related to information inflow: $\mu_v \geqslant \mu_v^{at}$.

The model given in (2.6) and (2.7) is called the MDH model with liquidity frictions, henceforth the MDHL model, and forms the basis of our empirical work. The particularity of this model is that it considers both information and liquidity shocks. Based on the MDHL model, we can exploit the volume-volatility correlation to decompose the traded volume for a given stock into two components and thus separate the information from the liquidity trading impact on the observed daily volume.

4. The data

Our sample consists of all FTSE 100 stocks listed on July 10, 2007. We consider the period from January 4, 2005 to June 26, 2007, i.e., 636 observation dates. We exclude stocks with missing observations, resulting in 93 stocks. The closing prices and daily transaction volumes are extracted from Bloomberg databases. The daily returns R_t are measured by the daily (log) price change.

The amount of shares on the market impacts the traded volume, as measured by the number of traded shares. Thus, a firm having an important float – as measured by the difference between annual common shares outstanding and closely held shares for any given fiscal year – may artificially appear to be more liquid (higher traded volume) than a more liquid firm having a smaller float. For instance, trading 100 000 shares of stock in a trading day does not provide the same information concerning the liquidity of the stock for a blue chip stock or a small cap stock. To address this point, we use the turnover ratio – instead of the number of traded shares – as a measure for volume. In fact, the turnover ratio is measured by the number of traded shares divided by the float. It allows us to control for the dependency between the number of traded shares and the float. Note that the turnover ratio appears to be appropriate when studying the market volume [Smidt (1990), LeBaron (1992),

[13] If there were 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 market makers and noise traders would occur at a disadvantageous price for the latter, who would bear, in that case, the liquidity premium perceived by the former. Because we do not allow for the presence of noise traders in our model, the liquidity arbitragers must 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 will arrive at date 2, causing the date-2 equilibrium price to change in a direction that is disadvantageous for them.

[14] For further details, see Tauchen and Pitts (1983).

[15] An independent draw means that for each $i = 1, \ldots, I_t, \delta_i$ 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 $Var(\delta_i) = p(1-p)$.

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 first row reports the average, the dispersion, the minimum, and the maximum of the means of the returns and the volume across the 93 stocks. Rows 2 to 4 provide the same cross-section statistics for the volatilities of the returns and the volume, and for the skewness, kurtosis, and the correlation between the squared returns and the volume. We perform a Pearson test to check the significance of the correlation coefficients.

	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

Campbell et al. (1993), Lo and Wang (2000)] or when comparing individual asset volumes [Morse (1980), Bamber (1986), Bamber (1987), Lakonishok and Smidt (1986), Richardson et al. (1986)]. However, Bialkowski et al. (2008) replace the number of shares outstanding by the float. Common and closely held shares are extracted from Factset databases. More precisely, let q_{kt} be the number of shares traded for asset $k, k = 1, \ldots, K$ on day t, $t = 1, \ldots, T$, and N_{kt} is the float for asset k on day t. The individual stock turnover for asset k on day t is $V_{kt} = \frac{q_{kt}}{N_{kt}}$.

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 is summarized in Table 1. The first row reports the average, the dispersion, the minimum, and the maximum of the means of the returns and the volume across the 93 stocks. Rows 2 to 4 provide the same cross-section statistics (average, dispersion, minimum, and maximum) for the volatilities of the returns and the volume, and for the skewness, kurtosis, and the correlation between the squared returns and the volume. We perform a Pearson test to check the significance of the correlation coefficients. These correlation coefficients are statistically significant for 92 out of 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 the squared returns and the volume.

The implications of the MDH for the joint distribution of daily returns and volume are examined in detail by Clark (1973), Westerfield (1977), Tauchen and Pitts (1983), Harris (1986), and Harris (1987), Board and Sutcliffe (1990), Kawaller et al. (1990), Andersen (1996), Chen and Daigler (2008) among others. They assume that both variables (the daily (log) price change and the 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, the above empirical patterns would not be obtained, and the daily returns and volume would be mutually independent and normally distributed. The results reported in Table 1 are thus consistent with the MDH. The average and minimum statistics of the volume skewness and the squared return's correlation with volume are positive, and the average and minimum statistics of return and volume kurtosis are greater then 3, as predicted by the MDH. Moreover, these cross-security statistics are larger than their corresponding values, which would be expected if the mixing variable were not random.[16]

The previous literature relates stock liquidity to total traded volume and considers that illiquid equity assets have a low traded volume or turnover [see Datar et al. (1998) and Chordia et al. (2000) among others]. Thus, the total traded volume is commonly used as a proxy for liquidity. We now provide some statistical evidence that the total traded volume is not an appropriate proxy. For this purpose, we form single-sorted portfolios from the 93 stocks in our sample. We use 3 criteria to sort the stocks: (i) the stock average size computed over the test period;[17] (ii) the standard deviation of stock returns observed over the sample period; and (iii) the average of the mean-adjusted illiquidity measure of Amihud (2002) computed over the test period.[18] These criteria are widely used to characterize stock liquidity. First, using the stock size as a proxy for stock liquidity is a common practice in financial markets; small stocks are assumed to face more liquidity problems than blue chip stocks. Second, it is usually observed that markets with high liquidity imperfections are often characterized by higher volatility. Third, the liquidity measure of Amihud (2002) represents the daily price impact of order flow, illiquid markets being characterized by a higher price impact from trades. We rank all FTSE 100 stocks by increasing order for the value of the considered criterion: average size, standard deviation of the returns or average Amihud measure. We then split the FTSE 100 stocks into three groups, small, medium and big, based on the breakpoints of the bottom 30% (Small or S), middle 40% (Medium or M), and top 30% (Big or B) of the ranked values of the considered criterion.[19] We compute the equally-weighted returns and volumes of these single-sorted portfolios. Panels A, B, and C of Table 2 present time-average returns and volumes for the portfolios of stocks sorted on size, return volatility, and Amihud illiquidity measure, respectively. Columns 2 to 4 of each panel report the average returns and volumes for the three portfolios (Small, Medium, and Big). The differences between the Big-Portfolio and the Small-Portfolio trading characteristics are reported in the last column of each panel. If total traded volume were a good proxy for liquidity, the portfolios composed of small, highly volatile or highly illiquid stocks in the sense of Amihud (2002) should present lower average volumes than the portfolios of respectively big, less volatile or liquid stocks. The results reported in Table 2 show that this is not the case. First, the difference between the average daily traded volume for the portfolios of big and small stocks is negative and statistically significant. Second, the average volume of highly volatile stocks is higher than the average volume of less volatile stocks.[20] Third, as reported in

[16] The expected value for 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.

