INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2003-46

Trading Volume and Arbitrage*

S. DAROLLES[1]
G. LE FOL[2]

---

* We thank Cécile Boyer and Christian Gouriéroux for helpful comments. The research reported here was supported by a grant from Europlace Institute of Finance.

[1] Société Générale Asset Management, Hedge Funds Quantitative Research and Laboratoire de Finance-Assurance, CREST-INSEE, Paris.

[2] GEAPE, University of Angers, CREST – Laboratoire de Finance-Assurance and Europlace Institute of Finance.

Trading volume and Arbitrage[1]

Serge Darolles[2]   Gaëlle Le Fol[3]

December 2003.

[1] We thank Cécile Boyer and Christian Gouriéroux for helpful comments. The research reported here was supported by a grant from Europlace Institute of Finance.

[2] Société Générale Asset Management, Hedge Funds Quantitative Research and Laboratoire de Finance - Assurance, CREST-INSEE, Paris.

[3] GEAPE, University of Angers, CREST - Laboratoire de Finance - Assurance and Europlace Institute of Finance.

Abstract

Decomposing returns into market and stock specific components is common practice and forms the basis of popular asset pricing models. But what about volume ? Can volume be decomposed in the same way as returns ? Lo and Wang (2000), in a recent paper, suggest such a decomposition. Our paper is in this line of work and, despite the similarity of the statistical approach, our contribution is twofold. First, we provide a theoretical model to explain the decomposition of volume. Our model is the first, to our knowledge, to justify the strategies of new generation of traders, that we call liquidity arbitrageurs. Second, we propose a new efficient screening tool that allows practitioners to extract specific information from volume time series. We provide an empirical illustration of the relevance and the possible uses of our approach on daily data from the FTSE index from 2000 to 2002.

Keywords: Volume, Market portfolio, Arbitrage, Liquidity.

Résumé

La décomposition des rentabilités en une composante de marché et une composante spécifique est une pratique courante et constitue la base de modèles très populaires d'évaluation en finance. Mais cette décomposition peut-elle s'appliquer aux volumes ? Lo et Wang (2000), dans un article récent, suggèrent une telle décomposition. Notre papier est dans la lignée de ce travail. Si notre approche statistique est similaire, notre contribution est double. Tout d'abord, nous proposons un modèle théorique pour expliquer la décomposition du volume. Notre modèle est le premier, à notre connaissance, à justifier les interventions d'une nouvelle génération d'investisseurs, que nous appelons les arbitrageurs de liquidité. Nous proposons également une nouvelle méthode de filtrage efficace qui permet aux praticiens d'extraire de l'information spécifique à partir des séries de volume. Nous présentons une application empirique à partir de données journalières de l'indice du FTSE en 2000-2002 qui illustre l'intérêt de notre approche.

Mots-Clés: Volume, Portefeuille de Marché, Arbitrage, Liquidité.

JEL Classification: G11, G14, C21.

1 Introduction

If volumes like prices are unquestionably central in all equity strategies, financial theory traditionally focuses on prices, volatility and price formation analysis. Recently however, volume has received a considerable attention and forms now a dynamic and rapidly growing field of the literature. Volume has to be taken into account because it conveys information, and because it is an important measure of market feelings concerning one particular stock, one sector or one stock exchange.

For example, a large stock index rise in low or large volumes is not interpreted similarly by the market. A rise in low volumes is usually considered as fragile or temporary ; on the contrary a rise in large volumes seems strong and durable. If this use of volume, and hence, these interpretations are intuitive in the case of market or sector index, it is not as clear when the analysis concerns individual stocks. To see this, consider an individual stock included in a market index. Large traded volumes on this stock can either be due to investors interest for the market or for that particular stock. To understand why the volume of a particular stock is suddenly raised we must understand who trades stocks and why they do it ?

Llorente, Michaely, Sean and Wang (2002), following Wang (1994), in an equilibrium model framework, propose such an analysis of volume for individual stocks. They say that investors trade either to rebalance their portfolio for risk sharing (hedgers), or to speculate on their private information (speculators). The price and volume behaviours on the following days depend on the trading reason. As a consequence, the observation of price behaviour allows to discriminate between the two different trading purposes. In particular, small price variations in low volumes reflect hedging strategies, whereas large price variations in large volumes reflect speculative trading.

Our main goal is also to analyse volume and to discriminate between traders strategies, but our focus is on arbitrageurs vs hedgers and not on speculators vs hedgers. By arbitrageurs, we mean imperfection based traders. This trading motive has never been modelled, to our knowledge, even if it is an important reason to trade, probably even more important than speculation for example. In fact, few investors have private information and their trades are not only infrequent and small, compared to the overall traded volume, but also they are pinpoint events. On the contrary, some large investors, like hedge funds, permanently track market imperfections to enter the market. These opportunistic traders, observe a price variation due to market imperfections and trade to correct them. For example, in the case of a liquidity problem, the arbitrageurs enter the market to provide the missing liquidity, which reverses prices and allows them to cash a liquidity premium.

Understanding and decomposing volume can give some insight concerning the role and impact of market imperfection arbitrageurs' strategies on market characteristics. Moreover, it can provide valuable information to market participants concerning the state of the market, i.e. the level of imperfection. As a by product, we propose a more accurate measure of volume to empirically test volume-return and volume-volatility relations.

In this paper, we propose a simple equilibrium model to derive the traded volume dynamics generated by hedgers and liquidity arbitrageurs. Our model justifies the strategies of a new generation of traders and illustrates that such strategies can be identified through the analysis of volume. We show that hedging trades are common across assets whereas liquidity trades are asset specific. Price variation or price volatility signals to opportunistic traders when entering the market. As a consequence, they react to market volatility rather than being responsible for it. The volume they trade adds to the volume that would be traded if there were no imperfection, the "*normal volume*", and is proportional to the price volatility. It is also worth noting that, in our framework, the volatility/co-volatility of volume is directly linked to the existence of a liquidity/co-liquidity risk.

This theoretical result forms the basis of our empirical approach. We apply our methodology to individual stocks traded on the London Stock Exchange. We conduct a principal component analysis (PCA) to decompose volume into common and specific components. We examine the statistical properties of the two volume components and focus on the asset-specific component. More particularly, we test for illiquidity clustering through the dynamic analysis of the specific component of volume.

Note that the use of a volume decomposition is not new. Technical analysis proposes an increasing/decreasing volume decomposition and some theoretical papers decompose volume into a normal component - usually an historical average - and an abnormal or unexpected component [see e.g. Easley and O'Hara (1987), Andersen (1996)]. However, Easley - O'Hara (1987) study the informational contents of volume. They show that volume, like price, contains useful information concerning future price. In our work, volume does not contain information about future price, but instead valuable information about market liquidity.

More recently, decompositions of volume, into common and specific components, have emerged and appear to be a growing interest of the literature [see e.g. Hasbrouck and Seppi (2001), Lo and Wang (2001), Cremers and Mei (2001)]. If our statistical approaches show some similarity, our contribution lies in the modelling of investor's behaviour which justify the decomposition. Moreover, Hasbrouck and Seppi (2001), as well as He, Velu and Chen (2003), conduct a principal component analysis (PCA) to both volume and volatility to focus on the volume and volatility factors. Our interest is different as we focus on the specific component, i.e. the asset-specific part of volume. The reason for this comes from the non-stationarity property of volume. As a consequence, the PCA should be able to capture volume non-stationarity in the first factor and the analysis of the evolution of the idiosyncratic part of the volume should be informative. Our results show that this is really the case. In this sense, our method gives the way to filter the stock specific component of volume.

The stock specific component of volume developped here is a signed measure. It is positive if the stock is over-traded, compared to the market, and negative if the stock is under-traded, compared to the market. Therefore, it represents the relative market interest for a stock.

Our paper is organized as follow. In Section 2, we first discuss the volume measure that we consider, namely the individual turnover. We then propose un new theoretical model for volumes that justify the decomposition of turnovers and introduce explicitly the link between traded volumes and investing strategies. Section 3 presents the empirical methodology. Section 4 provides the empirical results using daily data for eight stocks from the FTSE index from 2000 to 2002. Finally, section 5 concludes the paper.

2 Turnover and market portfolio

After a short review of the different notions and measures of volume proposed in the literature, we develop some basic notations. We then formally define the volume measure we retain.

2.1 Analysis and measures of volume

In active markets – high volume markets, and hence liquid markets, the information flow is rapidly incorporated into prices through trading and trading volume. Volume has essentially been considered from this perspective in the financial literature with three main research directions. In the first two approaches, volume conveys information into prices and as such, has been considered through the analysis of volume-price relationship [see Easley and O'Hara (1987), Foster and Wisvanathan (1990), Boyer and Le Fol (1999)] or volume-volatility relationship [see e.g. Tauchen and Pitts (1983), Karpoff (1987), Foster and Wisvanathan (1993), Andersen (1996)]. In the latter, volume stands for a measure of liquidity or market quality [see Gallant, Rossi and Tauchen (1992), Domowitz and Wang (1994), Gouriéroux and Le Fol (1998) among others].

In this large body of literature, the first studies take the number of transactions as a proxy for volume, mainly for data availability reasons [Ying (1966), Epps and Epps (1976), Gallant, Rossi and Tauchen (1992), Heimstra and Jones (1994)]. Since then, numerous – aggregated as well as individuals - measures of volume have been proposed [see Lo and Wang (2000) for a review of the literature]. Turnover as a measure of volume was first introduced to account for the dependency between the traded volume and the total number of shares outstanding. As such, the turnover ratio, that is the traded volume corrected by the number of shares outstanding, seems to be appropriate when studying the market volume [Smidt (1990), LeBaron (1992), Campbell, Grossman and Wang (1993)] or when comparing individual asset volumes [Morse (1980), Bamber (1986), Bamber (1987), Lakonishok et Smidt (1986), Richardson, Sefcik and Thompson (1986), Stickel and Verrechia (1994)].

Following Lo and Wang (2000), we also retain the turnover ratio replacing however, the number of shares outstanding by the float[1]. First from a numerical point of view, as said above, turnover ratios, by pulling back assets volumes on a common scale, allow for comparison between assets. Second, from a financial point of view, under the regular hypotheses required for the CAPM to be valid, turnover measures must all be identical [see Lo and Wang (2000)]. This implication leads to a simple empirical test of the model. Moreover, the intuition of this first result is simple. All the agents hold the market portfolio and any transaction is linked to a buy or a sell of part of this portfolio ; as a consequence, all turnover ratios have to be identical as shown in section 3.1.

