UMR 8184
THéorie Économique, Modélisation et Applications

THEMA Working Paper n°2016-08
Université de Cergy-Pontoise, France

Financial Market Liquidity: Who Is Acting Strategically?

Serge Darollesa, Gaëlle Le Fol, Gulten Mero

June, 2016

UNIVERSITÉ de Cergy-Pontoise
CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Financial Market Liquidity: Who Is Acting Strategically? *

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

June 30, 2016

Abstract

In a new environment where liquidity providers as well as liquidity consumers act strategically, understanding how liquidity flows and dries-up is key. In this paper, we propose a dynamic extension of the seminal model of Tauchen and Pitts (1983)' Mixture of Distributions Hypothesis (MDH), that specifies the impact of information arrival on market characteristics in the context of liquidity frictions. In our model, the daily price change and volume processes are represented by a bivariate mixture of distributions, conditioned by two latent time-persistent variables I_t and L_t. Since the price change and volume equations are nonlinear functions of the first latent variable I_t, we use an Extended Kalman Filter (EKF) to filter the two latent variables and estimate the model parameters, simultaneously. This procedure enables us to: (i) capture the impact of long-lasting liquidity frictions on the daily price change and volatility dynamics; (ii) separate out the impact of both long and short-lasting liquidity frictions, on the serial correlation of the daily volume. Our results show that, 48% (44/92) of the stocks of the FTSE100 are actually facing liquidity problems. Amongst these stocks, 28% (26/92) of them are also facing a slow-down in the information propagation in prices due to long-term investors' strategic behavior.

JEL classification: C51, C52, G12, G14

Key words: Strategic liquidity trading, market efficiency, mixture of distribution hypothesis, information-based trading, extended Kalman Filter.

*An earlier version of the paper was circulated under the title "Tracking illiquidities in intradaily and daily characteristics".The authors thank T. Andersen, R. Engle, C. Hurlin, A. Monfort, G. Tauchen, J.M. Zakoian as well as the participants of the 2012 European Meeting of the Econometric Society (ESEM), the 8th International Conference on Computational and Financial Econometrics (CFE 2014), the 2014 French Econometric conference, the 2015 AFFI Conference, and the 2016 SoFIE conference. 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).

[a]Université Paris-Dauphine, PSL Research University, CNRS, UMR 7088, DRM, Finance, 75016 Paris, France, and CREST-INSEE, France.

[b]THEMA - Université de Cergy-Pontoise, France, gulten.mero@u-cergy.fr, 33 Bd du Port 95011 Cergy-Pontoise cedex, Tel. +33 (0)6 30 01 33 70.

1 Introduction

Financial markets liquidity is a latent characteristic. For decades now, market participants as well as academics have been trying to filter liquidity from market summaries such as prices, volatility and volume. The vast literature on that topic [see e.g. Aitken and Comerton-Forde (2003), and Goyenko et al. (2009)] shows that there was no consensus at that time on how liquidity should be measured. Moreover, are these summaries still able to reflect market liquidity with the significant changes that occurred in the recent years? In fact, while automation was taking over the trading world and spreads were falling, the historical market maker[1] tends to disappear to the benefit of strategic liquidity providers. This new category of investors has accompanied and even favored the decrease in transaction costs by competing against each other and providing more liquidity as in the past [See for example, Chaboud et al. (2014), Hendershott et al. (2011), and Menkveld (2013)]. However, they also have been accused to provide only illusory liquidity, to weaken financial markets, and even to be responsible for crashes and illiquidity spirals by massively taking liquidity out of the market during periods of turmoil [Kirilenko et al. (2014)]. Jarrow and Protter (2013) show that their profits are at the expense of the other market participants, i.e. the medium to long-term investors who must act strategically by splitting their orders in response.[2] As a consequence, the link between market characteristics and liquidity is even more ambiguous now that liquidity providers as well as liquidity consumers are both acting strategically.

In this paper, we propose a dynamic extension of the seminal model of Tauchen and Pitts (1983)' Mixture of Distributions Hypothesis (MDH), that specifies the impact of information arrival on market characteristics in the context of liquidity frictions. We distinguish between two types of liquidity frictions. On the one hand, the liquidity frictions can be short-lasting and are represented by the intra-day order imbalances that are resorbed within the trading

[1]A market maker stands ready to buy and sell stocks listed on a stock exchange. He quotes bid and ask prices continuously and provide a required amount of liquidity to the security's market. He is supposed to take the other side of trades when there are short-term imbalances in customer orders.

[2]For Jarrow and Protter (2013), these other traders are ordinary traders who submit market orders and sophisticated traders who submit limit orders or trade large quantity splitting up their orders.

day: they correspond to the liquidity provision - or arbitrage, provided by the "new market makers" à la Menkveld (2013), who end the trading day flat. They impact the daily and intra-daily traded volume but only the intra-day price volatility. These liquidity frictions are a source of trade for the market makers who liquidate their positions once prices are back to the equilibrium level in order to cash the liquidity premium. On the other hand, the liquidity frictions can be long lasting, when information is not completely incorporated into prices within the day. Such a situation can arise when investors are strategically timing their trades, and trade at prices that do not fully reflect the information they possess [see Anderson et al. (2013)]. This strategic behavior could explain the increase in the number of daily transactions that comes with the decrease of the average trade size observed in all financial markets in the past 10 years [See for example Lehalle and Laruelle (2013), Fig. 5 p. 17 for the components of the FTSE]. It is also responsible for the autocorrelation of prices that we observe on financial markets [See Anderson et al. (2013), and Toth et al. (2014)].

This paper contributes to the literature in several directions. First, we give a theoretical ground to the empirical financial literature dedicated to the measure of illiquidity through the analysis of positive returns autocorrelation [see e.g. Getmansky et al. (2004)]. In our model, these positive serial correlations are directly linked to the lack of liquidity provision at the intra-day level. This liquidity deficit generates long-lasting liquidity frictions, and in turn positive returns autocorrelation. The short-term liquidity frictions are responsible for intra-day dynamics of return and volatility, while the long-lasting ones result in daily positive serial correlation of stock returns and squared returns.

Second, we extend the analysis of liquidity frictions to the case of information arrival. On no information days, liquidity frictions should push prices away from their equilibrium level (temporary effect) before returning to their previous level. On information days, liquidity frictions should also temporary prevent prices from becoming fully revealing before they revert (permanent effect). In the first case, the inefficiency is considered as over whenever the price revers, while in the second case, the interaction between information and liquidity frictions clearly complicates the issue [see Waelbroeck and Gomes (2013) for a discussion on the permanent effect and information contents of trades]. Working with triangular arbitrage relations is a way to concentrate on frictions without taking care of the information problem [see for example Foucault et al. (2014)]. However, this approach cannot be generalized to single risky assets as it is only relevant for arbitrage relations. Our approach allows to disentangle short versus long-lasting liquidity frictions in the general case where information arrival has an impact on prices.

Third, we propose new liquidity measures that add to the extensive literature [see e.g. Aitken and Comerton-Forde (2003), and Goyenko et al. (2009)]. Just like the new growing literature mostly coming from the trading cost analysis [see for example Almgren and Chriss (2001), Almgren et al. (2005), Criscuolo and Waelbroeck (2013), and Lehalle (2014)], these new measures are dynamic, they separate long-term from short-term effects and are linked to the autocorrelations of returns, volume and volatility. However, conversely to that stream, our measures do not need intra-day data.

Most long-term investors are only considering daily data even if intra-day strategies are impacting their performances. Here, we take the perspective of such investors and explain how intra-day strategies, information arrival, and liquidity frictions are affecting intra-daily and in turn, daily returns and volume evolutions. In our model, the daily price change and volume processes are represented by a bivariate mixture of distributions conditioned by two latent time-persistent variables I_t and L_t. Since the price change and volume equations are nonlinear functions of the first latent variable I_t, we use an Extended Kalman Filter (EKF) to filter the two latent variables and to estimate the model parameters, simultaneously. This procedure enables us to: (*i*) capture the impact of long-lasting liquidity frictions on the daily price change and volatility dynamics; (*ii*) separate out the impact of both long and short-lasting liquidity frictions, on the serial correlation of the daily volume.

Our results show that, over the 48% (44/92) of stocks, taken from the FTSE100, and that are actually facing liquidity problems, the liquidity providers are the only investors acting strategically on a group of stocks accounting for nearly 20% (18/92). For another 25% (23/92) of the stocks, the long-term investors are strategic while the liquidity providers are inactive. Finally, for the last 3% (3/92) group of stocks, the liquidity providers as well as the liquidity consumers are both acting strategically. It means that 28% (26/92) of the stocks of our sample are facing a slow-down of the information propagation to prices, and thus a decrease of (daily) price efficiency, due to the strategic behavior of the long-term investors. Based on some goodness of fit tests, as well as on some simple strategies based on our structural model, we show the superiority of our approach against other MDH' specifications.

The paper is organized as follows. Section 2 introduces our model. Section 3 presents our statistical methodology to extract information and liquidity latent factors from the daily time series of returns and volume. In Section 4, we apply our econometric set up to individual stocks belonging to the FTSE100 and discuss the empirical results. Section 5 explores the validity of our approach. Section 6 concludes the paper.

2 The model and its implications

In this section, we consider a dynamic extension of Tauchen and Pitts (1983), where liquidity frictions can be both short-term and/or long lasting. We first characterize stocks liquidity frictions' types based on the parameters that form the basis of the hypothesis testing. We then use this extension to show how the dynamic relationship between daily returns and traded volume as well as the serial correlation of price changes, squared price changes and traded volume depend on these frictions.

2.1 The model

In our scheme, the economy is characterized by a trading sequence in a single risky asset by two types of market participants: Liquidity Consumers (LC) who trade in response to information flow, and Liquidity Providers (LP) who trade in response to liquidity frictions. Within the trading day, the market passes through a sequence of distinct equilibria due to the arrival of new pieces of information. These information arrivals generate trades that impact the intra-daily price increments and transaction volumes, and in turn, the daily price changes and volumes.

In the absence of liquidity frictions, the intra-daily as well as daily price changes and volumes fully reflect the incoming information instantaneously. In this case, we get the standard Mixture of Distribution Hypothesis (MDH) framework of Tauchen and Pitts (1983). This model suggests that a bivariate mixture of distributions, in reduced form, can represent daily price changes and volumes, with information being the only mixing variable. Let us consider one single asset and J active participants with heterogeneous reservation prices who trade in response to information only. At each information arrival time i, each trader j revises his reservation price $\bar{P}_{ij}$ and ask for the quantity $Q_{ij} = a(P_{ij} - \bar{P}_{ij})$, $j = 1, ..., J$. Hence, the price change ΔP_i and the traded volume V_i only come from the information arrival at time i. A day is a sequence of i consecutive equilibria initiated by a random number I_t of information. We get at the daily level:

$$\Delta P_t = \sum_{i=1}^{I_t} \Delta P_{it}, \qquad V_t = \sum_{i=1}^{I_t} V_{it}.$$

At the end of the day, and under a gaussian assumption: $\Delta P_{it} \sim N(0, \sigma_p^2)$ and $V_{it} \sim N(\mu_{v1}, \sigma_v^2)$, Tauchen and Pitts (1983) obtain that the daily price change and volume processes are represented by a bivariate mixture of independent Gaussian distributions with the same mixing variable:

$$\Delta P_t = \sigma_p \sqrt{I_t} Z_{1t}, \tag{2.1}$$

$$V_t = \mu_{v1} I_t + \sigma_v \sqrt{I_t} Z_{2t}, \tag{2.2}$$

with Z_{1t} and Z_{2t} being two mutually-uncorrelated and serially-uncorrelated Gaussian variables with zero means and unit variances. In this simple model, we note that the variance of the two processes are proportional to the same latent variable I_t. The more information reaches the market during the day, the more volatile returns and traded volumes get.

Because trading might be, and usually is, asynchronous, short-term liquidity frictions exist creating at least temporary order imbalances at the intra-day frequency. These order imbalances prevent prices from being fully revealing. The price deviations they generate are resorbed within the trading day thanks to the intervention of liquidity providers. These "new market makers" à la Menkveld (2013), who end the trading day flat, are strategic. As discussed in Darolles et al. (2015), liquidity providers trade on short-term liquidity frictions and increase the daily traded volume by doing so. However, they do not impact the daily price changes; at the end of the trading day prices reflect all the incoming information during that day. The Mixture of Distribution Hypothesis with Liquidity (MDHL) model of Darolles et al. (2015) suggests that the daily price changes and volumes can be represented by a bivariate mixture of distributions, in reduced form, with two latent variables, I_t and L_t supposed to be i.i.d.. Compared to the Tauchen and Pitts (1983) specification corresponding to Equations (2.1)-(2.2), the second equation corresponding to the traded volume includes an additional term capturing the increase of volumes due to short-term liquidity frictions denoted L_t:

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

$$V_t = \mu_{v1} I_t + \mu_{v2} L_t + \sigma_v \sqrt{I_t} Z_{2t}, \tag{2.4}$$

with Z_{1t} and Z_{2t} being two mutually-uncorelated and serially-uncorrelated Gaussian variables with zero means and unit variances. In this formulation, the joint model for price changes and volumes captures short term liquidity frictions but does not capture potential autocorrelations on the time series of price changes and traded volumes.