[17] The size is measured by the float.

[18] For each day t and for a given stock i, we compute the ratio between the absolute value of stock returns and the daily GBP traded volume: $Illiq_{it} = \frac{|R_{it}|}{V_{it}}$. Then, for each date, we average $Illiq_{it}$ across stocks: $Ailliq_t = \frac{1}{N}\sum_{i=1}^{N} Illiq_{it}$. Finally, for each stock i, we adjust the time series $Illiq_{it}$ by dividing by $Ailliq_t$ to obtain the mean-adjusted illiquidity measure: $IlliqMA_{it} = \frac{Illiq_{it}}{Ailliq_t}$.

[19] For robustness checks, we also use alternative breakpoints – 10%/80%/10%, 20%/60%/20% and 35%/30%/35% – and obtain similar results. These results are not reported here but are available upon request.

[20] The difference between average volumes is statistically significant.

Table 2
In this table we relate liquidity to trading characteristics and try to asses: (i) whether there exists a liquidity premium in the cross-section of returns and (ii) whether illiquid stocks are characterized by low traded volume. To do so, three criteria widely considered as proxies for liquidity – average size, standard deviation of the returns or average Amihud (2002) illiquidity measure – are used to form single sorted portfolios. We rank all FTSE 100 stocks by increasing order for the value of the considered criterion and then split them into three groups, small, medium and big, based on the breakpoints of the bottom 30% (P_S portfolio), middle 40% (P_M portfolio), and top 30% (P_B portfolio) of the ranked values of the considered criterion. We compute the equally-weighted returns and volumes of these single-sorted portfolios for each criterion. The average values are reported in the three first columns of respectively panel A, B and C. The differences between the P_B portfolio and the P_S portfolio trading characteristics (on average) are reported in the last column of each panel. "*" and "**" indicate significance at 90% and 95% confidence levels, respectively.

	P_S	P_M	P_B	$P_B - P_S$
Panel A: Sorting on stock size				
$E(R_{P_i})$	0.00084**	0.00075**	0.00049*	−0.00035*
$E(V_{P_i})$	0.01152**	0.00839**	0.00639**	−0.00513**
Panel B: Sorting on standard deviation of returns				
$E(R_{P_i})$	0.00047**	0.00052*	0.00115**	0.00068**
$E(V_{P_i})$	0.00634**	0.00803**	0.01205**	0.00570**
Panel C: Sorting on Amihud illiquidity measure				
$E(R_{P_i})$	0.00068**	0.00065**	0.00077**	0.00009
$E(V_{P_i})$	0.00635**	0.00994**	0.00951**	0.00316**

panel C, the portfolio of the 30% most illiquid stocks presents a higher average volume than the portfolio of the 30% least illiquid stocks in the sense of Amihud (2002). A higher traded volume for illiquid stocks can be explained by the fact that liquidity pressures motivate liquidity arbitrage trading. It is then obvious why the total traded volume can increase in response to either information flow or liquidity frictions. Thus, decomposing the total traded volume into information- and liquidity-based components would allow us to explain the determinants of traded volume and to infer the presence of liquidity frictions for individual stocks. Finally, considering the average portfolio returns, size and volatility are priced in the cross-section of stock returns; in fact, in panels A and B, the difference between the average returns of P_B and P_S is statistically significant. These results confirm the previous literature in which stock volatility is positively correlated with stock returns whereas stock size has a negative effect on returns.

5. MDHL validity, performances and contributions

5.1. 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. Because our bivariate mixture with two latent variables imposes restrictions on the unconditional joint moments of the observables as a function of the model parameters, it is possible to form overidentifying restrictions on the data. Optimization methods can then be used to simultaneously estimate the coefficients and test the global validity of the model.

Let $X_t = (R_t, V_t)$ be the vector of return and volume observations prevailing at day t for a given stock, and let $\theta = (\mu_v^{at}, \mu_v^{la}, \sigma_p^2, \sigma_v^2, m_{2I}, p)$ 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 Eqs. (2.6) and (2.7); 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 the H unconditional moment conditions implied by our model. Because 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:

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

To derive nine moment restrictions implied by the MDHL model [i.e., Eqs. (2.6) and (2.7)], we focus on the first four moments of the return and volume time series and on some of their corresponding cross-moments.

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 to ensure its positiveness. Lognormality has also been suggested by several authors, such as Clark (1973) and Foster and Viswanathan (1993). Richardson and Smith (1994) tested several distribution functions for the information inflow and concluded 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

$$R_t \sim N(0, [\sigma_p^2 a] I_t^* | I_t, L_t), \tag{4.3}$$
$$V_t \sim N([a\mu_v^{at}] I_t^* + [\mu_v^{la}] L_t, [a\sigma_v^2] I_t^* | I_t, L_t), \tag{4.4}$$

is empirically the same as the MDHL model given in (2.6) and (2.7). By setting $E[I_t^*] = 1$, we can identify the transformed parameters that 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)]:

$$m_{3I} - m_{2I}^3 - 3m_{2I}^2 = 0$$
$$m_{4I} + 4(1+m_{2I})^3 + 3 - (1+m_{2I})^6 - 6(1+m_{2I}) = 0, \tag{4.5}$$

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 ($L_t|I_t \sim B(I_t, p)$), the sample moment vector $g_T(\theta)$ is given by[21]

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

[21] The functional forms of the sample moments (1)-(9) are provided by the authors upon request.

