

# Short-, Long- and Cross-Term Comovement of OMXH25 Stocks

Agnieszka Jach and Karl Felixson

## Abstract

Using novel, nonparametric comovement measures based on the Thick Pen Transform, we study the OMXH25 stocks in the post-financial-crisis decade. The new measures allow us to work with stationary returns and with nonstationary volumes. The comovement can be monitored in time, it is possible to distinguish between comovement on different time scales, and even cross-term comovement can be quantified. The approach is visually-interpretable and multivariate in nature. The results indicate the presence of a cyclical pattern in the relatively strong comovement of returns on semi-annual and annual time scales, with more oscillations in the comovement on quarterly and monthly time scales, and the presence of a slight increasing pattern in the relatively weak comovement of volumes on semi-annual and annual time scales. Cross-term dependence between Nokia's weekly and monthly features in returns and longer-term features in returns of other stocks is more variable than that based on volumes.

## Keywords:

Comovements, time scale, codependence, time-varying, returns, volumes, Nasdaq Helsinki

---

Agnieszka Jach is an Associate Professor of Statistics at Hanken School of Economics, Finland.

Karl Felixson is an Associate Professor of Finance at Hanken School of Economics, Finland.

A. Jach gratefully acknowledges the financial support of the Finnish Foundation for Share Promotion (Börsstiftelsen).

## 1. Introduction

The decade of 2008-2017, which began with the global financial crisis, represents an interesting time period in the global and also in the Finnish economy. During this period, we observed the collapse of the financial and real-estate sectors in the US, numerous bank failures and bailouts in the US and Europe, the Eurozone crisis, unprecedented quantitative easing programs by central banks, wide-spread austerity, high unemployment rates, low and even negative interest rates (see e.g., International Monetary Fund, 2018; Valencia and Laeven, 2012). We also saw the booming economies of Brazil and China, which then de-accelerated. In the era of highly-integrated international stock markets, the consequences of these events were also echoed by the Finnish economy (e.g., a big drop in Finland's GDP in 2009, Statistics Finland, 2018), which in addition was coming to terms with a gradual decline in its pulp and paper industry as well as the sudden demise of Nokia's global dominance in the mobile phone industry.

In this article, we focus on the Finnish economy in the decade of 2008-2017 as seen through the performance and behaviour of its largest companies included in the OMX Helsinki 25 (OMXH25) index. We assess the trends by carrying out a detailed analysis of standard indicators such as log returns, bid-ask spreads, volumes of OMXH25 stocks, but also by examining the comovement of these stocks over time using daily data. The former can be used as direct measures of stock (and hence market) performance, while the comovement as an indirect measure of impact of global or country-specific events on the Finnish market. We contribute to the academic literature by empirically examining OMXH25 stocks' comovement with respect to different time scales and across time scales, dynamically. This study could be of potential interest to investors for the formulation and implementation of their investment and risk-management strategies. Our analysis could also provide new insights into the dynamic behaviour of market participants, and the novel comovement metrics we employ could assist in shaping regulation of securities' markets.

Our study is related to the work of Nyberg and Vaihekoski (2014), which contains summary statistics of monthly Finnish stock market returns and volatilities and of monthly trading volumes between 1913-2009 (grouped by year), and which uses rolling-window cross-correlation as a time-varying comovement metric between Finnish and US stock market returns. Our article also complements the large body of Finnish-based studies compiled in Kallunki et al. (1996). Among those studies, Bos et al. (1995) employ vector autoregression (VAR) to describe tri-variate comovement of US, Swedish and Finnish market returns, and a regression coefficient to quantify bivariate dependence of individual stocks' returns in Finland on US, Swedish and Finnish market returns. Martikainen et al. (1995) incorporate long-term perspective into the analysis of comovement of (nonstationary) prices of six size-sorted portfolios of Finnish stocks by performing a bivariate cointegration analysis. A more recent study of Graham and Nikkinen (2011) examines bivariate comovement dynamically and with respect to different time scales by using wavelet coherence, whereby Finnish stock market returns and volatilities are paired with their international counterparts (see also, Aguiar-Conraria and Soares, 2014 for a survey of wavelet-coherence applications, In and Kim, 2012 for a collection of wavelet applications in finance, and Percival and Walden, 2000 for an introduction to wavelets).

The comovement metrics previously used in connection with the Finnish stock market are ubiquitous in finance literature on comovement. Other approaches to measuring comovement, including contagion as its subclass, might involve 1) a different metric, e.g., tail-dependence coefficient, wavelet multiple cross-correlation (Fernandez-Macho, 2012; Fernandez-Macho, 2018), 2) a model that embeds the notion of comovement, e.g., GARCH-type (Engle, 2002),

copula (Rodriguez, 2007), regime switching (Gallo and Otranto, 2008), latent component (Berger and Pozzi, 2013), or 3) a combination of the two, e.g., Lehkonen and Heimonen (2014) use wavelet decomposition and DCC-GARCH-based correlation on international stock market returns, while Bu et al. (2019) employ Granger causality in conjunction with a network model on Chinese stocks' returns. Given that the literature on comovement is enormous, we do not attempt to provide a systematic overview. Comprehensive discussions of the comovement measures can be found e.g. in Karolyi and Stulz (1996), Forbes and Rigobon (2002), Kaminsky et al. (2003), Bae et al. (2003), Dungey et al. (2003).

In contrast to the previous comovement studies based on the Finnish markets, we work with individual stocks rather than stock index, which automatically requires a multivariate measure of comovement to avoid computing 300 bivariate comovement values for 25 entities. Resorting to a multivariate model such as VAR for more than 5 sequences is practically infeasible. We also consider nonstationary volumes as well as stationary returns, so a metric not constrained to stationary sequences is needed. Monitoring comovement in time and distinguishing between short-, medium-, long-term codependence calls for a wavelet-based approach. Nevertheless, wavelet-based metrics are either bivariate or time scales are limited to being dyadic (2 raised to an integer power). Likewise, cross-correlation of multi-period returns lacks the time-varying property, rolling-window cross-correlation loses its multi-scale perspective, it is difficult to consolidate the two paradigms, and in either scenario a large number of observations is needed. Finally, existing approaches do not account for cross-term comovement (Gorroitieta et al., 2019) while maintaining all the features mentioned above.

For these reasons, we analyze the comovement of OMXH25 stocks using a novel approach called the Thick Pen Transform (Fryzlewicz and Oh, 2011) and two metrics derived from it called the Thick Pen Measure of Association (Fryzlewicz and Oh, 2011) and multi-thickness Thick Pen Measure of Association (Jach, 2019). These metrics have a number of desirable properties. They are applicable to more than two time series, which need not be stationary. They are time-evolving, so that codependence can be monitored in time. The Thick Pen Measure of Association is capable of capturing comovement with respect to a given time scale, for a range of time scales from small to large. In addition, the multi-thickness Thick Pen Measure of Association is capable of quantifying codependence across different time scales, allowing one to cross-correlate short-term components of some time series with long-term components of others.

The questions that we are trying to address in this article can be summarized as follows. In the decade of 2008-2017, what was the general level of stock comovement with respect to daily, weekly, monthly, quarterly and annual time horizons on the Finnish market? Were there periods of time of high level of comovement with respect to those different time horizons? Are these periods linked with particular events? How do up-ward and down-ward shocks in a given stock propagate across the market and are they short- or long-lasting? We are able to address these questions by calculating the Thick Pen Measure of Association and the multi-thickness Thick Pen Measure of Association in various configurations.

The remainder of this paper is organized as follows. In Section 2, we describe the comovement metrics together with the necessary background about the Thick Pen Transform of Fryzlewicz and Oh (2011). In Section 3, we describe the data, define the relevant variables and present their summary statistics in a series of graphs. We quantify long-, short- and cross-term comovement of OMXH25 stocks based on returns and volumes in Section 4. In Section 5, we present several ways in which TPMA and MTPMA could be used in finance and then provide a short summary in Section 6.

