Characterizing Commonality on NYSE Euronext Brussels

Catherine D'HONDT
Professor of Finance, FUCaM & LouvainSchool of Management

Christophe MAJOIS
Postdoctoral researcher – BNP Paribas Fortis Chair, Louvain School of Management

Based on intra-daily data for 74 continuously-traded stocks (including those belonging to the BEL20 index), we provide evidence of common variations in returns, order flow and liquidity on NYSE Euronext Brussels. We find that commonality in order flow is the strongest for principal trading. We also show that the opening of US markets plays a key role as commonality reaches its highest value at that time.

Introduction

Commonality refers to cross-security comovements in liquidity, returns, or order flow, and that appear within or across markets. Such common variations are important to consider, as they represent non-diversifiable sources of risk that portfolio managers cannot avoid. Gaining insight into their determinants will help better manage this risk. A growing body of literature documents the existence of commonality but empirical evidence on what really drives it is still scarce. The traditional view relates it to comovements in underlying fundamentals. However, from inconclusive empirical evidence have emerged friction-based theories[1] which relate commonality to correlated decisions within specific groups of investors (retail traders, institutional traders, mutual funds, etc.).

In this paper, we aim at characterizing commonality on NYSE Euronext Brussels. For this purpose, we base our approach on both Hasbrouck & Seppi (2001) and Corwin & Lipson (2009). Hasbrouck & Seppi (2001) use principal components analysis to address commonality for the 30 Dow Jones stocks. They find that both returns and order flow are characterized by common factors and show that commonality in order flow explains about two-thirds of common factors in returns. The authors also report comovements in liquidity proxies but of weaker significance. Building on Hasbrouck & Seppi (2001), Corwin & Lipson (2009) investigate the relative importance of different types of traders in driving commonality. Using electronic order flow data for a sample of 100 NYSE-listed stocks, they report that common factors in several measures of order flow, liquidity and returns are driven by program trading and other institutional trading, retail traders playing only a small role.

In line with the aforementioned papers, we compute principal components to address the magnitude and the sources of commonality on Belgian stocks. Using a unique dataset over a three-month period, our purpose is threefold. First, we document the presence of commonality in returns, liquidity, and order flow, for the 74 stocks in our sample. Second, our dataset includes trader type identification that allows us for each market member not only to distinguish between *principal* and *agency* trading but also to characterize trading as *domestic* or *foreign* based on the market member's location. These dataset features give us the opportunity to check whether common factors are primarily driven by principal trading, and most importantly, to assess the extent to which market members' location matters in driving commonality. Third, we investigate how commonality evolves throughout the trading day.

The remainder of the paper is structured as follows. Section 1 provides a short review of the extant literature on commonality

1. A review of these friction-based theories is available in Barberis *et al.* (2005).

and its determinants. The next section describes our dataset and the sample characteristics. The methodology we applied is briefly presented in Section 3. We report our empirical analysis and its results in Section 4. Section 5 concludes.

1. Literature

The emphasis from individual firms to global or market-wide determinants of liquidity has been shifted with Chordia *et al.* (2000). For a sample of 1169 NYSE stocks on 254 trading days, these authors estimate a market model for different liquidity measures and obtain significant positive betas for a majority of stocks, even after controlling for individual determinants of liquidity. Since then, a growing body of literature has brought other evidence that a firm's liquidity is sensitive to liquidity aggregated at the market-level, the industry-level, the index-level, or even at the global-level. We shortly review the main results hereafter.

Huberman & Halka (2001) analyze 254 daily observations for 240 NYSE stocks and document the presence of a systematic, time-varying component of liquidity. Hasbrouck & Seppi (2001) extend the search for commonality to both order flow and returns for the 30 Dow Jones stocks and report that commonality in order flow explains about two-thirds of commonality in returns. Extending the approach of Chordia *et al.* (2000), Brockman & Chung (2002) examine commonality in liquidity and order flow for all the stocks listed on the Stock Exchange of Hong Kong. They find a significant market-wide component in both liquidity and order flow as well as an industry-wide component but of smaller magnitude.