We obtain a system of nine equations and only six parameters to be estimated, which yields three overidentifying restrictions to test.[22]

5.2. Test results

We apply the GMM procedure described in the previous section to the 93 stocks of our sample using the entire data history. Recall that, for each stock, we estimate a vector of six parameters, $\theta = (\mu_v^{at}, \mu_v^{la}, \sigma_p^2, \sigma_v^2, m_{2I}, p)$. Note that, to restrict the Bernoulli parameter p to evolve between 0 and 1, we use a logistic transformation with x as the unconstrained parameter.[23] In addition, the GMM methodology allows us to compute the test statistics of Hansen (1982) in order to asses the global validity of the MDHL model. With three over-identifying restrictions, the test statistics are asymptotically distributed as χ_3^2. For 83% of the stocks, the test statistic values do not exceed their critical value of 7,82. Consequently, we cannot reject the MDHL model at the 95% level of significance.[24]

Because 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. However, $p\mu_v^{la}$ 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 that 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 is, the higher both its frequency and its liquidity-arbitrage-based traded volume, which, in turn, results in a higher volume and thus a higher $p\mu_v^{la}$.

According to the MDHL model, information provokes market movements from one equilibrium to the next and liquidity shocks occur within some of these equilibria. This implies that we should observe a statistically significant μ_v^{la} parameter only for the stocks that also have a significant μ_v^{at}. Our results confirm this intuition. The 42 stocks for which we obtain significant μ_v^{la} also have a μ_v^{at} parameter that is statistically different from zero. Note that for these stocks, we obtain statistically significant x parameters as well. We also compute the relative values of the average liquidity volume as measured by $p\mu_v^{la}$ divided by the sum of μ_v^{at} and $p\mu_v^{la}$, henceforth *relative* $p\mu_v^{la}$.[25] At this stage of the analysis, two additional remarks can be made:

(i) A significantly positive $p\mu_v^{la}$ implies that the stock faces (on average during the test period) intraday liquidity frictions. These frictions motivate the liquidity arbitragers to enter the market and thus increase the average traded volume. Because we do not observe liquidity shocks, we can infer their occurrence from liquidity arbitrage trading, which directly impacts volume. The MDHL model helps identify the intraday impact of this type of market participant on the traded volume using daily data: 39 of the 43 stocks with a significantly positive $p\mu_v^{la}$ are concerned with significant liquidity problems.[26]

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

5.3. The global validity of MDHL relative to TP's and Andersen's versions

Since our model is a direct extension of the standard MDH of TP, in this section, we compare the empirical performances of both models. In addition, we also confront our model to that of Andersen (1996) who proposes an alternative liquidity-based extension of the standard MDH of TP.

First, we estimate the standard MDH model of TP using the Richardson and Smith, 1994 procedure.[27] The estimation results show that the standard MDH model is valid for 89% of stocks versus 83% for the MDHL model.[28] Three primary reasons provide a possible explanation of the slight underperformance of the MDHL specification relative to the MDH model.

(i) Richardson and Smith (1994) estimate unbounded parameters, whereas we restrict the values of p to evolve between 0 and 1.
(ii) Richardson and Smith (1994) modify the TP's price change equation by artificially introducing a mean parameter μ_p to obtain much simpler moment conditions than in the absence of μ_p. In contrast, our model is directly derived from the standard MDH of TP, and we do not add any mean parameter in the price variation equation.
(iii) The MDHL model has a two-dimensional structure and provides a deeper comprehension of the determinants of volume compared to its one-dimensional counterpart, the standard MDH model. In our framework, we separate information from the liquidity components of volume while maintaining a similar level of global validity relative to the standard MDH, which considers the raw volume to be explained by information flow alone. When we compare the mean volume parameters obtained by the two models, we find that the μ_v parameter of the standard MDH is approximately equal to the sum of the μ_v^{at} and $p\mu_v^{la}$ parameters obtained in the case of the MDHL model. 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 traded volume into information-based and liquidity-based components.

Second, we compare the empirical validity of the MDHL model with that of the modified MDH specification proposed by Andersen (1996). Because the latter specification accounts for both the information and the noise trading components of volume, it can be considered to present the same degree of complexity as the MDHL model. We implement the GMM methodology described in Andersen (1996) to estimate the model parameters that minimize the distance between the theoretical and the empirical moment conditions.[29] With nine parameters and 12 moment conditions, three over-identifying restrictions are used to test the global validity of the model. Note that the number of degrees of freedom in

[22] Note that overidentification is used to build the goodness of fit tests of different model specifications. It allows us to easily compare our approach with the different models existing in the literature. It implies the definition of a distance matrix that is used to compute the goodness of fit test. In particular, 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 solution proposed in Newey and West (1987). We are aware that this choice has an impact on the asymptotic distribution of the test statistic.

[23] Note that the values of p are not directly estimated. They are obtained from $p = 1/(1 + e^x)$, with x estimated using the GMM procedure.

[24] The empirical results are not reported here, but are available upon request.

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

[26] The four remaining stocks, Cadbury Plc (stock 18), Man Group Plc (stock 27), Sabmiller Plc (stock 72), and Unilever Pcl (stock 85), have negligible relative $p\mu_v^{la}$ characterized by both p and μ_v^{la} values evolving in the neighborhood of zero.

[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 nine moment conditions and only six parameters to be estimated, there are three over-identifying restrictions to be tested. For more details, see Richardson and Smith, 1994.

[28] The results are not reported here but are available upon request.

[29] To account for serial correlation and heteroscedasticity, Andersen (1996) uses Newey and West (1987) methodology to estimate the weighting matrix in the GMM estimations.

Andersen's econometric setup is the same as for the MDHL model, implying that both models have directly comparable empirical performance.

We estimate Andersen's modified MDH model for each of the 93 stocks in our sample.[30] In terms of global validity, Andersen's model is rejected for more than half of the stocks of our sample. More precisely, for only 38 out of 93 stocks, the estimated χ^2_3 statistics are lower than the critical level of 7.82. The overall goodness of fit of Andersen's MDH specification is much lower than that of the standard MDH of TP as well as that of the MDHL model proposed in this paper. The poor performance of Andersen's version is also documented by Li and Wu (2006) based on individual stocks listed on the Dow Jones 30 Index. Moreover, Anderson's MDH version seems less attractive for practitioners because the estimated parameters are highly sensitive to the starting points relative to the standard MDH or MDHL models.