## 2. Methodology

In this section, we give a brief overview of the Thick Pen Transform of Fryzlewicz and Oh (2011) which serves as a basis for two nonparametric comovement measures employed in this study: 1) Thick Pen Measure of Association of Fryzlewicz and Oh (2011), which can be used for monitoring comovement on a given time scale; 2) the extension of measure 1 called multi-thickness Thick Pen Measure of Association of Jach (2019), which is suitable for monitoring comovement across time scales.

We begin by introducing the notation for a generic time series  $X = (X_t)_{t=1}^T$ , e.g., a series of daily returns, and a set of thickness parameters  $\Lambda$ , i.e., a set of positive values denoted by  $\tau_i, i = 1, 2, \dots, |\Lambda|$  ( $|\Lambda|$  is the number of elements in  $\Lambda$ ). To lighten the notation, we will use  $\tau$  for one of the elements of  $\Lambda$  ( $\tau = \tau_i$  for some  $i$ ).

To analyze  $X$  using the Thick Pen Transform (TPT) means to plot  $X_t$  against  $t$  using pens of varying thickness  $\tau \in \Lambda$  and to collect all the boundaries produced in the drawing process, denoting the set of  $2 \cdot |\Lambda|$  sequences of length  $T$  by

$$TP_\Lambda(X) = \{(L_t^\tau(X), U_t^\tau(X))_{t=1}^T\}_{\tau \in \Lambda}. \quad (1)$$

The lower and upper boundaries of the area marked by a (square) pen of a given thickness appearing in equation (1) are defined as

$$\begin{aligned} L_t^\tau(X) &= \min(X_t, X_{t+1}, \dots, X_{t+\tau}), \\ U_t^\tau(X) &= \max(X_t, X_{t+1}, \dots, X_{t+\tau}). \end{aligned}$$

These quantities extract features of  $X$  associated with a time scale of (approximately)  $\tau$  units of time. This means that if we are interested in the short-term features of  $X$  (and later on, short-term comovement), we should include small values of  $\tau$ , e.g., one day,  $\tau=1$ , or one week,  $\tau=5$ , for daily data, into our analysis. Likewise, if we are interested in the long-term features of  $X$ , we should include larger values of  $\tau$ , e.g.,  $\tau=250$ . An example of the TPT for several choices of  $\tau$  is given in Figure 2A; different subplots correspond to different  $\tau$ 's,  $\tau \in \Lambda = \{1, 5, 20, 60, 120, 250\}$ , and different colours correspond to different OMXH25 stock tickers (the data will be introduced in the next section). Now, having several TPT-transformed time series such as those shown in Figure 2A, how can we quantify their comovement for a given time scale? It turns out that we can do this by measuring their relative overlap, in other words, by calculating the Thick Pen Measure of Association. Specifically, indexing the time series with  $k = 1, 2, \dots, K$ ,  $X^{(1)}, X^{(2)}, \dots, X^{(K)}$ ,  $X^{(k)} = (X_t^{(k)})_{t=1}^T$  and having their respective TPTs available, we calculate the Thick Pen Measure of Association (TPMA) of Fryzlewicz and Oh (2011) via

$$\rho_t^\tau(X^{(1)}, X^{(2)}, \dots, X^{(K)}) = \frac{\min_k(U_t^\tau(X^{(k)})) - \max_k(L_t^\tau(X^{(k)}))}{\max_k(U_t^\tau(X^{(k)})) - \min_k(L_t^\tau(X^{(k)}))}. \quad (2)$$

This metric is restricted to the interval  $(-1, 1]$  (the lower end-point is excluded) and is time-varying, so that the comovement can be monitored in time. It takes on negative values when there is a 'gap' between the areas marked by the TPT-transformed time series and equals zero when the areas touch each other (in the vertical direction). An example of the TPMA for a given thickness/time scale is provided in Figure 2B for  $K=23$  time series of returns referred to earlier.

Although TPMA allows us to quantify comovement of many time series with respect to a

given time scale (and monitor such comovement in time), it is not capable of assessing cross-scale comovement. For example, we might be interested in correlating short-term features of one series with long-term features of other time series, rather than limiting our comparisons to a single term, either short or long. To facilitate such cross-scale comparisons, Jach (2019) proposed an extension of the TPMA termed multi-thickness Thick Pen Measure of Association (MTTPMA) defined as

$$\rho_t^{(\tau^{(1)}, \tau^{(2)}, \dots, \tau^{(K)})}(X^{(1)}, X^{(2)}, \dots, X^{(K)}) = \frac{\min_k(U_t^{\tau^{(k)}}(X^{(k)})) - \max_k(L_t^{\tau^{(k)}}(X^{(k)}))}{\max_k(U_t^{\tau^{(k)}}(X^{(k)})) - \min_k(L_t^{\tau^{(k)}}(X^{(k)}))}, \quad (3)$$

for  $\tau^{(k)} \in \Lambda$ . The extension boils down to allowing each time series  $X^{(k)}$  to be transformed using its own thickness value  $\tau^{(k)}$  rather than transforming each time series using a common thickness  $\tau$ . This is incorporated into the notation by replacing the scalar superscript  $\tau$  on the left-hand side of equation (2) by a vector superscript  $(\tau^{(1)}, \tau^{(2)}, \dots, \tau^{(K)})$  on the left-hand side of equation (3), with appropriate superscript adjustments on the right-hand side.

### 3. Data and summary statistics

#### 3.1 Data

Data consists of daily historical records of 23 constituents of OMX Helsinki 25 index spanning the decade of 01.01.2008-31.12.2017. These 23 companies formed part of the index over the entire time period under consideration; their tickers are (in alphabetical order): AMEAS, CGCBV, ELISA, FORTUM, HUH1V, KCR, KESKOB, KNEBV, METSB, METSO, NDA1V, NESTE, NOKIA, NRE1V, ORNBV, OTE1V, OUT1V, SAMPO, STERV, TELIA1, UPM, WRT1V, YIT. The openly-available records (data can be downloaded from <http://www.nasdaqomxnordic.com/shares/historical-prices>) include: bid price (variable  $P^{bid}$ ), ask price ( $P^{ask}$ ), opening price, high price, low price, adjusted closing price ( $P$ ), average price, total volume (variable  $V$ ), turnover, and the number of trades. In addition to these, several other variables were computed to obtain an overall picture of the Finnish stock market in the decade of 2008-2017 with respect to stock returns, liquidity/illiquidity, and activity. These additional variables are the log return ( $R$ ), bid-ask spread ( $Spr = P^{ask} - P^{bid}$ ), and euro-volume ( $EV$ ), as well as the Amihud (2002) measure of illiquidity ( $Amihud$ ) and a recently-introduced measure of information asymmetry called (euro) volume coefficient of variation ( $EVCV$ ) of Lof and van Bommel (2018). To define these variables precisely, we introduce two types of subscript notation, using price  $P$  as an example, with similar convention for other variables: a) double subscript, where  $P_{s,t}$  denotes the closing price of stock,  $s = 1, 2, \dots, S$ ,  $S = 23$ , for the  $t$ th trading day over the entire time period,  $t = 0, 1, 2, \dots, T = 2516-1$ ; b) triple subscript, where  $P_{s,d,y}$  denotes the closing price of the stock  $s$ ,  $s = 1, 2, \dots, S = 23$ , for the  $d$ th trading day of year  $y$ ,  $d = 1, 2, \dots, D_y$  ( $D_y$  is the number of trading days in year  $y$ ,  $\approx 252$ ),  $y = 1, 2, \dots, 10$ , with  $y = 1$  corresponding to year 2008 (we will occasionally use  $y = 2008$  for convenience). Using the first notational system, the log-return for stock  $s$  and day  $t$  is defined as  $R_{s,t} = \log(P_{s,t}/P_{s,t-1})$ , for  $s = 1, 2, \dots, S$  and  $t = 1, 2, \dots, T$ , the bid-ask spread is  $Spr_{s,t} = P_{s,t}^{ask} - P_{s,t}^{bid}$  and the euro volume equals  $EV_{s,t} = P_{s,t} V_{s,t}$ . Amihud measure of illiquidity is the ratio of the absolute return to the euro volume,  $Amihud_{s,t} = |R_{s,t}| / EV_{s,t}$ . The second convention is needed to define the coefficient of variation (for euro volume), which is only given on the annual basis. For stock  $s$  and year  $y$ , this metric is given as the standard deviation of the euro volume (calculated over trading days falling into that particular year) divided by the mean of the euro volume (calculated over trading days falling into that particular year). In symbols,  $EVCV_{s,y} = \hat{\sigma}_{s,y} / \hat{\mu}_{s,y}$ , where  $\hat{\mu}_{s,y} = \frac{1}{D_y} \sum_{d=1}^{D_y} EV_{s,d,y}$  and  $\hat{\sigma}_{s,y}^2 = \frac{1}{D_y-1} \sum_{d=1}^{D_y} (EV_{s,d,y} - \hat{\mu}_{s,y})^2$ .