We now go further in the analysis. We allow for the presence of long lasting liquidity frictions reflecting situations where information is not completely incorporated into prices within the trading day. This can arise when only part of market participants are able to trade for any reason. This may also happen when long-term investors, as liquidity consumers, are slicing efficiently in order not to be spotted by the strategic liquidity providers. By doing so, they also become strategic and disseminate slowly, over several days or even weeks, the information they possess to the market.

Let ΔP_t and V_t be the daily price change and traded volume, respectively. The statistical model proposed here follows Richardson and Smith (1994) and Andersen (1996). Richardson and Smith (1994) assume that information has an impact on the price variation level, while Andersen (1996) consider a time-persistent information levels. Combining these two features to the short-lasting liquidity frictions model leads to a specification that covers the consequences of slicing on the price change and the volume dynamics. The model corresponds to a bivariate mixture of distributions, in reduced form, conditioned by two latent variables, I_t and L_t, which are supposed to be time-persistent:

$$\Delta P_t = \mu_p I_t + \sigma_p \sqrt{I_t} Z_{1t}, \tag{2.5}$$

$$V_t = \mu_{v1} I_t + \mu_{v2} L_t + \sigma_v \sqrt{I_t} Z_{2t}, \tag{2.6}$$

where

$$\ln I_t = \beta \ln I_{t-1} + \eta_t, \tag{2.7}$$

$$L_t = a L_{t-1} + \omega_t, \tag{2.8}$$

with η_t and ω_t being two i.i.d. mutually independent variables. The intuition behind this model, and in particular the first equation relative to price change, can be found in the ARCH-in-mean model [Engle et al. (1987)] or the stochastic volatility-in-mean model [Koopman and Upensky (2002)] developed to capture the positive relationship between the expected returns and the volatility level. If μ_p is positive, then $\mu_p I_t$ increases with $\sigma_p \sqrt{I_t}$:

the expected reward increases with the risk.

Note that, I_t and L_t capture the effects of long-term and short-term liquidity frictions, respectively. Indeed, the first latent variable I_t represents the portion of the incoming information which is absorbed by the market at day t. We allow I_t to be time-persistent in order to capture the effect of long lasting liquidity frictions on daily price change and traded volume. The more important the slicing of orders, the smaller each bit of I_t and the longer the persistence time.

2.2 Implications in terms of liquidity frictions

To give a basic outline, and following our specification, short-term liquidity frictions create excess volume but have no particular impact on prices: the key parameter is μ_{v2}. Long term liquidity frictions create some directional price movement but no excess volume: the key parameters are μ_p and β. Going into details, we can picture four different stock environment types:

(i) The first type consists in stocks with short-term liquidity frictions. They are characterized by $\mu_{v2} \neq 0$ and either $\mu_p = 0$, or if $\mu_p \neq 0$ then $\beta = 0$. We call this stock category, the pure short-term liquidity friction case. Note that, this short-term liquidity friction case differs from that of Darolles et al. (2015) since in our framework, the L_t process is supposed to be time-persistent. For these stocks, liquidity providers are strategic while liquidity consumers are not. The latters are trading at once or at least within the day and the formers are providing the missing liquidity. By the end of the day, there is no more liquidity friction left in the market and the prices contain all the information possessed by market participants.

(ii) The second category concerns stocks that are facing long-term liquidity frictions. We have $\mu_p \neq 0$, $\beta \neq 0$ and $\mu_{v2} = 0$. For these stocks, the liquidity providers are inactive, maybe because liquidity consumers are strategic and efficient in hiding from the liquidity arbitragers. As a consequence, information does not disseminate to the prices accurately.

(iii) Stocks belonging to the third category are affected by both long lasting and short-term liquidity frictions (the mixed liquidity friction case). For these stocks, $\mu_p \neq 0$, $\beta \neq 0$ and $\mu_{v2} \neq 0$. The short-term liquidity providers' activity (μ_{v2}) is not enough to ease the trading so that the incoming information is not completely revealed on the day of its arrival. It means that both type of participants are acting strategically but none of them is efficient in doing so. The short-term liquidity frictions diminish because liquidity providers are trading. However, because the liquidity consumers are splitting up their orders, information propagates slowly to the prices over several days.

(*iv*) The last category is made of stocks which do not suffer from liquidity frictions. If $\mu_p = 0$, $\beta = 0$, and $\mu_{v2} = 0$, we get the Tauchen and Pitts (1983) case. When $\mu_p = 0$, $\beta \neq 0$ and $\mu_{v2} = 0$, we get the Andersen (1996) case.

Finally, estimating the parameters of our model enables us to test, the four hypotheses presented above in order to characterize the liquidity profile of each stock during a given period of time. We are able to test whether a stock has been facing liquidity frictions during the test period. And if this is the case, we can go further in the analysis by specifying the type of liquidity frictions (long-lasting, short-term liquidity frictions or both at the same time) affecting the stock trading characteristics.

2.3 Implications in terms of serial correlations

The empirical financial literature reports that returns are highly autocorrelated and this serial correlation is linked to liquidity [see for example Getmansky et al. (2004) for hedge fund investments, and Säfvenblad (2000) for review of the empirical literature for stocks]. Our model allows for a better understanding of serial correlations of returns, volume and volatility: these autocorrelations are due to information and liquidity, which translates into short-term and/or long lasting liquidity frictions. As we will see in this section, any classification based on the empirical autocorrelations will not allow to disentangle between the different sources.

Equations (2.5) and (2.7), taken jointly, define a stochastic volatility model. The presence of long-lasting liquidity frictions is responsible for the persistence of the stochastic volatility in the daily price change time series, as captured by I_t, and more particularly by β measuring its persistence. Estimating the parameters in (2.5) and (2.7) for individual stocks allows us to infer the presence of long-lasting liquidity frictions; a stock affected by long-term liquidity frictions should present statistically significant μ_p. In the absence of long-lasting liquidity frictions, μ_p should equal to zero, and equations (2.5) and (2.7) reduce to the daily price change equation of Darolles et al. (2015), with I_t representing an i.i.d. information process.

Equations (2.5) and (2.7) imply that daily returns and squared returns are serially correlated. Their respective autocovariances are given by:

$$Cov(\Delta P_t, \Delta P_{t+1}) = \mu_p^2 Cov(I_t, I_{t+1}), \tag{2.9}$$

$$Cov(\Delta P_t^2, \Delta P_{t+1}^2) = \sigma_p^4 Cov(I_t, I_{t+1}) + \mu_p^4 Cov(I_t^2, I_{t+1}^2) + \mu_p^2 \sigma_p^2 Cov(I_t, I_{t+1}^2) + \mu_p^2 \sigma_p^2 Cov(I_t^2, I_{t+1}). \tag{2.10}$$

Equation (2.9) shows that the presence of serial correlation in daily price changes results from the interaction of the drift effect of I_t on prices (μ_p parameter) with the serial correlation pattern of the I_t process ($Cov(I_t, I_{t+1})$). As for equation (2.10), the dynamics of daily volatility is entirely due to the presence of long lasting liquidity frictions; it results from a combination of information intensity μ_p and precision σ_p with the time-persistence of I_t process. Note also that, the short-term liquidity frictions have no impact on the dynamics of neither the daily price change, nor the squared price change.

Since the daily traded volume depends on both the I_t and the L_t processes, its dynamics will be impacted by both the occurrence of the long-lasting liquidity frictions, through μ_{v1}, and the persistence of short-term liquidity frictions, through μ_{v2}. More precisely, equations (2.6) and (2.8) imply that:

$$Cov(V_t, V_{t+1}) = \mu_{v1}^2 Cov(I_t, I_{t+1}) + \mu_{v2}^2 Cov(L_t, L_{t+1}). \tag{2.11}$$

This equation shows that the unconditional autocovariance of daily traded volume results from the interaction of long lasting and short-term liquidity frictions. Based on our model, it is possible to disentangle the two types of liquidity frictions, which cannot be assessed by simply computing the serial correlation coefficients from daily traded volume time-series. In this context, and , based on the estimated parameters of our model, we can separate the impacts of both types of liquidity frictions on volume serial correlation.

Our model encompasses the MDH version of Andersen (1996), and thus the seminal MDH model of Tauchen and Pitts (1983). We have:

(a) Andersen (1996) proposed an *ad hoc* dynamic version of his static modified MDH model in order to account for the time-persistence of daily squared price change and traded volume series. In Andersen's modified MDH version, daily price change and traded volume series are generated by a unique latent variable, the information flow process. Andersen's dynamic MDH model explains the time-persistence of daily squared returns and traded volumes by the serial correlation of the information flow process. However, the author does not account for drift effects of I_t on the daily price change. Our model reduces to Andersen's dynamic MDH version for $\mu_p = 0$, and $\mu_{v2} = 0$. In particular, Andersen's framework implies, that Equations 2.9-2.11 resum to:

$$
\begin{aligned}
Cov(\Delta P_t, \Delta P_{t+1}) &= 0, \\
Cov(\Delta P_t^2, \Delta P_{t+1}^2) &= \sigma_p^4 Cov(I_t, I_{t+1}), \\
Cov(V_t, V_{t+1}) &= \mu_{v1}^2 Cov(I_t, I_{t+1}).
\end{aligned}
$$

As compared to our model, Andersen's dynamic MDH version does not account for time-persistence of daily price changes while this stylized fact is empirically well-known. In addition, the intuition behind the presence of stochastic volatility is not the same in both models. In Andersen (1996), the information flow time-persistence is responsible for the presence of stochastic volatility. In our framework, although information process is considered to be i.i.d., the presence of long-lasting liquidity frictions modifies the way news are incorporated into daily price changes. Thus, the stochastic volatility is the consequence of the presence of long-lasting liquidity problems. Finally, in Andersen's model, the serial correlation of daily traded volume is due to the time-persistence of information flow process. In our framework, volume serial correlation is due to the presence of both long-lasting and short-term liquidity frictions.

(b) If we disregard the effect of both types of liquidity frictions on daily trading characteristics by setting $\mu_p = 0$, $\beta = 0$, and $\mu_{v2} = 0$, we get the standard MDH model of Tauchen and Pitts (1983) implying that:

$$
\begin{aligned}
Cov(\Delta P_t, \Delta P_{t+1}) &= 0, \\
Cov(\Delta P_t^2, \Delta P_{t+1}^2) &= 0, \\
Cov(V_t, V_{t+1}) &= 0.
\end{aligned}
$$

Note that the MDH model of Tauchen and Pitts (1983) shares the same implications in terms of autocorrelation with some alternative MDH versions, like the MDHL model of Darolles et al. (2015). While the former is a no liquidity friction model, the latter accounts for short-term liquidity frictions. As a consequence, serial correlations can only be rough indicators of stock liquidity frictions, and this advocate for our structural model.

3 Estimation

In the model (2.5)-(2.8), the daily price change and the volume processes are represented by a bivariate mixture of distributions conditioned by two latent time-persistent variables I_t and L_t. Since the price change and the volume equations are nonlinear functions of the first latent variable I_t, we use an Extended Kalman Filter (EKF) to filter the two latent variables and estimate the model parameters, simultaneously. The EKF is a widely used class of algorithms to approximate the non-linear measurement or transition equations using linearization technique [see e.g. Anderson and Moore (1979)]. We first provide the state-space formulation of the model based upon a first-order Taylor expansion of the measurement equations. Then, we present the associated algorithm together with the estimation procedure.