Domowitz *et al.* (2005) use simulations and Australian Stock Exchange data to show that order-type (*i.e.* market or limit orders) comovement leads to commonality in liquidity while order flow (*i.e.* order direction and size) comovement leads to commonality in returns. They conclude that some stocks can have little return correlation but high liquidity commonality, and vice versa. These authors also highlight the downward risk of liquidity commonality as it reveals itself more through deterioration than improvement across stocks. Harford & Kaul (2005) compare commonality in order flow, returns and trading costs for S&P and non-S&P stocks. Their findings emphasize the importance of indexing in driving commonality in order flow and returns. Kamara *et al.* (2008) study the evolution of liquidity commonality in the cross-section of US stocks from 1963 through 2005 and find that commonality has significantly increased for large firms but declined for small firms. These authors directly relate this result to the prevalence of both institutional investing and index trading in large stocks. Furthermore, they show that market volatility, market return, and market liquidity affect stock systematic liquidity as well as systematic return. Brockman *et al.* (2009) analyze intraday spreads and depths on a sample of 47 exchanges across the globe and document that local sources of commonality represent about 39% of the firm's total commonality while global sources account for an additional 19%.

Most of the time, the above empirical findings are consistent with friction-based theories, suggesting that commonality is driven by correlated trading decisions within specific groups of investors (retail traders, institutional traders, mutual funds, etc.). However, studies that attempt to assess the link between commonality and decisions within different groups of traders are not numerous. This scarcity is due to the difficulty of directly attributing trading activity (trades or orders) to a specific group of investors. Such piece of information is often not available in public market data and having access to data from brokerage firms is not always possible. Our paper contributes to the extant literature on this aspect, as do Corwin & Lipson (2009).

In a recent working paper, Corwin & Lipson (2009) investigate the relative importance of different types of traders in driving commonality. Using electronic order flow data for NYSE-listed stocks, they are able to estimate the common factors in various categories of order flow that are identified by account type, namely program traders, non-program institutional traders, retail traders and exchange members. The authors report that order flow common factors are significantly related to commonality in returns and are primarily driven by program traders. Furthermore, commonality in professional order flow provides incremental explanatory power for returns on large firms, thereby confirming the predominant role of professional traders in large stocks' activity.

In this paper, we follow the approach of Corwin & Lipson (2009) to address the magnitude and the sources of commonality on Belgian stocks.

2. Data

Our sample contains 74 Belgian stocks for the period from February through April 2006 (61 trading days). All these stocks are traded in continuous and have at least one trade per trading day. Among them, 18 belong to the BEL20 index during the three-month period.

For each stock, our dataset includes public market data about both orders and trades (price, total size, hidden size, submission/execution time-stamp, etc.) as well as additional private information such as market members' identification and the account for which the market member is trading. Indeed, NYSE Euronext market members can act either as brokers when submitting orders on behalf of clients (agency trading), or as dealers when trading for their own account (principal trading)[2]. Furthermore, the market member's identification allows us to characterize trading as domestic or foreign based on the market member's location[3]. Finally, this rich dataset about orders and traders allowed us to reprogram the continuous market algorithm and get accurate order book data[4].

Summary statistics for the sample are provided in Table 1, respectively for the 74 stocks (All), the BEL20 stocks (BEL20), and the 56 non-BEL20 stocks (Others). Cross-section means and medians are reported for market capitalization (estimated on February 1, 2006), price, number of shares per trade, daily number of shares traded, and daily number of trades. These descriptive statistics unsurprisingly highlight the strong difference in trading activity between the BEL20 stocks and the others.

Table 1: Summary Statistics

		All	BEL20	Others
Market Capitalisation (€ million)				
	mean	3.611	11.721	1.005
	median	512	5.655	360
Price				
	mean	112,48	65,33	127,63
	median	38,50	59,68	29,58
Volume/trade (#shares)				
	mean	496	372	536
	median	268	285	260
Daily volume (#shares)				
	mean	132.039	434.827	34.715
	median	20.940	212.032	11.707
Daily number of trades				
	mean	294	939	87
	median	75	848	47

3. Methodology

Following Hasbrouck & Seppi (2001) and Corwin & Lipson (2009), we use the principal components analysis (hereafter PCA) to extract and analyze common factors in returns, order flow, and liquidity. PCA is basically a descriptive method that relies on the variance-covariance structure of the original variables to create new variables that should better represent the original dataset. This method has the advantage that it does not require any distributional assumption on the data, and that it does not impose any structure on the common components.