### 3.2 Summary statistics

Having introduced the variables of interest, we next present summary statistics for the daily return, volume, bid-ask spread, Amihud and of the annual coefficient of variation, tracking them over the years to give an overview of the Finnish stock market in the decade of 2008-2017. The statistics are presented graphically in Figures 1A-1E, with x-axis showing the year and y-axis giving the cross-sectional average of the temporal summary statistic in a given year  $\pm$  the standard error of the average. For example, based on the top left panel of Figure 1A, we see that in year 2008, the mean daily return  $\bar{R}_{s,2008} = \sum_{d=1}^{D_{2008}} (R_{s,d,2008}) / D_{2008}$  averaged out over the 23 stocks to  $-0.0033$ , give-or-take  $0.0003$ . In the same year, the standard deviation of the daily return  $\sqrt{\frac{1}{D_{2008}-1} \sum_{d=1}^{D_{2008}} (R_{s,d,2008} - \bar{R}_{s,2008})^2}$  (bottom left) averaged out to  $0.0357$  over the 23 stocks, give-or-take  $0.0013$ , while the average excess kurtosis of the daily return (bottom right) was  $3.0053 \pm 0.3524$ . Note that for the euro volume coefficient of variation (Figure 1E) there are no error bars ( $\pm$  standard error), because this metric is computed on the annual and not daily basis - for a given year, the summary statistics are calculated in the cross-section and not in time as for the other variables.

We begin the discussion of the summary statistics with the return in Figure 1A. We observe, on average, negative mean daily return in 2008 and 2011, the largest values in 2009 and 2010 and otherwise values just above or at zero (top left). Daily risk, measured by the standard deviation of the daily returns (bottom left) was on average highest in 2008, 2009, 2011. In 2017, it was the lowest. Short error bars in both plots indicate homogeneity in mean daily return and daily risk with respect to the stocks. Similar patterns can be observed in the second column, which is based on more robust measures of center and dispersion compared to those used in the first column (mean  $\Rightarrow$  median, sd  $\Rightarrow$  iqr). Looking at the average values for the daily maximum and minimum return (penultimate column) we observe a slight asymmetric effect (the daily max is, on average, between  $0.05$  and  $0.15$ , while the min between  $-0.16$  and  $-0.07$ ) and a narrowing range (the distance between the typical daily max and min) as we move through the years, with some deviations in 2010, 2015, 2016 at the lower extreme values. There is also a drop in the typical daily max value in 2017. The analysis of averages between 2008 and 2017 of the skewness in daily returns (top right) confirms these observations. For the average excess kurtosis of daily returns (bottom right) - excess kurtosis is a proxy for the thickness of the tail of the distribution of daily returns - we observe an increase from the average of about  $2.5$  to about  $10$ , with high level of heterogeneity (large error bars) among the stocks in 2013 and 2017. In the cross-section, skewness was also variable in those two years.

When it comes to the daily volume (Figure 3B), it is immediately apparent how much more the OMXH25 stocks differ from each other (large error bars) compared to when the non-volume variables are used to represent these stocks. A notable case is Nokia (results not shown), whose mean daily volume was about 18-35 times bigger than the average of the mean daily volumes of the remaining stocks, for all the years between 2008 and 2017. The heterogeneity in volumes is also partly due to the fact that the magnitude of this variable is much bigger than the magnitude of the non-volume variables. For the mean daily volume (top left), the averages for the years 2008-2017 are about 2 mln, and somewhat below that level when the median is used instead of the mean (second top). For the standard deviation in the daily volume (bottom left), the typical values are about 1.5 mln and about 1 mln when the iqr is employed instead of the sd to gauge the variability (second bottom). The maximum and the minimum daily volumes (penultimate column) are, on average, about 1.5 mln and 0.5 mln, respectively. For the daily maximum volume, there is a bigger spread among the stocks in 2013 and for the daily

minimum, this is the case in 2012. The average values for the skewness of the daily volume, the average values for the excess kurtosis of the daily volume, and the variability around those averages (right column) highlight years 2013 and 2014 as unusual (higher values and larger variability).

For the mean daily bid-ask spread (top left of Figure 1C), overall we report a declining trend in the average values. The average variation in the daily spread (bottom left) reflected by the average sd's also declined and so did the variability in the cross-section. These trends prevail to some extent in the median- and iqr-based subplots (second column). When it comes to the maximum daily spread (penultimate column, top), the averages for 2010-2012 of  $\approx 35$  cents are much higher than those for 2008 and 2009 ( $\approx 15$  cents), which in turn, are higher than those for the latter half of the decade 2008-2017 ( $\approx 5$  cents), with very similar values across all the stocks (extremely short error bars). The situation is quite different for the minimum daily spread (penultimate column, bottom). There is a declining trend from about 0.9 cents in 2008 to about 0.75 cents between 2011 and 2017 with an exception of 2015. From the last column, showing the typical skewness and excess kurtosis of the daily spread, it is easy to pick up 2011 and 2012, which are marked by higher skewness and excess kurtosis.

Illiquidity, proxied by the Amihud measure, is depicted in Figure 1D. The mean daily values (top left) were, on average, higher in 2008-2009, with a 'bump' in 2012, though below the 2008-2009 levels, after which the illiquidity has been declining. Similar pattern can be observed in the standard deviation of the daily Amihud values (bottom left), and likewise, in the median and the inter-quartile range (second column), but with lower values compared to the mean and the standard deviation. The maximum daily Amihud values (penultimate top) are, on average, also in line with the first four sets of results. On the other hand, the minimum daily Amihud values (penultimate bottom) were, on average, high and quite dispersed in 2008, 2009, 2011, 2012, 2016. The skewness and excess kurtosis of the daily Amihud values (last column), have been overall declining, with an exception of year 2010.

Euro volume coefficient of variation, which reflects information asymmetry, is shown in Figure 1E. In contrast to Figures 1A-1D, there are no error bars in this figure because EVCV is quoted on an annual basis and hence the summary statistics are calculated in cross-section and not in time. The mean EVCV (top left) was largest in 2008, 2009, 2013, 2014 indicating higher proportion of informed traders in those years compared to other years. The variability among the stocks' EVCV measured by the standard deviation (bottom left) was highest in 2008, 2009, 2014. The patterns are somewhat different when median and iqr are used (second column), with the last three years exhibiting the lowest values and decreasing tendency. There seem to be three extreme maximum values (penultimate top) in 2008, 2009, 2014, while skewness and excess kurtosis (right column) indicate that in those three years plus 2017 there were some stocks in the right tail of the distribution of the EVCV.