The primary difference between the MDHL model and Andersen's model that can explain the above results relies on the definition of liquidity trading. In Andersen's framework, liquidity trading is due to noise traders who trade in response to personal liquidity needs. These noise traders, reacting to non-information, are assumed to have inelastic demand and supply schedules. Thus, a liquidity trader, in the sense of Andersen (1996), is supposed to either buy or sell a unit of the risky security with probability one half, indifferent to the market needs for immediacy.

The liquidity traders considered in our framework are the same as those of GM. They act as market makers who track liquidity frictions and enter the market to provide liquidity. They have elastic demands, and they do not trade randomly or independently of market needs for immediacy. The liquidity traders provide liquidity when needed by selling high and buying low and require a liquidity premium for bearing the risk that future information shocks will contradict current information forcing them to liquidate their positions at a disadvantageous price. Following GM, we do not consider the impact of noise traders on trading characteristics. In fact, noise traders do not impact equilibrium prices, and they impact trading volume independently of the presence or lack of liquidity frictions. Thus, the impact of noise traders on volume cannot help to discriminate between situations with liquid markets, in which information flow alone drives prices and volumes, and situations that are characterized by immediacy needs, in which information shocks as well as liquidity frictions determine prices and volumes. From this point of view, we extend our analysis to isolate the liquidity portion of volume, i.e., the volume component that is due to the intervention of market makers in response to liquidity frictions. Identifying illiquid stocks based on daily returns and volumes represents one of the primary empirical contributions of the MDHL model.

5.4. MDHL contributions

Our framework extends the past literature toward two main directions.

First, in Table 3 we compare our results to those of different MDH versions proposed in the literature. The main interest of the MDHL model relative to the existing literature is related to the number of the determinants of the positive relation between volatility and volume. We use two latent factors, information flow and liquidity frictions, and consider that the randomness of the occurrence of liquidity frictions modifies the way information is incorporated into daily price changes and traded volume. Some other MDH versions, such as that of Andersen (1996) or Li and Wu (2006) also propose to estimate the non-informative part of volume.[31]

However, in Andersen (1996) model (and consequently in Li and Wu (2006)) the traded volume follows a Poisson distribution with only one latent variable, the information process K_t [see for instance equation (8) in Andersen (1996)]. The constant parameter m_0 is used to capture the noise trading volume. Alternatively, in our framework the traded volume distribution is driven by two latent variables, the information process and the liquidity friction process. The mu_v^{la} parameter represents the mean parameter related to the presence of liquidity frictions. However, the liquidity part of volume in our framework, $mu_v^{la}L_t$, is not constant in time since it is driven by the random latent variable L_t.

Empirically, in their static version, both models aim to separate, on average, the total traded volume into two components, due to information and liquidity shocks.[32] But if we consider the dynamic versions of both models, by assuming a dependent stochastic processes for K_t [see Section 6 in Andersen (1996)] and, alternatively, for I_t and L_t, the difference between the two models becomes more obvious. In fact, in Andersen's model we can only filter, at each point of time, the impact of information on daily traded volume (m_0 being only a constant), whereas in our model we could be able to filter and separate, at each point of time, the impacts of information as well as liquidity shocks on daily trading characteristics.[33]

Second, we complete the existing literature by proposing a new stock-specific liquidity measure based on the liquidity part of volume. Recall that, the use of total traded volume as a proxy for liquidity is well documented in the literature [see Gallant et al. (1992), Domowitz and Wang (1994), and Amihud (2002)among others]. However, recent studies support the idea that stocks with a high traded volume are not necessarily the most liquid ones. Indeed, the total traded volume can increase in response to both information and liquidity shocks. For example, Borgy et al. (2010) note that price-impact-based indicators are more accurate than raw traded volume for identifying liquidity problems in the currency exchange (FX) market. The flash market crash of May 6, 2010 is a good illustration of how the presence of liquidity frictions can increase the traded volume, implying that situations with high traded volume are not necessarily due to the arrival of new information to the market.

From this point of view, our framework provides a more accurate liquidity measure based on the liquidity part of volume, instead of total volume, thus allowing us to distinguish between situations of liquid from those of less liquid markets more efficiently.

6. Analyzing the liquidity part of volume

6.1. The MDHL-based liquidity measure

We use the MDHL 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 MDHL model is particularly attractive in practice because it provides a static, stock-specific liquidity measure $p\mu_v^{la}$, which helps identify the presence of intraday liquidity

[30] The estimation results are not reported here but are available upon request.

[31] Note that Li and Wu (2006) extend the MDH model of Andersen (1996) by accounting for the impact of noise trading on return volatility as well. Despite this difference, the theoretical background and the econometrical formulation of both models are the same.

[32] The m_0 parameter of Andersen serves the same empirical purpose as the mu_v^{la} parameter estimated by our model, even if the structural and microstructure interpretation of both parameters is different in both frameworks, and even if the empirical specification differs in terms of the number of latent variables used (only one in Andersen versus two in our framework).

[33] The dynamic version of our MDHL model is out of the scope of this paper and is part of current research.

Table 3
MDHL model contributions compared to other MDH versions

	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)	5 U.S. stocks	Non-informed part of volume; Noise traders have inelastic demand schedules	Unfavorable to standard MDH; Modified MDH does better	Volume decomposition: informed versus non-informed part of volume
Roskelley (2001)	Dow Jones30 stocks	$Cov(I_t, I_{t-1}) \neq 0$	Unfavorable	Moment simplification
Li and Wu (2006)	Dow Jones30	Extend Andersen (1996); Non-informed part of return volatility	Rejection of Andersen (1996): Validation of their own model	Non-informed traders have a negative impact on $Cov(R_t^2, V_t)$
MDHL model	FTSE 100 Stocks	Extend TP (1983):	Favorable to standard MDH	Liquidity arbitragers act as market makers and have elastic demand functions;
		Information and liquidity shocks	and to MDHL	Inferring the presence of liquidity frictions using daily data; Volume decomposition; Proposes a new liquidity measure

frictions using daily data. Based on the μ_v^{at}, μ_v^{la}, and p parameters, we can distinguish stocks affected 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 C.1 and C.2, 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 frequent and large liquidity accidents. For example, Hammerson Plc (stock 32), Segro Plc (stock 77), Scottish & Southern Energy Plc (stock 81), and Xstrata Plc (stock 92), which exhibit the four highest relative $p\mu_v^{la}$ of our sample, are characterized by both important μ_v^{la} parameters (two to three times higher than the corresponding μ_v^{at}) and important probabilities of trade asynchronization p whose values fall in the sample's highest decile. However, some other equity assets, such as Lonmin Plc (stock 50) and Mitchels & Butlers Plc (stock 51), face liquidity shocks characterized by a much higher amplitude of trade asynchronization than in the former cases (of an order seven to nine times higher then the corresponding μ_v^{at}), but much lower p values.