If our sample contains N stocks and T periods, then the original dataset can be represented by a matrix X with T lines and N columns: $X = [X_1; X_2; \ldots; X_N]$ where X_i is a vector of length T. For example, X_1 represents the return computed for the first stock of our sample. The length T of each vector depends on the sampling frequency chosen. If we compute only one value per trading day, then T will be equal to 61 as our sample period counts 61 trading days. If we compute 2 values per day, then T will equal 122, etc. The objective of PCA is to create new variables that we denote F_j, which are called "principal components". The number of principal components equals the number of original variables, N in our case. So we get a new matrix $F = [F_1; F_2; \ldots; F_N]$[5].

The principal components are ranked according to their explanation power. The first principal component is built in such a way that it maximizes its explanation power, *i.e.* that it maximizes the variance along its axis. The second principal component has the second highest variance, that is, it explains the maximum of what is not explained by the first component, and so on. The importance of a principal factor is thus represented by its variance, which we usually denote by the greek letter lambda. λ_1 corresponds to the variance of the first principal factor, λ_2 equals the variance of the second principal factor with $\lambda_1 > \lambda_2$, etc. The total variance of the new dataset is exactly equal to the total variance of the original dataset. The importance of a principal factor F_j is then measured through the proportion of total variance that it explains, that is by:

$$\frac{\lambda_j}{\sum_{j=1}^{N} \lambda_j}$$

It is usual in PCA to standardize the original variables. Given that each of the standardized variables has a variance equal to 1, the total variance of the dataset equals the total number of variable, N. The proportion of variance

2. For some illiquid stocks, they can also act as liquidity providers. We drop from our analysis all the orders and trades referring to these liquidity provision agreements.
3. Since our sample contains only Belgian stocks, this characterization is quite straightforward: when the market member is located in Belgium, trading is flagged as domestic.
4. Based on both public and private data, we developed a program rebuilding the limit order book second by second. Within this program, the state of the order book is updated whenever a new order is submitted or a standing limit order is modified or cancelled.
5. Mathematically, the principal components are obtained through the extraction of eigenvalues and eigenvectors from the variance-covariance matrix of the original dataset.

explained by a principal factor F_j then becomes $\frac{\lambda_j}{N}$. If data were obtained randomly and the principal factors would not give new information, their variances would all equal unity. Hence, $\frac{1}{N}$ is a natural point of comparison. For instance, when we focus on the BEL20 stocks, the original table has 18 variables, one for each stock. A principal component F_j will be relevant only if its variance represents more than 1/18 = 5,56% of the total variance. That value comes down to 1,35% when we consider all the 74 Belgian stocks of our sample. In the following, we consider at most the first three principal components, as usually done in the literature.

In order to provide an interpretation to the principal components, we also look at the "factor loadings", *i.e.* the correlations between the new variables F_j and the original variables X_i. Accordingly, commonality implies factor loadings that all have the same sign: it means that all original variables (all the stocks in our case) evolve in the same direction, which is indeed what commonality refers to.

4. Empirical Results

We next proceed to the empirical results. A PCA is applied to every variable of interest. In a first step, we compute values for the variables at the hourly interval. Given that trading starts at 9:00 and ends at 17:25, this means that we have 9 observations per day[6] and the number of lines T equals: 61 trading days x 9 observations per day = 549.

To analyze commonality in returns, we focus on the midpoint return, *i.e.* the return computed on quote midpoints at the end of each interval[7].

We consider several variables related to the order flow. We focus either on trades or on orders submitted, and also consider the number of shares implied in those trades/orders. Imbalances are computed as: (buy-sell)/(buy+sell).

Finally, liquidity is assessed on two dimensions. Tightness is measured by the relative spread computed either at the best quote or at the five best quotes. Depth is measured by the sum of shares displayed at the five best quotes on both sides of the order book.

