MPRA
Munich Personal RePEc Archive

A Microstructure Model for Spillover Effects in Price Discovery: A Study for the European Bond Market

Perlin, Marcelo and Dufour, Alfonso and Brooks, Chris

Reading Universty

June 2010

Online at https://mpra.ub.uni-muenchen.de/23380/
MPRA Paper No. 23380, posted 20 Jun 2010 04:27 UTC

A Microstructure Model for Spillover Effects in Price Discovery: A Study for the European Bond Market

Marcelo Perlin (m.perlin@icmacentre.ac.uk)[1]
Alfonso Dufour (a.dufour@icmacentre.ac.uk)[2]
Chris Brooks (c.brooks@icmacentre.ac.uk)[3]

Last Revision: 19 June 2010

Abstract: This paper is set to investigate the existence of spillover effects for the trading process of correlated financial instruments. While the main literature in price impact models has focused mainly on multivariate processes for a unique asset, we argue that transitory spillover effects in such class of models should exist as a simple biproduct of explicit relationships among prices of different (but correlated) financial instruments. Firstly we assess the theoretical implications of a transitory spillover effect in an extended microstructure model and then we investigate our different hypothesis in the European bond market with a formal econometric model. The results showed that the estimated parameters of the econometric models do conform to what we expect in the theoretical derivations, where the trades of one instrument would be correlated to the trades in others. But, even though the results are positive, they could also be explained by traders splitting orders across different instruments or joint periods of intensive trading. Further analysis also showed that the trading intensity in other instruments does affect the trading process of the particular bonds. We found that a buy (sell) order is less likely to be followed by a buy (sell) order if the market is trading intensively. We explain such effect as an inventory problem, where volatility of prices forces market makers to improve trades in the opposite direction from the current order flow. The main conclusion of this study is that we find inconclusive results towards the particular microstructure model set in the theoretical part of the paper, but positive results for a general spillover effect in the trading process of European fixed income instruments.

Keywords: market microstructure, spillover effect, commonalities, liquidity, price impact of a trade.

[1] PhD Student at ICMA Centre, Henley Business School, Reading University (United Kingdom).
[2] Lecturer in Finance, ICMA Centre, Henley Business School, Reading University (United Kingdom).
[3] Professor in Finance, ICMA Centre, Henley Business School, Reading University (United Kingdom).

INDEX

Introduction

The study of trading processes in the high frequency dimension has attracted a lot of attention of the academic community for the past few years. One can settle this attention on the advance of computational power and on the new availability of high frequency data for different instruments. While the advent of tick by tick trades and quotes data has forced an adaptation from the usual modelling tools used in discrete time models, the intuition is the same, explain the factors that move prices.

Microstructure models of trading draw explicit relationships between efficient prices and tradable (quoted) prices given particular stylized effects observed in the market. These effects are labelled as transitory effects (or microstructure frictions) and include the inventory problem and the asymmetric information issue. Summarizing, the first is given by the need of the market maker (dealer) to avoid excessive risk in his trading operations and the second is simply the result of (better) informed economic agents impacting prices with their trades.

In general, these theoretical models are represented by a set of recursive equations describing the evolution process for efficient prices, mid-quotes, trades and possibly market maker's inventory. The deviations of the tradable prices from the true (or efficient) price are the result frictions in the trading process. Mathematically, it is possible to show that these structural models retain an autoregressive structure in the reduced form equations, which then motivates the econometric analysis in empirical data. See (Hasbrouck, 1995) for further details on microstructure models of price discovery.

The present paper is related to a general microstructure model. Our main argument is towards the introduction of spillover effects in such a representation. More precisely, we are interested in gaining a better understanding how the trades and quotes of other instruments affect the trading process of a particular financial asset. We can draw motivations on this idea by realizing that traders are not restrained to trade only a single asset class. Standard finance textbooks make significant points on describing arbitrage and hedging strategies which involve operations on different asset classes. On the market maker side, the behaviour of other assets classes can also provide information regarding the true value of a particular instrument. For instance, futures contracts and underlying spot prices are intrinsically related so it's not hard to imagine the actions of a futures dealer being conditional on the behaviour of the underlying asset.

The idea of correlated trading processes over different instruments is not novel. Empirical examinations of co movements in aggregated measures of liquidity can be found in (Hasbrouck, et al., 2001), (Chordia, et al., 2000) and (Chordia, et al., 2001). Also, closer to our idea of a explicit microstructure representation, a simple model of multiple prices for two assets in which there was zero correlation across assets in the random walk error term but trades were correlated is described in (Hasbrouck, 2007) p. 94 . One can imagine such an event for instance if a trader is to proxy a particular stock index by buying all (or some of) the underlying shares. We use the same intuition as in (Hasbrouck, 2007) multiple assets model but we further detail the structural model by defining asset (and non asset) specific drivers of trading. The novelty in the theoretical microstructure model defined in the paper is that trades in asset i are correlated to the mispricing in asset j and not directly to trades in j itself. This is what we call a spillover effect and we argue that this effect should be classified as transitory and trade related.

Our reason for testing such a hypothesis in the European fixed income market is that while this market is composed of instruments for different maturities and countries, they have similar risk factors. Therefore, it can be argued that microstructure effects will be spilled over across similar bonds. Another effect which we are particularly interested is the spillover of trading intensity in the market upon the trading process if a particular bond. More precisely, we want to know whether the amount of trades (or quotes changes) in similar instruments is affecting the price impact of trades of a respective bond. The intuition is similar to (Dufour, et al., 2000), that is, to test whether trading intensity in other instruments also conveys information. These are the main arguments behind this paper and represent sufficient arguments to demand a scientific investigation.