To sum up, the overall tendencies reported for the decade of 2008-2017, include: expected return stabilizing just above the zero level, decreasing risk (Figure 1A), decreasing illiquidity and information asymmetry (Figures 1D and 1E), somewhat decreasing volume and bid-ask spread (Figures 1B and 1C).

#### 4. Comovement of the OMX Helsinki 25 stocks

In this section, we present comovement results based on the methodology of Section 2. We restrict our analysis to two variables, return and volume. We begin by normalizing the sequences in order to put them on the same scale, which is necessary when working with the

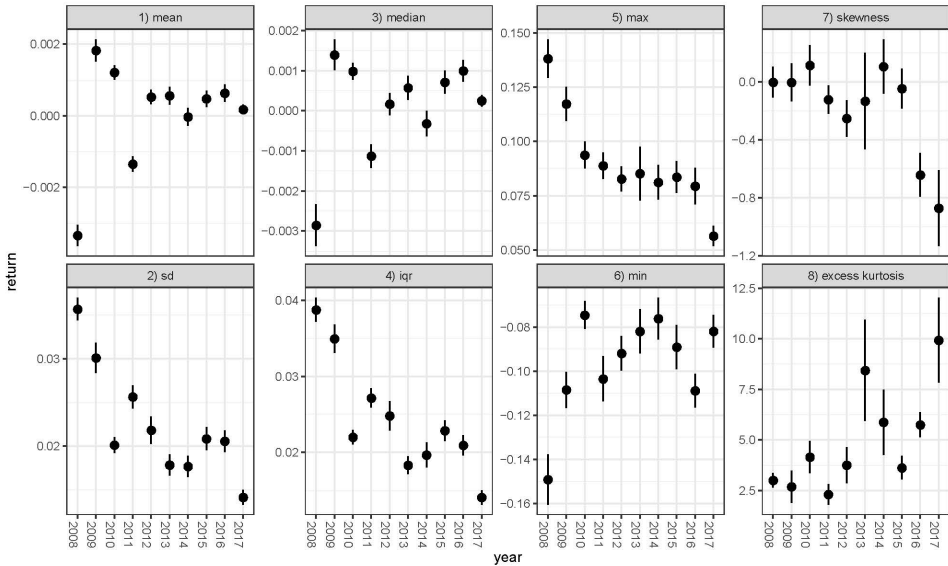


Figure 1A: Return: cross-sectional average of temporal summary statistic in a given year  $\pm$  standard error of the average.

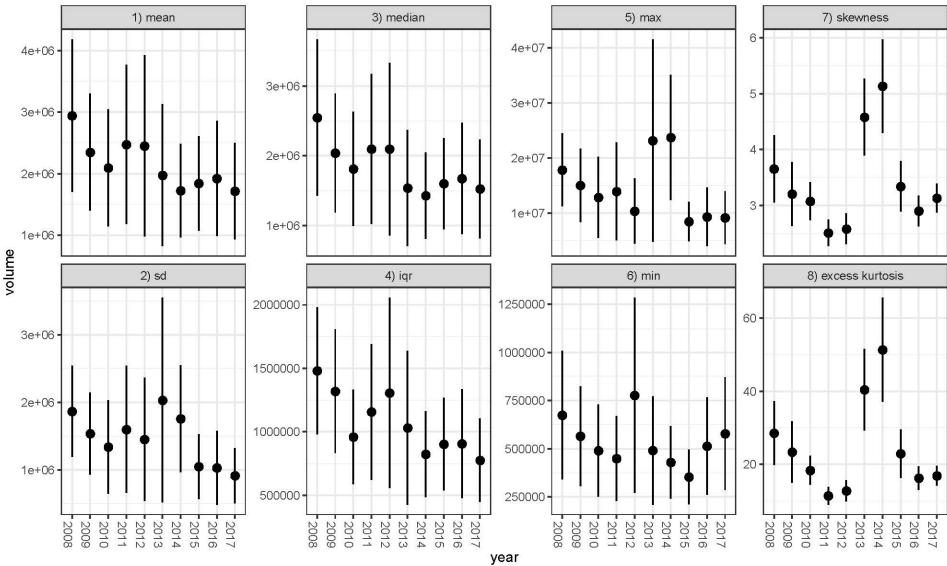


Figure 1B: Volume: cross-sectional average of temporal summary statistic in a given year  $\pm$  standard error of the average.

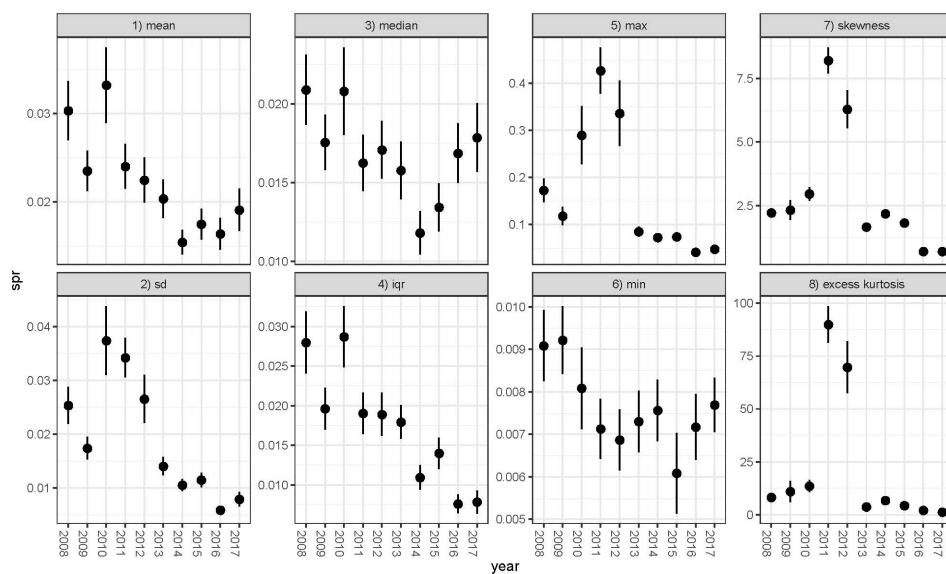


Figure 1C: Spread: cross-sectional average of temporal summary statistic in a given year  $\pm$  standard error of the average.

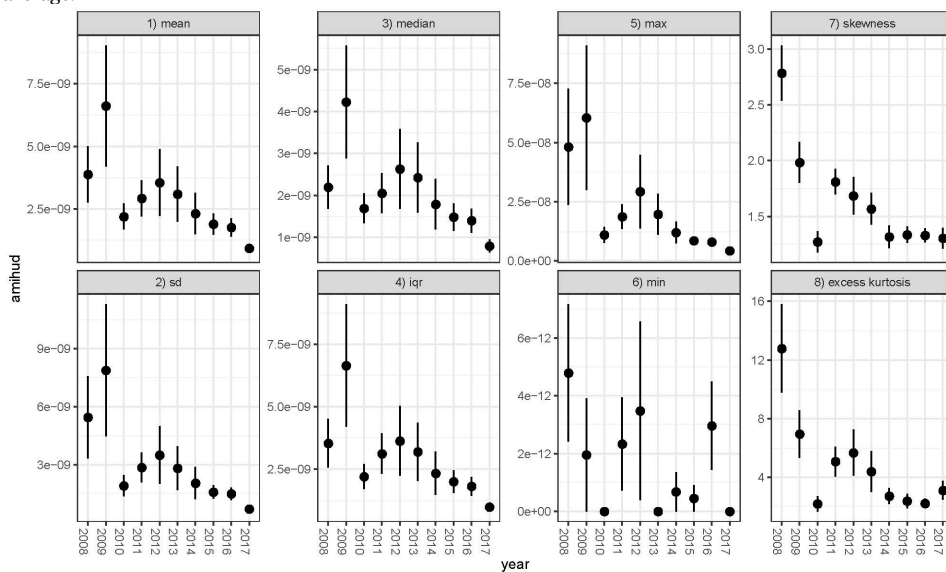


Figure 1D: Amihud: cross-sectional average of temporal summary statistic in a given year  $\pm$  standard error of the average.