(a) Commonality on Returns, Order Flow, and Liquidity

We begin by analyzing the importance of common factors in returns, order flow and trading variables. Results are provided in Table 2.

Table 2: PCA for Returns, Order Flow and Liquidity Variables

Variables		BEL20 (1/18=5,56%)				All (1/74=1,35%)			
		1	2	3	Σ	1	2	3	Σ
Return	midpoint return	12,83%	7,67%	6,78%	27,28%	4,03%	2,70%	2,53%	9,26%
Order flow									
Trades	# trades	16,78%	8,67%	7,56%	33,00%	5,93%	5,15%	3,64%	14,72%
	# shares	12,00	8,28%	7,50%	27,78%	4,46%	3,46%	3,15%	11,07%
	imbalance (# trades)	13,06%	7,22%	6,50%	26,78%	3,86%	2,84%	2,51%	9,22%
	imbalance (# shares)	9,28%	7,44%	7,22%	23,94%	3,12%	2,61%	2,36%	8,09%
Orders	# orders	14,39%	11,06%	7,28%	32,72%	6,85%	4,91%	4,35%	16,11%
	# shares	11,22%	9,00%	8,06%	28,28%	4,34%	3,92%	3,61%	11,86%
	imbalance (# orders)	12,33%	8,67%	7,44%	28,44%	4,84%	3,77%	3,31%	11,92%
	imbalance (# shares)	7,56%	7,39%	7,11%	22,06%	3,04%	2,51%	2,46%	8,01%
Liquidity									
Relative Spread	at the BQ	8,89%	8,33%	7,50%	24,72%	4,14%	3,41%	3,07%	10,61%
	at the 5 BQ	11,61%	8,56%	7,89%	28,06%	5,77%	4,11%	3,96%	13,84%
Displayed Depth	at the 5 BQ	11,61%	10,67%	8,00%	30,28%	6,49%	4,72%	4,11%	15,31%
	imbalance at the 5 BQ	9,39%	8,33%	7,50%	25,22%	5,46%	4,91%	4,46%	14,82%

For each variable, the table lists the proportion of total variance explained by the first three principal components. Under the assumption of no commonality, the variance explained by each component would equal 5,56% for the BEL20 stocks and 1,35% for the whole sample. All variables were calculated at one-hour intervals from 9:00 a.m. through 17:25 p.m., and were standardized using firm and time-of-day specific means and standard deviations. Returns are defined with quote midpoints at the end of each interval. Order flow imbalances are defined in number of trades or orders as well as in number of shares. In all cases, they are computed following the approach $\frac{\text{Buy} - \text{Sell}}{\text{Buy} + \text{Sell}}$. The relative spread at the 5 best quotes is defined as $\frac{\text{Ask}^5 - \text{Bid}^5}{\frac{(\text{Ask}^5 + \text{Bid}^5)}{2}}$. Displayed depth at the 5 best quotes is the sum of displayed volume available at the 5 best quotes. Imbalances at the 5 best quotes are defined as buy volume minus sell volume divided by total volume.

6. Note that the last interval actually only covers a slightly less than an half hour period, between 17 and 17:25.
7. Since we exclude inter-day returns, we lose one observation per day.

Strong evidence of common components appears for all variables.

For returns, the first common factor explains 12,83% of total variation for the BEL20 stocks and 4,03% for the whole sample. This may give the impression that commonality is higher for BEL20 stocks than for all Belgian stocks as a whole, but remember that the right point of comparison is the 1/N value, which is 5,56% for the BEL20 stocks and 1,35% for the whole sample. The values for the first component are clearly bigger than those cutoff values for both samples. We highlight that there is a clear gap between the first component and the second and third components. Although this result is not shown in the table, we can add that the factor loadings are positive for all stocks in the case of the BEL20 sample, and for almost all stocks for the larger sample. The first component thus provides evidence of real commonality among stocks.

The results for order flow variables provide even stronger evidence of commonality. While present for all measures, commonality appears stronger for trades than for orders. Comovements are the strongest for the number of trades or the number of orders while they appear weaker for imbalances. For the BEL20 stocks, the first principal component explains up to 16,78% (14,39%) of total variation in the number of trades (orders).