2.2 Volume and benchmarked volume

Let V_{it} be the number of shares traded for asset i on day t and N_{it} the float for asset i, $i = 1, ..., N$. We assume that the float for each asset is constant over time, i.e. $N_{it} = N_i$ for all t. The individual stock turnover for asset i on day t is given by:

$$x_{it} = \frac{V_{it}}{N_i}. \tag{1}$$

For a given asset, the individual turnover can equivalently be calculated in number of shares or in value, i.e. in euro volume. In the latter, one just have to multiply numerator and denominator, in the previous definition, by the stock price. More precisely, for daily turnover, the stock price to use is the daily volume weigthed average price,

$$P_{it} = \frac{\sum\limits_n P_{itn} \times V_{itn}}{\sum\limits_n V_{itn}},$$

where n is the index of the number of transactions . For a portfolio, however these definitions lead to different aggregation properties.

From the definition of the portfolio average turnover, or market index, we introduce the notion of benchmarked volume. To do so, we must take into account the individual asset price, and define the average turnover as the index euro volume – or the index traded value – divided by the index value:

$$x_t^I = \frac{\sum_i P_{it} V_{it}}{\sum_k P_{kt} N_k} = \frac{\sum_i P_{it} N_i \frac{V_{it}}{N_i}}{\sum_k P_{kt} N_k} = \sum_i w_{it} x_{it}, \tag{2}$$

[1] The float is the number of shares that are freely bought and sold buy the public.

where

$$w_{it} = P_{it}N_i \Big/ \sum_k P_{kt}N_k$$

is the weight of asset i in the market index.

3 The model

Let us consider an economy with N risky assets and a riskfree asset where two classes of investors trade. All investors have homogenous and rational expectations concerning the future value of risky assets. The first class of investors are J classical investors. They behave as portfolio managers and their trading desks. This means that they decide how much to trade at some decision dates (portfolio managers), and eventually trade sequentially this amount at some trading dates (trading desks). The second class of investors are opportunistic investors who enter the market only when there is a riskfree liquidity premium to cash. We consider a one period model, where classical investors decide to adjust their holdings in date 1 and to trade in date 2 and/or in between dates 1 and 2. The riskfree asset price is normalized to one in date 1, and its return is denoted by r_f. The risky asset i price and return in date t are $p_{i,t}$ and $r_{i,t}$, respectively, with $i = 1, ..., N$ and $t = 1, 2$. Moreover, we impose that returns are Gaussian, between two decision dates, and independent of the price level. Conditional on the information available at date t, we have :

$$\begin{aligned} E_t(r_{t+1}) &= \mu, \\ Var_t(r_{t+1}) &= \Omega, \end{aligned}$$

where $\mu = (\mu_1, ...\mu_N)'$ and Ω is an $N \times N$ positive definite matrix of variance-covariance of returns. In our framework, there is no incoming information and the floats are constant for all stocks.

3.1 Preferences and Investment strategies

The classical investors have negative exponential preferences and are endowed by w_1^j, $j = 1, ..., J$. In the first date, they hold an optimal portfolio in the Markowitz sense, that is, they hold $a_{1,0}^j, a_1^j$ which maximize their expected utility over their next period wealth of the following form :

$$E_1\left(-e^{-\lambda_1^j w_2^j}\right),$$

where λ_1^j is investor j's risk aversion coefficient in period 1. This program simplifies to :

$$\begin{cases} \underset{a_0^j, a_1^j}{Max} E_1\left(w_2^j\right) - \frac{\lambda_1^j}{2} V_1\left(w_2^j\right) \\ s.t. \quad w_1^j = a_{1,0}^j + \left(a_1^j\right)' p_1 \end{cases}$$

where the riskfree asset price is 1, and $p_1 = (p_{1,1}, ..., p_{N,1})'$ are the prices of the risky assets. The value of her portfolio in period 2 becomes :

$$w_2^j = a_0^j(1 + r_f) + \left(a_1^j\right)' diag\, p_1(1 + r_t),$$

where $r_{i,t} = \frac{p_{i,t}-p_{i,t-1}}{p_{i,t-1}}, i = 1, ..., N$, is the rate of return of the asset i in period t. $diag\, p_1$ is a diagonal matrix with $p_{i,1}$ prices on the diagonal. The maximisation of the program above gives the risky asset demand :

$$a_1^j = \frac{1}{\lambda_1^j}(diag\, p_1)^{-1}\Omega^{-1}\left[\mu + (1 + r_f)e\right], \tag{3}$$

where e is a unitary vector of size N. This is a classical result where the demand in risky assets is a function of the information and the risk aversion coefficient. In the context of no incoming information, the only parameter of interest if the risk aversion coefficient.

3.2 Equilibrium and turnover equality

If we suppose an exogenous constant a^M supply of shares, in equilibrium we have $a^M = a_1 = \sum_j a_1^j$, and we get the following relation :

$$a^M = \frac{1}{\lambda_1}(diag\, p_1)^{-1}\Omega^{-1}\left[\mu + (1 + r_f)e\right], \tag{4}$$

where $\lambda_1 = \left(\sum\limits_j \frac{1}{\lambda_1^j}\right)^{-1}$ is the harmonic mean of the individual risk aversion coefficients.

However, still in date 1, classical investors want to modify their positions due to an exogenous change in their risk aversion from $\left(\lambda_1^1, ..., \lambda_1^J\right)'$ to $\left(\lambda_2^1, ..., \lambda_2^J\right)'$. Since the overall numbers of shares remains constant, the supply is unchanged and in equilibrium, we have :

$$a^M = \frac{1}{\lambda_2}(diag\, p_2)^{-1}\Omega^{-1}\left[\mu + (1 + r_f)e\right]. \tag{5}$$

From equations 4 and 5, we get the equilibrium prices relation :

$$p_2 = \frac{\lambda_1}{\lambda_2}p_1.$$

For simplification purposes, we suppose that the overall risk aversion - across agents - remains constant[2]: $\lambda_1 = \lambda_2 = \lambda$. As a consequence, we also have

[2]Note that this simplification does not imply that individual λ_t^j are constant.

$p_2 = p_1 = p^*$. The relative price variation is zero. The allocation, $a_{i,1}^j$, at date 1, and the demand, $a_{i,2}^j$, set at date 1 to be hold at date 2, are :

$$a_{i,1}^j = \frac{\lambda}{\lambda_1^j} a_i^M, \quad a_{i,2}^j = \frac{\lambda}{\lambda_2^j} a_i^M. \tag{6}$$

The classical investors trade $a_{i,2}^j - a_{i,1}^j$ at date 2. Note that, they do not trade at date 1, as date 1 is only a decision date. In fact, they do not trade until date 2. The asset i turnover ratio x_i, being half the sum of all the volume bought or sold in stock i compare to the float of stock i on the market :

$$x_i = \frac{1}{2} \sum_j \left| \frac{a_{i,2}^j - a_{i,1}^j}{a_i^M} \right| .$$

From equation 6 and the definition of the turnover, we get the following proposition.

Proposition 1 *: Asset i observed turnover is equal to :*

$$x_i = \frac{\lambda}{2} \sum_j \left| \frac{1}{\lambda_2^j} - \frac{1}{\lambda_1^j} \right| , \tag{7}$$

is independent of i.

The proposition shows that the turnover ratio is a measure of the activity due to a modification of the risk aversion heterogeneity among investors. Moreover, it shows that moving from one equilibrium to another, the turnovers remains all equal across stocks. This result is consistent with the Lo and Wang (2000) one factor model, even if they do not explain why investors trade. Here, classical investors are trading because their risk aversion changes over time and independently from other traders. In Llorente, Michaely, Sean and Wang (2002), this equality does not hold anymore due to some heterogeneity stemed from different beliefs of investors. We show in the following section, that even when investors have homogenous beliefs, turnovers can vary from one stock to another.

3.3 Liquidity problem and excess in turnover

In this section, we introduce liquidity problems in our modelling by relaxing the assumption that all the trades are contemporaneous. We allow traders to trade sequentially.

3.3.1 Classical investors behaviour

Suppose that, if the classical investors are still deciding their portfolio composition, as previously at date 1, they can now trade sequentially. As a consequence, instead of trading all at once at date 2, they can trade between dates 1 and 2 in addition.

Again, this assumption is supported by market practices. In fact, portfolio managers decide to adjust their porfolio at some decision dates, but do not trade at these dates. Then, they send the number of shares to buy or sold to their trading desk, who realizes it in one or several trades if she wants to minimize the market impact generated by individuals portfolio adjustments. In fact, if the trading desk chooses to trade all at once, we are back in the previous set up. On the contrary, if she splits her trade to realize an average price over the period, she has to set a maximum quantity to trade per trading date. In such a case, only part of the trading is done at the final trading date. In fact, the trading desks reveal preferences for the present by anticipating the realization of part of the trading of the period.

To go further in our modelling, we need the following assumptions on the classical investors behaviour.

A_1 : Trade can only occur at the intermediate and final dates, but still not at the decision date.

A_2 : The maximal trading quantity at the intermediate date is drawn by a purely random mechanism.

A_3 : The final clearing of residual demands is done at the final date 2, at the previously defined equilibrium price p^*.

Assumption A_1, states that classical traders can only split their orders between two dates. Following A_2, let $X^j = \left(X_1^j, ..., X_N^j\right)'$ be a vector of independent uniform variables on $[0, 1]$, she trades, at this date, at most,

$$\overline{d}^j = diag\left(X^j\right)\left(a_2^j - a_1^j\right).$$

This mecanism allows to capture the splitting behavior of the classical investors. Summing up over the J classical investors, we get :

$$\overline{d} = \sum_j diag\left(X^j\right)\left(a_2^j - a_1^j\right).$$

This vector is a measure of illiquidity due to the classical investors' splitting scheme. Note that if $X_i^j = 0 \ \forall i, j$, we are in the particular case where investors trade all at once in date 2 - presented in the previous section - and there is no liquidity problem. When $X_i^j = 1 \ \forall i, j$, investors trade all at once but at the intermediate date and again, there is no liquidity problem in the market.