The paper is organized as follows. First we briefly look at previous papers in the topic, second we build our theoretical microstructure model and derive the econometric implications of the spillover effect. This is followed by description of the data, methodology and results. We finish the paper with the usual concluding remarks.

Literature Review

The main background of this study is in the price discovery of financial assets, more precisely how economic agents' behaviour drives the process behind trades and quotes. The main reference in this area of market microstructure goes back to (Kyle, 1985), which was one of the first to formalize the economic situation where a group of risk neutral market makers face insiders (informed traders) and liquidity traders in the observed order flow. In this setup, the market maker is aware of the insider trader strategy but not its identity. He will therefore, based on the observed order flow, update his beliefs about the true value of the asset. This model was then further extended to an order book structure in the (Glosten, et al., 1985) paper. In this study the authors showed that the existence of asymmetric information among traders motivates an increase in the quoted spread from the market maker's side. This larger spread is then the premium for the market maker for the uncertainty of trading with informed traders. The (Glosten, et al., 1985) setup was then further extended for variable trading sizes in the work of (Easley, et al., 1987).

For the case of empirical models of price discovery, in which the interest is in estimating parameters of the underlying microstructure model, the seminal paper in the area is (Hasbrouck, 1991). In this work, it was shown that in a theoretical framework trades and quotes affect each other in an autoregressive fashion. In the paper, the author showed that the incoming of a buy/sell order in the market will move the mid quote price in the same direction as the trade (e.g. buy trades moving prices up). This is related to the informational content of trades, that is, trades convey information regarding the true price of an asset. Also, there was significant evidence that volume of trades and the spread also contribute to the price impact of a trade. This model made a significant contribution to the field and the extension of this basic model soon drew the attention of the academic community. The subsequent paper on this topic is (Brennan, et al., 1996), which made an analysis comparing the measures of illiquidity and expected returns for NYSE stocks. The authors find that the most illiquid stocks, measured by a high frequency econometric model, also present higher expected returns. This result is intuitive as the lack of liquidity of a stock creates costs for uninformed investors, and therefore, on average, they should be compensated by higher returns.

In the paper of (Dufour, et al., 2000), Hasbrouck's model was extended to allow for the effect of time on the price impact of a trade. This research was based on theoretical foundations[4]

[4] See (Hasbrouck, 1995) and (Easley, et al., 1995).

which stated that the time between events should indirectly indicate the present of news in the market. This study was conducted for 18 stocks in the TORQ database and the main result is that the time (measured in durations) does affect the price impact of a trade in a negative way. Higher (lower) the duration between trades, lower (higher) the price impact of a trade. Such result is intuitive as long time intervals with no trade are usually associated with the lack of news in the market. Therefore, if the trading frequency is high (and durations are small), there is a higher likelihood of informed traders which will result in the market maker moving quotes more sensibly with respect to the observed order flow.

More recently, in the work of (Furfine, et al., 2005), the interest was in analyzing how price discovery in the US treasury market would change for stressful time periods. Using different concepts of stressful times, the authors find that the impact of trades in quote changes is bigger and spreads are wider across different maturities when the market is experiencing stress. This is an intuitive result as the cost to carry inventory is higher for stressful times and therefore, for these time periods, the market makers would have the right motivation for increasing implicit trading costs such as impact of trades and quoted spread.

In the common ground of microstructure of fixed income markets, we have the work of (Brandt, et al., 2003) and (Cheung, et al., 2005). In the first of these, authors looked at the empirical relationship between orderflow, liquidity and yield curve for a high number of instruments. The main conclusion of this study was the evidence of price discovery, where the order flow impacts the yield curve. In the second paper, (Cheung, et al., 2005), the idea was to understand how the different trading platforms for European bonds differ in terms of microstructure issues. More precisely how different are the implicit trading costs from one platform to another. The authors find that there were differences in the trading aspects of both platforms in terms of observed spread and price impact of trades. But, these were small. Therefore the authors argue that both platforms (local and European) are integrated.

Switching the literature review to the case of the investigation of common factors on high frequency data, we have the work of (Hasbrouck, et al., 2001). Using principal component and canonical correlation analysis, the authors studied the co-movements between order flow and returns for thirty stocks in the Dow Jones Index for the period of 1994. The authors find that a significant proportion (21%) of the variation in returns can be explained by an underlying common factor. But, the results for the existence of common factor in order flow dynamics were weaker.

(Chordia, et al., 2001) studied the aggregated measures of liquidity for a wide range of NYSE stocks. The authors find that the measures of liquidity are negatively autocorrelated with strong day-of-week effects and there are also clear patterns for increase in trading activity just prior to major economic announcements. Further results on liquidity co movements are given in (Chordia, et al., 2000) where the focus in on the dependency of individual asset's liquidity with respect to the market and industry wide liquidity. Using different measures of liquidity and controlling for stylized effects (such as volatility), the study shows that a great part of the variation in the liquidity proxies for individual stocks is due to variations in the liquidity of the market as a whole.

The motivations behind the last batch of studies are very similar to the motivations for the research in this paper. In general we are trying to better understand how the trading process in the market as whole can relate to the trading process of the assets individually. While the previous literature has studied the impact of different explanatory variables such as dummies for stressful periods and heterogeneous trading platforms of the instrument in question, so far there is no formal empirical study on possible spillover effects (or co-movements) across different instruments in the joint process of trades and quotes (and not liquidity measures). Next we describe the theoretical foundations behind the study.