The previous literature relates stock liquidity to total traded volume and suggests that illiquid equity assets have low traded volume or turnover [see Datar et al. (1998) and Chordia et al. (2000), among others]. 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 two measures with the MDHL-based liquidity indicator $p\mu_v^{la}$.

Figs. 1 and 2 focus on the 39 stocks of our sample that present a significantly positive relative $p\mu_v^{la}$ and show the relative liquidity volume against the average raw daily volume and the average market capitalization over the estimation period, respectively.[34] The first graph shows 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

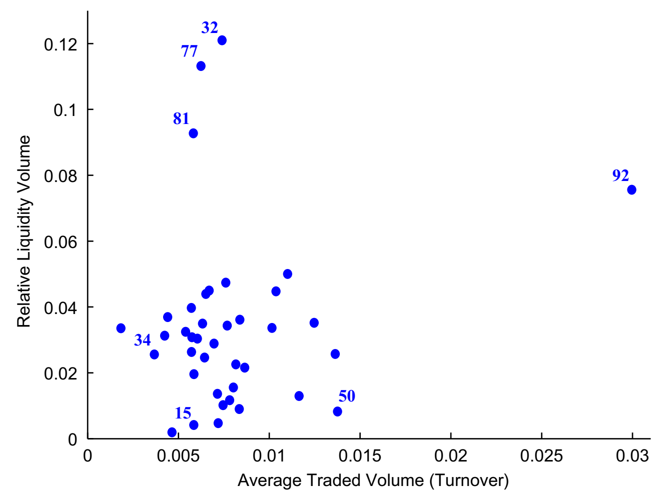

Fig. 1. Relative liquidity volume versus average daily traded volume. Stocks 15, 32, 34, 50, 77, 81 and 92 represent respectively the following companies: BP Plc, Hammerson Plc, HSBC Holdings, Lonmin Plc, Segro Plc, Scottish & Southern Energy and Xstrata Plc.

relative $p\mu_v^{la}$ than some others with lower MDHL-based liquidity measures, such as HSBC Holdings (stock 34) and BP Plc (stock 15). More generally, within the groups of high 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 to discriminate stocks facing liquidity shocks on the basis of their degree of illiquidity.

These results confirm the findings of Borgy et al. (2010) regarding the inability of the traded volume and the number of transactions to correctly measure market illiquidity. For example, a higher number of transactions may occur because of a higher liquidity risk, which induces market participants to split their trades, or an increasing market liquidity due to a larger number of liquidity providers being present in the market. Similarly, in our framework, an increasing total traded volume for a given stock may be explained by an increase 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. These possibilities imply that decomposing the total traded volume into two components due to information and liquidity shocks provides more precise indications on market liquidity.

[34] Note that the traded volume is measured by the turnover, whereas the market capitalization is measured by the float.

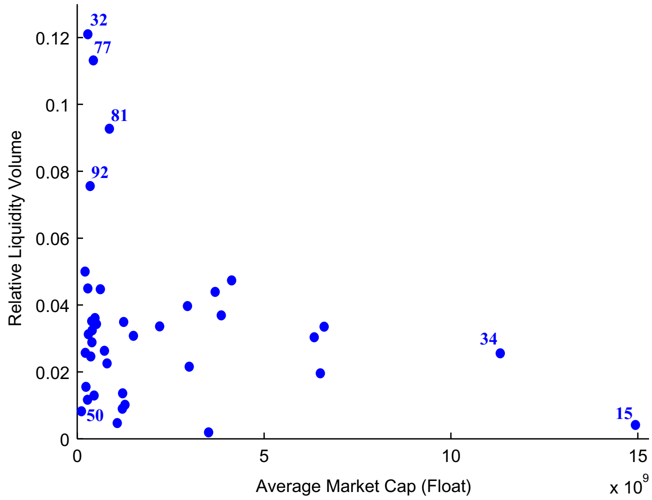

Fig. 2. Relative liquidity volume versus average market cap measured by the float. Stocks 15, 32, 34, 50, 77, 81 and 92 represent respectively the following companies: BP Plc, Hammerson Plc, HSBC Holdings, Lonmin Plc, Segro Plc, Scottish & Southern Energy and Xstrata Plc.

Fig. 2 shows that the largest companies among the 39 stocks are also the most liquid. For companies with large market capitalizations, there is quite a strong negative relation between firm size and illiquidity level. However, within the group of companies with 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) and Segro Plc (stock 77), as well as some of the less illiquid ones, such as Lonmin Plc (stock 50), belong to the lowest size deciles. These findings highlight the fact that market size is not a good proxy for liquidity shocks. In particular, considering small firms illiquid may be misleading because market size fails to distinguish small companies on the basis of their illiquidity level.

Finally, assessing the stock liquidity level simultaneously using total traded volume and market capitalization is quite disconcerting. To illustrate this point, Xstrata Plc (stock 92) is considered to be among the less illiquid of the 39 firms according to the total traded volume criterion but one of the most illiquid firms as reported by the market capitalization indicator. Conversely, HSBS Holdings (stock 34) and BP Plc (stock 15) appear to be highly illiquid when focusing on the total traded volume, whereas their (large) size ranks them among the less illiquid of the 39 equity assets considered here. These results highlight the relevance of a structural liquidity measure such as the $p\mu_v^{la}$ 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 by reconciling 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.