3.1 State space formulation of the model

Let $X_t = (X_{1,t},\, X_{2,t})$ be the vector of state variables implied by Equations (2.7)-(2.8). The variables $X_{1,t} = \ln I_t$ and $X_{2,t} = L_t$ represent the dynamic linear processes given by (2.7) and (2.8), respectively. The transition equation can then be formulated as follows:

$$X_t = f(X_{t-1}, w_t) = FX_{t-1} + w_t, \tag{3.1}$$

where f is a linear function defined by the transition matrix:

$$F = \begin{pmatrix} \beta & 0 \\ 0 & a \end{pmatrix},$$

and $w_t = (\eta_t,\, \omega_t)'$ is a vector of two mutually-uncorrelated sequences of temporally-uncorrelated Gaussian additive white noise with covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_\eta^2 & 0 \\ 0 & \sigma_\omega^2 \end{pmatrix}.$$

Let $Y_t = (Y_{1,t},\, Y_{2,t})'$ be the measurement vector implied by Equations (2.5)-(2.6). The variables $Y_{1,t} = \Delta P_t$ and $Y_{2,t} = V_t$ correspond to the observed daily price changes and observed daily volumes. Based on (2.5)-(2.6), the measurement equations relating each element of Y_t to the state variables can be formulated as follows:

$$Y_t = h(X_t, \delta_t) = \left[\begin{array}{c} \mu_p \exp(X_{1,t}) + \sigma_p \sqrt{\exp(X_{1,t})} Z_{1t} \\ \mu_{v1} \exp(X_{1,t}) + \mu_{v2} X_{2,t} + \sigma_v \sqrt{\exp(X_{1,t})} Z_{2t} \end{array} \right], \tag{3.2}$$

where h is a function of the state variable X_t and $\delta_t = (Z_{1t}\ , Z_{2t})$ is a vector of two mutually-uncorrelated and serially-uncorrelated Gaussian variables with zero mean and unit variances. These measurement equations are non-linear functions of the state variables X_t with non-additive noises, preventing us from using the simple Kalman Filter (KF here after). Instead, we implement the EKF procedure consisting in linearizing the measurement equations (3.2) before applying the KF. Indeed, applying the first-order Taylor expansion to $h(X_t, \delta_t)$ around $X_{t|t-1}$ and 0, we get:[3]

$$h(X_t, \delta_t) \approx h(X_{t|t-1}, 0) + H_{t|t-1}(X_t - X_{t|t-1}) + D_{t|t-1}\delta_t, \tag{3.3}$$

where A_t and H_t are the Jacobian matrices of the non-linear function h with respect to the state vector and the system noise respectively:

$$H_{t|t-1} = \frac{\partial h}{\partial X_t'}(X_{t|t-1}, 0) = \begin{pmatrix} \mu_p \exp(X_{1,t|t-1}) & 0 \\ \mu_{v1} \exp(X_{1,t|t-1}) & \mu_{v2} \end{pmatrix}, \tag{3.4}$$

$$D_{t|t-1} = \frac{\partial h}{\partial \delta_t'}(X_{t|t-1}, 0) = \begin{pmatrix} \sigma_p \sqrt{\exp(X_{1,t|t-1})} & 0 \\ 0 & \sigma_v \sqrt{\exp(X_{1,t|t-1})} \end{pmatrix}. \tag{3.5}$$

We can now apply the KF algorithm to the system including the transition equations (3.1) and the linearized measurement equations (3.3), which yields the EKF. The following subsection presents the EKF algorithm as well as the statistical procedure for estimating the model parameters and filtering the state variables simultaneously.

[3]Note that the best estimate of the state vector at time t, $X_{t|t-1}$ is the expectation of X_t given the available information up to $t-1$: $X_{t|t-1} = E[X_t|\underline{Y_{t-1}}]$, where $\underline{Y_{t-1}} = Y_{t-1}, Y_{t-2,...}$. As it will be presented in the next subsection, $X_{t|t-1}$ is computed via the prediction phase of the EKF algorithm.

3.2 Parameter estimation

Let us denote $X_{t-1|t-1} = E[X_{t-1}|\underline{Y_{t-1}}]$, where $\underline{Y_{t-1}} = Y_{t-1}, Y_{t-2,\ldots}$, the optimal estimator of X_{t-1} based on the observations up to Y_{t-1} and $P_{t-1|t-1}$ the corresponding 2×2 covariance matrix of the estimation error: $P^X_{t-1|t-1} = E[(X_{t-1} - X_{t-1|t-1})(X_{t-1} - X_{t-1|t-1})^{'}|\underline{Y_{t-1}}]$. We can now details the different steps involved in the KF algorithm.

State prediction. Given $X_{t-1|t-1}$ and $P_{t-1|t-1}$, the optimal estimator of X_t is given by:

$$X_{t|t-1} = E[X_t|\underline{Y_{t-1}}] = FX_{t-1|t-1}, \tag{3.6}$$

while the covariance matrix of the estimation error is:

$$P^X_{t|t-1} = E[(X_t - X_{t|t-1})(X_t - X_{t|t-1})^{'}|\underline{Y_{t-1}}] = FP^X_{t-1|t-1}F^{'} + \Sigma, \tag{3.7}$$

where F and Σ are defined in the previous subsection. Equations (3.6)-(3.7) correspond to the state prediction equations.

Measurement prediction. From the optimal estimator of X_{t-1}, we can predict the observable variables:

$$Y_{t|t-1} = h(X_{t|t-1}, 0) = \begin{bmatrix} \mu_p \exp(X_{1,t|t-1}) \\ \mu_{v1} \exp(X_{1,t|t-1}) + \mu_{v2} X_{2,t|t-1} \end{bmatrix}, \tag{3.8}$$

and the covariance matrix of the prediction error:

$$M_{t|t-1} = E[(Y_t - Y_{t|t-1})(Y_t - Y_{t|t-1})^{'}|\underline{Y_{t-1}}] = H_{t|t-1}P^X_{t|t-1}H^{'}_{t|t-1} + D_{t|t-1}D^{'}_{t|t-1}. \tag{3.9}$$

where $H_{t|t-1}$ and $D_{t|t-1}$ are defined in the previous subsection. Equations (3.8)-(3.9) corresponds to the measurement prediction equations.

State updating. Once a new observation Y_t becomes available, the state equation can be updated using the two following updating equations:

$$X_{t|t} = X_{t|t-1} + P^X_{t|t-1} H'_{t|t-1} M^{-1}_{t|t-1} (Y_t - Y_{t|t-1}), \tag{3.10}$$

and

$$P^X_{t|t} = P^X_{t|t-1} - P^X_{t|t-1} H'_{t|t-1} M^{-1}_{t|t-1} H_{t|t-1} P^X_{t|t-1}, \tag{3.11}$$

Taken together, equations (3.6)-(3.11) are the EKF, whose recursive application yields the time series of the state variables X_t conditional on $\theta = (\mu_p, \sigma_p, \beta, \sigma_\eta, \mu_{v1}, \mu_{v2}, \sigma_v, a, \sigma_\omega)$. Given the vector of parameters θ, the calibration can be carried out via a Maximum Likelihood Estimator (MLE). We first apply the EKF (3.6)-(3.11) to the observed daily price change and traded volume time series, and get, for each t, $t = 1, ..., T$, the one-step-ahead error vector $Y_t - Y_{t|t-1}$ and the associated covariance matrix $M_{t|t-1}$. The next step is to use these two time series to construct the log likelihood function:

$$L(\theta) = -T \ln(2\pi) - \frac{1}{2} \sum_{t=1}^{T} \ln |M_{t|t-1}| - \frac{1}{2} \sum_{t=1}^{T} (Y_t - Y_{t|t-1}) M^{-1}_{t|t-1} (Y_t - Y_{t|t-1})'. \tag{3.12}$$

Finally, numerical methods can be used to maximize the log-likelihood function (3.12) with respect to the unknown vector of parameters θ.

4 Empirical applications

4.1 The data

Our sample consists in all FTSE100 stocks listed on March 27, 2014. We consider the period from January 1^{st}, 2010 to December 31, 2013, i.e., 1009 observation dates. We exclude stocks with missing observations ending up with 92 stocks. Daily returns and traded volumes are extracted from Bloomberg databases. Following Bialkowski et al. (2008) and Darolles and

Le Fol (2014), we retain the turnover ratio[4] as a measure of volume which controls for the dependency between the traded volume and the float. The float represents the difference between annual common shares outstanding and closely held shares for any given fiscal year. Common and closely held shares are also extracted from Bloomberg databases.

For each of the 92 stocks, we compute: (*i*) the empirical four first moments (mean, volatility, skewness and kurtosis) of volume and returns; (*ii*) the correlation between squared returns and volume; (*iii*) the first-order serial correlation of returns, volume and squared returns. The distributions of these statistics are summarized in Table 1 in the appendices. The first row reports the average, the dispersion, the minimum, and the maximum of the means of returns and volume across the 92 stocks. The second row gives the same cross-section statistics (average, dispersion, minimum and maximum) for the volatilities of returns and volume, and so on for the skewness, kurtosis, the correlation between squared returns and volume as well as the serial correlation coefficients. Finally, note that we perform a Pearson test to check the significance of the correlation coefficients computed in Table 1. The statistics reported in the three last rows are computed using the statistically significant coefficients only.

On the one hand, the results reported in Table 1 are consistent with the empirical implications of the standard mixture of distribution hypothesis. In fact, (*i*) volume is positive and present a positive skewness; (*ii*) the return and volume distributions are fat tails since thekurtosis statistics are greater than 3; (*iii*) the correlation between the squared returns and volume is always positive. Note that, the standard MDHL model of Darolles et al. (2015) explains this positive correlation by the effects of both information flow and short-term liquidity frictions, while in other competing MDH versions [such as that of Tauchen and Pitts (1983) or Andersen (1996)] this positive correlation is strictly due to the information flow, disregarding the presence of liquidity frictions.

On the other hand, the results on the serial correlation of returns, volumes and squared

[4] Let q_{kt} be the number of shares traded for asset k, $k = 1, ..., K$ on day t, $t = 1, ..., T$, and N_{kt} the float for asset k on day t. The individual stock turnover for asset k on day t is $V_{kt} = \frac{q_{kt}}{N_{kt}}$.

returns do not match the standard MDH predictions. The related descriptive statistics, reported in the last two rows of Table 1, indicate that: (*i*) the serial correlations of daily returns range from -0.0978 to 0.0866, exhibiting a dispersion of 0.0762 across stocks; (*ii*) for the daily traded volume, we get larger values lying in between 0.1781 and 0.7539 with a cross-sectional average of 0.4305 and dispersion of 0.1017; (*iii*) the first order serial correlation coefficients of squared returns evolve between 0.0619 and 0.2616, with a cross-sectional average of 0.1267 and a dispersion of 0.0556.

4.2 Characterizing stock liquidity profile

As discussed in section 3, we apply the Extended Kalman Filter methodology to our dynamic MDHL model given in (2.5)-(2.8) in order to filter the two latent variables I_t and L_t and estimate the related parameters, simultaneously.

We report in columns 2 to 4 of Tables 2 and 3 in the appendices, the estimated μ_p, β and μ_{v2} parameters for the 92 stocks of our sample.[5] These results show that at the 90% level of confidence almost 48%, i.e., 44 out of 92 of the stocks[6] present statistically significant μ_p, β and/or μ_{v2} parameters so as to show that they are facing liquidity frictions, as specified in Section 2. More precisely, and following the description of the liquidity profiles given on page 9, we have:

(*i*) Table 4 contains the 18 stocks (19.6% of our sample) belonging to the pure short-term liquidity friction case. Note that for these 18 stocks, μ_{v1} is significant[7]. The liquidity providers act strategically; they enter the market to provide liquidity and cash the liquidity premium supported by non-strategic (impatient) liquidity consumers.

[5]The remaining parameters of the price change and the volume equations, not reported in this article, are available upon request.

[6]At 95% this percentage falls to 41%, i.e. 38 stocks out of 92. Note that, all the results presented hereafter in the text corresponds to a level of confidence of 90% even if the results at 90% are displayed in all the tables of the appendices.

[7]Darolles et al. (2015) suggests that stocks presenting statistically significant μ_{v2} parameter are expected to present a statistically significant μ_{v1} parameter as well since the occurrence of short-term liquidity frictions is closely related to the arrival of information flow to the market modifying the way it is incorporated into prices. The estimations of the μ_{v1} parameters are not reported here but are available upon request.

(*ii*) Table 5 reports the estimated parameters for the 23 stocks (25% of our sample) of the pure long-term liquidity friction category, with $\mu_p \neq 0$, $\beta \neq 0$ and $\mu_{v2} = 0$. For these stocks, liquidity consumers act strategically by splitting their trades in order to hide from the liquidity providers. It may be because the strategic liquidity consumers are efficient in remaining out of the radar of the strategic liquidity providers that the latter are not trading. A by-side effect of their splitting is a slowdown of the propagation of information in the prices that can last for days and even weeks inducing a deterioration of the market (price) efficiency.

(*iii*) The 3 stocks (3% of our sample) that are of the mixed liquidity friction type are reported in Table 6. As for these stocks, liquidity providers as well as liquidity consumers act strategically without being fully efficient. In fact, liquidity consumers are splitting but their splitting scheme is detected by the strategic liquidity providers who lower the impact of short-term liquidity frictions on price changes. However, they are not able to detect, or to correct these short-term frictions completely, and the end-of-the-day prices are revealing only part of the information possessed by market participants.

(*iv*) There are 48 remaining stocks (52% of our sample) that do not belong to any of the above liquidity profile categories. The results relative to these stocks are displayed in Table 7 in the appendices. We find 6 stocks following a standard MDH model, and 34 stocks following Andersen's (1996) modified MDH model. The eight last stocks suffer from some frictions ($\mu_p \neq 0$) but not linked to liquidity problems ($\beta = 0$, and $\mu_{v2} = 0$).

4.3 Empirical serial correlation coefficients

We here confront the results obtained by our structural approach based on the dynamic MDHL framework to those resulting from the purely empirical *ad hoc* approach based on serial correlation coefficients of returns, squared returns and traded volumes. We show that both approaches do not always feature the same liquidity profile.