At the intermediate date, a new trading price $\widetilde{p}$ is set, and classical investors trade $\left(a^j(\widetilde{p}) - a_1^j\right)$, such that :

$$\begin{cases} \sum_j \left[a^j(\widetilde{p}) - a_1^j\right] = 0 \\ s.t. \; \left|a^j(\widetilde{p}) - a_1^j\right| \leq \left|\overline{d}^j\right| \end{cases} \tag{8}$$

where the demand $a^j(\widetilde{p})$ is of the same form as in equation (3) :

$$a^j(\widetilde{p}) = \frac{1}{\lambda_2^j}(diag\ \widetilde{p})^{-1}\Omega^{-1}\left[\mu + (1+r_f)e\right].$$

Finally, we get :

$$a^M = \sum_j \widetilde{a}^j(p), \tag{9}$$

$$\text{where } \widetilde{a}^j(p) = \begin{cases} a^j(p), & \text{if the quantity constraint is binding,} \\ \overline{d}^j + a_1^j, & \text{otherwise.} \end{cases} \tag{10}$$

In this situation, $\widetilde{p}$ is not exactly the set of prices that offset demand and supply since we have quantity constraints. As a consequence, the intermediate date can be interpreted as a disequilibrium date. In this framework, liquidity problems arise because of investors' heterogeneity splitting scheme behaviour, through the variable X^j. Moreover, because of this heterogeneity, their behaviour cannot be captured by the one of a representative agent.

Finally, the consequences of liquidity problems are not surprising. First, agents trade less than they want. Moreover, they pay an extra cost for immediacy since part of their trades (the anticipated trades) are concluded at a price $\widetilde{p} \neq p^*$. There are no strategic behaviour; some of them win $p^* - \widetilde{p}$ that others loose.

3.3.2 Opportunistic investors behaviour

We introduce a new class of investors whose aim is to gain from market imperfections, of the kind previously described. The opportunistic investors all have the same behaviour as a representative agent. We consider a stock (signed) demand, at the intermediate date, say $a^l(.)$, coming from a representative investor willing to cash the liquidity premium (liquidity arbitrageur). The liquidity arbitrageur strategically acts on liquidity undergone by classical investors. Her demand function satisfies the following assumptions :

$A_4 : a^l(p^*) = 0$.

$A_5 : a^l(.)$ is a monotonous increasing function of the difference of $p - \widetilde{\widetilde{p}}$.

$A_6 : a^l\left(\widetilde{\widetilde{p}}\right) < \overline{d}$, for all $\widetilde{\widetilde{p}}$.

Assumption A_5 ensures that the opportunistic investor will not trade if there is no liquidity problem. Assumption A_6 defines the strategical behaviour of the liquidity arbitrageur. She buyes when the market is selling, and sells otherwise.

Here again, the price mechanism is the same as in the previous section. Classical investors decide in date 1 the amount to trade and effectively trade at two posterior dates, under some quantity constraints. The new trading price, when the liquidity arbitrageur trades, satisfies :

$$\begin{cases} \sum_j \left[a^j\left(\widetilde{\widetilde{p}}\right) - a_1^j \right] + a^l\left(\widetilde{\widetilde{p}}\right) = 0, \\ s.t.\ a^j\left(\widetilde{\widetilde{p}}\right) - a_1^j \leq \overline{d}^j. \end{cases} \tag{11}$$

On the market level, the system (11) can be written as :

$$a^M - a^l\left(\widetilde{\widetilde{p}}\right) = \sum_j \widetilde{a}^j\left(\widetilde{\widetilde{p}}\right), \tag{12}$$

where $\widetilde{a}^j(.)$ is the function defined in 10.The additional demand $a^l\left(\widetilde{\widetilde{p}}\right)$ partly mitigates the effect of the splitting behaviour, by partly relaxing the quantity constraints. The traded volume increases, the impatient traders are less constraint and the impact of illiquidity on price is lower. We have :

$$p^* - \widetilde{\widetilde{p}} \leq p^* - \widetilde{p}.$$

Note that, if the liquidity arbitrageur provides all the missing liquidity trading, $a^l\left(\widetilde{\widetilde{p}}\right) = \overline{d}$, the quantity constraint disappears, the equilibrium price is p^*, and the arbitrageur gain is zero.

The volume of asset i, traded by the opportunistic trader, is $2 \times a^l\left(\widetilde{\widetilde{p}}\right)$; one half at the illiquid intermediary date and the same amount at date 2.

Proposition 2 *: Summing up the classical investors' trades, given in 7, and the liquidity arbitrageur's ones, we get the following turnover ratio between dates* 1 *and* 2 *:*

$$x_i^* = \frac{\lambda}{2} \sum_j \left| \frac{1}{\lambda_2^j} - \frac{1}{\lambda_1^j} \right| + \left| \frac{a_i^l\left(\widetilde{\widetilde{p}}\right)}{a_i^M} \right|, \tag{13}$$

for all stocks i. The turnover ratio has both a common and a specific components.

During market illiquid periods, new investors enter the market to provide liquidity where missing. The total volume traded between the two equilibria is raised by their trades and the turnover ratios are no longer the same as shown in equation 13 and in Figure 1.

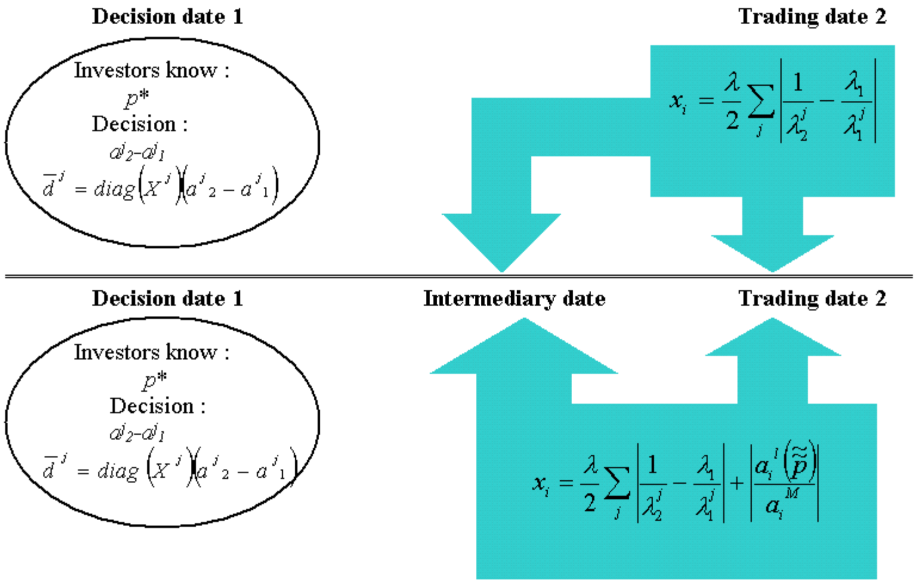

Figure 1: Trading scheme and turnover ratios.

3.4 Model discussion and testable empirical implications

Our model is much simpler than Wang (1994) and Llorente, Michaely, Sean and Wang (2002) models since we do not have to suppose heterogenous beliefs to explain variability in the turnover across stocks. Our investors are trading for hedging and liquidity arbitrage purposes. Besides, in their scheme each investor holds one stock and a nontraded asset. Since the returns of the two assets are correlated, any shock on the nontraded asset will force investors to adjust their stock holdings to maintain an optimal risk profile even when their risk aversion coefficient remains constant. This is qualitatively equivalent to a change in the risk aversion coefficient that we choose in our framework.

The price variation is usually the signal for opportunistic investors to enter the market in the intermediate period. The response to this signal depends on the risk behaviour of the opportunitics investors. If the signal can be misleading - saying that it may also be due to a change in the expectations of the future value - then they will only partially compensate the disequilibrium and potentially only cash out part of the liquidity premium. Taking several arbitrage positions, across signals, allow to diversify their risk. At one point, the risk can even be two high to get enough opportunistic traders entering the market. However, because we suppose no incoming information, opportunistic investors suffer no risk. Moreover, the signal and the opportunistic trader's intervention are simultaneous in our model.

The empirical analysis of turnover ratios of multiple assets traded on a single market leads to the rejection of the turnover equality property. This stylised fact brings Lo and Wang (2000) to reject the one factor model in favour of a two-factor model suggested by a principal component analysis. They show the existing conformity between the risk factors of pricing models and the factorial structure of volume series. Lo and Wang (2001) suppose the existence of only two types of risk : a market risk and a risk of market conditions modification. As a consequence at equilibrium, investors hold and trade only the market portfolio and a hedging portfolio providing the interpretation of their two factors linear model. Our model shows that liquidity problems can explain the rejection of the turnovers equality property without implying the failure of one-factor models.

The two-steps trading scheme of classical investors is very simple, and of course, affects the results on turnovers. More precisely, the independence of the probabilities to trade each asset implies that the only co-movement in turnovers are coming from hedging demands (common component). If the hypothesis is not supported by the data, the cross correlation of the specific component of turnovers should not be zero.

Finally, the liquidity arbitrage activity comes from the splitting scheme of the classical investors whereas the common component in the turnover comes from adjustments in the risk coefficients $\lambda_t^j, j = 1, ..., J$. Therefore, repeating our trading sequence independently would imply stationnarity of the specific component. On the contrary, the dynamics properties of the first component can hardly be described since λ_t is exogenous.

3.5 Numerical illustration

As an illustration of volume decomposition of any portfolio into its two components, we give a simple numerical example.

Here again, agents are deciding the amount of the market portfolio to trade at the first date and realize their trades eventually at two trading dates. Consider an economy with three stocks - 1, 2 and 3 - and four agents - A, B, C and D, where D is an arbitrageur. The characteristics of the stocks are summarized in the following table :

Stock	Float	Price	Index weight
1	40	1	0.50
2	20	1	0.25
3	20	1	0.25

Let agents A, B, C and D hold the following combinations of the three stocks :

$$A : \begin{bmatrix} 10 & 5 & 5 \end{bmatrix}, \quad B : \begin{bmatrix} 10 & 5 & 5 \end{bmatrix}, \quad C : \begin{bmatrix} 20 & 10 & 10 \end{bmatrix}, \ D : \begin{bmatrix} 0 & 0 & 0 \end{bmatrix};$$

saying that A and B hold 10 shares of stock 1 and 5 of stock 2 and 3 and agent C holds 20 shares of stock 1 and 10 of stocks 2 and 3, so that their investment in stocks 1, 2 and 3 in relative proportion is :

$$\begin{bmatrix} 0.5 & 0.25 & 0.25 \end{bmatrix}.$$

Hence, A, B and C hold pure benchmarked portfolios while agent D holds nothing.