6.2. Is the MDHL-based liquidity measure priced in the cross-section of returns?

We perform a cross-sectional analysis to check whether the MDHL-based liquidity measure, $p\mu_v^{la}$, is priced across stocks. If this is the case, it is of interest to assess which of the two determinants of our liquidity indicator is responsible for commonalities in the cross-section of assets. In the first paragraph, we conduct a portfolio analysis to explore the relation between stock illiquidity, as measured by our measure, and average returns. In the second paragraph, we implement cross-sectional regressions and confront our liquidity indicator with a set of control variables.

6.2.1. Portfolio analysis

The idea is to check whether a zero-investment strategy that is long on the most illiquid stocks and short on the less illiquid stocks yields positive returns. We first rank the 42 stocks of our sample, having $p\mu_v^{la}$ statistically significant, by increasing the order of their $p\mu_v^{la}$, μ_v^{la}, and p values, respectively. Then, for each sort, we split the stocks into three groups, small, medium, and big, based on the breakpoints of the bottom 30%, middle 40%, and top 30% of the ranked values of the considered criterion. We then compute the equally-weighted returns and volumes of these single-sorted portfolios. The time-series averages of the portfolio returns and volumes are reported in the three panels of Table 4 (columns 2 to 4 of each panel). The last column of each panel presents the difference between the big-portfolio and small-portfolio average returns and volume.

Panel A shows that $p\mu_v^{la}$ is not a priced factor because the difference between the average returns of P_B and P_S is not statistically significant. However, as reported in panel B, the amplitude of trade asynchronization is positively correlated with stock returns; in fact, the μ_v^{la}-sort portfolio strategy generates a positive daily return of 0,038%, which is 9,5% annually. However, panel C shows that the return differential between the high p and the low p portfolios is not statistically significant. These results show that the MDHL-based liquidity indicator as a whole is not priced because commonality across stock μ_v^{la} is blurred by the absence of commonality across stock p. This result is explained by the fact that liquidity arbitragers are attracted by large and rare rather than small and frequent liquidity events. Indeed, if order imbalances are not important enough to cover trading costs, there is no incentive to act as the liquidity arbitrager. Analyzing the average traded volume confirms these remarks. Portfolios with a high MDHL-based liquidity measure or a high amplitude of order imbalance have a higher average volume than the portfolios with a low MDHL measure or order imbalance size. When controlling for the probability of occurrence, there is no statistically significant difference between portfolios with a high p and those with a low p. Thus, our data are consistent with the fact that the presence of liquidity frictions can increase the daily traded volume. In addition, our

Table 4
This table presents the results of preliminary investigations on whether the MDHL-based liquidity measure, $p\mu_v^{la}$, as well as its two determinants – the probability of occurrence of liquidity frictions (p) and the amplitude of trade asynchronization (μ_v^{la}) – are priced in the cross-section of returns. We rank the 42 stocks of our sample, having $p\mu_v^{la}$ statistically significant, by increasing the order of their $p\mu_v^{la}$, μ_v^{la},and p values, respectively. We then split them into three groups, small, medium and big, based on the breakpoints of the bottom 30% (P_S portfolio), middle 40% (P_M portfolio), and top 30% (P_B portfolio) of the ranked values of the considered criterion. We also compute the equally-weighted returns and volumes for the single-sorted portfolios based on each of the three criteria. These values are reported in columns 1 to 3 of respectively panels A, B and C. The differences between the P_B portfolio and the P_S portfolio trading characteristics are reported in the last column of each panel. "*" and "**" indicate significance at 90% and 95% confidence levels, respectively.

	P_S	P_M	P_B	$P_B - P_S$
Panel A: Sorting on stock $p\mu_v^{la}$				
$E(R_{P_i})$	0.00062*	0.00070**	0.00076**	0.00014
$E(V_{P_i})$	0.00658**	0.00708**	0.0103**	0.00369**
Panel B: Sorting on stock μ_v^{la}				
$E(R_{P_i})$	0.00052*	0.00066**	0.00090**	0.00038**
$E(V_{P_i})$	0.00621**	0.00661**	0.01120**	0.0050**
Panel C: Sorting on stock p				
$E(R_{P_i})$	0.00090**	0.00057*	0.00063*	−0.00027
$E(V_{P_i})$	0.00791**	0.00790**	0.00792**	−0.00001

Table 5
This table reports cross-sectional regression results performed in order to investigate whether the MDHL-based liquidity measure, $p\mu_v^{la}$, as well as its two determinants – the probability of occurrence of liquidity frictions (p) and the amplitude of trade asynchronization (μ_v^{la}) – are priced in the cross-section of returns. This table presents risk premiums estimated using cross-sectional regressions for different specifications of Eq. (5.2). The first column reports the factors used in the regression, whereas the values reported in columns 2 to 8 represent risk premium coefficients associated to different factors. "*" and "**" indicate significance at 90% and 95% confidence levels, respectively. T-statistics are based on Newey and West (1987) robust standard errors. The last row presents the regression R^2 coefficients.

	CAPM	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Intercept	0.0002	0.0002	0.0002	0.0002	0.0001	0.0001	0.0001
β_m	0.0005**	0.0004**	0.0004**	0.0004**	0.0004**	0.0004**	0.0004**
β_{smb}	0.0009*	0.0010**	0.0008*	0.0010*	0.0006	0.0009*	0.0007
β_{hml}	0.0012**	0.0012**	0.0012**	0.0012**	0.0010**	0.0013**	0.0011**
$p\mu_v^{la}$		0.2084**		0.2039**			
μ_v^{la}					0.0040**		0.0041**
p						0.0019	0.0026
$Illiq(MA)$			0.00004	0.00004	0.00003	0.00004	0.00002
R^2 (in %)	59	61	60	61	65	60	65

results show that the positive effect of liquidity events on the total traded volume is captured by the size of the order imbalances at the expense of their probability of occurrence.