First, as we see in Tables 4 to 7 (columns 5 to 7), the results on the serial correlation of returns, volumes and squared returns do not match the standard MDH predictions. Recall that, serial correlations of returns should be zero for all the competing models but our, and the serial correlation of volumes and squared returns should be zero for Tauchen and Pitts (1983). At the 90% confidence level, the first order serial correlations of returns, volumes and squared returns are jointly statistically significant for 18 stocks; there are individually significant for 23, 91 and 70 over 92 stocks, respectively.[8]

Our modeling implies that, in the presence of long-lasting and short-term liquidity frictions, the daily return, squared return and volume time-series should simultaneously exhibit statistically significant (and positive) serial correlations. The empirical implications of our model are satisfied by 10 stocks for which the serial correlation coefficients of returns, volumes and squared returns are jointly statistically positive. Among these 10 stocks, two of them belong to long-term liquidity friction case (GKN LN and SL LN stocks), three others belong to the short-term liquidity friction case (ADM LN, RBS LN and RIO LN stocks), another one verifies the standard MDH of Tauchen and Pitts (IMI LN stock), while the remaining four of them verify the Andersen's (1996) case (LAND LN, RDSA LN, WEIR LN and IAG LN stocks). In addition, amongst the 51 stocks verifying Andersen's (1996)' model, only 23 of them appear to verify the model when we consider our model parameters μ_p, β and μ_{v2}. More importantly, 22 out of these 51 stocks belong to the three first liquidity profile categories pictured by our dynamic MDHL model (i.e., the pure short-term, the pure long-lasting and the mixed liquidity friction case, respectively).

Besides, the use of the empirical serial correlation coefficients to understand the dynamic patterns of the daily traded characteristics in the context of information and liquidity frictions presents two main limitations. First, these coefficients, as well as the related Pearson test, are computed under the homoscedasticity assumption. They may be biased in the presence of heteroscedasticity, and so is the test of ther statistical significance. Second, the serial correlation coefficients do not allow us to either separate the impact of information

[8]Note that, at the 95% confidence level, the first order serial correlations are jointly statistically significant for 12 stocks, and individually significant for 17, 91 and 64 over 92 stocks, respectively.

and liquidity frictions on the dynamics of the daily trading characteristics or to assess how the interaction of both types of liquidity frictions (short versus long-lasting) drives these dynamics. In particular, we cannot tell whether the statistically significant serial correlation coefficients of daily traded volume time-series are implied by the time-persistence of information process alone (as in Andersen (1996)) or by a combination of both information and liquidity shock processes (as assumed in our framework). In this context, our generalized MDHL framework improves the comprehension of daily time-series dynamics by putting enough structure on the data in order to separate the impacts of both information and liquidity shock time-persistent processes on daily trading characteristics.

4.4 Extracting the I_t and L_t time series

Our approach enables us to extract two time series of interest: the information process I_t and the liquidity friction process L_t. The information process is more volatile than the liquidity friction one and as expected, we see in the Figure 1 that whatever the I_t evolution, L_t is almost zero for no liquidity friction stocks, like AHT LN [Panel *a*)] and BATS LN [Panel *b*)]. L_t then gets bigger for the liquidity friction cases, ABF LN [Panel *c*)] and ADN LN [Panel *d*)].

Looking specifically at ADN LN, we get an interesting example of the possible use of the two time series. Zooming over September 2011 for that stock (see Figure 2), we can see that the two factors are reacting to some common shocks but do not behave the same. Aberdeen Asset Management PLC (ticker ADN LN) is a global investment management group which have offices around the world, including in New York. The shock on September 11, 2011 is huge and by far, the largest over the entire period. As we can see in the two graphs, the pikes in the information factor are more pronounced than in the liquidity factor. Moreover, if the information process come back to its "normal level" quite quickly, the decay in the liquidity friction process is much slower. The detailed analysis of I_t and L_t will definitely be key, and we leave it for further research. Another possible use of these processes, is to propose some strategies based on I_t or L_t. In the next session, we build an efficient (liquidity-based) strategy based on I_t.

5 Model comparison and robustness checks

To formally assess the validity of our approach, we first follow Andersen (1996) and implement Likelihood Ratio tests to compare the goodness of fit of our model to those of Tauchen and Pitts (1983) and Andersen (1996). We then build two strategies to capture the liquidity risk premium: one based on empirical correlations, and another one based on our estimated coefficients.

5.1 Goodness of fit

In order to compare our model to the two rival models, we estimate the two restricted MDH versions and test each of the corresponding null hypothesis, independently: (i) H_0^{TP}: $\{\mu_p = 0,\ \beta = 0,\ \mu_{v2} = 0,$ and $a = 0\}$ (Tauchen and Pitts case); (ii) H_0^A: $\{\mu_p = 0,\ \mu_{v2} = 0,$ and $a = 0\}$ (Andersen case). The likelihood ratio is then computed based on the difference in the (log) objective function of the constrained H_0^{TP} (or H_0^A) and unconstrained models. Under the null H_0^{TP} (or H_0^A), the test statistic is asymptotically drawn from a χ^2 distribution with 4 (or 3) degrees of freedom. The p-values associated to the likelihood ratio, denoted $p-val_{(TP)}$ for Tauchen and Pitts case and $p-val_{(A)}$ for Andersen case, are displayed in the two last columns of Tables 4 to 7.[9]

Four main remarks can be drawn. First, overall our results provide strong evidence of the weakness of the standard MDH version of Tauchen and Pitts (1983); at the 95% level of confidence, the null H_0^{TP} is rejected by 89 out of the 92 stocks of our sample (i.e., 97% of stocks). These results are coherent with those of subsections 4.2 and 4.3, and confirm the superiority of our dynamic MDHL specification relative to its static counterpart.

[9]The bolded p-values mean that the null (H_0^{TP} or H_0^A) is rejected at the 95% level of confidence.

Second, the Andersen MDH version is rejected by 52 stocks (i.e., 57% of our sample) at the 95% level of confidence, meaning that for these stocks our dynamic MDHL model fits better than Andersen's MDH specification with the statistical patterns of the data. More specifically, 20 of these stocks (i.e., almost 40% of them) belong to our second category, i.e. they present a pure long-term liquidity friction profile.[10]

Third, we cannot reject the null, H_0^A, for 40 stocks of our sample. It is important to note that these stocks are concentrated in our fourth category, the no liquidity friction case (i.e., 30 out of the 40 stocks or 75% of them) as reported in Table 7 in the appendices. Among these stocks, 23 (4) have parameters that verify Andersen's (Tauchen and Pitts's) model. The last 3 stocks seem to present other frictions that cannot be assimilated to liquidity problems according to our framework.

Finally, The restricted Andersen case is rejected for most of the stocks of our sample affected by short-term and/or long-term liquidity frictions.[11] This result highlights the importance of using our unrestricted approach especially for stocks affected by short-term and/or long-lasting liquidity frictions.

5.2 Liquidity premium strategies

Another way to assess the accuracy of our approach is to compare strategies investing in the most illiquid stocks to capture the liquidity premium: some of them built on the characteristics of our estimated model (μ_p and/or β for our application), and another one based on the empirical correlation of returns.

With this aim, we first build four competing hedged portfolios based on different stock-picking criteria estimated on data from January 1^{st}, 2010 to December 31, 2013. All portfolios are equally weighted. Then, we analyze the out of sample performance of these portfolios by assuming that they are held unchanged for the following two years, i.e., from January 2^{nd},

[10]For the stocks of the pure long-term category, H_0^A is rejected for 20 out of the 23 stocks.

[11]More precisely, the null H_0^A is rejected for all the stocks belonging to categories (i) to (iii) except for 7 stocks presenting a short-term liquidity profile and only 3 belonging to the pure long-term liquidity type.

2014 to December 31, 2015. More precisely, the four competing hedged portfolios considered here present long/short portfolios obtained as follows:

(*i*) The first one, called PF 1, consists in a long position on a portfolio including all stocks with a statistically significant and positive first-order serial correlation of returns, and a short position on the Index made of our 92 stocks.[12]

(*ii*) The second portfolio, PF 2, consists in a long position on a portfolio including all stocks with a statistically significant and positive μ_p parameter (i.e., 29 stocks of our sample), and a short position on the Index.

(*iii*) The third portfolio, PF 3, is long position on stocks whose μ_p and β parameters are jointly statistically significant and positive (i.e., 20 stocks), and a short position on the Index.

(*iv*) The fourth portfolio, PF 4, is long on stocks verifying Andersen (1996) framework (i.e., the 34 stocks presenting $\mu_p = 0$ and $\beta \neq 0$), and short on the Index.

For each portfolio, we compute the out of sample returns from January 2^{nd}, 2014 to December 31, 2015. In Table 8 in the appendices, we display for each portfolio, as well as for the equally-weighted Index, the average returns, standard deviations as well as the Sharpe ratios in annualized terms: Panel A of Table 8 for the overall two-year out-of sample period, and Panels B and C for each of the one-year periods.

As we can see in Table 8, for all considered out-of-sample periods, all the strategies based on our structural approach outperform the strategy based on the empirical serial correlation of daily returns. Moreover, PF 3 - based on μ_p and β- is quite stable (and by far the more stable) and always outperform both the empirical correlation strategy (PF1) as well as the Andersen-based strategy (PF 4).

Generally speaking, as shown in Figure 3 in the appendices, our results highlight the superiority of a structural approach as compared to the empirical *ad hoc* one based on empirical

[12]This portfolio contains 14 stocks. For eleven of them, the serial correlation coefficient of squared returns is also statistically positive. The other 3 stocks exhibit positive serial correlation of daily returns while the first-order serial correlation of squared returns is not statistically different from zero. We also dropped these 3 stocks from the portfolio 1 and obtain similar results. For this reason we here present only the results inherent to portfolio 1 comprising the 14 stocks of our sample presenting positive first-order serial correlation of returns.

serial correlation coefficients in inferring the presence of long-lasting liquidity frictions. In addition, the selection criteria based on μ_p and β jointly does a better job than that based on either μ_p or β parameter alone in identifying stocks to be included in long-term liquidity premium strategies.

5.3 Momentum strategies

In the toy example of the previous (static) strategy, we showed that strategies based on μ_p and β are outperforming strategies based on classical returns' serial correlation coefficients. In this section, we add to the strategies based on the μ_p and/or β parameters, the use of the extracted I_t process to finally end-up with a dynamic liquidity based strategy, or momentum strategy. We compare this strategy to the empirical one, combining the serial correlation of returns with realized pas returns. Here again, all the portfolios are equally weighted. To simplify, we do not account for the transaction costs related to the dynamic rebalancing of the portfolios. This assumption does not biased the comparison of the two strategies as they are both associated with the same turnover level. We have:

(i) The first portfolio (PF^M 1) consists of stocks presenting statistically significant serial correlation coefficient of daily returns over the period ranging from January 1^{st}, 2010 to December 31, 2013 (i.e., 13 stocks). Based on these stocks, we build a momentum portfolio following a simple trading rule: at the end of the trading day t, we take a long (short) position on a portfolio of stocks with positive (negative) returns for that day, and compute its performance at day $t+1$. This portfolio is rebalanced at the end of each trading day from January 2^{nd}, 2014 to December 31, 2015. Then for each trading day, we compute the daily return of our portfolio in excess to the index return to control for market risk and isolate the momentum effect.

(ii) The second portfolio (PF^M 2) consists of stocks presenting jointly statistically significant and positive μ_p and β parameters over our sample period, Jan. 2010 - Dec. 2013 (i.e., 20 stocks). For each of these stocks, we assume that the parameters of our model remain unchanged during the out of sample period (2014-2015) and use the Extended Kalman Filter methodology to filter the latent variable I_t given these parameters up to December 2015. We then build a momentum strategy by taking a long (short) position on an equally-weighted portfolio of stocks with I_{t-1} variable higher (lower) than a given threshold, say the moving average of a considered stock-specific I_t time-series computed during the previous year (i.e., the last 260 days). We rebalance this portfolio at the end of each trading day from Jan. 2014 to Dec. 2015 and again compute its daily returns in excess to the index returns.

The average returns, standard deviations as well as the Sharpe ratios in annualized terms for these two hedged portfolios are given in Table 9 in the appendices: Panel A of Table 9 for the overall two-year out-of sample period, and Panels B and C for each of the one-year periods. These results show that $\mathrm{PF}^M 2$ combining our μ_p and β parameters with the signal extracted from the stock-specific latent variable I_t outperforms the standard momentum strategy combining the empirical serial correlation coefficient with the past daily returns. This is true for the overall out of sample period as well as the two one-year sub-periods (see also the Figure 4 in the appendices).

To conclude, in subsections 5.2 and 5.3 we provide strong evidence of the superiority of our structural approach with respect to the empirical one based on serial correlation coefficients to capture the risk premium inherent to the presence of long-lasting liquidity frictions as well as to improve the efficiency of the momentum strategies.

6 Concluding remarks

In this article, we distinguish between two types of liquidity problems: short-term and long-lasting liquidity frictions. The former correspond to the liquidity provision and impact stock returns at the intra-day frequency while affecting the traded volume at the daily periodicity. The later are responsible for the dynamics of daily return and volume time-series because information is not completely incorporated into prices within the day. These frictions come from the strategic behavior of liquidity consumers, i.e. the long-term investors. We propose a statistical model that specifies the impact of information arrival on market characteristics in the context of liquidity frictions. It accounts for both the long lasting liquidity frictions and the time-persistence of short-term liquidity frictions whose presence can be inferred from the daily return dynamics. Our modelling allows us to disentangle both types of liquidity frictions on stock return and volume dynamics.