Theoretical Foundations

A Simple Microstructure Model

As a starting point to derive the ideas behind our microstructure model, we present the simplest (benchmark) case, also found in (Hasbrouck, 1991). This model incorporates information asymmetry and inventory control. We start with the process for efficient prices.

$$m_t = m_{t-1} + z v_t^x + v_t^m \tag{1}$$

For Equation (1), m_t is the efficient price of the instrument conditional on all available information. The variable v_t^x is the unexpected amount of trades for this particular instrument, z is the response to unexpected trades and v_t^m is the idiosyncratic random behaviour for the price of the asset (e.g. changes in public available information). This disturbance is usually assumed to have zero expectation, constant variance σ^2 and zero autocorrelation in all lags (see (Hasbrouck, 2007)). From (1) it is easy to see that the expected changes in the efficient price $E\left(\Delta m_t\right)$ are not predictable. The process for the mid quote prices will be given by:

$$q_t = m_t + \alpha\left(q_{t-1} - m_{t-1}\right) + b x_t \tag{2}$$

For (2), α is the degree of adjustment of the mid quote price with respect to the lagged difference from the efficient price. This parameter is inventory control related. For example, if $q_{t-1} - m_{t-1}$ is positive then quotes will be raised to motivate a sell and de-motivate a buy order. The coefficient b in (2) is the adjustment of quotes to incoming trades or the fixed price cost per unit of trade. The final equation that describes this system is the trade equation:

$$x_t = -c\left(q_{t-1} - m_{t-1}\right) + v_t^x \tag{3}$$

For (3), the parameter c measures how trades would respond to a mispricing from the existing quote with respect to the efficient price. It is straightforward to prove that (1)-(3) implies a lagged regression for the difference in mid quote prices ($\Delta q_{i,t}$). This lagged regression has the form[5]:

$$\Delta q_t = \left(z + b\right).x_t + \sum_{k=1}^{\infty} \alpha^{k-1}\left(zbc - b\left(1-\alpha\right)\right) x_{t-k} + v_t^m \tag{4}$$

Therefore, the changes in mid quote price can be represented as regression on the level and lagged values of x_t. Clearly, if $\alpha < 1$, the impacts of further lags will decrease over time so that the effect of a shock is transitory. For the trade equation, the autoregressive representation according to Hasbrouck microstructure model will follow:

[5] See Appendix 1 for a derivation.

$$x_t = -\sum_{k=1}^{\infty} \alpha^{k-1} cb.x_{t-k} + v_{2,t} \tag{5}$$

Therefore, we also see that the process for the trades can be represented by an autoregressive formula with infinite lags and decaying memory as long as α is lower than one. Also, since c, b and α are expected to be positive, then a negative autocorrelation is expected for the trades. Following the formulae in (4) and (5), we have an econometric argument for the use of autoregressive models for trades and quotes. While the previous formulas in (4) and (5), do not imply that trades will affects quotes (and vice versa), we still have theoretical reasons to include them in the specifications. This motivates the use of a standard VAR model to analyze trades and quotes.

A Microstructure Model with Spillover Effects

The econometric model used in the study has a theoretical foundation which is defined here. We start with the simple model as before but change the notation in order to follow our arguments. We are interested in defining a process for different assets in the market. Therefore, we will incorporate an index *i* for the different instruments so that $m_{i,t}$ is the efficient price of instrument *i* at time *t*. Consider the following full process for trades and quotes:

$$m_{i,t} = m_{i,t-1} + z_i.v_{i,t}^{x} + v_{i,t}^{m} \tag{6}$$

$$q_{i,t} = m_{i,t} + \alpha_i \left(q_{i,t-1} - m_{i,t-1}\right) + b_i.x_{i,t} \tag{7}$$

$$x_{i,t} = -c_i \left(q_{i,t-1} - m_{i,t-1}\right) + \sum_{j=1, j\neq i}^{M} f_{i,j} \left(q_{j,t-1} - m_{j,t-1}\right) + v_{i,t}^{x} \tag{8}$$

Equation (6) to (8) basically define the time evolution of a whole market of different instruments. The extension of the microstructure model defined before relates to the trade equation (8). The main point is that the mispricing of other instruments $j = 1..N$ also motivates trades in the instrument *i*. This will have a shape of a (so far) generic function $f_{i,j}$. Such a relationship is not difficult to imagine if one thinks of the non-arbitrage pricing argument. Consider that $q_{i,t}$ is the tradable price of a futures contract. The fair tradable price of this instrument is also a function of the spot price and the interest rate[6]. Suppose now that interest rates stay still and a strong transitory effect impacts the quoted value of the spot price (e.g. an excessive number of liquidity traders on one side of the market or a market maker improving quotes to motivate a buy/sell, given his inventory risk limits). While there was no news and no change in the fundamental value of both assets, the quoted spot prices are implying the existence of a risk free[7] profit some time in the future. This will motivate

[6] This relationship is $F = S\left(1+r\right)$ holds for the non existence of transaction costs and extra cash flows (e.g. dividends). When trading costs are added, the true price is bounded with a minimum and a maximum.
[7] Assuming, of course, no margins for trading futures and no transaction costs.

opposing trades in both instruments which will move prices until the no arbitrage bounds holds once again.