6.2.2. Cross-sectional regressions

The portfolio analysis is further backed by cross-sectional regressions to control for the standard well-known variables that drive commonalities in stock returns. We first perform time-series regressions to estimate the Fama–French factor betas for each stock with $p\mu_v^{la}$ statistically significant, using the entire sample period:

$$R_{it} = \alpha_i + \beta_{i,m}R_{t,m} + \beta_{i,smb}R_{t,smb} + \beta_{i,hml}R_{t,hml} + \epsilon_{i,t}, \tag{5.1}$$

where R_{it} represents the day t return for stock i, whereas $R_{t,m}, R_{t,smb}$, and $R_{t,hml}$ represent the daily returns of the market portfolio, size, and book-to-market factors of Fama and French, respectively. Second, we perform the following cross-sectional regression:

$$\overline{R}_i = \gamma_0 + \gamma_m\beta_{i,m} + \gamma_{smb}\beta_{i,smb} + \gamma_{hml}\beta_{i,hml} + \lambda Z_i + \varepsilon_i. \tag{5.2}$$

In this equation, $\overline{R}_i$ represents the average daily return for stock i computed over the entire sample period, γ_j; $j = (m, smb, hml)$ represents the risk premium associated with risk factor j, and λ is a vector of k risk premiums associated with the vector Z of additional stock-specific characteristics, such as $p\mu_v^{la}, p, \mu_v^{la}$ or the Amihud adjusted liquidity measure, $Illiq(MA)_i$. Note that our liquidity measure and its two determinants are static and estimated using the entire sample period. For this reason, we cannot implement a cross-sectional out-of-sample methodology, as in Fama and Macbeth (1973). We perform a single in-sample cross-sectional regression using the average values of different variables computed through the entire sample period. Table 5 presents the estimated coefficients proceeding from the cross-sectional regressions of different specifications for Eq. (5.2) with $Z_i = [p_i\mu_{i,v}^{la}\mu_{i,v}^{la}p_iIlliq(MA)_i]$. The coefficient t-statistics are based on Newey and West (1987) robust standard errors.

Table 5 shows that our liquidity measure is priced in the cross-section of stock returns. This result is robust when controlling for Fama–French factors as well as for the Amihud adjusted liquidity measure. In addition, our results confirm those of the previous subsection, showing that commonality in the MDHL liquidity indicator is explained by its amplitude determinant; as expected, the probability of occurrence of order imbalances is not a priced factor.

Table 6 reports results for the cross-sectional regressions with an extended set of stock-specific variables: $Z_i = [p_i\mu_{i,v}^{la}\mu_{i,v}^{la}p_iIlliq(MA)_iSTD\ Volume\ ln(size)]$. Among the additional variables, only the variable related to stock size is statistically significant. Its negative risk premium confirms the size effect discussed by Fama and French (1992), among others. Table 6 shows that μ_v^{la} is always statistically significant, whereas the effect of $p\mu_v^{la}$ vanishes. This result confirms our findings in the previous subsection implying that commonality in the MDHL liquidity measure is explained by common variations across stock-specific order imbalance amplitudes as measured by μ_v^{la}.

Table 6
This table presents risk premiums estimated using cross-sectional regressions for different specifications of Eq. (5.2) with an extended set of control variables: $Z_i = [p_i\ \mu_{i,v}^{la}\ \mu_{i,v}^{la}\ p_i\ Illiq(MA)_i\ STD\ Volume\ ln(size)]$. The first column reports the factors used in the regression, whereas the values reported in columns 2 to 5 represent the risk premium coefficients associated with different factors. "*" and "**" indicate significance at 90% and 95% confidence levels, respectively. T-statistics are based on Newey and West (1987) robust standard errors. The last row presents the regression R^2 coefficients.

	Model 7	Model 8	Model 9	Model 10	Model 11
Intercept	0.0026**	0.0024**	0.0024**	0.0025**	0.0023**
β_m	0.0000	0.0000	0.0001	0.0000	0.0001
β_{smb}	0.0006	0.0006	0.0004	0.0006	0.0005
β_{hml}	0.0008**	0.0008**	0.0007*	0.0008**	0.0007*
$p\mu_v^{la}$		0.0999			
μ_v^{la}			0.0030*		0.0031*
p				0.0011	0.0017
$Illiq(MA)$	−0.0001	−0.0001	−0.0001	−0.0002	−0.0001
STD	0.0750	0.0757	0.0639	0.0768	0.0663
$Volume$	−0.0000	−0.0000	−0.0000	−0.0000	−0.0000
$ln(size)$	−0.0001**	−0.0001**	−0.0001**	−0.0001**	−0.0001**
$R^2(in\%)$	69	70	70	69	72

Overall, the portfolio and the factor model analyses provide consistent evidence that the amplitude of order imbalances commands a premium in the cross-section of returns. This premium is economically and statistically significant.

7. Concluding remarks

In this paper, we separate information from liquidity shock impact on the daily traded volume and measure the liquidity portion of volume. To do so, we focus on the theoretical framework of Grossman and Miller (1988) and develop a new version of the MDH that accounts for the presence of liquidity frictions. The MDHL model offers a better understanding of the determinants of the daily traded volume. The increase in volume due to liquidity providers helps infer the presence of liquidity frictions corresponding to order imbalances driven by the 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 liquidity measure using daily data.[35]

[35] Note that the MDHL liquidity-based indicator is analogical to that of Getmansky et al. (2004), who provide a static measure of the illiquidity affecting hedge fund returns, as measured by the serial correlation of fund returns. The time-series of hedge fund returns can then be used to estimate the serial correlation of individual funds, which helps to separate liquid from illiquid hedge funds for a given period.

This measure helps distinguish, for a given period, the liquid (non-significant $p\mu_v^{la}$) from the less liquid stocks (significant $p\mu_v^{la}$). Note that, separately estimating the probability of occurrence of liquidity frictions, p, and their amplitude, μ_v^{la}, provides a better comprehension of the stock liquidity profile as determined by the amplitude of order imbalances and the probability of their occurrence. This understanding may be useful for building high frequency trading stock-picking strategies.

In addition, the portfolio and the factor model analysis performed in Section 6 provide consistent evidence that the amplitude of order imbalances commands a premium in the cross-section of returns. This premium is economically and statistically significant. We also find that the presence of liquidity frictions can increase the daily traded volume and that the positive effect of liquidity events on total traded volume is captured by the size of order imbalances. Moreover, in terms of global validity, the MDHL model is strongly validated by the data in the GMM estimations, and outperforms the modified MDH version of Andersen (1996).