In our model, a bivariate mixture of distributions conditioned by two latent time-persistent variables - information and liquidity-, represents the daily price change and the volume processes. We use an Extended Kalman Filter (EKF) to filter the two latent variables and estimate the model parameters, simultaneously.

Our results show that, over the 48% (44/92) of stocks from the FTSE100 that are actually facing liquidity problems, the liquidity providers are the only investors acting strategically on a group of stocks accounting for nearly 20% (18/92). For another 25% (23/92) of the stocks, the long-term investors are strategic while the liquidity providers are inactive. Finally, for the last 3% (3/92) group of stocks, the liquidity providers as well as the liquidity consumers are both acting strategically. It means that 28% (26/92) of the stocks of our sample are facing a slow-down of the information propagation to prices and thus a decrease of (daily) price efficiency due to the strategic behavior of the long-term investors. Based on some goodness of fit tests, as well as on some simple strategies based on our structural model, we show the superiority of our approach against other MDH specifications.

Moreover, based on some goodness of fit tests, as well as on some simple strategies based on our structural model, we show the superiority of our approach.

References

Aitken, M. and Comerton-Forde, C. (2003). How should liquidity be measured? *Pacific-Basin Finance Journal*, 11(1):45–59.

Almgren, R. and Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40.

Almgren, R., Thum, C., Hauptmann, E., and Li, H. (2005). Direct estimation of equity market impact. *Risk*, 18(7):58–62.

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

Anderson, B. and Moore, J. (1979). Optimal Filtering. *Prentice-Hall*.

Anderson, R. M., Eom, K. S., Hahn, S. B., and Park, J. (2013). Autocorrelation and partial price adjustment. *Journal of Empirical Finance*, 24(C):78–93.

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

Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., and Vega, C. (2014). Rise of the Machines: Algorithmic Trading in the Foreign Exchange Market. *The Journal of Finance*, 69(5):2045–2084.

Criscuolo, A. and Waelbroeck, H. (2013). Optimal execution in presence of short-term trading. *Available at SSRN: http://ssrn.com/abstract=2227139*.

Darolles, S. and Le Fol, G. (2014). Trading Volume and Arbitrage. *GSTF Journal of Business Review*, 3:30–39.

Darolles, S., Le Fol, G., and Mero, G. (2015). Measuring the Liquidity Part of Volume. *Journal of Banking and Finance*, 50:92–105.

Engle, R., Lilien, D., and Robins, R. (1987). Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model. *Econometrica*, 55:391–407.

Foucault, T., Kozhan, R., and Tham, W. W. (2014). Toxic Arbitrage. *Working Paper SSRN*.

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

Goyenko, R., Holden, C. W., and Trzcinka, C. A. (2009). Do Liquidity Measures Measure Liqudity? *Journal of Financial Economics*, 92(2):153–181.

Hendershott, T., Jones, C. M., and Menkveld, A. J. (2011). Does Algorithmic Trading Improve Liquidity? *The Journal of Finance*, 66(1):1–33.

Jarrow, R. A. and Protter, P. (2013). Liquidity Suppliers and High Frequency Trading. *Working paper SSRN*.

Kirilenko, A. A., Kyle, A. S., Samadi, M., and Tuzun, T. (2014). The Flash Crash: The Impact of High Frequency Trading on an Electronic Market. *Working Paper SSRN*.

Koopman, S. J. and Upensky, E. H. (2002). The Stochastic Volatility in Mean Model: Empirical Evidence from International Stock Markets. *Journal of Applied Econometrics*, 17:667–689.

Lehalle, C. A. (2014). Towards Dynamic Measures of Liquidity. *AMF Scientific Advisory Board Review*, 1:55–62.

Lehalle, C. A. and Laruelle, S. (2013). Market Microstructure in Practice. *World Scientific Publishing Company*.

Menkveld, A. J. (2013). High Frequency Trading and the New-Market Makers. *Journal of Financial Markets*, 16:712–740.

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

Säfvenblad, P. (2000). Trading Volume and Autocorrelation: Empirical Evidence from the Stockholm Stock Exchange. *Journal of Banking and Finance*, 24(8):1275–1287.

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

Toth, B., Palit, I., Lillo, F., and Farmer, J. D. (2014). Why Is Equity Order Flow so Persistent? *Working paper SSRN.*

Waelbroeck, H. and Gomes, C. (2013). Is Market Impact a Measure of the Information Value of Trades? Market Response to Liquidity vs. Informed Trades. *Working paper SSRN.*

Appendices

	Returns				Volume			
	Average	Dispersion	Min	Max	Average	Dispersion	Min	Max
Mean	0.0006	0.0005	-0.0006	0.0023	0.0033	0.0011	0.0009	0.0067
Volatility	0.0167	0.0046	0.0103	0.0285	0.0023	0.0013	0.0004	0.0080
Skewness	-0.1659	0.7764	-4,3833	0.7056	5,3963	4,8583	1,1732	27,8285
Kurtosis	8,4371	10.2007	3,6762	67,3546	82,5710	136,3502	5,3676	845,0944
$(\text{Return})^2$ with Volume Correlation	0.4075	0.1327	0.1175	0.7358	-	-	-	-
1^{st} Order Serial Correlation	0.0200	0.0762	-0.0978	0.0866	0.4305	0.1017	0.1781	0.7539
1^{st} Order Serial Correlation $(\text{Return})^2$	0.1267	0.0556	0.0619	0.2616	-	-	-	-

Table 1: Summary statistics for return and turnover time series from January 1^{st}, 2010 to December 31, 2013.

Bloomberg Code	μ_p	β	μ_{v2}	$\rho_{(R_t,R_{t-1})}$	$\rho_{(R_t^2,R_{t-1}^2)}$	$\rho_{(V_t,V_{t-1})}$
AAL LN	**0,0001**	0,9000	0,0000	-0,0017	-0,0029	**0,4938**
ABF LN	**0,0010**	**0,2000**	**0,0010**	**-0,0732**	**0,1123**	**0,4704**
ADM LN	-0,0005	0,2000	**0,0004**	***0,0570***	**0,0711**	**0,4947**
ADN LN	**0,0014**	0,4000	0,0041	0,0173	0,0360	**0,0574**
AGK LN	0,0003	**0,9002**	0,0001	**0,0700**	0,0084	**0,3328**
AHT LN	**0,0024**	0,7000	-0,0001	0,0340	**0,0931**	**0,3578**
AMEC LN	-0,0002	**0,8997**	-0,0014	0,0047	**0,1846**	**0,3934**
ANTO LN	-0,0009	**0,8997**	0,0038	-0,0261	**0,0859**	**0,5216**
ARM LN	**0,0019**	**0,9000**	0,0010	0,0026	**0,0764**	**0,4881**
AV/ LN	-0,0004	**0,4000**	**0,0014**	0,0059	**0,1141**	**0,4540**
AZN LN	0,0001	**0,9000**	-0,0002	-0,0131	***0,0531***	**0,4287**
BA/ LN	0,0006	**0,9000**	-0,0002	-0,0506	**0,2616**	**0,5417**
BAB LN	0,0002	**0,9001**	0,0001	-0,0239	**0,2505**	**0,3992**
BARC LN	0,0000	0,9000	0,0053	0,0284	**0,0884**	**0,4019**
BATS LN	0,0002	**0,2012**	0,0004	**-0,0653**	**0,0619**	**0,3880**
BG/ LN	0,0001	0,2000	0,0006	0,0281	**0,0906**	**0,3697**
BLND LN	0,0000	**0,4000**	0,0019	0,0062	**0,2393**	**0,4367**
BLT LN	-0,0003	**0,9000**	0,0005	0,0368	**0,0778**	**0,5039**
BNZL LN	**0,0009**	**0,7000**	0,0011	0,0180	**0,0764**	**0,2839**
BP/ LN	-0,0002	0,9000	0,0009	0,0506	**0,1994**	**0,7118**
BRBY LN	-0,0006	**0,8998**	**0,0004**	**0,0761**	0,0010	**0,4769**
BSY LN	0,0005	**0,9002**	**0,0055**	0,0208	0,0357	**0,4094**
BT/A LN	**0,0011**	0,9000	0,0000	-0,0384	**0,0701**	**0,5520**
CCL LN	***-0,0009***	0,4000	0,0015	**0,0816**	0,0212	**0,4280**
CNA LN	0,0003	0,9033	**0,0005**	***-0,0585***	**0,0850**	**0,4106**
CPG LN	**0,0037**	0,2000	0,0020	**-0,0975**	**0,1221**	**0,4504**
CPI LN	-0,0001	**0,9000**	0,0002	-0,0197	**0,1196**	**0,4331**
CRH LN	-0,0004	**0,5001**	0,0017	0,0174	0,0458	**0,2658**
DGE LN	***0,0006***	0,5023	0,0213	**-0,0741**	0,0252	**0,3285**
EXPN LN	0,0003	**0,5000**	0,0008	-0,0143	0,0182	**0,3832**
EZJ LN	**0,0020**	**0,6999**	0,0012	0,0151	-0,0017	**0,3537**
GFS LN	**-0,0020**	**0,6999**	0,0013	***-0,0525***	**0,1219**	**0,4807**
GKN LN	**0,0017**	**0,5093**	0,0000	**0,0736**	**0,0878**	**0,3531**
GSK LN	0,0003	**0,5000**	**0,0009**	**-0,0674**	**0,1023**	**0,3624**
HMSO LN	0,0002	0,2000	0,0064	**0,0624**	**0,1762**	0,0472
HSBA LN	0,0001	0,9000	**0,0003**	-0,0409	***0,0599***	**0,5154**
IAG LN	0,0004	**0,9003**	-0,0001	**0,0748**	**0,0655**	**0,5890**
IHG LN	0,0005	0,9000	***0,0013***	0,0148	**0,1662**	**0,4107**
IMI LN	0,0008	0,7000	0,0016	**0,0721**	**0,2552**	**0,2672**
IMT LN	0,0003	**0,2000**	0,0011	0,0004	***0,0573***	**0,4044**
ITRK LN	***0,0011***	**0,9000**	0,0010	-0,0123	**0,0683**	**0,1781**
ITV LN	**0,0016**	**0,6997**	0,0083	-0,0053	**0,0626**	**0,4077**
JMAT LN	***0,0008***	**0,2000**	0,0014	-0,0105	**0,1585**	**0,4463**
KGF LN	0,0002	**0,9000**	-0,0009	-0,0010	***0,0600***	**0,5231**
LAND LN	0,0003	**0,9000**	0,0003	***0,0536***	**0,1735**	**0,4577**
LGEN LN	**0,0011**	**0,9000**	-0,0005	0,0430	**0,2228**	**0,5751**

Table 2: Estimated μ_p, β and μ_{v2} parameters (Columns 2 - 4) for the first 48 FTSE 100 stocks. The bolded (bolded and italic) parameters are statistically significant at the 95% (90%) level of confidence. Estimated first order serial correlation coefficients (Columns 5 - 7) of returns, squared returns and volume time series, respectively. The bolded (bolded and italic) coefficients are statistically significant at the 95% (90%) level of confidence, following a Pearson correlation test.