Now suppose that due to some (exogenous) changes in their risk aversion coefficient, A, B and C decide to change their position. A wants to close her position, B wants to triple her risk exposure and C to lower hers by 50%. If they post their orders simultaneously to the market, B buys 20 shares of stock 1 (10 to A and 10 to C), 10 shares of stock 2 and 3 (5 to A and 5 to C). Their final positions are :

$$A : \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \quad B : \begin{bmatrix} 30 & 15 & 15 \end{bmatrix}, \quad C : \begin{bmatrix} 10 & 5 & 5 \end{bmatrix}.$$

In this situation, there is no tension in the market and the prices remain unchanged. Agents A, B and C's trades can be summarized as :

$$A : \begin{bmatrix} 10 & 5 & 5 \end{bmatrix}, \quad B : \begin{bmatrix} 20 & 10 & 10 \end{bmatrix}, \quad C : \begin{bmatrix} 10 & 5 & 5 \end{bmatrix}.$$

and there is no arbitrage since we observe $\begin{bmatrix} 20 & 10 & 10 \end{bmatrix}$ trades.

Consider now a situation where all the agents want to end up with the same positions as above, but if agent A still sell all her shares at once, agent B and C split their trades. Suppose that agent B trades first on stock 1 and 2 and postpones her trades on stock 3, while agent C trades on stock 2 and 3 and postpones her trades on stock 1.

A sells $\begin{bmatrix} 10 & 5 & 5 \end{bmatrix}$, B is willing to buy $\begin{bmatrix} 20 & 10 & 0 \end{bmatrix}$ and C is willing to sell $\begin{bmatrix} 0 & 5 & 5 \end{bmatrix}$. If the matching is instantaneous on stock 2, there is a lack of liquidity on the sell side for stock 1 (there are 20 shares to buy for only 10 to sell) and on the buy side for stock 3 (there are 10 shares to sell for not even one share to buy). These unbalances cause price pressures which will rise the price of stock 1 and lower the price of stock 3. These price movements incline agent D to enter the market to provide liquidity. Buying and selling the remaining quantities, she brings back the prices to their previous level until the equilibrium recovery. Doing so, she plays the role of a liquidity purveyor or a market maker.

Their positions between the intermediate trading date are :

$$A : \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \quad B : \begin{bmatrix} 30 & 15 & 5 \end{bmatrix}, \quad C : \begin{bmatrix} 20 & 5 & 5 \end{bmatrix}, \quad D : \begin{bmatrix} -10 & 0 & 10 \end{bmatrix}.$$

Agents A, B, C and D trades between the decision date and the intermediary trading date can be summarized as :

$$A : \begin{bmatrix} 10 & 5 & 5 \end{bmatrix}, \quad B : \begin{bmatrix} 20 & 10 & 0 \end{bmatrix}, \quad C : \begin{bmatrix} 0 & 5 & 5 \end{bmatrix}, \quad D : \begin{bmatrix} 10 & 0 & 10 \end{bmatrix}.$$

Once back to equilibrium, the arbitrageur will sell back the shares of stock 3 to agent B and buy the shares of stock 1 to agent C. The trading motives of agent C are to cash the liquidity premium, while agent B and C are adjusting their portfolio to end up with a pure market portfolio as in the case where there is no tension in the market. Hence, the agents trade $\begin{bmatrix} 20 & 10 & 10 \end{bmatrix}$ to move from the initial position to the intermediate position, and $\begin{bmatrix} 10 & 0 & 10 \end{bmatrix}$ once back at the equilibrium. The observed traded volume is $\begin{bmatrix} 30 & 10 & 20 \end{bmatrix}$ which represents $\begin{bmatrix} 20 & 10 & 10 \end{bmatrix}$ benchmark trades and $\begin{bmatrix} 10 & 0 & 10 \end{bmatrix}$ arbitrage trades. Note that the arbitrage represents in our example $\frac{10+0+10}{30+10+20} = 33\%$ of the activity observed in the volume. This example clearly shows that the reason for turnovers to be different comes from arbitrage.

Note that in this example, because prices are constant and because the liquidity purveyor provides all the missing liquidity, her profit is zero.

4 Empirical methodology

Our model explains positive deviation from the common component of turnovers, i.e. the market index turnover. In practice, the identification of arbitrage is straightforward. The idea is to isolate the lowest turnover among all stocks. The arbitrage is the excess in turnovers observed on the other stocks and the sum of all these extra turnovers is a measure of the overall illiquidity of the market. However, empirically this approach is not satisfying. In fact, it depends on only one observation (the lowest turnover) and thus cannot be robust. We prefer to base the empirical approach on an average instead of an extreme observation. Doing so, we implicitly assume that the overall illiquidity can be spread out over all stocks. The deviations from the average turnover can now be either positive or negative which fits better what we usually observed on markets and will be justified in this section.

4.1 Link with market practices

Our approach comes from asset management practices, in which any portfolio can be decomposed into a market portfolio and an arbitrage portfolio. Applied to volume, we get a market component and an arbitrage component of the trading volume. The first factor in our volume factorial analysis can be identified as the market component whereas the remaining part will represent the arbitrage component. In our one-factor approach, stock turnovers inequality comes from the existence of arbitrage behaviours.

Consider a market where I assets, indexed by $i = 1, \ldots, I$, are traded, by J market participants. The float for all asset is fixed and denoted by N_i. Knowing the prices of all the asset at date t, we get the market value, say $\sum_k P_{kt} N_k$. The relative weight of asset i compared to the market value is $w_{it} = P_{it} N_i \,/ \sum_k P_{kt} N_k$, where P_{it} is the price of asset i at date t. This weight also stands for its weight in the market portfolio. Consider that agent j portfolio differs from the market portfolio at date t. This portfolio depends on the weights α_{ijt}, $i = 1, \ldots, I$, of all assets. These weights can be decomposed in the following way:

$$\alpha_{ijt} = I_{ijt} + A_{ijt}, \tag{14}$$

i.e. a market component (or index component) I_{ijt} plus an arbitrage component A_{ijt}. The arbitrage component can either be positive if asset i is over weighted in agent j portfolio, or negative in the reverse case.

The same reasoning appliesto trading volume for a particular asset. When an investor adjusts her portfolio, she buys or sells a risky portfolio fairly close to the market portfolio. If her behaviour is the one of an agent in equilibrium, she trades exactly the market portfolio. If her goal is to trade on her private information – concerning one asset or more – she will trade a quite different portfolio from the market portfolio. The extreme situation being agent j buying or selling one only asset. Therefore, the volume V_{ijt} traded by agent j on asset i at date t is the result of adjustments of both her index portfolio and her arbitrage portfolio. In terms of individual turnover, we can write:

$$x_{ijt} = x^I_{ijt} + x^A_{ijt}, \tag{15}$$

where x^I_{ijt} stands for the index – or market – turnover and x^A_{ijt} for the arbitrage turnover. Summing over all agents j, we get an aggregate measure of the activity derived from risky positions adjustments on asset i, say:

$$x^I_{it} = \frac{1}{2} \sum_j x^I_{ijt}, \tag{16}$$

and an aggregate measure of the activity derived from arbitrage strategies:

$$x^A_{it} = \frac{1}{2} \sum_j x^A_{ijt}. \tag{17}$$

Note that the coefficient $\frac{1}{2}$ corrects for the double counting when summing the shares over all investors.

Finally, at an aggregate level, we get for any asset, the following turnover decomposition:

$$x_{it} = \frac{1}{2} \sum_j x_{ijt} = x^I_{it} + x^A_{it}. \tag{18}$$

The practical interest of such a decomposition is obvious and will be detailed in the following. On a theoretical point of view, the question is to identify the two components of the turnover from the observation of the sum.

Without any constraint, this identification cannot be done. For any agent i, the problem can be set up as the resolution of a one-dimensional linear system with two variables, where the variables are the market portfolio and the arbitrage portfolio weights. This system has an infinite number of solutions and uniqueness can only be reached by imposing a constraint to the arbitrage portfolio.

4.2 Components identification

At an individual level, say for any agent j, the solution is straightforward and comes from portfolio management practices. A fund manager willing to invest in a pure arbitrage portfolio must have an identical risk exposure both on her long – the stock she buys - and short positions – the stocks she sells. The risk exposure notion is not obvious but, to make it simple, we assume that it can be captured by the invested value. Under this assumption, an arbitrage portfolio is thus said *dollar neutral*[3] as opposed to beta-neutral portfolios where the fund manager adjusts the betas of the long and short positions.

The constraint to impose in order to obtain a unique arbitrage portfolio is then obvious: any arbitrage portfolio must be dollar-neutral, and hence for all date t and agent j, it must satisfy:

$$\sum_i P_{it} N_i x^A_{ijt} = 0. \tag{19}$$

From this constraint, we recover identification: if the portfolio is risk-neutral, then agent j uses the total value she trades at date t to adjust her market component. Knowing the market portfolio weights, and from the total number of shares traded by agent j, one can easily get her traded market portfolio. The deviations from this virtual portfolio gives the arbitrage portfolio of agent j in traded volume. We get the decomposition in terms of turnover dividing the volumes by the number of float shares.

At an aggregate level – when we only observe the total number of traded share, and without imposing any additional constraint, identification is not either possible. Here again we will follow the same reasoning. We suppose that the arbitrage activity satisfies a dollar neutral constraint saying that the value invested to buy is equal to the one received from selling. In all date t, the constraint is:

$$\sum_i P_{it} N_i x^A_{it} = 0, \tag{20}$$

and we get back the identification of the two components of the traded volume for stock i.

[3] The term *dollar neutral* refer to a zero-cost portfolio, i.e. a portfolio composed of an equal dollar amounts of long and short investments.

The decomposition between the benchmark portfolio and the arbitrage activity is as simple as in the individual case. The identification constraint imposes to the total traded value – or dollar volume – to be equal to the value traded on the market component in all date t. From the stocks weights in the market portfolio, we derive the number of shares traded for the benchmarked activity. The difference between this number of shares and the observed number of shares traded gives the level of the arbitrage activity.

4.3 Principal component analysis and first factor identification

In this section, we propose some empirical tests to discriminate between Lo and Wang (2000) interpretation of observed differences across stocks-turnover and ours. This can be done by studying the dynamic characteristics of the factors which summarize the joint evolution of stock turnovers.

This data reduction can be done by conducting a principal component analysis on the stock turnovers. Then the factors are analysed and we carry on a simple empirical test to identify the first component to the market average turnover. Hence, any stock turnover, at any date, depends on an average term and a deviation term. The average part corresponds to trading volume coming from market portfolio adjustments. Our interpretation is that the deviation part is due to the opening and closure of arbitrage positions.

A second test consists in analysing the dynamic properties of factors of order greater than one to discriminate between Lo and Wang (2000) interpretation and ours.