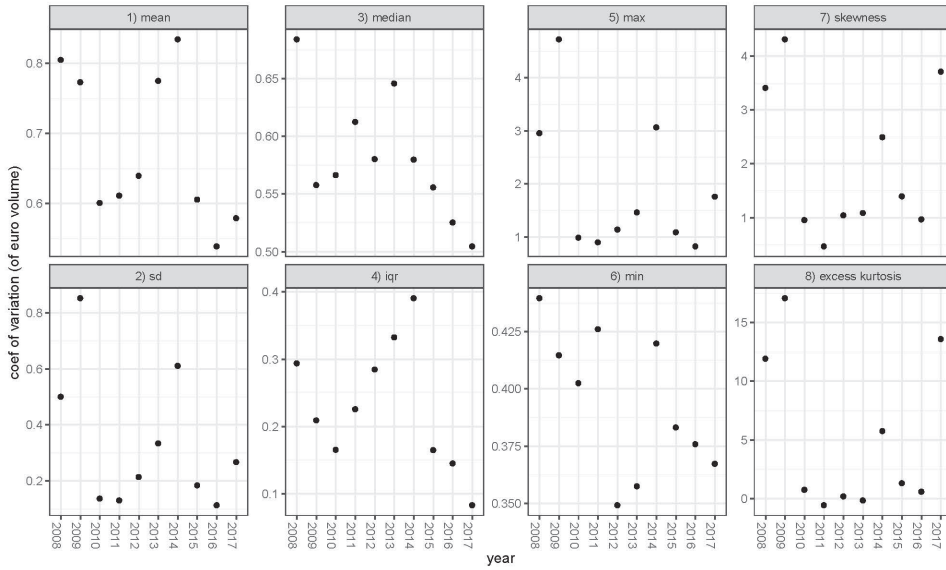


Figure1E: Eurovolume coefficient of variation: cross-sectional summary statistic in a given year.

TPT-based metrics as these quantify the overlap/gap between the TPTs of many sequences. For the stationary series of returns, we choose the z-score normalization  $(R_{s,t} - \hat{\mu}_s) / \hat{\sigma}_s$ , where  $\hat{\mu}_s, \hat{\sigma}_s$  are stock's sample mean and sample standard deviation based on the daily observations for the entire period under consideration. This converts the returns to a zero-mean sequence with the sample standard deviation of 1. For the nonstationary volume, we apply the min-max normalization, i.e., we subtract stock's grand minimum from all the observations and then divide the difference by the stock's range; in symbols  $(V_{s,t} - \min_{1 \leq t \leq T} \{V_{s,t}\}) / (\max_{1 \leq t \leq T} \{V_{s,t}\} - \min_{1 \leq t \leq T} \{V_{s,t}\})$ . This operation scales the volume to the interval  $[0, 1]$ . If the analysis needs to be done in real time, then such normalizations should be carried out on a stretch of data within a window of time ending on the current date. To study the comovement on daily, weekly, monthly, quarterly, semi-annual and annual time scales (time scale in our case equals the number of days), we choose thicknesses of 1, 5, 20, 60, 120, 250, respectively. Such choice of  $\Lambda$  covers a range of time scales from small to large. The TPT-transformed time series of normalized returns for the 23 OMX-Helsinki stocks are shown in Figure 2A (different colours represent different tickers and different panels correspond to different thicknesses) and those of normalized volumes in Figure 2C. Within a given panel, each ticker is represented by a 'tube' whose lower and upper boundaries come from the TPT (equation (1)).

#### 4.1. Short-, medium- and long-term comovement of returns and volumes

The comovement of all 23 stocks based on returns for a given time scale, where time scale ranges from one day to one year is presented in Figure 2B and is quantified by the TPMA (equation (2)). Similarly to Figures 2A and 2C, different panels correspond to different time scales, however this time there is only one curve per panel. This curve corresponds to the values of the TPMA based on all 23 series of (normalized) returns, in other words, the curve gives the proportional overlap/gap of the 23 'tubes' from the respective panel of Figure

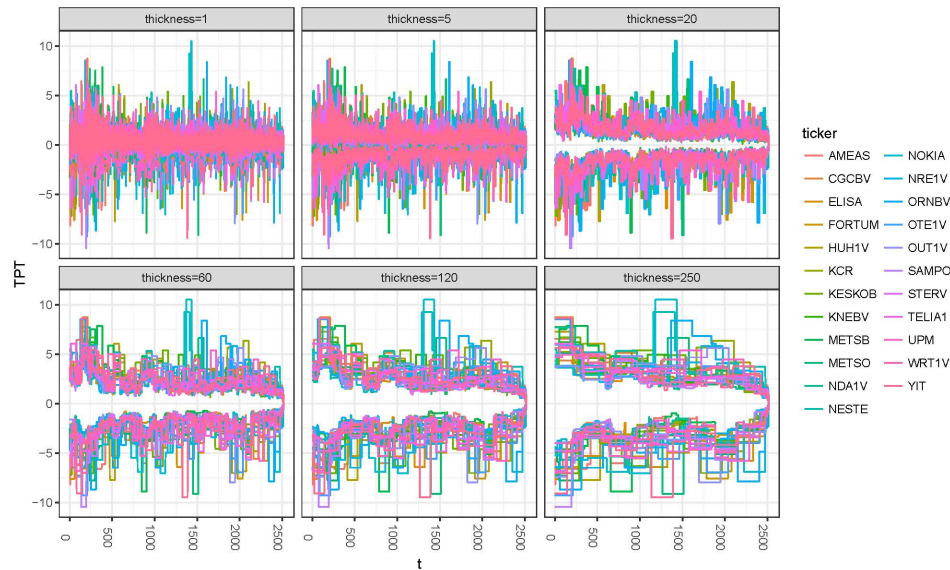


Figure 2A: Thick pen transform of daily (normalized) returns of 23 OMX-Helsinki stocks for several thickness values.

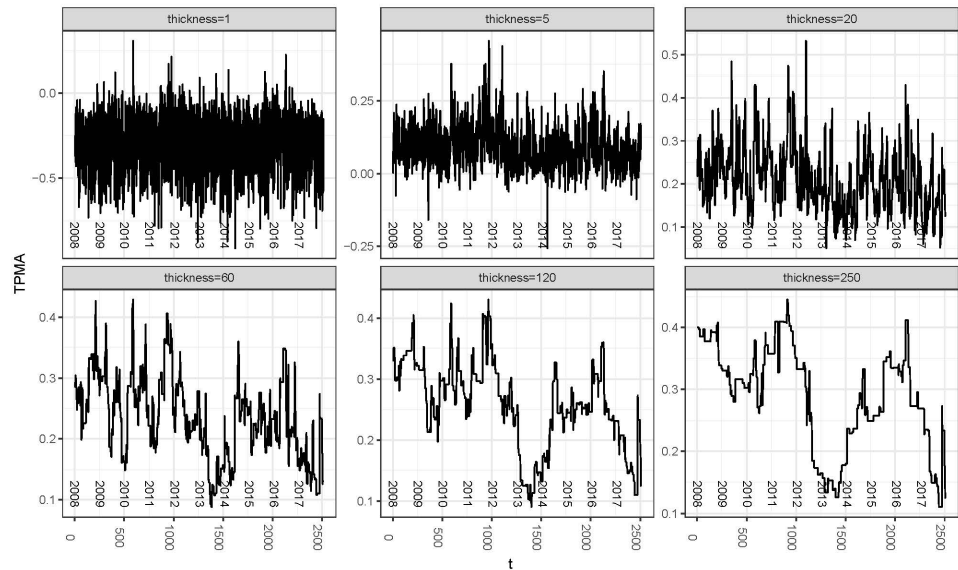


Figure 2B: Thick pen measure of association of daily (normalized) returns of 23 OMX-Helsinki stocks for several thickness values.