As one can see, the intuition behind the example is simple: transitory mispricing in one instrument implies the existence of a risk free profit, therefore motivating trades in both instruments where the direction of the trades is related to the sign of the mispricing and the arbitrage operation itself. From formulas (6) to (8) it should also be possible to see a cascade effect, where the mispricing in instrument A for *t*-1 influences the mispricing in B for *t*, which will then influence the quotes for C in *t*+1 and so forth. The strength of this effect would clearly be dependent on the parameters of the model.

In this extended microstructure model, the spillover effects will also be classified as a transitory and mean reverting change. The mispricing of the quotes with respect to the fundamental value, $q_{j,t-1} - m_{j,t-1}$, which is the driver of the spillover effect, is simply the result of microstructure frictions such as the inventory problem. This is easy to see since $q_{j,t-1} - m_{j,t-1} = \alpha_j \left(q_{j,t-2} - m_{j,t-2} \right) + b_j . x_{j,t}$. Also, the mispricing of other instruments will not affect the efficient price of instrument *i*. One can see this by simply observing that the new term, $\sum_{j=1, j \neq i}^{M} f_{i,j} \left(q_{j,t-1} - m_{j,t-1} \right)$, is part of the expected trade equation, while the impact on the fundamental price (information asymmetry) is only for the unexpected trades, $v_{i,t}^{x}$. But, as said before, in order for the function $f_{i,j}$ to be true there must be a common risk factor among the instruments. Therefore our microstructure model implies the existence of a covariance between the disturbances $v_{i,t}^{m}$ and $v_{j,t}^{m}$ in the efficient price process.

The microstructure model described in (6) to (8) implies[8] the following autoregressive formula for the mid quote change and trades:

$$\Delta q_{i,t} = \left(z_i + b_i \right) . x_{i,t} + \sum_{k=1}^{\infty} \alpha_i^{k-1} \left(z_i b_i c_i - b_i \left(1 - \alpha_i \right) \right) x_{i,t-k} + v_{i,t}^{m} \tag{9}$$

$$x_{i,t} = -\sum_{k=1}^{\infty} \alpha_i^{k-1} c_i b_i . x_{t-k} + \sum_{j=1, j \neq i}^{M} f_{i,j} \left(\sum_{k=1}^{\infty} \alpha_j^{k-1} b_j . x_{j,t-k} \right) + v_{i,t}^{x} \tag{10}$$

Clearly, Equation (9) is the same as Equation (4), so we have the same lagged regression for the mid quote changes in between Hasbrouck and the spillover model. For Equation (10), we see that trades in the other instruments will also affect the trades in asset *i,* given values of parameters α_i, b_i and the function $f_{i,j}$. Again, if α_i and α_j are less than one, then the weights of this autoregressive model with infinite lags will decline geometrically.

The mathematical representations given in (9) and (10) provide some intuition of what are the consequences in the introduction of spillover effects in a simple microstructure model. When following Hasbrouck model, Equations (1)-(3), the autoregressive representation of the trade process didn't imply any spillover effect but the formula in (10) clearly indicates that our microstructure model suggests a correlation between present trades in each instrument and lagged trades in others. This is a similar result to the multi asset microstructure model

[8] See Appendix 2 for derivations.

presented in (Hasbrouck, 2007). The difference is that for our model we show that this correlation among trades would be a function of the parameters from the structural equations.

So far we have stated that $f_{i,j}$ is a generic unknown function. For the rest of the paper, we will address $f_{i,j}$ as a linear function with some (estimated) weight parameter. The intuition behind such simplification is that the data used in the study is composed of European bonds of different maturities and countries. These instruments have similar risk factors in terms of default risk and discount rate. Therefore, it can be expected that the prices of different bonds are correlated in a linear fashion. This simplification greatly facilitates the econometric analysis taken in the paper.

A Simulation Exercise

We build an example on our simple microstructure model by considering just two assets, A and B. Both are driven by the same risk factors but Asset A is more liquid and less volatile than asset B. For simplicity, we assume that the link function $f_{i,j}$ in between these assets is linear, given a weight parameter. Following Equations (6)-(8), the simulated full process for asset A is:

$$m_{A,t} = m_{A,t-1} + 0.2 v_{A,t}^{x} + \frac{2}{10} v_t^m$$

$$v_t^m = \begin{cases} 1 \text{ if } \mathrm{u}_t > 0.5 \\ -1 \text{ if } \mathrm{u}_t < 0.5 \end{cases}$$

$$q_{A,t} = m_{A,t} + 0.5\left(q_{A,t-1} - m_{A,t-1}\right) + 0.2.x_{A,t}$$

$$x_{A,t} = -0.5\left(q_{A,t-1} - m_{A,t-1}\right) + 0.2\left(q_{B,t-1} - m_{B,t-1}\right) + v_{A,t}^{x}$$

$$v_{A,t} = \begin{cases} 1 \text{ if } \mathrm{u}_t > 0.5 \\ -1 \text{ if } \mathrm{u}_t < 0.5 \end{cases}$$

And the process for B will be given by:

$$m_{B,t} = m_{B,t_{-1}} + 0.5 v_{B,t}^{x} + \frac{5}{10} v_t^m$$

$$q_{B,t} = m_{B,t} + 0.2\left(q_{B,t-1} - m_{B,t-1}\right) + 0.5.x_{B,t}$$

$$x_{B,t} = -0.2\left(q_{B,t-1} - m_{B,t-1}\right) + 0.5\left(q_{A,t-1} - m_{A,t-1}\right) + v_{B,t}$$

$$v_{B,t} = \begin{cases} 1 \text{ if } \mathrm{u}_t > 0.5 \\ -1 \text{ if } \mathrm{u}_t < 0.5 \end{cases}$$