As for liquidity, Table 2 also provides evidence of common components in all measures but, consistent with Corwin & Lipson (2009), it is less strong than for returns or order flow. The strength of commonality seems to increase as more information from the order book is incorporated. The highest proportion of variance explained by the first principal component is for the relative spread at the 5 best quotes and for displayed depth at the 5 best quotes. This result may however be misleading. For the BEL20 stocks, factor loadings are positive for 17 out of 18 stocks in the case of the spread at the best quote, while they are positive for only 13 out of 18 stocks when we consider the spread at the 5 best quotes. This phenomenon is similar for depth measures. Consequently, commonality, in the sense that variables evolve in the same direction for all stocks, concentrates at the best quote of the order book.

(b) Commonality on Order Flow: Principal vs. Agency Trading and Domestic vs. Foreign Trading

We have shown above the existence of commonality among Belgian stocks for returns, order flow and, to a lesser extent, for liquidity. Prior research suggests that commonality may be driven by correlated trading decisions within specific groups of investors. For a sample of NYSE-listed stocks, Corwin & Lipson (2009) find that comovements in order flow are mainly driven by program trading and other institutional trading. To test this phenomenon in our sample, we estimate principal components for order flow variables defined separately for principal and agency trading. Furthermore, we also compute principal components for order flow variables when making a distinction between domestic and foreign trading. As mentioned earlier, the purpose is to assess whether market members' location affects comovements in order flow.

For each order flow variable and the BEL20 stocks[8], Table 3 reports the total variation explained by the first principal component.

Table 3: PCA for Principal vs. Agency Trading and Domestic vs. Foreign Trading

Variables		BEL20 (1/18=5,56%)			
		Principal	Agency	Domestic	Foreign
Trades	# trades	18,83%	12,44%	15,61%	16,83%
	# shares	12,22%	10,00%	12,11%	12,06%
	imbalance (# trades)	12,89%	8,11%	-	-
	imbalance (# shares)	9,28%	7,61%	-	-
Orders	# orders	16,33%	12,94%	17,00%	14,11%
	# shares	11,78%	11,89%	9,72%	11,83%
	imbalance (# orders)	13,06%	9,94%	15,22%	11,11%
	imbalance (# shares)	8,94%	7,89%	9,17%	8,28%

For each variable, the table lists the proportion of total variance explained by the first principal component. Under the assumption of no commonality, this proportion would equal 5,56%. All variables were calculated at one-hour intervals from 9:00 a.m. through 17:25 p.m., and were standardized using firm and time-of-day specific means and standard deviations. Imbalances are defined in number of trades or orders as well as in number of shares. In all cases, they are computed following the approach $\frac{\text{Buy} - \text{Sell}}{\text{Buy} + \text{Sell}}$. Principal (agency) trading refers to orders or trades initiated by market members for their own account (on behalf of clients). Domestic (foreign) trading refers to orders or trades initiated by market members who are located in Belgium (abroad).

From Table 3 (and in comparison with Table 2), it is clear that commonality is the strongest for order flow from principal trading. For the number of principal trades, the first principal component explains up to almost 19% of total variation, while this proportion is about 12% for agency trades. A slightly weaker difference is observed between the number of principal and agency orders. These findings are consistent with Corwin & Lipson (2009) and

8. Results for the whole sample are quite similar.

suggest that commonality in order flow is mainly driven by professional traders.

As for the distinction between domestic and foreign trading, Table 3 provides results for which interpretation seems less straightforward. When focusing on variables built on trades, no obvious difference appears between domestic and foreign trading. For orders-based variables, commonality seems to be the strongest for domestic trading when we consider the number of orders submitted, but the opposite seems to be true for the number of shares submitted. The results are therefore somewhat inconclusive.

(c) Intraday Commonality

In a final step, we examine how commonality evolves over the trading day. To our knowledge, this has never been done before in the literature. For this purpose, we compute variables at a higher frequency for the BEL20 stocks only[9]. We divide the trading day in 17 periods of 30 minutes and apply a PCA for each period. Although we performed the analysis for all the variables considered in Table 2, we only present the results for a subset of them.