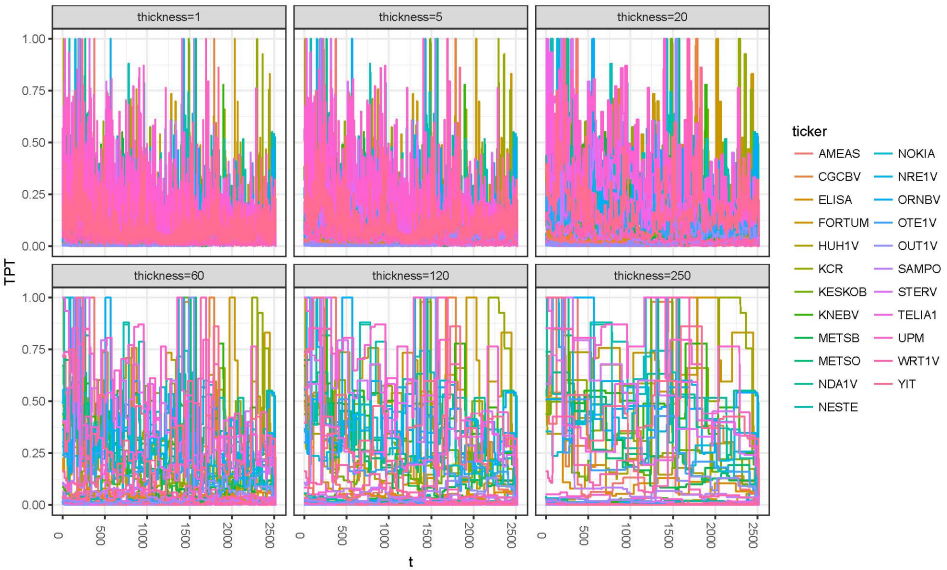


Figure 2C: Thick pen transform of daily (normalized) volumes of 23 OMX-Helsinki stocks for several thickness values.

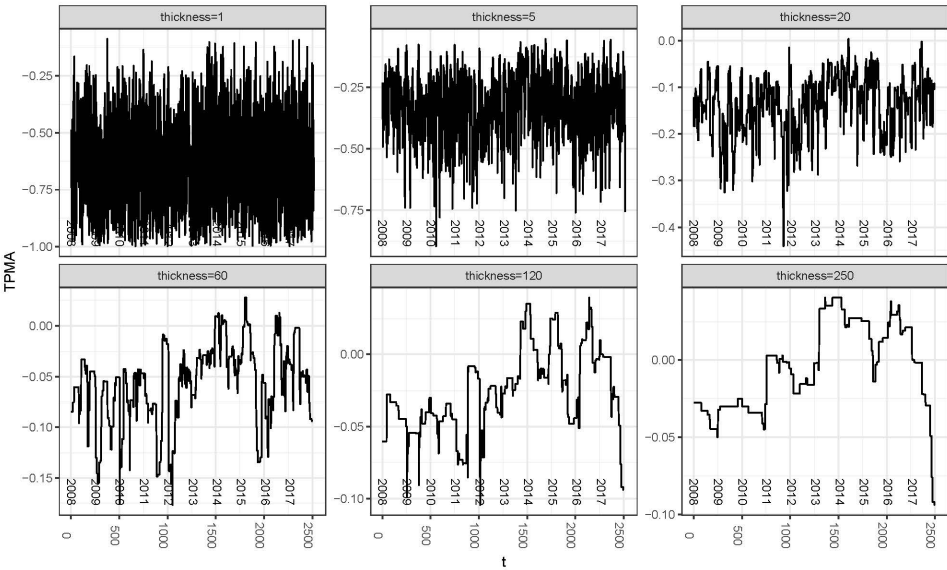


Figure 2D: Thick pen measure of association of daily (normalized) volumes of 23 OMX-Helsinki stocks for several thickness values.

2A. Note the variable range on the y-axis. As the time scale increases from one day to one year, the TPMA (Figure 2B) becomes more and more smooth as we correlate less and less noisy features of the returns. At the same time, the TPMA moves from generally negative values (around  $-0.25$ ) to the positive ones (around  $0.25$ ), which is consistent with our expectations of higher level of long-term comovement compared to the short-term comovement. For the intermediate- and long-term comovement (bottom panels) some trends start to appear. The TPMA for the quarterly time scale (bottom left) has several peaks and troughs, with a big drop around 2013-2014. This drop becomes more pronounced on the semi-annual and annual time scales (bottom middle and right), while two peaks, one around 2011-2012 and one around 2016-2017, start to emerge. It seems that the end of the sample coincides with another drop (subject to some boundary effects as we run out of the observations for  $t$  close to  $T$ ).

The comovement of all 23 stocks based on volumes for a given time scale is depicted in Figure 2D. The layout of the plots is the same as in Figure 2B, and likewise the range of the y-axis is variable. The TPMA is much more negative here compared to the TPMA of return, though similar smoothing-effect can be reported as before. In particular, for the daily time scale (top left) the volumes of the 23 stocks are particularly out of sync leading to the TPMA between  $-1$  and  $-0.25$ . For the intermediate- and long-term comovement there seems to be an increasing, flattening out trend, if we discard the end-of-sample effect.

#### 4.2 Cross-term comovement of returns and volumes

In this section we study the cross-term comovement of the 23 OMXH25 stocks instead of a single-term comovement of the stocks. For that we switch from the TPMA, which uses a single thickness, to the MTPMA (equation (3)), which uses more than one thicknesses and hence facilitates the cross-term comparisons. First, we need to decide which thickness value will be used with each time series. To consider all possible cross-term comovement cases we would need to calculate  $6^{23} = 78973020 \cdot 10^{10}$  MTPMAs (for six  $\tau$ s and  $K = 23$  time series there are  $6^{23}$  sequences  $(\tau^{(1)}, \tau^{(2)}, \dots, \tau^{(K)})$ ). Due to this prohibitively large number of instances, a different strategy is needed. To that end, we choose to study the cross-term comovement of Nokia, the most liquid stock (Section 3) on the Finnish market, and the remaining 22 OMXH25 stocks. In particular, we are interested in how the short-term features of Nokia's returns/volumes correlate with longer-term features of returns/volumes of the remaining stocks. We focus on six cases: in the first three, we correlate Nokia's weekly features ( $\tau = 5$ ) with monthly, quarterly and semi-annual features of the other stocks ( $\tau = 20, 60, 120$ ); in the other three cases, we correlate Nokia's monthly features ( $\tau = 20$ ) with quarterly, semi-annual and annual features of the other stocks ( $\tau = 60, 120, 250$ ). Remembering that Nokia's ticker appears in the 13th position on the alphabetical list, the six MTPMAs are denoted as

$$\begin{aligned} & \rho_t^{(20,20,\dots,20,5,20,\dots,20,20)}(X^{(1)}, X^{(2)}, \dots, X^{(12)}, X^{(13)}, X^{(14)}, \dots, X^{(22)}, X^{(23)}), \\ & \rho_t^{(60,60,\dots,60,5,60,\dots,60,60)}(X^{(1)}, X^{(2)}, \dots, X^{(12)}, X^{(13)}, X^{(14)}, \dots, X^{(22)}, X^{(23)}), \\ & \rho_t^{(120,120,\dots,120,5,120,\dots,120,120)}(X^{(1)}, X^{(2)}, \dots, X^{(12)}, X^{(13)}, X^{(14)}, \dots, X^{(22)}, X^{(23)}), \\ & \rho_t^{(60,60,\dots,60,20,60,\dots,60,60)}(X^{(1)}, X^{(2)}, \dots, X^{(12)}, X^{(13)}, X^{(14)}, \dots, X^{(22)}, X^{(23)}), \\ & \rho_t^{(120,120,\dots,120,20,120,\dots,120,120)}(X^{(1)}, X^{(2)}, \dots, X^{(12)}, X^{(13)}, X^{(14)}, \dots, X^{(22)}, X^{(23)}), \\ & \rho_t^{(250,250,\dots,250,20,250,\dots,250,250)}(X^{(1)}, X^{(2)}, \dots, X^{(12)}, X^{(13)}, X^{(14)}, \dots, X^{(22)}, X^{(23)}). \end{aligned}$$