For the disturbances in the simulation, we choose to use Bernoulli variables (v_t^m and $v_{B,t}^x$) for simplicity but it is clear that any distribution would fit. When choosing the values of the parameter in the simulation, we followed the argument that asset A is more liquid than asset

B. For instance, the reaction of the efficient price to unexpected trades (parameter z) is lower for A when comparing to the value for B, meaning that A is more resilient to unexpected trades than B. Also, it should be clear that A and B have the same disturbance in the price process (v_t^m), but with different weights. The idea is that these assets have a common risk factor that drives the efficient price process, but with distinct reactions to this random component. For our simulation we assumed the tick (minimum price increment) is equal to 1/10, therefore asset A has a weight factor in the efficient price disturbance of 2 ticks against 5 ticks of asset B. Clearly, given this setup, prices changes in asset B are going to be more volatile than price changes in asset A.

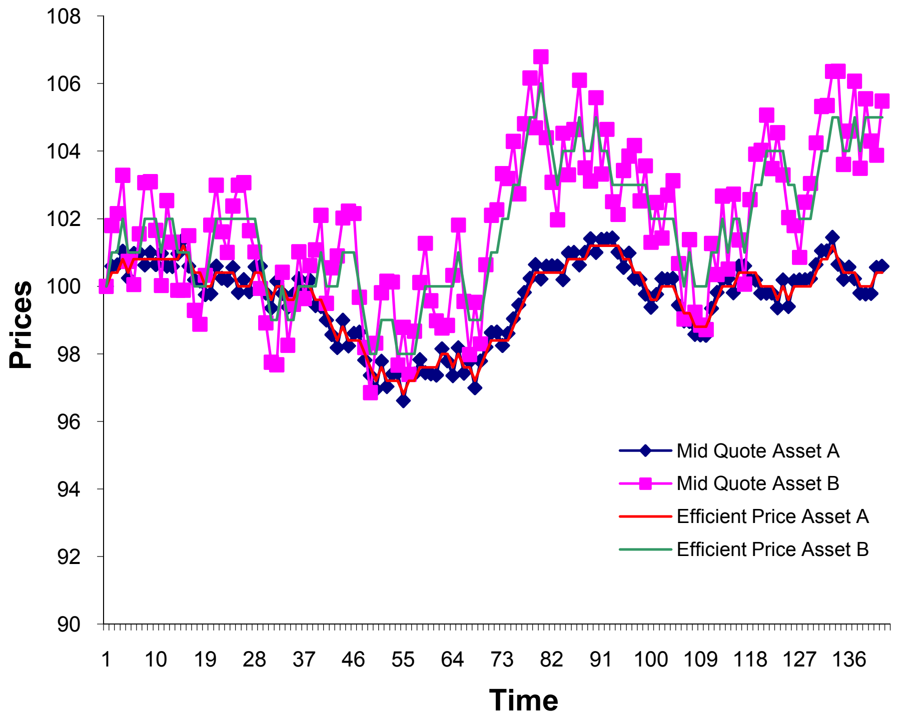
For the trade equation in our simulations, we imply that A has a higher reaction[9] to mispricing than B. The idea is that since A is more liquid, it is reasonable to say that more people will be trading that stock when a mispricing occurs. Therefore the reaction to a mispricing in the trade equation of A should be higher than for a less traded asset B. On the spillover effect we have that asset B reacts to the mispricing in A with a rate of 0.5, which is higher than the reaction of A to B (0.2). The idea is that when traders trade on the mispricing of B, they hedge it against the risk factor by trading with opposite signs in A. Notes that if $q_{A,t} - m_{A,t}$ is positive implying that the quotes are underpriced, then it is likely that a sell sign will be generated for asset A and the spillover part of the trade equation in B will generate a buy sign. The same argument holds for spillover effect in the trade equation of B but with a higher weight since the difference in volatility in between the instruments has to be taken into account in the hedging process[10].

Next, we show the mid quote price and the efficient price behaviour for one of our simulations.

[9] See parameter c_i.

[10] These weights were chosen arbitrarily, but it is clear that one can also use the structural equation in order to find the optimal weight for hedging asset B with A.

Figure 1 – Simulated Efficient and Quoted Prices for Asset A and B



From Figure 1, it is possible to see that the quoted prices of Asset A (pink squares) has lower discrepancy to its efficient prices (red line), when comparing to asset B. This is mainly the consequence of parameter b_i, which increases the sensitivity of the quotes with respect to the incoming of a trade (buy or sell). For Figure 1, we can also see an intrinsic linear correlation between prices in B and A. This is the effect of the common risk factor v_t^m. The simulation we present is simply an illustration of our ideas, but it already points out some intuition on how our microstructure model implies correlation on trades and efficient prices.

Methodology

The methodology of this paper is divided into two steps. First we define the origin of the data and also data handling issues. Second, we describe the models estimated in the research.

The MTS Platform

Before explaining the data, it is worthwhile to describe the platform on which the European bonds are traded. Such a system is composed of a primary and a secondary market. The primary market is the local platform of each member of the Eurozone (e.g. Italy, Germany, etc). Each of these members defines locally the ways to finance their own operations with debt instruments. In general, the local platform (primary market) will present a higher variety of fixed income asset classes when compared to the European platform (secondary market). Both platforms have the same trading hours following central European time (CET). Each day starts with a pre-market phase (from 7:30 -8:00), a pre-open phase (8:00-8:15) and open phase (8:15 – 17:30). For the pre-market and pre-open part of the day, participants can submit orders and post proposals but no trades are effectively executed. These proposals are then ranked following price-time priority and are matched when the market officially opens (8:15 CET).