Let x_{it}, $i = 1, \ldots, I$, $t = 1, \ldots, T$ denote the turnover series, i.e. the number of traded shares divided by the number of float shares. Since the aim of principal component analysis is to explain the variance-covariance structure of the data through a few linear combinations of the original data, the first step is to calculate the $I \times I$ dimension variance-covariance matrix of the data. The spectral decomposition of this matrix leads to I orthogonal vectors, $C_t^k = x'_{it} u_k$, with dimension T, where u_k is the k^{th} eigenvector. Each eigenvector is associated with a positive eigenvalue λ_k such that:

$$Cov(C_t^k, C_t^l) = \lambda_k \delta_{kl}, \tag{21}$$

where δ_{kl} stands for Kroneker symbol. The standardized turnover times series can be decomposed as:

$$\frac{x_{it} - \overline{x}_i}{\sigma_i} = \sum_k u_k^i C_t^k.$$

Since $corr(x_{it}, C_t^k) = \sqrt{\lambda_k} u_k^i$, the previous equation can be rewritten as:

$$\begin{aligned} x_{it} - \overline{x}_i &= \sigma_i \sum_k \frac{corr(x_{it}, C_t^k)}{\sqrt{\lambda_k}} C_t^k \\ &= \sigma_i \sum_k \frac{corr(x_{it}, C_t^k)}{\sqrt{var\left(C_t^k\right)}} C_t^k \\ &= \sum_k \frac{Cov(x_{it}, C_t^k)}{var\left(C_t^k\right)} C_t^k \end{aligned}$$

Finally, we get the centered turnovers :

$$\begin{aligned} x_{it} - \overline{x}_i &= \sum_k \frac{Cov(x_{it}, C_t^k)}{var\left(C_t^k\right)} C_t^k & (22) \\ &= \sum_k \frac{1}{\lambda_k} Cov(x_{it}, C_t^k) C_t^k, & (23) \end{aligned}$$

Isolating the first factor, we get:

$$x_{it} - \overline{x}_i = \frac{1}{\lambda_1} Cov(x_{it}, C_t^1) C_t^1 + \sum_{k>1} \frac{1}{\lambda_k} Cov(x_{it}, C_t^k) C_t^k. \qquad (24)$$

To see if the market turnover, as defined in Section 2.2, is a good candidate for the first factor, we compare the first component of the sum in Equation (24) to the market turnover. This comparison can only be done after correcting for the mean and the variance, thus we compare the following times series:

$$\overline{x}_i + \sigma_i \frac{1}{\lambda_1} C_t^1, \qquad (25)$$

to the market turnover.

4.4 Dynamic properties of stock specific component

The empirical analysis of the dynamic properties of the factors, derived from the aforementioned approach, leads to a different interpretation from the one by Lo and Wang (2000). In fact, any observed non stationarity in the joint analysis of volume is due to the existence of nonstationary common factors. If the number of such factors is greater than one, the Lo and Wang type of analysis is the most accurate. On the contrary, if there is exactly one nonstationary factors, deviations from the first factor must be interpreted differently.

Whenever the first factor has been identified as the index – or benchmark – component of volume, we can focus on the analysis of the second component of the sum in Equation (24):

$$\sum_{k>1} \frac{1}{\lambda_k} Cov(x_{it}, C_t^k) C_t^k. \tag{26}$$

Once again, different interpretations are possible. Lo and Wang (2000) see this term as an hedging strategy against a risk associated with market conditions modifications. In this view, the first factor and the others are associated with investment decisions of the same kind. This implies that they should both present the same dynamic characteristics.

On the contrary, we suppose that the second component is due – or linked – to some short term arbitrage activity. Then the two components of the decomposition should feature very different dynamic behaviours. In particular, the first component should capture all the trend observed in the turnover series whereas the second should be stationary. A standard stationarity test can then be a validation test of either one of the two approaches.

5 Empirical results

In this section, we apply the approach presented above to daily data from the eight most important stocks from the FTSE index, namely AstraZeneca (AZN), Barclays Bank (BARC), GlaxoSmithKline (GSK), HSBC Holdings (HSBA), Lloyds TSB Group (LLOY), Royal Bank of Scotland Group (RBS), Shell Transport and trading co (SHELL) and Vodaphone Group (VOD) from May 17, 2000 to December 5, 2002 (648 trading days). Note that, if intraday data seems to be a more appropriate choice when working on investment practices, the high intraday seasonality of volume and the associated intraday seasonal adjustment problems encourage us to work on daily data.

5.1 Data description

Table 1 gives some summary statistics about the eight aforementioned stocks from May 17, 2000 to December 5, 2002. Over this period and for all the stocks we have 648 trading days, i.e. 648 daily observations. Table 1 displays the mean, the standard deviation, the minimum and the maximum of the traded volume and the number of float shares in millions of shares.

Variables	Traded volume			Number of float shares		
Stat Stocks	*Mean*	*Std*	*Min/Max*	*Mean*	*Std*	*Min/Max*
AZN	5.1	3.0	0.27/28.6	1755.1	14.5	1730.1/1770
BARC	24.1	12.1	2.23/94.1	6508.1	264.7	5912/6628
GSK	14.2	10.4	1.15/193.2	5570.9	1077.3	3646/6218
HSBA	25.8	18.5	1.19/267.9	9268.5	221.5	8530/9431.1
LLOY	19.8	10.7	1.13/83.0	5528.1	31.3	5497/5561.2
RBS	9.2	7.3	0.55/141.4	2778.1	99.2	2586/2887.5
SHELL	33.4	19.2	3.19/133.9	9856.9	81.7	9733.2/9942
VOD	314.7	171.5	18.6/1458.8	65949.4	2695.9	61443/67895.1

Table 1 : Descriptive statistics, daily data from Mai 17, 2000 to December 5, 2002.

Over this period, volumes and outstanding number of shares are very different among stocks. In addition, stock ranking is roughly the same when considering daily averages of volume or number of shares outstanding. These observations justify the choice of turnovers instead of traded volumes.

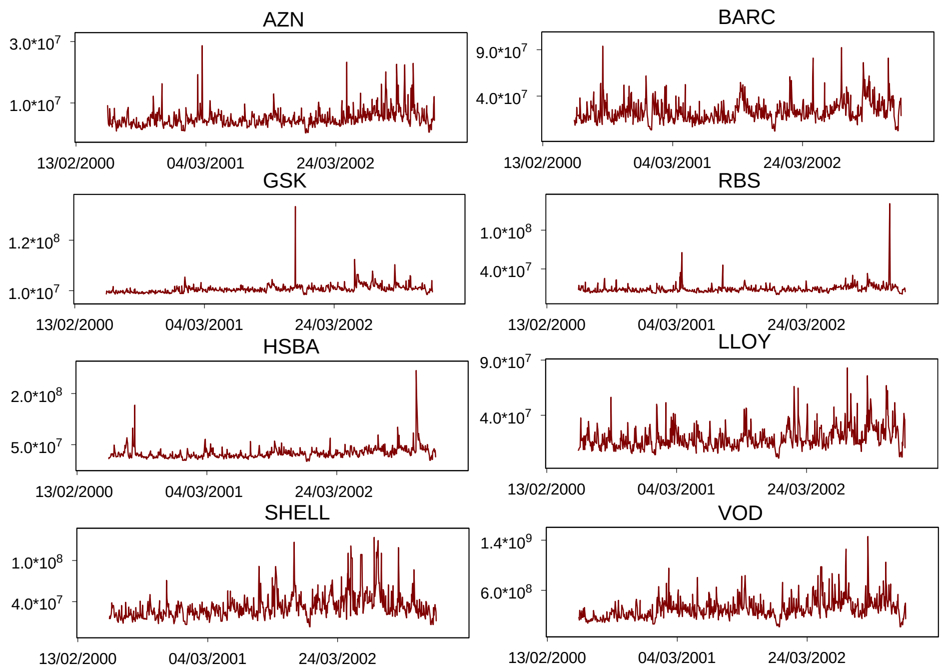

Figure 2: Daily traded volume evolution from May 17, 2000 and December 5, 2002.

A visual inspection of Figure 2, which gives the evolution of volumes in daily number of traded shares, shows that some rises in volumes appear in all stocks, like at the end of September 2002 for example, whereas some other ones seem to be stock specific. These large jumps can even hide common rises of volume. This is the case for GSK, on November 30, 2001, where the daily traded volume reaches 193 200 000 shares compared to an average of 33 400 000 traded shares per day over the period.

This first analysis shows that the analysis of the traded volume must account for the total number of shares outstanding. Moreover, there seems to be two components in volume : a common component and a specific component.

From these observations, we first propose a measure of volume corrected from the outstanding number of shares as previously described, that is the daily turnover in percentage. This measure is 100 times the daily traded volume divided by the float. Figure 3 displays the evolution of the observed daily stock turnover ratio in percentage for the eight stocks from May 17, 2000 to December 20, 2002.

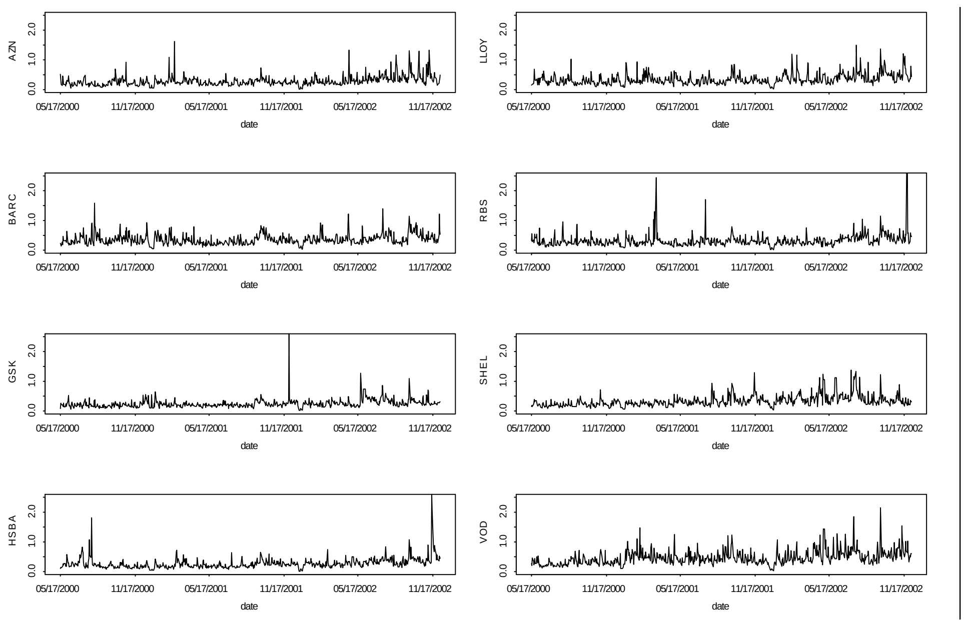

Figure 3: Turnover evolutions, from Mai 17, 2000 to December 6, 2002.