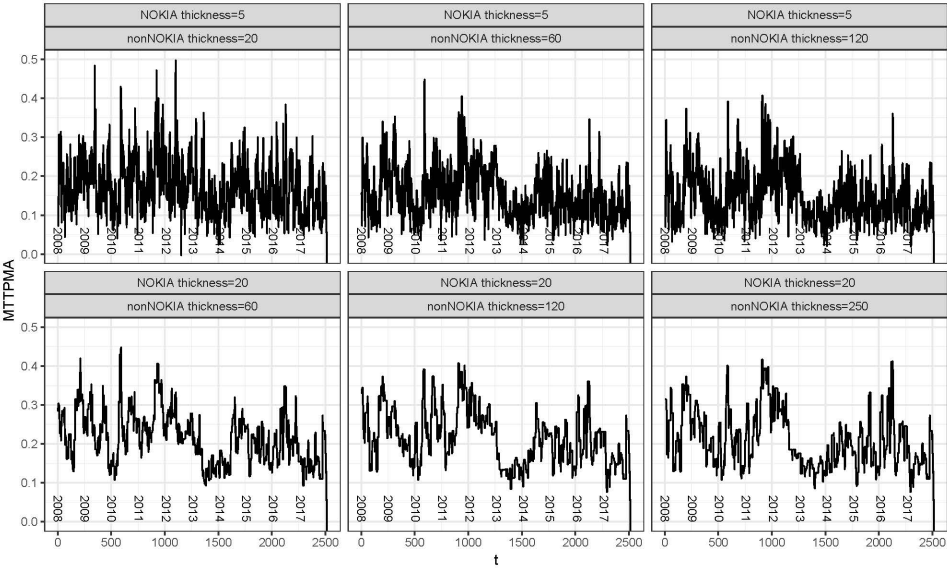


Figure 3A: Multi-thickness thick pen measure of association of daily (normalized) returns of 23 OMX-Helsinki stocks. Top: Nokia's 5-day features correlated with 20-day, 60-day, 120-day features of other stocks. Bottom: Nokia's 20-day features correlated with 60-day, 120-day, 250-day features of other stocks.

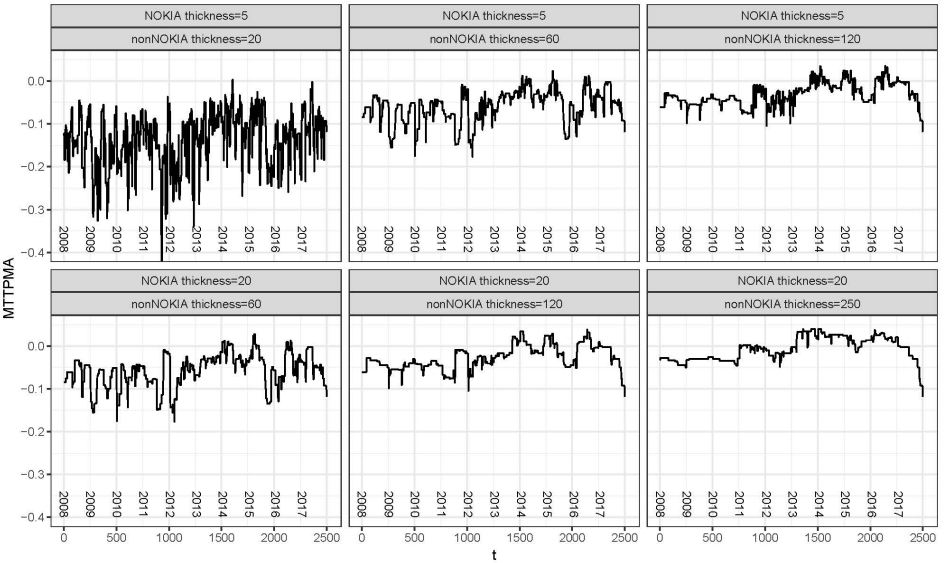


Figure 3B: Multi-thickness thick pen measure of association of daily (normalized) volumes of 23 OMX-Helsinki stocks. Top: Nokia's 5-day features correlated with 20-day, 60-day, 120-day features of other stocks. Bottom: Nokia's 20-day features correlated with 60-day, 120-day, 250-day features of other stocks.

The MTTPMAs based on (normalized) returns are given in Figure 3A and those based on (normalized) volumes in Figure 3B. In both figures, the range of the y-axis is common across all subplots ([0, 0.5] in Figure 3A and [-0.4, 0.05] in Figure 3B).

For the returns, we observe overall positive values of the MTTPMA, with more extreme values for the cross-term comovement involving thickness of 5 (top subplots) compared to that involving thickness of 20 (bottom subplots). For the former ( $\tau = 5$  for Nokia), the cross-term comovement from 2013 onwards seems to be somewhat lower than that prior to 2013 and with fewer oscillations. For the latter ( $\tau = 20$  for Nokia), somewhat similar conclusions hold.

For the volumes, the values of the MTTPMA are generally negative. The MTTPMA correlating Nokia's weekly features with monthly features of other stocks (top left of Figure 3B) is more variable and more negative compared to other MTTPMAs in that figure. A slight increasing trend can be observed in the MTTPMAs.

## 5. Beyond monitoring comovement

In the previous section, we witnessed how the TPMA and MTTPMA can be used to monitor comovement in real time - monitoring comovement can be regarded as an end in itself, e.g., to guarantee the correct functioning of the financial systems or for risk management. Nevertheless, we discuss other specific applications of these measures, and of the thick pen transform, in finance.

First application, closely linked to comovement, is portfolio diversification, whereby an investor aims to reduce cross-dependence between the assets or asset classes of their portfolio and therefore lower portfolio's risk. This can be achieved by computing TPMA (equation (2)) between assets' returns in portfolio evaluation period of length  $E$  (in units of time) and averaging it over time to obtain a one-number summary of the bivariate or multivariate comovement. A big advantage of using TPMA rather than say cross-correlation coefficient, is that TPMA is automatically linked with a given time scale, and hence investor's targeted investment horizon of length  $H$  (in units of time) can naturally be incorporated into their asset allocation decision based on the comovement metric that matches such horizon  $H$ . For example, a long-term investor, who seeks to invest over a long period of time, can use TPMA with large thickness value (large thickness corresponds to long term) to choose assets which comove least in the evaluation period. An example of such a selection rule can be found in Jach (2017).