Finally, our liquidity indicator presents two primary limitations. First, it is a static indicator; as such, it fails to capture the time-varying dynamics of liquidity frictions. The second limitation concerns the impossibility of building a common (market-wide) liquidity factor using a stock-specific MDHL-based liquidity measure. Several recent studies are based on the commonality and time-varying properties of liquidity risk. In the same spirit, Patton and Li (2007) extend Getmansky et al. (2004)'s analysis by allowing serial correlation parameters to vary over time, thus providing a dynamic time-dependent proxy of liquidity for individual hedge funds. Alternatively, Nagel (2009) uses the profitability of contrarian strategies as a proxy for returns that compensate for 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 this type of indicator is that it provides information on how market liquidity evolves over time and the factors that determine its evolution.

Therefore, it would be interesting to expand our stock-specific approach to extract time-varying latent liquidity factors for individual stocks. For this purpose, the MDHL version developed in this paper can be extended to allow for time-persistence of liquidity problems, which may explain the dynamics of daily returns and volume. Several studies show that liquidity shocks are not isolated events in time but rather appear to be time persistent [see, for example, Acharya and Pedersen (2005)]. Signal extraction methods can then be used to filter the latent liquidity process for individual assets to provide a time-varying, stock-specific liquidity indicator. Thus, factor decomposition analysis can be applied to the panel of individual liquidity indicators to build market-wide liquidity factors and thus to separate, for a given stock, the common from the specific liquidity components. These points are outside of the scope of this paper and are part of our ongoing research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbankfin.2014.09.007.

References

Acharya, V.V., Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 375–410.

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.

Bamber, L., 1986. The information content of annual earning releases: a trading volume approach. Journal of Accounting Research 24, 40–56.

Bamber, L., 1987. Unexpected earnings, firm size and trading volume around quarterly earnings announcements. The Accounting Review, 510–532.

Bialkowski, J., Darolles, S., Fol, G.L., 2008. Improving VWAP strategies: a dynamic volume approach. Journal of Banking and Finance 32, 1709–1722.

Board, J., Sutcliffe, C., 1990. Information, volatility, volume and maturity: an investigation of stock index futures. Review of Futures Markets 9, 532–549.

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

Brunnermeier, M.K., Pedersen, L.H., 2009. Market liquidity and funding liquidity. The Review of Financial Studies 22 (6), 2201–2238.

Campbell, J., Grossman, S., Wang, J., 1993. Trading volume and serial correlation in stock returns. Quarterly of Journal of Economics 108, 905–939.

Chen, Z., Daigler, R.T., 2008. An examination of the complementary volume–volatility information theories. The Journal of Future Markets 28 (10), 963–992.

Chordia, T., Subrahmanyam, A., 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.

Darolles, S., Fol, G.L., 2014. Trading volume and arbitrage. Journal of Business Review 3 (4).

Datar, V.T., Naik, N.Y., Radcliffe, R., 1998. Liquidity and stock returns: an alternative test. Journal of Financial Markets 1, 203–219.

DeLong, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990a. Noise Trader Risk in Financial Markets.

DeLong, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990b. Positive Feedback Investment Strategies and Destabilized Rational Speculation.

Domowitz, I., Wang, J., 1994. Auctions as algorithm: computerized trade execution and price discovery. Journal of Economic Dynamics and Control 18, 29–60.

Foster, D.F., 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.

Fama, E.F., Macbeth, J.D., 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81 (3), 607–636.

Fama, E.F., French, K.R., 1992. The Cross-Section of Expected Stock Returns. Journal of Finance 47 (2), 427–465.

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

Getmansky, M., Lo, A.W., Makarov, I., 2004. An econometric analysis of serial correlation and illiquidity in hedge-fund returns. Journal of Financial Economics 74, 529–609.

Grossman, S.J., 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.

Harris, M., Raviv, A., 1993. Differences of opinion make a horse race. Review of Financial Studies 6, 473–506.

Hendershott, T., Riordan, R., 2013. Algorithmic trading and the market for liquidity. Journal of Financial and Quantitative Analysis 48, 1001–1024.

Kawaller, I.G., Koch, P.D., Koch, T.W., 1990. Intraday relationships between volatility in S&P 500 futures prices and volatility in the S&P index. Journal of Banking and Finance 14, 373–398.

Lakonishok, J., Smidt, S., 1986. Volume for Winners and Losers: Taxation and Other Motives for Stock Trading.

LeBaron, B., 1992. Persistence of the Dow Jones Index on Rising Volume.

Li, J., Wu, C., 2006. Daily return volatility, bid-ask spreads and information flow: analyzing the information content of volume. Journal of Business 79 (51), 2697–2739.

Lo, A., Wang, J., 2000. Trading volume: definition, data analysis and implication of portfolio theory. The Review of Financial Studies 13, 257–300.

Menkveld, A.J., 2013. High frequency trading and the new-market makers. Journal of Financial Markets 16, 712–740.

Morse, D., 1980. Asymmetric information in securities markets and trading volume. Journal of Financial and Quantitative Analysis 15, 1129–1148.

Nagel, S., 2009. Evaporating Liquidity, Working Paper NBER.

Newey, W.K., West, K.D., 1987. Hypothesis testing with efficient method of moments estimation. International Economic Review 28, 777–787.

Patton, A.J., Li, S., 2007. Time-Varying Liquidity in Hedge Fund Returns, Working Paper, The Oxford-Man Institute – University of Oxford.

Richardson, M., 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.

Richardson, G., Sefcik, S., Thompson, R., 1986. A Test of Dividend Irrelevance Using Volume Reaction to a Change in Dividend Policy.

Roskelley K., 2001. Using Information Flow to Explain Volume, Volatility and Conditional heteroscedasticity in Returns, Working Paper, Louisiana Tech University.

Shalen, C.T., 1993. Volume, volatility and the dispersion of beliefs. Review of Financial Studies 6, 406–434.

Smidt, S., 1990. Long run trends in equity turnover. Journal of Portfolio Management, 66–73.

Tauchen, G., 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.