Because of this preliminary treatment, the volume series become comparable. The analysis of their dynamics shows the existence of a trend, whatever the stock. This trend appears also in the one-year daily average of stock turnover evolutions as shown in Table 2.

	Year		
Stocks	**2000**	**2001**	**2002**
AZN	0.226	0.260	0.368
BARC	0.364	0.327	0.422
GSK	0.208	0.233	0.302
HSBA	0.234	0.232	0.356
LLOY	0.288	0.321	0.440
RBS	0.288	0.312	0.383
SHELL	0.229	0.317	0.449
VOD	0.303	0.480	0.595

Table 2 : One-year Daily turnover average evolutions from year 2000 to year 2003.

In fact, volume is rising at rates of about 16% to 96%. As this trend is observed on every stock, it should be captured by the common component of our decomposition. As a consequence, we will focus on the specific component of the decomposition and there is no need to correct from the initial series trend. Our decomposition is a natural answer to this problem.

We report summary statistics for the daily stock turnover ratio in percentage in Table 3.

	Mean	**Std. Dev**	**Skweness**	**Kurtosis**
AZN	0.29462	0.17424	2.91100	13.67123
BARC	0.37516	0.18373	1.80527	5.87145
GSK	0.25350	0.16400	9.08276	144.52370
HSBA	0.28050	0.20084	5.64628	55.75478
LLOY	0.35950	0.18862	1.82406	5.27974
RBS	0.33506	0.26048	9.74110	152.69574
SHELL	0.34444	0.19172	2.04292	6.15997
VOD	0.48319	0.24852	1.70851	5.58066

Table 3 : Descriptive statistics of stocks turnover.

The average turnover is quite different from one stock to another ; the largest turnover being almost twice the lowest. However, this difference is mitigated by the use of turnover instead of traded volume series. In fact, the average traded volume for VOD (314 700 000 shares) was greater than 60 times the AZN average traded volume (5 100 000 shares). We can also note that the distributions of AZN, GSK, HSBA and RBS are the ones with the larger skewness and kurtosis and hence the most asymmetric with the largest tails.

5.2 Principal component analysis

As seen before, the approach allows us to summarize the behaviour of stock volume series over the entire period 2000-2002. The first factor explains

40.6% of the variability-covariability of the stocks turnover. In this section and the following one, the principal component analysis approach is conducted using daily data for the aforementioned stocks. However for presentation purposes, we report only the figures for three stocks : GSK, AZN and RBS.

Figure 4 reports the evolution of monthly averages of the first factor as well as the market turnover. We can see that the activity is quite stable

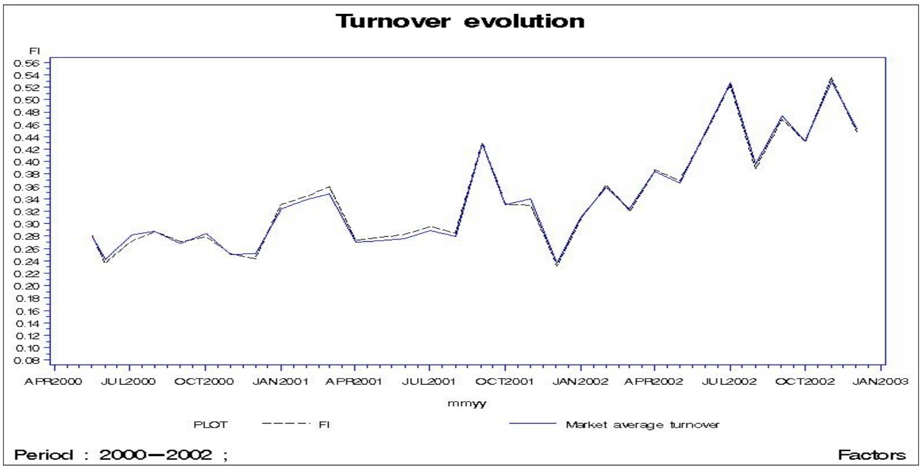

Figure 4: Turnover (Market average turnover) and Index component (FI), month-by-month average evolution, May 2000-December 2002.

during the year 2001. In October 2001, this activity sharply falls before recovering and starting to sharply increase during the rest of the period. A visual inspection of the fit between the market turnover and the first factor confirms the identification of this factor to the benchmark - or index - component of volume.

We give in table 4, the correlation analysis between AZN, GSK and RBS.

	AZN/RBS	AZN/GSK	RBS/GSK
Turnover	0.25*	0.307*	0.158*
Specific component	-0.057	-0.0766	-0.136*

Table 4 : Cross correlations of turnovers and of the specific component of turnovers.

Most of the turnover co-movements have been captured by the one-factor model. However, some correlations remain significant[4] in the specific component of the RBS and GSK's turnovers. As mentioned before, non-zero correlations imply that the theoretical model with independent trading probabilities is rejected. There seems to be some clustering in the preferred group of stocks investors trade first.

[4]The * symbol means that the correlation is significative different from zero at 5% or lower.

Figure 5 illustrates the nature of the second component which shows an erratic behaviour around zero but no trend. The stock specific component appears to be stationnary unlike the common component.

To confirm these results, we calculate and report in Figures 6, 7 and 8 the autocorrelation and the partial autocorrelation functions of the common and the specific components for GSK, AZN and RBS, respectively.

If the index component is non stationary and features long memory or changes in regime, the second component presents only short term memory. Moreover, the individual analysis of the seven factors sum, i.e. the second component, shows that none of them is nonstationary. Hence, the nonstationarity is completely absorbed by the common component. This result fits the financial interpretation presented in Section 4. Moreover, the first component features the same seasonality as the turnover series. At an aggregate level, this also confirms that the two components can not reflect the same type of portfolio management. The first component reflects long term management strategies whereas the second component reflects short term management strategies. We conclude that the variability of stocks trading volume is well approximated by a two components model (a one factor model) : an index component and an arbitrage component.

5.3 Uses and limits of the methodology

In this section, the principal component analysis is applied to each stock separately and each turnover - each volume measure - is decomposed in a common and a specific components. The idea is to discriminate between investors interest for the market and a specific activity on the stock itself from the volume analysis. Over the entire period, and for all stocks, the activity is mostly driven by its index component.

The first example concerns the stock GSK. We give in Figure 9, the evolutions of the turnover components over a two month period at the very end of the year 2001. We clearly see the common component driving the evolution of the turnover almost all the time except during few days at the turn of the month. This phenomenon is a direct consequence of large fund management companies intervention. In fact, these companies are sometimes modifying their holding in some groups in order to clear their position before the end of the year. These large transfers are usually done using applications[5] at the end of the year. These trades which do not correspond to any underlying market activity are part of the traded volume. Figure 9 illustrates such practices. Here, illiquidity is not the source of trade, and there is no arbitrage. In such a case however, our measure still helps. In fact, one should replace the traded volume of these few days by the index component to correct from this portfolio rebalancing clearing practice. This

[5] An application - block trade - is a buy and sell agreement concluded outside the electronic market and reintroduced into the system with a delay.

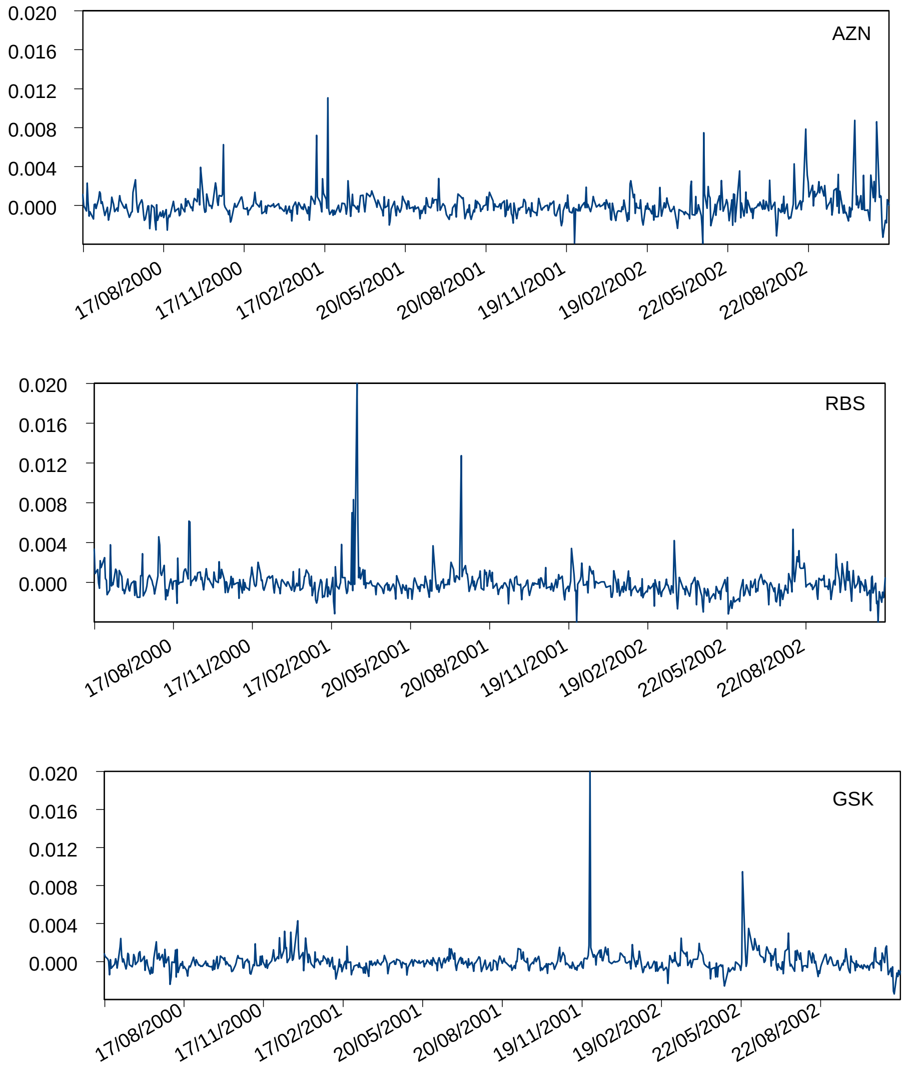

Figure 5: Evolution of the the specific components of volume, from May 17, 2000 to December 6, 2002.