Another way of exploiting the multi-scale character of the thick pen transform is through building a time-scale-dependent trading strategy, and momentum-based literature lends itself to be a good starting point in that respect. Momentum-based trading strategy (cross-sectional or time-series, Goyal and Jagadeesh, 2017) aims at identifying past 'winners' and 'losers' in hope that they will remain being winners and losers in the future, and by constructing a long-minus-short portfolio of these two groups, respectively; such portfolio is then held over a certain horizon. By combining different lengths  $E$  of the evaluation period and of the holding period  $H$ , one obtains a collection of  $E/H$  variants of this strategy, which implicitly is multi-scale in nature as the identification of winners and losers is performed on past multi-period ( $E$ -period) returns. One modification of this strategy, to incorporate the methodology of the current article, would be to replace the  $E$ -period return by the TPT-based mean,  $M_t^\tau(X) = \frac{1}{2}(U_t^\tau(X) + L_t^\tau(X))$  with  $\tau$  equal to  $E$ , to classify the assets into winners and losers.

Third example of how TPMA or MTTPMA can be exploited in finance is related to lead/lag relationship, where the objective is to assess if some assets drive others. This was demonstrated on simulated (synchronous and regularly-sampled) data using TPMA in Example 3 of Section 3.2.3 in Fryzlewicz and Oh (2011), where the rationale behind the use of this method for phase detection can be

found. Another variant could be to apply TPMA and MTPMA in high-frequency set-up (Huth and Abergel, 2014). High-frequency trading is characterized by asynchronous and irregularly-sampled tick-by-tick data and as such poses difficulties when applying standard cross-correlation coefficient or even its modifications (Hayashi and Yoshida, 2005) if the normality assumption is not met. This is not the case with TPMA and MTPMA, which remain applicable under such circumstances.

These are just few examples where TPMA and MTPMA can be employed. In general, any data-analysis problem which requires a general, multivariate, dynamic, term-wise or cross-term comovement measure, can be addressed with these metrics.

## 6. Summary

Between 2008 and 2017, the stocks included in the OMX Helsinki 25 index had negative returns in 2008 and 2011 and exhibited high level of risk in 2008, 2009, and 2011 compared to other years, with the lowest risk in 2017. This indicates a high level of integration of the Finnish market and the international markets, which were subjected to the global financial crisis and the Eurozone crisis. The activity on the Finnish stock market, proxied by volume, decreased somewhat over the years and the market became more liquid, as reflected by the decreasing bid-ask spread and decreasing Amihud measure of illiquidity. Information asymmetry, quantified by the coefficient of variation, also seemed to go down. These indicators suggest that the Finnish economy was doing relatively well towards the end of the decade of 2008-2017.

Longer-term comovement of returns of the OMXH25 stocks followed a cyclical trend with ups around 2011-2012 and 2016-2017 and downs around 2013-2014 and possibly at the end of the decade. Somewhat similar patterns appeared in the intermediate-term comovement of returns; this comovement was also more variable. Long-term and intermediate-term return comovement was generally lower in the second half of the decade. The comovement of volumes was overall weaker, though exhibited an increase in its strength. The cross-term interactions of Nokia's weekly/monthly features and non-Nokia-stocks' longer-term features based on returns were different than those based on volumes.

## References

- Aguiar-Conraria, L. & Soares, M. (2014). The continuous wavelet transform: moving beyond uni- and bivariate analysis. *Journal of Economic Surveys* 28, 344–375.
- Amihud, Y. (2002). Illiquidity and stock return cross-section and time-series effects. *Journal of Financial Markets* 5, 31–56.
- Bae, K., Karolyi, G., & Stulz, R. (2003). A new approach to measuring financial contagion. *The Review of Financial Studies* 12, 851–861.
- Berger, T. & Pozzi, L. (2013). Measuring time-varying financial market integration: An unobserved components approach. *Journal of Banking & Finance* 37, 463–473.
- Bos, T., Fetherston, T., Martikainen, T., & Perttunen, J. (1995). The International Co-Movements of Finnish Stocks. *European Journal of Finance* 1, 95–111.
- Bu, H., Tang, W., & Wu, J. (2019). Time-varying comovement and changes of comovement structure in the Chinese stock market: A causal network method. *Economic Modelling* 81, 181–204.
- Dungey, M., Fry, R., Hermosillo, B. G., & Martin, V. (2003). Empirical modelling of contagion: A review of methodologies. *The Review of Financial Studies* 5, 9–24.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20, 339–350.
- Fernandez-Macho, J. (2012). Wavelet multiple correlation and cross-correlation: A multiscale analysis

- of eurozone stock markets. *Physica A* 391, 1097–1104.
- Fernandez-Macho, J. (2018). Time-localized wavelet multiple regression and correlation. *Physica A* 492, 1226–1238.
- Forbes, K. & Rigobon, R. (2002). No contagion, only interdependence: Measuring stock market comovements. *The Journal of Finance* 57, 2223–2261.
- Fryzlewicz, P. & Oh, H.-S. (2011). Thick-pen transformation for time series. *Journal of the Royal Statistical Society Series B* 73, 499–529.
- Gallo, G. & Otranto, E. (2008). Volatility spillovers, interdependence and comovements: A markov switching approach. *Computational Statistics and Data Analysis* 52, 3011–3026.
- Gorrostieta, C., Ombao, H., & von Sachs, R. (2019). Time-dependent dual-frequency coherence in multivariate non-stationary time series. *Journal of Time Series Analysis* 40, 3–22.
- Goyal, A. & Jagadeesh, N. (2017). Cross-sectional and time-series tests of return predictability: What is the difference? *The Review of Financial Studies* 31, 1784–1824.
- Graham, M. & Nikkinen, J. (2011). Co-movement of the Finnish and international stock markets: A wavelet analysis. *European Journal of Finance* 17, 409–425.
- Hayashi, T. & Yoshida, N. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli* 11, 359–379.
- Huth, N. & Abergel, F. (2014). High frequency lead/lag relationships - empirical facts. *Journal of Empirical Finance* 26, 41–58.
- In, F. & Kim, S. (2012). *An Introduction to Wavelet Theory in Finance: A Wavelet Multiscale Approach*. World Scientific, Singapore.
- International Monetary Fund (2018). World Economic Outlook: Challenges to Steady Growth. Available at: <https://www.imf.org/en/Publications/WEO/Issues/2018/09/24/world-economic-outlook-october-2018> (accessed 18 November 2018).
- Jach, A. (2017). International stock market comovement in time and scale outlined with a thick pen. *Journal of Empirical Finance*, 43:115–129.
- Jach, A. (2019). A general comovement measure for time series. Working paper, Hanken School of Economics.
- Kallunki, J.-P., Martikainen, M., Martikainen, T., & Yli-Olli, P. (1996). The Finnish stock market: A survey of some empirical evidence and its practical relevance. *Finnish Journal of Business Economics* 4, 474–495.
- Kaminsky, G., Reinhart, K., & Vegh, C. (2003). The unholy trinity of financial contagion. *Journal of Economic Perspectives* 17, 51–74.
- Karolyi, G. A. & Stulz, R. M. (1996). Why do markets move together? An investigation of U.S.-Japan stock return comovements. *The Journal of Finance* 51, 951–986.
- Lehkonen, H. & Heimonen, K. (2014). Timescale-dependent stock market comovement: BRICs vs. developed markets. *Journal of Empirical Finance* 28, 90–103.
- Lof, M. & van Bommel, J. (2018). Asymmetric information and the distribution of trading volume. Research Discussion Paper, Bank of Finland.
- Martikainen, T., Perttunen, J., & Puttonen, V. (1995). The lead-lag effect between large and small firms: Evidence from Finland. *Journal of Business Finance and Accounting* 22, 449–454.
- Nyberg, P. & Vaihekoski, M. (2014). Equity premium in Finland and long-term performance of the Finnish equity and money markets. *Clometrica* 8, 241–268.
- Percival, D. & Walden, A. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.
- Rodriguez, J. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance* 14, 401–423.
- Statistics Finland (2018). GDP per capita. Available at: <https://findikaattori.fi/en/table/2> (accessed 10 July 2018).
- Valencia, F. & Laeven, L. (2012). Systemic banking crises database: An update. Working paper, International Monetary Fund.