Bloomberg Code	μ_p	β	μ_{v2}	$\rho_{(R_t,R_{t-1})}$	$\rho_{(R_t^2,R_{t-1}^2)}$	$\rho_{(V_t,V_{t-1})}$
LLOY LN	0,0004	0,9000	**0,0010**	0,0500	**0,1918**	**0,5176**
LSE LN	**0,0017**	**0,5002**	0,0033	0,0472	0,0297	**0,3402**
MGGT LN	0,0007	**0,9000**	**0,0015**	0,0221	**0,1065**	**0,3437**
MKS LN	0,0003	**0,8999**	0,0024	-0,0129	0,0050	**0,4990**
MRO LN	**0,0012**	**0,9555**	-0,0010	0,0376	**0,1170**	**0,3198**
MRW LN	-0,0003	**0,2000**	**0,0016**	0,0100	**0,0765**	**0,4245**
NG/ LN	-0,0003	**0,9000**	0,0011	0,0175	**0,0713**	**0,7539**
NXT LN	**0,0010**	**0,9144**	0,0001	-0,0008	0,0268	**0,5564**
OML LN	0,0002	**0,8998**	0,0001	-0,0175	**0,0938**	**0,4415**
PFC LN	-0,0004	**0,3999**	0,0019	-0,0268	0,0460	**0,4571**
PRU LN	0,0008	**0,9000**	0,0013	0,0163	**0,1786**	**0,7077**
PSN LN	0,0005	**0,7000**	0,0019	0,0477	**0,1163**	**0,4150**
PSON LN	0,0003	**0,4000**	0,0012	-0,0269	**0,0834**	**0,3799**
RB/ LN	0,0005	**0,2001**	**0,0027**	**-0,0978**	**0,1578**	**0,4010**
RBS LN	0,0009	**0,2000**	**0,0020**	**0,0866**	**0,1841**	**0,4834**
RDSA LN	0,0003	**0,9000**	0,0004	***0,0575***	**0,1424**	**0,2327**
RDSB LN	0,0004	**0,2000**	**0,0005**	0,0418	**0,1210**	**0,3676**
REL LN	0,0004	**0,9000**	0,0007	0,0377	***0,0537***	**0,5091**
REX LN	0,0001	**0,5022**	-0,0001	0,0487	0,0481	**0,4281**
RIO LN	0,0000	**0,9000**	**0,0005**	**0,0645**	**0,1253**	**0,4674**
RR/ LN	**0,0013**	**0,5000**	-0,0011	-0,0426	0,0379	**0,5008**
RRS LN	-0,0009	**0,8999**	-0,0001	-0,0402	**0,0728**	**0,5032**
RSA LN	**-0,0034**	**0,9001**	0,0020	-0,0174	0,0257	**0,3845**
SAB LN	0,0006	**0,9000**	0,0013	-0,0160	**0,1107**	**0,3353**
SBRY LN	0,0002	**0,7000**	0,0013	0,0074	**0,1383**	**0,3785**
SDR LN	0,0001	**0,8999**	**0,0009**	0,0278	***0,0528***	**0,3640**
SGE LN	**0,0002**	0,9000	**0,0013**	***-0,0571***	**0,1158**	**0,4947**
SHP LN	**0,0009**	**0,9000**	0,0027	-0,0334	0,0462	**0,3848**
SL/ LN	***0,0006***	**0,5001**	0,0001	**0,0747**	**0,1071**	**0,3841**
SMIN LN	0,0004	**0,4000**	0,0007	0,0327	**0,0682**	**0,2423**
SN/ LN	0,0004	**0,9000**	-0,0004	-0,0451	**0,2264**	**0,5926**
SSE LN	0,0002	0,2000	**0,0014**	-0,0341	0,0476	**0,3917**
STAN LN	**-0,0036**	**0,7000**	0,0059	-0,0266	**0,2485**	**0,3983**
SVT LN	**0,0012**	0,5026	-0,0012	0,0002	-0,0019	**0,6622**
TATE LN	**0,0010**	0,9035	0,0000	-0,0050	0,0444	**0,2465**
TLW LN	**-0,0015**	**0,9000**	0,0007	0,0021	**0,1120**	**0,4326**
TPK LN	0,0006	0,7000	0,0026	0,0143	**0,1357**	**0,3763**
TSCO LN	**-0,0015**	**0,2000**	0,0011	0,0461	0,0143	**0,3784**
ULVR LN	0,0001	**0,8999**	0,0000	-0,0350	**0,0638**	**0,3757**
UU/ LN	**0,0011**	**0,4000**	**0,0018**	0,0129	**0,0632**	**0,4550**
VOD LN	**0,0008**	**0,2000**	**0,0015**	-0,0480	**0,0888**	**0,3766**
WEIR LN	0,0007	**0,7001**	0,0024	**0,0793**	**0,1758**	**0,4176**
WMH LN	**0,0016**	**0,6999**	0,0000	0,0193	**0,1067**	**0,2822**
WOS LN	***0,0010***	**0,9003**	-0,0010	0,0275	**0,1083**	**0,4553**
WPP LN	0,0002	**0,9000**	-0,0003	-0,0377	**0,0954**	**0,5075**
WTB LN	**0,0010**	**0,4000**	0,0014	0,0463	**0,1657**	**0,4426**

Table 3: Estimated μ_p, β and μ_{v2} parameters (Columns 2 - 4) for the first 48 FTSE 100 stocks. The bolded (bolded and italic) parameters are statistically significant at the 95% (90%) level of confidence. Estimated first order serial correlation coefficients (Columns 5 - 7) of returns, squared returns and volume time series, respectively. The bolded (bolded and italic) coefficients are statistically significant at the 95% (90%) level of confidence, following a Pearson correlation test.

Bloomberg Code	μ_p	β	μ_{v2}	$\rho_{R_t,R_{t-1}}$	$\rho_{R_t^2,R_{t-1}^2}$	$\rho_{V_t,V_{t-1}}$	$p-val_{(TP)}$	$p-val_{(A)}$
ADM LN	-0,0005	0,2000	**0,0004**	***0,0570***	**0,0711**	**0,4947**	**0,0000**	**0,0000**
AV/ LN	-0,0004	**0,4000**	***0,0014***	0,0059	**0,1141**	**0,4540**	**0,0000**	**0,0000**
BRBY LN	-0,0006	**0,8998**	**0,0004**	**0,0761**	0,0010	**0,4769**	**0,0000**	0,6280
BSY LN	0,0005	**0,9002**	**0,0055**	0,0208	0,0357	**0,4094**	**0,0000**	**0,0000**
CNA LN	0,0003	0,9033	**0,0005**	***-0,0585***	**0,0850**	**0,4106**	**0,0000**	0,9998
GSK LN	0,0003	**0,5000**	**0,0009**	**-0,0674**	**0,1023**	**0,3624**	**0,0001**	**0,0002**
HSBA LN	0,0001	0,9000	**0,0003**	-0,0409	***0,0599***	**0,5154**	**0,0000**	0,9998
IHG LN	0,0005	0,9000	***0,0013***	0,0148	**0,1662**	**0,4107**	**0,0000**	**0,0337**
LLOY LN	0,0004	0,9000	**0,0010**	0,0500	**0,1918**	**0,5176**	**0,0000**	0,1393
MGGT LN	0,0007	**0,9000**	**0,0015**	0,0221	**0,1065**	**0,3437**	**0,0000**	0,9997
MRW LN	-0,0003	**0,2000**	**0,0016**	0,0100	**0,0765**	**0,4245**	**0,0000**	**0,0000**
RB/ LN	0,0005	**0,2001**	**0,0027**	**-0,0978**	**0,1578**	**0,4010**	**0,0000**	**0,0000**
RBS LN	0,0009	**0,2000**	**0,0020**	**0,0866**	**0,1841**	**0,4834**	**0,0000**	**0,0000**
RDSB LN	0,0004	**0,2000**	**0,0005**	0,0418	**0,1210**	**0,3676**	**0,0000**	**0,0000**
RIO LN	0,0000	**0,9000**	**0,0005**	**0,0645**	**0,1253**	**0,4674**	**0,0000**	**0,0021**
SDR LN	0,0001	**0,8999**	**0,0009**	0,0278	***0,0528***	**0,3640**	**0,0000**	0,9308
SGE LN	**0,0002**	0,9000	**0,0013**	***-0,0571***	**0,1158**	**0,4947**	0,9999	0,9999
SSE LN	0,0002	0,2000	**0,0014**	-0,0341	0,0476	**0,3917**	**0,0000**	**0,0000**

Table 4: 1^{st} stock category: the pure short-term liquidity friction case: $\mu_{v2} \neq 0$ and either $\mu_p = 0$ or, if $\mu_p \neq 0$ then $\beta = 0$.

Bloomberg Code	μ_p	β	μ_{v2}	$\rho_{R_t,R_{t-1}}$	$\rho_{R_t^2,R_{t-1}^2}$	$\rho_{V_t,V_{t-1}}$	$p-val_{(TP)}$	$p-val_{(A)}$
ARM LN	**0,0019**	**0,9000**	0,0010	0,0026	**0,0764**	**0,4881**	**0,0000**	**0,0050**
BNZL LN	**0,0009**	**0,7000**	0,0011	0,0180	**0,0764**	**0,2839**	**0,0000**	0,1379
CCL LN	***-0,0009***	**0,4000**	0,0015	**0,0816**	0,0212	**0,4280**	**0,0000**	**0,0000**
EZJ LN	**0,0020**	**0,6999**	0,0012	0,0151	-0,0017	**0,3537**	**0,0000**	**0,0010**
GFS LN	**-0,0020**	**0,6999**	0,0013	***-0,0525***	**0,1219**	**0,4807**	**0,0000**	**0,0000**
GKN LN	**0,0017**	**0,5093**	0,0000	**0,0736**	**0,0878**	**0,3531**	**0,0000**	**0,0334**
ITRK LN	***0,0011***	**0,9000**	0,0010	-0,0123	**0,0683**	**0,1781**	**0,0000**	**0,0000**
ITV LN	**0,0016**	**0,6997**	0,0083	-0,0053	**0,0626**	**0,4077**	**0,0000**	**0,0004**
JMAT LN	***0,0008***	**0,2000**	0,0014	-0,0105	**0,1585**	**0,4463**	**0,0000**	**0,0000**
LGEN LN	**0,0011**	**0,9000**	-0,0005	0,0430	**0,2228**	**0,5751**	**0,0000**	**0,0378**
LSE LN	**0,0017**	**0,5002**	0,0033	0,0472	0,0297	**0,3402**	**0,0000**	**0,0004**
MRO LN	**0,0012**	**0,9555**	-0,0010	0,0376	**0,1170**	**0,3198**	**0,0000**	**0,0000**
NXT LN	**0,0010**	**0,9144**	0,0001	-0,0008	0,0268	**0,5564**	**0,0000**	**0,0000**
RR/ LN	**0,0013**	**0,5000**	-0,0011	-0,0426	0,0379	**0,5008**	**0,0000**	**0,0003**
RSA LN	**-0,0034**	**0,9001**	0,0020	-0,0174	0,0257	**0,3845**	**0,0000**	**0,0000**
SHP LN	**0,0009**	**0,9000**	0,0027	-0,0334	0,0462	**0,3848**	**0,0000**	**0,0000**
SL/ LN	***0,0006***	**0,5001**	0,0001	**0,0747**	**0,1071**	**0,3841**	**0,0000**	0,9935
STAN LN	**-0,0036**	**0,7000**	0,0059	-0,0266	**0,2485**	**0,3983**	**0,0000**	**0,0000**
TLW LN	**-0,0015**	**0,9000**	0,0007	0,0021	**0,1120**	**0,4326**	**0,0000**	**0,0048**
TSCO LN	**-0,0015**	**0,2000**	0,0011	0,0461	0,0143	**0,3784**	**0,0000**	**0,0000**
WMH LN	**0,0016**	**0,6999**	0,0000	0,0193	**0,1067**	**0,2822**	**0,0000**	**0,0056**
WOS LN	***0,0010***	**0,9003**	-0,0010	0,0275	**0,1083**	**0,4553**	**0,0000**	0,0650
WTB LN	**0,0010**	**0,4000**	0,0014	0,0463	**0,1657**	**0,4426**	**0,0000**	**0,0000**

Table 5: 2^{nd} stock category: the pure long-term liquidity friction case: $\mu_{v2} = 0$ and $\mu_p \neq 0$ and $\beta \neq 0$

Bloomberg Code	μ_p	β	μ_{v2}	$\rho_{R_t,R_{t-1}}$	$\rho_{R_t^2,R_{t-1}^2}$	$\rho_{V_t,V_{t-1}}$	$p-val_{(TP)}$	$p-val_{(A)}$
ABF LN	**0,0010**	**0,2000**	**0,0010**	**-0,0732**	**0,1123**	**0,4704**	**0,0000**	**0,0000**
UU/ LN	**0,0011**	**0,4000**	**0,0018**	0,0129	**0,0632**	**0,4550**	**0,0000**	**0,0000**
VOD LN	**0,0008**	**0,2000**	**0,0015**	-0,0480	**0,0888**	**0,3766**	**0,0000**	**0,0000**

Table 6: 3^{rd} stock category: the mixed liquidity friction case: $\mu_{v2} \neq 0$ and $\mu_p \neq 0$ and $\beta \neq 0$