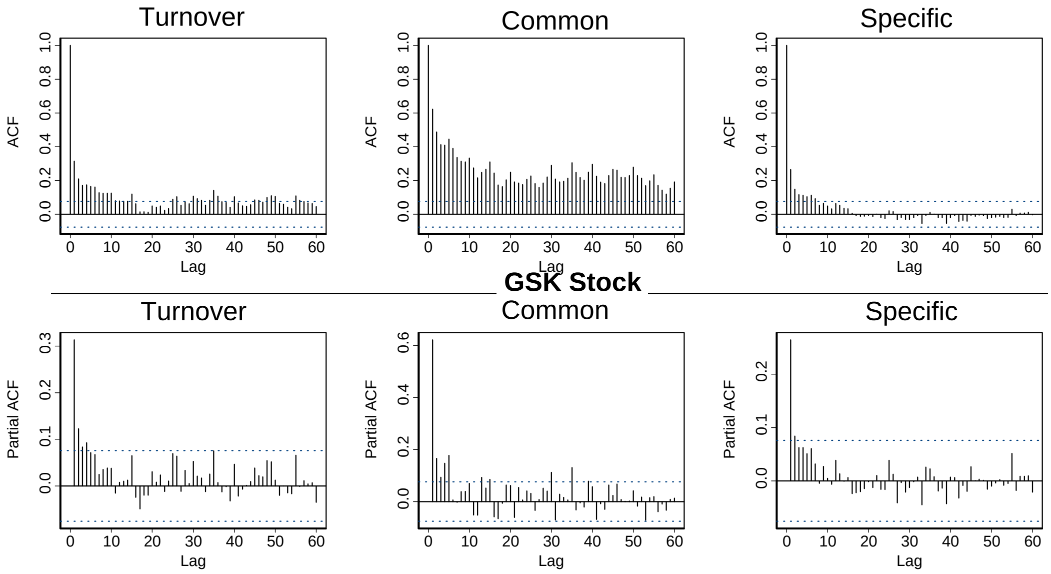

Figure 6: Autocorrelation and partial autocorrelation functions of the two components, GSK Stock.

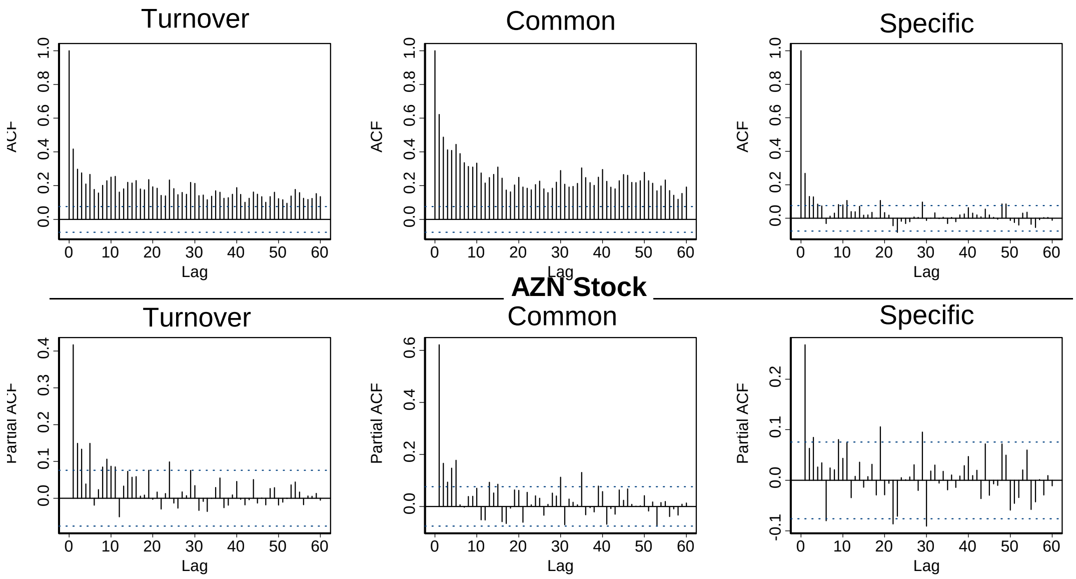

Figure 7: Autocorrelation and partial autocorrelation functions of the two components, AZN Stock.

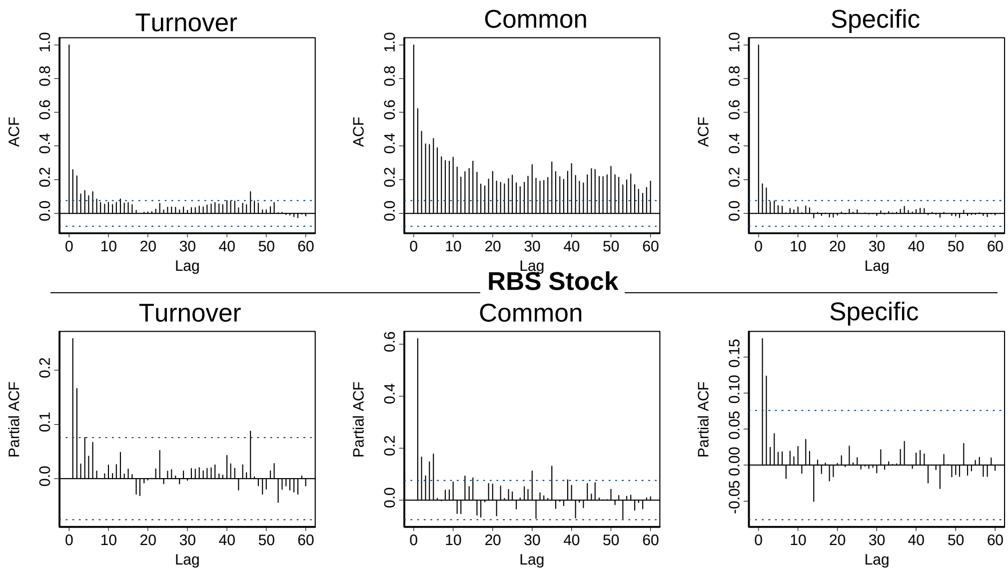

Figure 8: Autocorrelation and partial autocorrelation functions of the two components, RBS Stock.

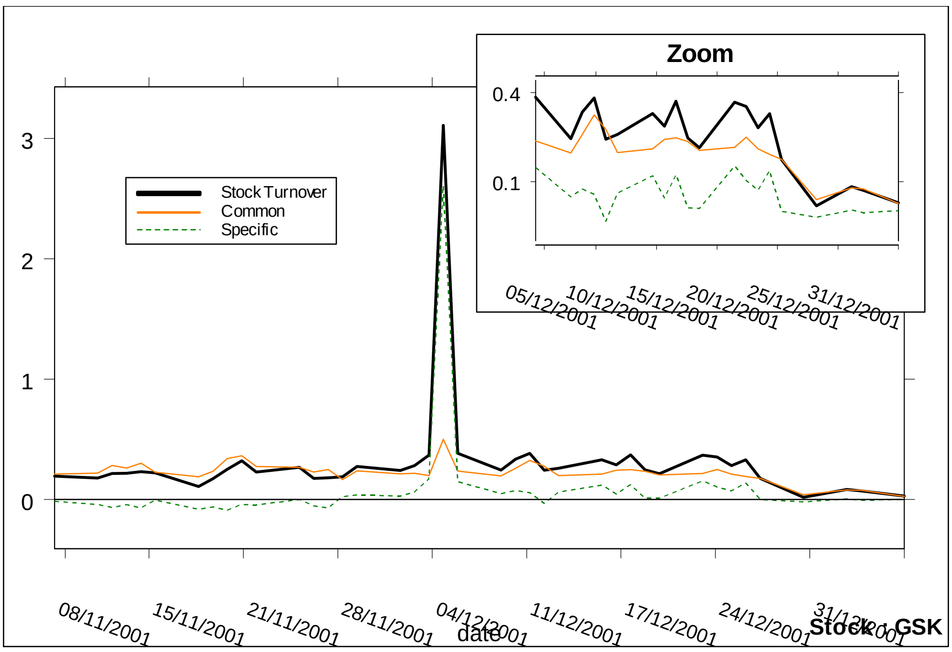

Figure 9: Turnover and common component, GSK stock.

example is a practical application of our approach. Using our data error correction method on the volume series can eliminate noise due to particular financial practices.

The next figure shows the capacity of our statistical method to accurately extract seasonalities as previously mentioned in Figures 6, 7 and 8. Figure 10 shows the classical end of the year drop in volumes. This effect is completely captured in the common component and explains the RBS volume evolution between December 24, 2001 and the very beginning of the year 2002.

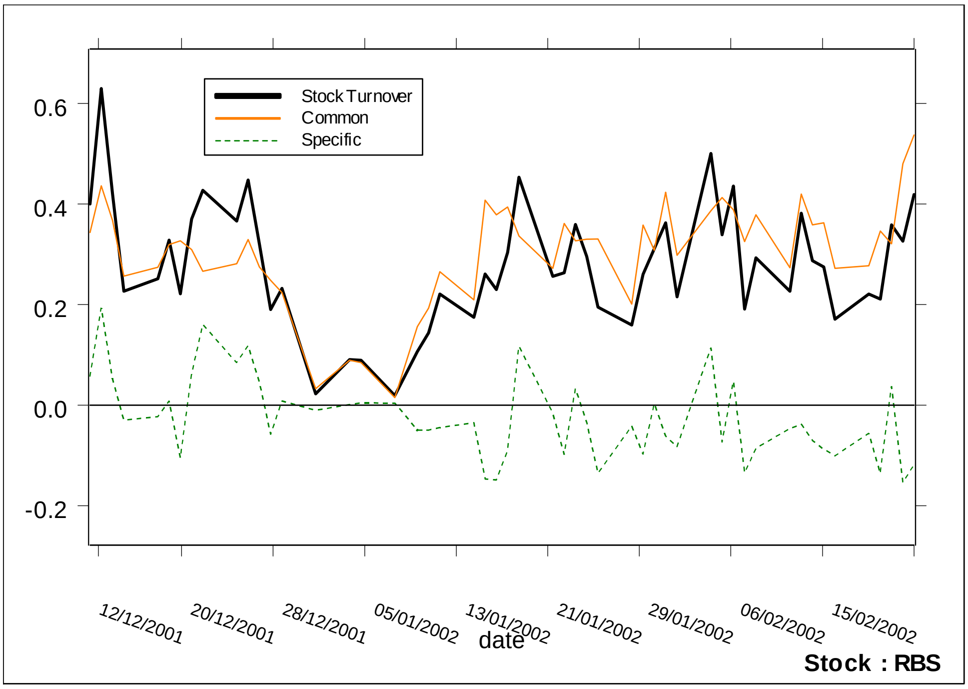

Figure 10: Turnover and turnover components, RBS Stock.