The secondary market for European bonds is the European platform. This is provided by MTS Global Markets[11]. This company provides the interdealer electronic platform (EuroMTS) for trading European benchmark bonds. It not only covers government bonds but it also provides trading for high quality non government bonds such as mortgages and public state loans. The members of this platform can either be a participant or a market maker. While the participants (dealers) have no particular obligation towards the system, the market makers (primary dealers) are obliged to provide quotes under specific restrictions for each asset class. It should also be pointed out that this is an exclusive interdealer market composed of highly capitalized banks; therefore an individual cannot participate directly.

Some of the bonds, depending on particular requirements such as the principal amount outstanding and the available number of dealers, may acquire the Euro “benchmark” status, meaning that it can be traded on the local and European platform. Arbitrage traders with access to both markets should eliminate price discrepancies. Also, it is a multi dealer’s market meaning that there might be several market makers for each asset class. Therefore, it is possible to have more than one market maker quoting each specific bond.

These market makers (or primary dealers) have an obligation to continuously provide two way quotes on the particular market they operate during trading hours. These quotes are anonymous[12] and remain valid for the day until cancelled, altered or matched by a trade. The maximum spread for each bond is also defined and set according to market and maturity. The proposals can be formulated for a minimum quantity equal to 10, 5, 2.5 or 1 million Euros. Odd volume lots can also be traded depending on market maker’s acceptance. The market takers (or simply dealers) are relatively passive in the trading operation and can only either hit the bid or ask quotes of the order book.

Further description of the trading platform for European bonds, including detailed information for the Italian and German local markets, can be found in the paper of (Cheung, et al., 2005) and (MTS Group, 2007).

The Data

The data for this study were kindly provided by the MTS group in association with the ICMA Centre. Along with full description of each instrument the original data includes fills (the trades) and posted quotes, which contains any change in the top 3 levels of prices on the order book. From this data, firstly we select the time period of 2004-2005 and also restrict it to only government bonds. We select these types of bonds and the year of 2004 given their higher liquidity (measured by number of trades) when comparing to their counterparts. The selected countries are Italy, Germany and France. These are selected given their higher trading volume (these 3 countries represent approximately 76% of all trades in the dataset for the period of 2004). For each of these, we make a further selection of bond types (fixed coupon bonds) and across different maturities (1, 2, 5, 10 maturity years (from 01/01/2004)). We use a band of a half year to classify the maturities. That is, a bond maturing in 01/06/2005 would

11 This is the result of the merger between MTS Spa and EuroMTS. The MTS stands for “Mercato dei Titoli de Stato”.

12 The identity of the counterparties, however, is revealed after the trade for clearing and settlement purposes. This anonymous system was implemented in 1997. Before this period, the market maker’s identity was visible by all participants.

be considered a 2 year maturity from 01/01/2004. When faced with 2 or more possible selections given a maturity band, we select the bonds with highest number of trades. Using these criteria, the final selection comprises the following bonds:

Table 1 – Bonds in Study

Bond Code (ISIN Code)	Bond Type	Market Code	Issue Date	Maturity Date	Maturity in Years (from 01/01/2004)*
IT0003248512	BTP	EBM&MTS	01/03/2002	01/03/2005	1
IT0001488102	BTP	MTS	15/06/2000	15/12/2005	2
IT0003652077	BTP	EBM&MTS	15/04/2004	15/04/2009	5
IT0003472336	BTP	EBM&MTS	01/02/2003	01/08/2013	10
DE0001137024	DEM	EBM&GEM	17/06/2003	17/06/2005	1
DE0001137057	DEM	EBM&GEM	10/03/2004	10/03/2006	2
DE0001141430	DEM	EBM&GEM	10/10/2003	10/10/2008	5
DE0001135242	DEM	EBM&GEM	31/10/2003	04/01/2014	10
FR0104756962	BTA	EBM&FRF	12/01/2002	12/01/2005	1
FR0106589445	BTA	EBM&FRF	12/03/2004	12/03/2006	2
FR0106589437	BTA	EBM&FRF	12/01/2004	12/01/2009	5
FR0010011130	OAT	EBM&FRF	25/10/2002	25/10/2013	10

* The number of years is given by using a band of half year around each maturity date.

For Table 1, the bond Code column is the ISIN[13] nomenclature for the different assets. One can also see that most of the bonds are traded in different markets. For instance, the one year Italian bond IT0003248512 is traded on the Italian platform (MTS) and also on the European market (EBM). With respect to the bond types, the BTP (*"Buoni del Tesoro Poliennali"*) bonds are bullet[14] bonds with different maturities paying a fixed coupon rate annually or semi-annually. The French BTA and the Deutshe DEM are also fixed coupon bonds. But, the French OAT bonds have coupons linked to the price index. This was the only 10 year French bond we could find in the database. While it would be desirable to have only fixed coupon paying bonds, we keep the French OAT in order to have a country wise symmetry for the different maturities in the dataset.

The raw data has a several issues which have to be dealt with before the estimation of any model. Next, we describe these issues and the steps taken in the paper.