Bloomberg Code	μ_p	β	μ_{v2}	$\rho_{R_t,R_{t-1}}$	$\rho_{R_t^2,R_{t-1}^2}$	$\rho_{V_t,V_{t-1}}$	$p-val_{(TP)}$	$p-val_{(A)}$
AAL LN	**0,0001**	0,9000	0,0000	-0,0017	-0,0029	**0,4938**	0,9999	0,9999
ADN LN	**0,0014**	0,4000	0,0041	0,0173	0,0360	**0,0574**	**0,0001**	**0,0000**
AGK LN	0,0003	**0,9002**	0,0001	**0,0700**	0,0084	**0,3328**	**0,0000**	**0,0000**
AHT LN	**0,0024**	0,7000	-0,0001	0,0340	**0,0931**	**0,3578**	**0,0000**	**0,0140**
AMEC LN	-0,0002	**0,8997**	-0,0014	0,0047	**0,1846**	**0,3934**	**0,0000**	0,3037
ANTO LN	-0,0009	**0,8997**	0,0038	-0,0261	**0,0859**	**0,5216**	**0,0000**	**0,0233**
AZN LN	0,0001	**0,9000**	-0,0002	-0,0131	***0,0531***	**0,4287**	**0,0000**	0,2520
BA/ LN	0,0006	**0,9000**	-0,0002	-0,0506	**0,2616**	**0,5417**	**0,0000**	0,4939
BAB LN	0,0002	**0,9001**	0,0001	-0,0239	**0,2505**	**0,3992**	**0,0000**	0,7973
BARC LN	0,0000	0,9000	0,0053	0,0284	**0,0884**	**0,4019**	**0,0000**	**0,0112**
BATS LN	0,0002	**0,2012**	0,0004	**-0,0653**	**0,0619**	**0,3880**	**0,0000**	**0,0000**
BG/ LN	0,0001	0,2000	0,0006	0,0281	**0,0906**	**0,3697**	**0,0000**	**0,0000**
BLND LN	0,0000	**0,4000**	0,0019	0,0062	**0,2393**	**0,4367**	**0,0000**	**0,0000**
BLT LN	-0,0003	**0,9000**	0,0005	0,0368	**0,0778**	**0,5039**	**0,0000**	0,4874
BP/ LN	-0,0002	0,9000	0,0009	0,0506	**0,1994**	**0,7118**	**0,0000**	0,9967
BT/A LN	**0,0011**	0,9000	0,0000	-0,0384	**0,0701**	**0,5520**	**0,0000**	0,6810
CPG LN	**0,0037**	0,2000	0,0020	**-0,0975**	**0,1221**	**0,4504**	**0,0000**	**0,0000**
CPI LN	-0,0001	**0,9000**	0,0002	-0,0197	**0,1196**	**0,4331**	**0,0000**	0,5058
CRH LN	-0,0004	**0,5001**	0,0017	0,0174	0,0458	**0,2658**	**0,0000**	0,9828
DGE LN	***0,0006***	0,5023	0,0213	**-0,0741**	0,0252	**0,3285**	**0,0000**	**0,0000**
EXPN LN	0,0003	**0,5000**	0,0008	-0,0143	0,0182	**0,3832**	**0,0000**	0,9997
HMSO LN	0,0002	0,2000	0,0064	**0,0624**	**0,1762**	0,0472	0,9998	0,9998
IAG LN	0,0004	**0,9003**	-0,0001	**0,0748**	**0,0655**	**0,5890**	**0,0000**	0,9451
IMI LN	0,0008	0,7000	0,0016	**0,0721**	**0,2552**	**0,2672**	**0,0000**	0,6148
IMT LN	0,0003	**0,2000**	0,0011	0,0004	***0,0573***	**0,4044**	**0,0000**	**0,0000**
KGF LN	0,0002	**0,9000**	-0,0009	-0,0010	***0,0600***	**0,5231**	**0,0000**	0,9631
LAND LN	0,0003	**0,9000**	0,0003	***0,0536***	**0,1735**	**0,4577**	**0,0000**	0,2712
MKS LN	0,0003	**0,8999**	0,0024	-0,0129	0,0050	**0,4990**	**0,0000**	0,6283
NG/ LN	-0,0003	**0,9000**	0,0011	0,0175	**0,0713**	**0,7539**	**0,0000**	0,9997
OML LN	0,0002	**0,8998**	0,0001	-0,0175	**0,0938**	**0,4415**	**0,0000**	0,5962
PFC LN	-0,0004	**0,3999**	0,0019	-0,0268	0,0460	**0,4571**	**0,0000**	**0,0000**
PRU LN	0,0008	**0,9000**	0,0013	0,0163	**0,1786**	**0,7077**	**0,0000**	0,9997
PSN LN	0,0005	**0,7000**	0,0019	0,0477	**0,1163**	**0,4150**	**0,0000**	**0,0000**
PSON LN	0,0003	**0,4000**	0,0012	-0,0269	**0,0834**	**0,3799**	**0,0000**	**0,0000**
RDSA LN	0,0003	**0,9000**	0,0004	***0,0575***	**0,1424**	**0,2327**	**0,0000**	0,9998
REL LN	0,0004	**0,9000**	0,0007	0,0377	***0,0537***	**0,5091**	**0,0000**	0,0711
REX LN	0,0001	**0,5022**	-0,0001	0,0487	0,0481	**0,4281**	**0,0000**	0,0842
RRS LN	-0,0009	**0,8999**	-0,0001	-0,0402	**0,0728**	**0,5032**	**0,0000**	0,1776
SAB LN	0,0006	**0,9000**	0,0013	-0,0160	**0,1107**	**0,3353**	**0,0000**	0,9391
SBRY LN	0,0002	**0,7000**	0,0013	0,0074	**0,1383**	**0,3785**	**0,0000**	**0,0000**
SMIN LN	0,0004	**0,4000**	0,0007	0,0327	**0,0682**	**0,2423**	**0,0000**	**0,0000**
SN/ LN	0,0004	**0,9000**	-0,0004	-0,0451	**0,2264**	**0,5926**	**0,0000**	0,4725
SVT LN	**0,0012**	0,5026	-0,0012	0,0002	-0,0019	**0,6622**	**0,0000**	0,9999
TATE LN	**0,0010**	0,9035	0,0000	-0,0050	0,0444	**0,2465**	**0,0000**	**0,0000**
TPK LN	0,0006	0,7000	0,0026	0,0143	**0,1357**	**0,3763**	**0,0000**	0,1415
ULVR LN	0,0001	**0,8999**	0,0000	-0,0350	**0,0638**	**0,3757**	**0,0000**	**0,0000**
WEIR LN	0,0007	**0,7001**	0,0024	**0,0793**	**0,1758**	**0,4176**	**0,0000**	0,2916
WPP LN	0,0002	**0,9000**	-0,0003	-0,0377	**0,0954**	**0,5075**	**0,0000**	0,2099

Table 7: Last stock category: No liquidity friction case (including Tauchen and Pitts (1983) and Andersen (1996) cases)

Panel A: Jan. 2014 - Dec. 2015	**PF 1**	**PF 2**	**PF 3**	**PF 4**	**EW Index**
Annualized Average Return	-0,0369	0,0428	0,0685	0,0219	-0,0424
Annualized Standard Deviation	0,0511	0,041	0,056	0,0267	0,1443
Annualized Sharpe Ratio	-0,7221	1,0439	1,2232	0,8202	-0,2938
Panel B: Jan. 2014 - Dec. 2014	**PF 1**	**PF 2**	**PF 3**	**PF 4**	**EW Index**
Annualized Average Return	0,0296	0,0567	0,0639	0,0126	-0,0149
Annualized Standard Deviation	0,0468	0,0398	0,0523	0,0263	0,1184
Annualized Sharpe Ratio	0,6325	1,4246	1,2218	0,4791	-0,1258
Panel C: Jan. 2015 - Dec. 2015	**PF 1**	**PF 2**	**PF 3**	**PF 4**	**EW Index**
Annualized Average Return	-0,1035	0,0288	0,073	0,0312	-0,0700
Annualized Standard Deviation	0,0548	0,0422	0,0596	0,0272	0,1665
Annualized Sharpe Ratio	-1,8887	0,6825	1,2248	1,1471	-0,4204

Table 8: Out of Sample performance statistics in annualized terms. Portfolios 1 to 4 (called PF 1, ..., PF 4, respectively) consist in a long position on a equally-weighted portfolio relying on four competing stock selecting criteria, and a short position on the equally-weighted portfolio including all the stocks of our sample (called EW Index). The stock selecting criteria are: the first-order serial correlation of returns (which should be positive) for portfolio 1; the μ_p parameter estimated by our structural approach (which should be positive) regardless of β parameter value for portfolio 2; the μ_p and β parameters obtained by our approach considered jointly (they both should be positive) for portfolio 3; the β parameter should be positive while $\mu_p = 0$ for portfolio 4 (Andersen (1996) case). All portfolios are formed based on the parameters estimated based on our overall test period extending from January 1^{st}, 2010 to December 31, 2013 and then are held unchanged for two consecutive years after their formation date. Panel A displays the performance persistence statistics for the two-year holding period $2014 - 2015$, while Panels B and C present the same performance statistics for each of the of the two one-year subperiods.

	Panel A		Panel B		Panel C	
	Jan. 2014 - Dec. 2015		**Jan. - Dec. 2014**		**Jan. - Dec. 2015**	
	PF^M 1	**PF^M 2**	**PF^M 1**	**PF^M 2**	**PF^M 1**	**PF^M 2**
Annualized Average Return	0,1245	0,3377	0,0626	0,1527	0,1864	0,5227
Annualized Standard Deviation	0,2281	0,3257	0,1857	0,2924	0,264	0,356
Annualized Sharpe Ratio	0,5458	1,0368	0,3371	0,5222	0,7061	1,4683

Table 9: Annualized out-of-sample performance statistics. The first portfolio (PF^M 1) consists of stocks presenting statistically significant serial correlation coefficient of daily returns estimated over January 1^{st}, 2010 to December 31, 2013 (i.e., 13 stocks). For these stocks, we build a momentum strategy based on the sign of the returns. The second portfolio (PF^M 2) consists of stocks presenting jointly statistically significant and positive μ_p and β parameters estimated over Jan. 2010 - Dec. 2013 (i.e., 20 stocks). For each of these stocks, we use the Extended Kalman Filter methodology to filter the latent variable I_t given these parameters from January 2010 to December 2015. We then build a momentum strategy by taking a long (short) position on a portfolio of stocks with I_{t-1} greater (lower) than the moving average of the considered stock-specific I_t computed during the previous year (i.e., the last 260 days). We rebalance this portfolio at the end of each trading day from Jan. 2014 to Dec. 2015 and compute its daily returns in excess of the equally-weighted index returns. Panel A displays the performance persistence statistics for the two-year holding period $2014-2015$, while Panels B and C present the same performance statistics for each of the of the two one-year subperiods.

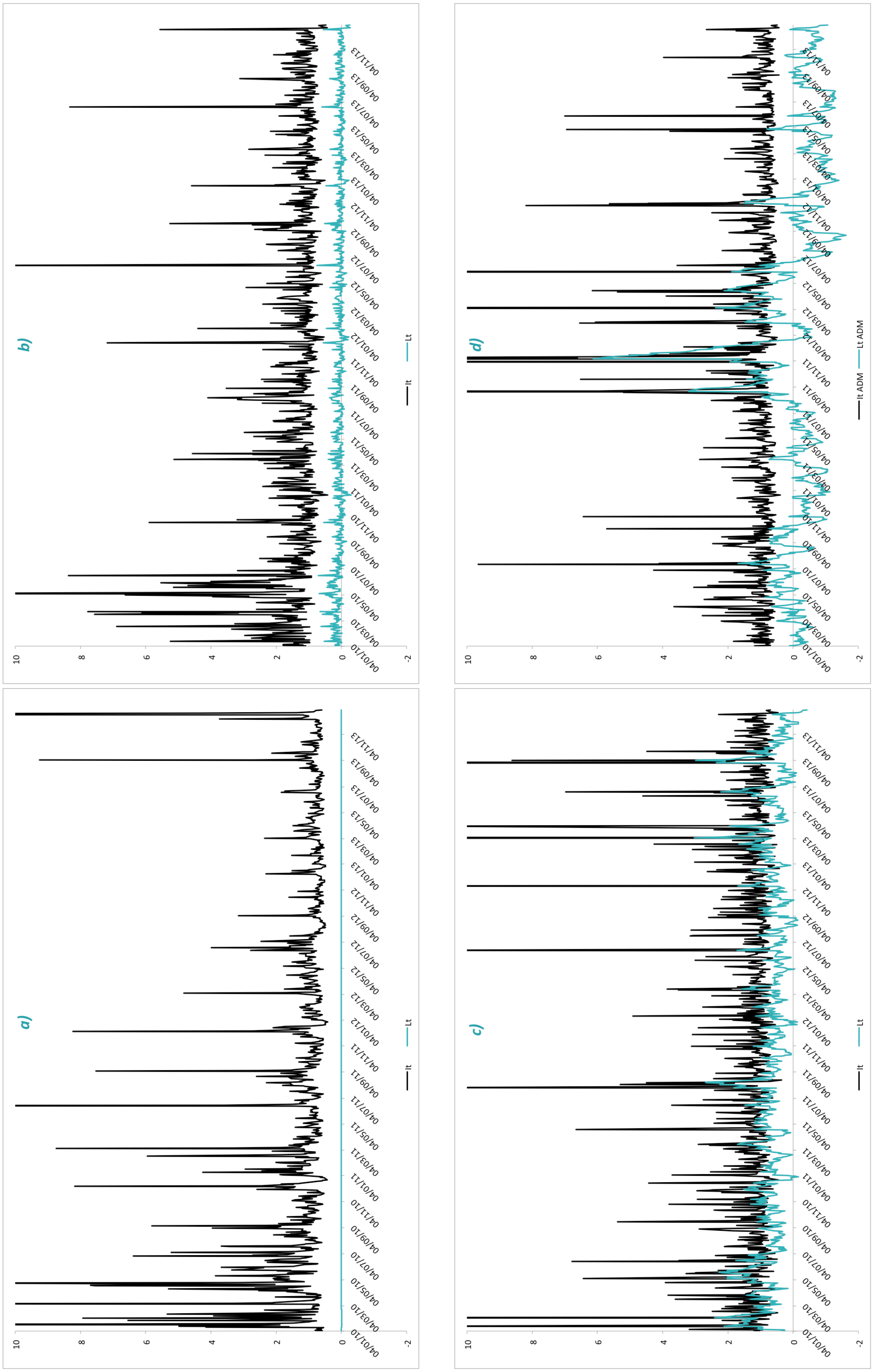
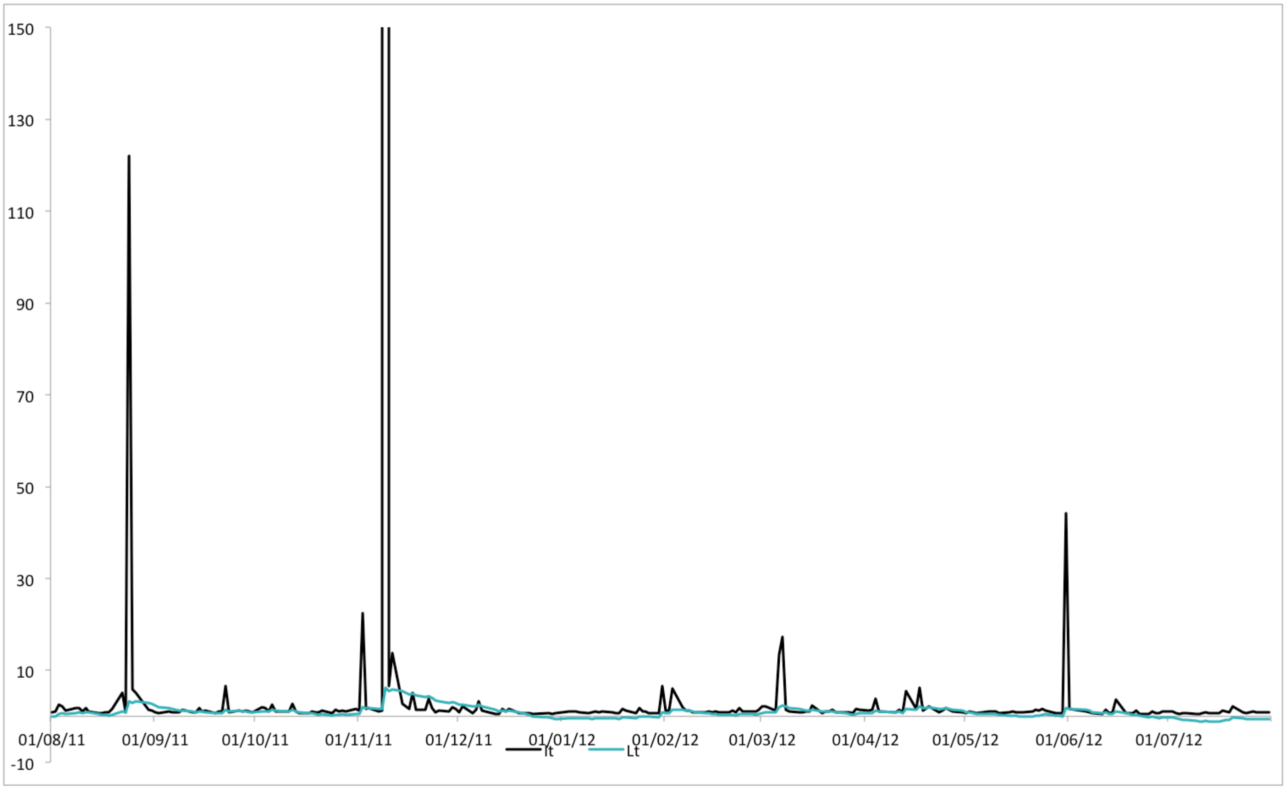