The methodology presented in this paper, is accurate for filtering purposes and in periods when no information hits the market as we will see in the following example of the AZN stock. The daily evolution of the AZN turnover and its component, given in Figure 11, is mostly driven by benchmarked strategies.

However, the two series are not that close in July due to the growing global arbitrage activity, i.e. liquidity arbitrage activity but also informational arbitrage activity, at that time. As we can see, the arbitrage activity on AZN stock is growing from the very beginning of July and displays a peak on the July 17, 2002. This period corresponds to a pessimistic period concerning future profitability of AstraZeneca. AstraZeneca is a pharmaceutical company whose earnings come mostly from the production of one

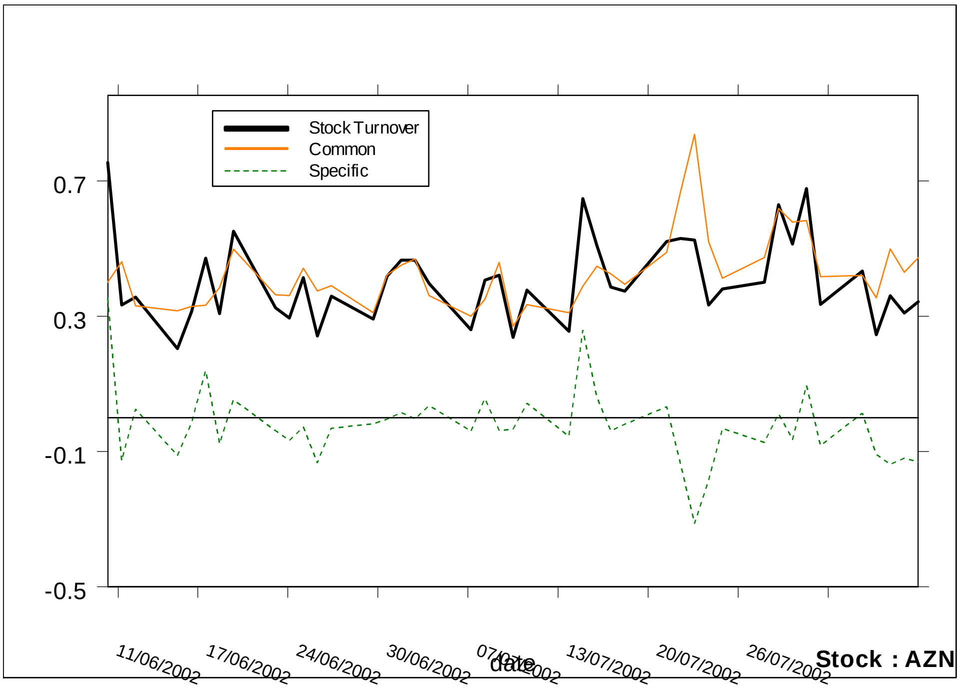

Figure 11: Turnover and turnover components, AZN stock.

particular drug[6]. During that period, generics were promised access the market very soon as the patent was expiring. The impending competition with their leading product was inducing a potential loss in AstraZeneca future earnings. Uncertainty was in favour of arbitrage strategies which increased greatly. This example is obviously not only a problem of liquidity, but rather a problem of information and uncertainty which creates in turn liquidity problems. It shows the need to incorporate information in the analysis when studying liquidity as information is an important source of liquidity variation.

6 Conclusion

In this paper, we propose a simple in which investors trade for hedging motives. Because they do not rebalance their portfolio at once, a liquidity problem arise creating a temporary disequilibrium and therefore a change in stock prices. This price variation is a signal to a new class of traders, called liquidity arbitrageurs, to enter the market to provide the missing liquidity. We show, that the volume they trade adds to the volume that would be

[6]Their leading product represents 50 % of their earnings in 2001.

traded if there were no imperfection - the "*normal volume*" . Moreover, this volume appears to be proportional to the price volatility.

Our model justify the strategies of a new generation of traders and illustrates that they can be identified through the analysis of volume. We show that hedging trades are common across assets whereas liquidity trades are asset specific.

We propose a one factor empirical model based on the theoretical model. We decompose the trading volume using a principal component analysis. The specific component of volume is a measure of stock trading volume corrected from trend, seasonalities and data errors. As such, it represents a valuable tool for investors. It can be seen as a preprocessing tool that allows market practitioners to extract information from trading volume time series. Moreover, we can now work on intraday data without fearing the strong intraday seasonalities.

One objection of PCA is that it does not exploit the dynamic of the data. However, since the theoretical model is static, PCA accurately works in implementing and testing it. Natural extensions from static to dynamics exist and include methods such as dynamic factor analysis [Harvey (1989)] or Min/Max Autocorrelation Factorial Analysis (MAFA) [see Box and Tiao (1977), Pena and Box (1987), Solow (1994)]. But the use of such extensions imposes the complete specification of the components dynamics. As discussed in section 3.4, in our memoryless world, the stock specific components are stationnary while the common component dynamics is exogenous. Hence, any choice of dynamics could lead to specification error.

Besides, the use of such measure of volume should lead to a more accurate analysis of the volume-volatility relation. In fact, if there is a relation between volume and volatility on one stock, the volume and volatility measures to consider must only reflect stock-specific volume and volatility components. This is a challenging task that we will undertake in the future.

Bibliography

Andersen T., « Return Volatility and Trading Volume : An Information Flow Interpretation », Journal of Finance, **51**, 1996, 169-204.

Box G.E.P. and TIAO G.C., « A Canonical Analysis of a Multiple Analysis », Biometrica, **70**, 1977, 57-65.

Bamber L., « The Information Content of Annual Earnings Releases: A Trading Volume Approach », Journal of Accounting Research, **24**, 1986, 40-56.

Bamber L., « Unexpected Earnings, Firm Size, and Trading Volume Around Quarterly Earnings Announcements », The Accounting Review, 1987, 510-532.

Boyer C. and Le Fol G., « Order Book Dynamics », Working paper, MSE - Paris I University - Panthéon - Sorbonne, **33**, 1999.

Campbell J., Grossman S., and Wang J., « Trading Volume and Serial Correlation in Stock Returns », Quarterly of Journal of Economics, **108**, 1993, 905-939.

Cass D., and Stiglitz J., « The Structure of Investor Preferences and Asset Returns, and Separability on Portfolio Allocation: A Contribution to the Pure Theory of Mutual Funds », Journal of Economic Theory, **2**, 1983, 122-160.

Cremers K.J.M. and Mei J, « Testing the Duo-factor Model of Return and Volume », , Working paper New-York University, 2001.

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

Easley, O'Hara, « Price, Trade Size, and Information in Securities Markets », Journal of Financial Economics, **19**, 1987, 69-90.

Epps T. and Epps T., « The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture of Distribution Hypothesis », Econometrica, **44**, 1976, 305-321.

Foster F. D. and S.Viswanathan « A Theory of Interday Variations in Volumes, Variances and Trading Costs in Securities Markets », Review of Financial Studies, 1990, 595-624.

Foster F. D. and S.Viswanathan « Variations in Trading Volume, Return Volatility and Trading Costs: Evidence on Recent Price Formation Models», Journal of Finance, **48,** 1993, 187-211.

Gallant R., Rossi P. and Tauchen G., « Stock Prices and Volume », Review of Financial Studies, **5**, 1992, 199-242.

Gouriéroux C. and Le Fol G., « Effet des Modes de Négociation sur les Echanges », Revue Economique, **49**, 1998, 795-808.

Harvey A.C., « Forcasting structuring time series models and the Kalman filter », 1989, Book, (Cambridge University Press).

Hasbrouck J. and Seppi D.J., « Common Factors in Prices Order Flows and Liquidity », Journal of Financial Economics, **59**, 2001, 383-411.

He X., Velu R. and Chen C., « Commonality, Information and Return/Return Volatility - Volume Relationship, Working Paper Syracuse University, 2003.

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

Karpoff J., « The Relation between Price Changes and Trading Volume: A Survey », Journal of Financial and Quantitative Analysis, 22, 1987, 109-126.

Lakonishok J. and Smidt S., « Volume for Winners and Losers: Taxation and Other Motives for Stock Trading », Journal of Finance, **41**, 1986, 951-974.

LeBaron B., « Persistence of the Dow Jones Index on Rising Volume », working paper, University of Wisconsin, 1992.

Llorente G, R. Michaely, G. Saar and J. Wang, « Dynamic Volume-Return Relation of Intividual Stocks », Review of Financial Studies, **15**, No. 4, 2002, 1005-1047.

Lo A. and Wang J., « Trading Volume: Definition, Data Analysis, and Implication of Portfolio Theory », The Review of Financial Studies, **13**, 2000, 257-300.

Lo A. and Wang J., « Trading Volume: Implications of an Intertemporal Capital Pricing Model», Working Paper, MIT, 2001.

Markovitz H., « Portfolio Selection », Journal of Finance, **7**, 1952, 77-91.

Merton R., « An Intertemporal Capital Asset Pricing Model », Econometrica, **41**, 1973, 867-887.

Morse D., « Asymmetric Information in Securities Markets and Trading Volume », Journal of Financial and Quantitative Analysis, **15**, 1980, 1129-1148.

Pena D. and Box G.E.P., « Identifying a simplifying structure in time series », Journal of the American statistical Association, **82**, 836-843.

Richardson G., Sefcik S. and Thompson R., « A Test of Dividend Irrelevance Using Volume Reaction to a Change in Dividend Policy », Journal of Financial Economics, **17**, 1986, 313-333.

Ross S., « Mutual Fund Separation in Financial Theory – The Separating Distributions », Journal of Economic Theory, **17**, 1978, 254-286.

Smidt S., « Long Run Trends in Equity Turnover », Journal of Portfolio Management, Fall, 1990, 66-73.

Solow, A.R. , « Detecting Change in the Composition of a Multispecies Community », Biometrics, **50**, 1994, 556-565.

Stickel S., and Verrechia R., « Evidence that Volume Sustains Price Changes », Financial Analysts Journal, 1994, 57-67.

Tauchen G. and Pitts M., « The Price Variability-Volume Relationship on Speculative Markets », Econometrica, **51**, 1983, 485-506.

Tobin J., « Liquidity Preference as behaviour Towards Risk », Review of Economic Studies, **25**, 1958, 68-85.

Wang J, « A Model of Competitive Stock Trading Volume », Journal of Political Economy, **102**, No.1, 1994, 127-168.

Ying C., « Stock Market Prices and Volume of Sales », Econometrica, **34**, 1966, 676-686.