Data Handling

Usually, raw high frequency data is very noisy and can incorporate different types of unwanted effect. Therefore, the handling of high frequency data is a very important part of empirical microstructure research. In this study we opt for using an event[15] time structure for the creation of all variables. A big part of the research was filtering and handling the quotes and trades database. This procedure was far from trivial. In order to make it clear, we organized it in a set of steps. These phases are run sequentially in the algorithm. The steps are:

[13] The ISIN code is a unique international label for each bond. Those are assigned by each country's numbering agency.
[14] These are bonds that cannot be redeemed prior to maturity.
[15] We also ran the models for aggregated data in calendar time (2 and 5 minutes intervals). The final conclusions are very robust.

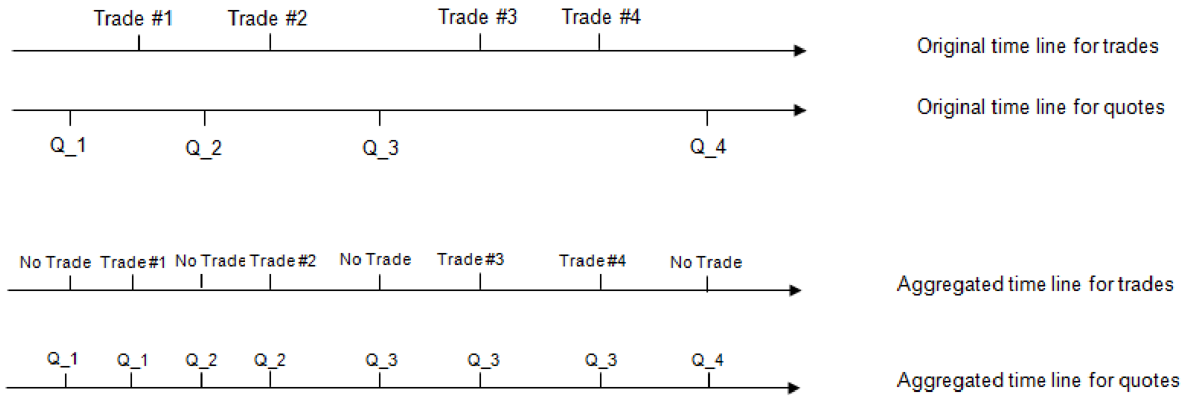
1. Avoiding overnight noise. The MTS market opens at 8:15 – 17:30 CET (central European time). For the pre-open phase of the market, an order can be send but it will not be matched. The automatic matching process occurs when the market opens so one might see a significant number of trades in the first minutes after the opening. For the research, in order to avoid noise in market opening hours we only use events between 8:30-17:30 CET.
2. Deleting events with no change in quote's value or volumes. The original data contains events in any part of the order book. We are particular interested in the movements in the top of the order book. Therefore, we delete any event where there is no price or volume improvement in the first level of bid and ask prices.
3. Finding best prevailing quotes. One of the problems in the data is that these bonds can be traded in the local or in the European platform. While some studies were concerned about different aspects of these markets (see Cheung et al (2005)) we treat this duality as a single market. Following this logic, for the cases where a bond is traded on different markets, we find the best prevailing quotes[16]. As an example, suppose an Italian bond is trading at 95 bid, 95.1 ask units on the MTS for time *t*. Now, a new quote comes to the EBM market at *t*+1 with the prices 95.05 bid, 95.2 ask. While this quote has improved the bid price (95.05>95), it has not improved the ask price. So, for time *t*+1, the prevailing quotes following the single market logic are 95.05 bid, 95.1 ask.
4. Filtering for large quote prices changes. The market makers in MTS are not obliged to provide quotes in all trading hours or the day. It may happen that for some time window, there are no primary dealers providing competitive quotes for a particular bond. If this event is true and a large trade comes through, it will consume the order book in the same direction of the trade. This joint event, if it happens, results in high price movements in the bid ask quotes (very low (high) for bid (ask)). Since the mid quote change is one of the dependent variables in our model, we are particularly interested in removing such noise. The approach here is to find any price movement that is higher than an arbitrary 5% percent threshold. When these cases are found, the "inadequate" prices are substituted by the previous prices[17]. This procedure produces a smooth behaviour for the bid/ask prices and consequently, a smooth behaviour for the mid quote price changes.
5. Filter for algorithmic quote changes. For the quotes data, when there is a change of market from t_i to t_{i+1} and the duration between the events is small (lower than 0.1 seconds, $t_i - t_{i-1} < 0.1$), we treat this as one quote change by the same market maker (algorithm quoting). The intuition here is that, even though this quote may be a genuine unique quote change within 0.1 second from the last one, we are trying to avoid an algorithmic quote change by the same market maker in different markets. This counts as a single quote change and should be treated as such.

[16] Curiously, using this scheme we find cases of negative spreads, meaning that one could make a riskless profit by buying/selling in one market and selling/buying in the other. But, understandably, the number of cases is small when compared to the sample size and these negative spreads tend to last for a short period of time before quotes are again changed.

[17] The consequence of this procedure is the assumption that if there was a market maker, he/she would not price the impact of these large trades in the quotes. While this is unrealistic, it is more intuitive (and simpler) than assigning an arbitrary impact in the prices.

6. Time aggregation of trades and quotes. For our high frequency data, the changes in quote and trade prices are naturally happening in different calendar times. This results in vectors of different sizes which are not suitable for standard VAR estimation. While it would be possible to simply aggregate[18] it for an arbitrary Δt, we argue that this would result in a great loss of information. In order to solve that, we build a pooled process for trades and quotes. In Figure 1, the procedure is illustrated.