Figure 1: I_t and L_t time series for a) a no liquidity friction case à la Tauchen and Pitts (1983) - AHT LN, b) a no liquidity friction case à la Andersen (1996) - BATS LN, c) a Mixed liquidity friction case - ABF LN, and d) a short term liquidity friction case - ADN LN, from January 4, 2010 to December 31, 2013.



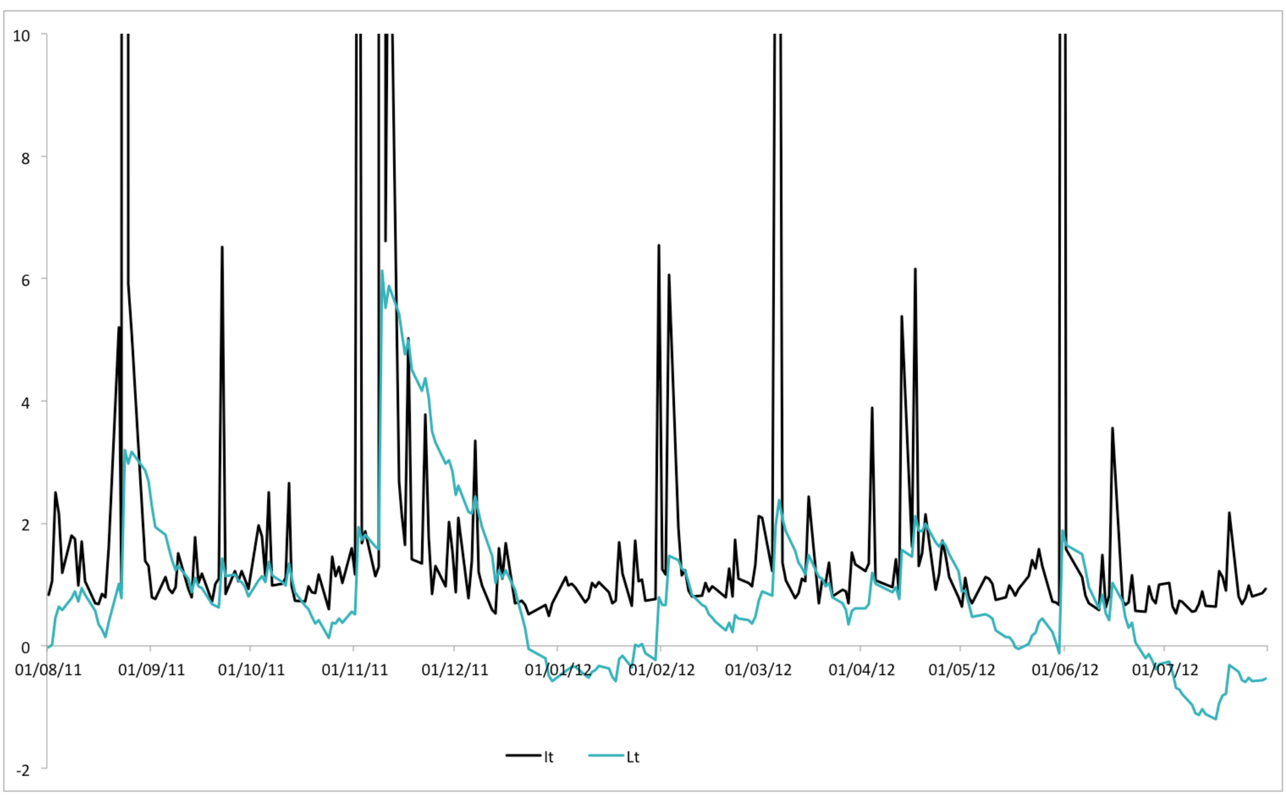


Figure 2: I_t and L_t time series for ADN LN, from August 1^{st}, 2011 to July 31, 2012.

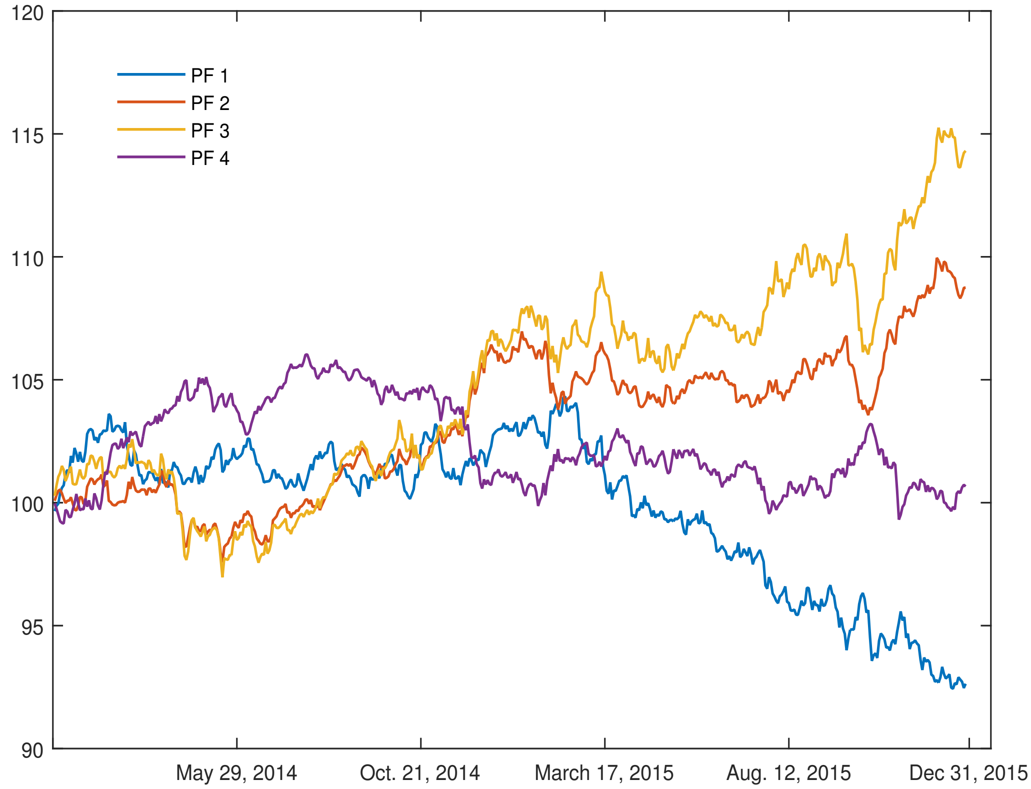


Figure 3: This figure plots the Net Asset Values of the four hedged portfolios considered here during the out of sample overall period extending from January 2014 to December 2015.

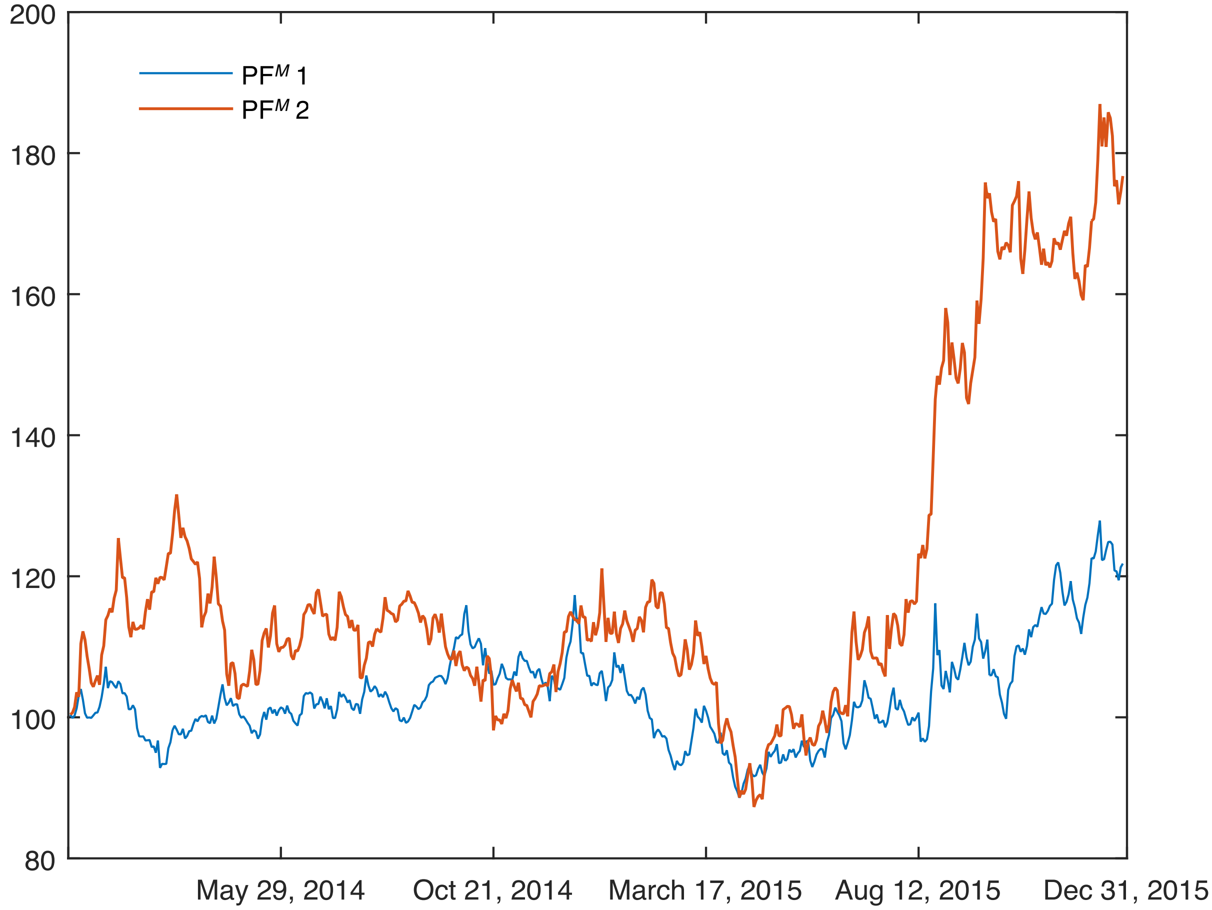


Figure 4: Net Asset Values of the two hedged momentum portfolios considered here during the out of sample overall period extending from January 2014 to December 2015.