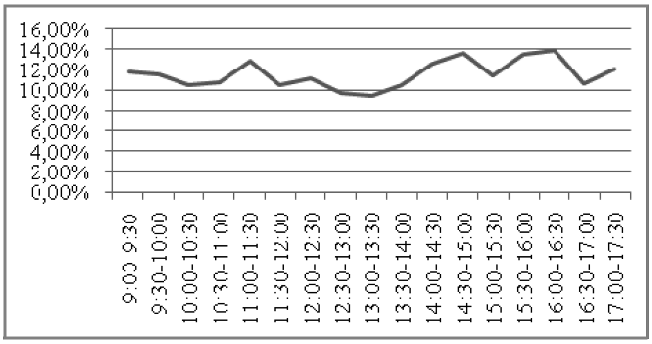
Figure 1 displays the intraday evolution of the variance explained by the first principal component for midpoint returns. We do not discern any time-specific pattern, as this value is very stable over the entire trading day.

Figure 2 provides interesting results for the number of trades initiated as principal or as agent. The variance explained by the first principal component is very stable over time for agency trades, at around 12%. For principal trades, the value also fluctuates around 12% most of the time, but we observe a first peak at 30% between 12:30 and 13:00 and, most strikingly, a second peak at 75% between 15:30 and 16:00. This last period corresponds to the opening of US markets and leads to important comovements for all the BEL20 stocks.

Looking at Figure 3, where we distinguish between domestic and foreign trading, we also observe a peak at the opening of US markets. Both series display the peak, but it is more pronounced for domestic trading. The opening of US markets seems to be a key driver of commonality, especially for domestic trading.

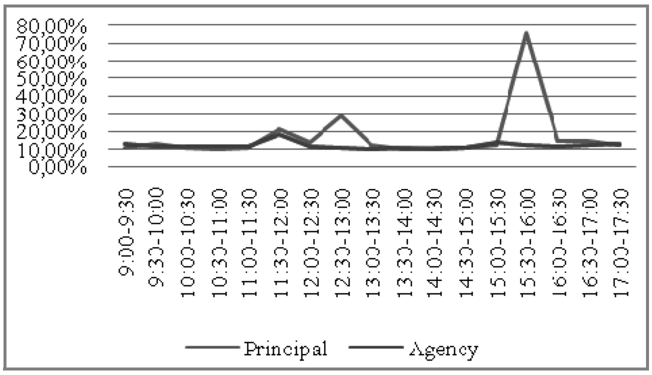
Finally, we provide in Figure 4 the results for two liquidity variables. Similarly to what we observe for midpoint returns, no intra-day pattern seems to emerge from this figure.

Figure 1: Intraday PCA for Midpoint Return



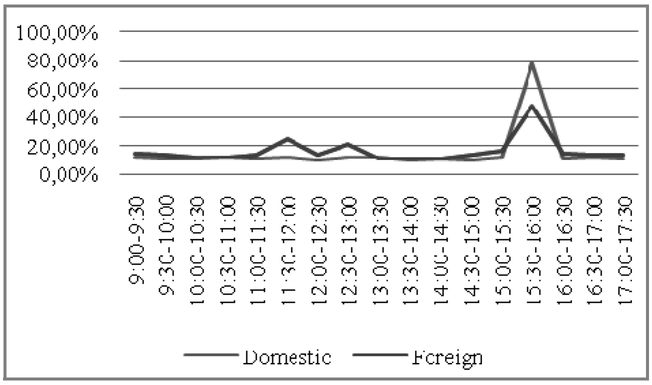
For each 30-minute interval, this figure plots the proportion of total variance explained by the first principal component. Under the assumption of no commonality, this proportion would equal 5,56%. Returns are defined with quote midpoints at the end of each interval, and were standardized using firm and time-of-day specific means and standard deviations.

Figure 2: Intraday PCA for # Trades in Principal vs. Agency Trading



For each 30-minute interval, this figure plots the proportion of total variance explained by the first principal component. Under the assumption of no commonality, this proportion would equal 5,56%. The number of trades was standardized using firm and time-of-day specific means and standard deviations. Principal (agency) trading refers to trades initiated by market members for their own account (on behalf of clients).

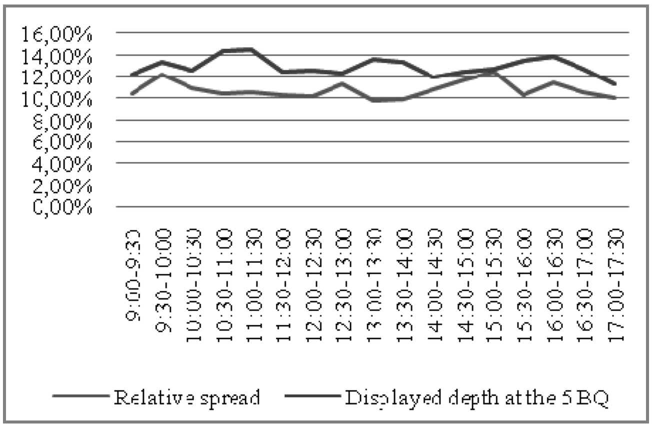
Figure 3: Intraday PCA for # Trades in Domestic vs. Foreign Trading



For each 30-minute interval, this figure plots the proportion of total variance explained by the first principal component. Under the assumption of no commonality, this proportion would equal 5,56%. The number of trades was standardized using firm and time-of-day specific means and standard deviations. Domestic (foreign) trading refers to trades initiated by market members who are located in Belgium (abroad).

9. Other stocks are less liquid and could not provide reliable results.

Figure 4: Intraday PCA for Relative Spread and Displayed Depth



For each 30-minute interval, this figure plots the proportion of total variance explained by the first principal component. Under the assumption of no commonality, this proportion would equal 5,56%. Relative spread is defined as $\frac{\text{Ask} - \text{Bid}}{\frac{(\text{Ask} + \text{Bid})}{2}}$. Displayed depth at the 5 BQ is the sum of displayed volume available at the 5 best quotes. Both variables were standardized using firm and time-of-day specific means and standard deviations.

5. Concluding Remarks

In this paper, we have characterized commonality in returns, order flow, and liquidity, for two samples of stocks traded on NYSE Euronext Brussels. Common variations are important to analyze, as they represent non-diversifiable sources of risk that portfolio managers cannot avoid. We have shown that there exists commonality in returns and order flow, and to a lesser extent, in liquidity for both samples. Looking at detailed data, we found that commonality in order flow mainly comes from principal trading. A focus on the intraday evolution of commonality has revealed that the opening of US markets is a key moment, where commonality may be up to 6 times bigger than during the rest of the day.

References

BARBERIS, N., SCHLEIFER, A. & WURGLER, J., 2005, "Comovement", *Journal of Financial Economics* 75, 283-317.

BROCKMAN, P. & CHUNG, D., 2002, "Commonality in liquidity: Evidence from an order-driven market structure", *Journal of Financial Research* 25(4), 521-539.

BROCKMAN, P., CHUNG, D. & PÉRIGNON, C., 2009, "Commonality in Liquidity: A Global Perspective", *Journal of Financial and Quantitative Analysis* 44, 851-882.

CHORDIA, T., ROLL, R. & SUBRAHMANYAN, A., 2000, "Commonality in liquidity", *Journal of Financial Economics* 56, 3-28.

CORWIN, S. A. & LIPSON, M. L., 2009, Order characteristics and the sources of commonality in prices and liquidity, Working Paper.

DOMOWITZ, I., HANSCH, O. & WANG, X., 2005, "Liquidity commonality and returncomovement", *Journal of Financial Markets* 8, 351-376.

HARFORD, J. & KAUL, A., 2005, "Correlated order flow: Pervasiveness, sources, and pricing effects", *Journal of Financial and Quantative Analysis* 40, 29-55.

HASBROUCK, J. & SEPPI, D., 2001, "Common factors in prices, order flows and liquidity", *Journal of Financial Economics* 59, 383-411.

HUBERMAN, G. & HALKA, D., 2001, "Systematic liquidity", *Journal of Financial Research* 14(2), 161-178.

KAMARA, A., LOU, X. & SADKA, R., 2008, "The divergence of liquidity commonality in the cross-section of stocks", *Journal of Financial Economics* 89, 444-466.