Figure 2 – Illustration of Data Aggregation



From Figure 2 it can be seen that trades and quotes are being sampled in event time. First, comparing the original time line for trades against the resulting aggregated time line, the rule is that, for the times where a quote has taken place (in a different time than a trade), we set a "no trade" flag. This means that the trading direction and the volume are both set to zero. For the second time line (quotes), the rule is based on searching for the best prevailing quotes. So, when there is a trade, the values for the quotes are set as the prevailing ones, which is just the value for the previous event.

After the filtering and handling of the original data, we are left with the following descriptive statistics of the sample, along with a report on the data handling:

[18] E.g. taking averages in each time interval.

Table 2 – Descriptive Statistics and Data Handling Report

Panel A - Descriptive Statistics

Bond Code	Proportion of Buy Trades (%)	Proportian of Sell Trades (%)	Mean of log Return (mid price)	Standard Deviation of log Return (mid price)	Skewness of log Returns (mid price)	Kurtosis of log Returns (mid price)	Mean of Quoted Spread	Standard Deviation of Spread
IT0003248512	0.4752	0.5248	-0.0013	0.3993	-5.5259	46,620.0800	0.0068	0.0419
IT0001488102	0.4606	0.5394	-0.0013	0.3102	-0.4751	436.1399	0.0148	0.1317
IT0003652077	0.5662	0.4338	0.0006	0.5797	-3.9947	6,559.2553	0.0184	0.0212
IT0003472336	0.5261	0.4739	0.0015	0.8541	20.1144	11,585.2035	0.0229	0.0258
DE0001137024	0.5313	0.4687	0.0005	0.2547	0.0023	429.1207	0.0220	0.0068
DE0001137057	0.4937	0.5063	-0.0002	0.2627	1.5626	533.3570	0.0183	0.0062
DE0001141430	0.6739	0.3261	0.0012	0.6296	4.3150	1,736.5686	0.0252	0.0079
DE0001135242	0.5061	0.4939	0.0023	0.8029	-2.4362	1,619.7376	0.0260	0.0145
FR0104756962	0.4792	0.5208	-0.0023	0.1356	-2.9457	311.6521	0.0178	0.0074
FR0106589445	0.6244	0.3756	-0.0001	0.2681	-5.7099	598.4070	0.0184	0.0072
FR0106589437	0.4163	0.5838	0.0009	0.7732	2.9861	22,085.2395	0.0255	0.0809
FR0010011130	0.3923	0.6077	0.0023	0.7715	-0.5302	1,883.1036	0.0261	0.0179

Panel B - Data Handling Report

Bond Code	Number of Trades (original Data)	Number of Quotes (original Data)	Number of Deletions (zero price/volume movement)	Number of Deletions (short duration)	Number of Replacements (High Price Moves)	Number of Negative Spreads	Number of Quotes (after data Handling)
IT0003248512	3,169	323,828	179,065	20,597	0	121	124,166
IT0001488102	15,617	232,320	76,528	0	4	0	155,792
IT0003652077	6,657	303,672	73,838	40,094	22	762	189,740
IT0003472336	7,385	555,148	116,365	82,706	31	590	356,077
DE0001137024	559	162,752	60,409	15,076	0	3	87,267
DE0001137057	557	182,687	57,601	17,845	58	12	107,241
DE0001141430	417	262,589	51,387	40,623	18	46	170,579
DE0001135242	575	347,313	70,730	56,208	99	431	220,375
FR0104756962	144	151,145	85,766	10,242	12	4	55,137
FR0106589445	631	218,033	77,095	21,110	62	61	119,828
FR0106589437	800	313,301	65,965	40,906	182	366	206,430
FR0010011130	826	366,970	73,748	49,545	86	395	243,677

Table 2 shows a statistical description of the data used in this study. Panels A describes the main statistics of the trade direction (buy/sell), log return from the mid quote price and also the spread, which are the most significant variables in the dataset. One can see from the first two columns of Panel A that the proportion of buys and sells is relatively comparable for each bond in the sample. When one looks at the standard deviation from the mid quote prices, its possible to see that the values are mostly increasing with the maturity of the bonds. Remember that the column bond code is sorted first by country (see two initial letters in ISIN column) and then by maturity (see Table 1). Therefore, for higher maturities we observe a higher variability in the mid quote price log return. This is intuitive as bonds with higher maturity are more sensitive (price wise) to interest rate changes (see (Martellini, et al., 2003)). So, if there are unexpected news regarding the effective interest rate, those bonds with higher maturity will present a higher variation in the price, increasing consequently the standard deviation of the mid quote log return. The last argument can also be extended to the kurtosis of the log returns, meaning that we see higher frequency of extreme price movements for bonds with higher maturity. But it is not as quantitatively clear for kurtosis as it is for the standard deviation.

On a liquidity space, we also see that on average the mean spread grows with maturity. This could also be explained by the fact that long term bonds are more sensitive to interest rate changes. Therefore, if a trader has privileged information regarding a change in the interest rate, he will maximize his profits by trading the longer maturities. The market maker is aware of this effect and he will, in average, increase the spread[19] in response to this natural motivation for informed traders to trade the longer maturities. Similar arguments can be made with respect to the inventory effect. Since longer maturities have higher volatility, then holding this inventory is riskier and should also be compensated in the form of a higher spread premium.

For Panel B, Table 2, we show the summary from the data handling stage of the research. The first columns show the number of trades and quotes in the original data and the following columns show how much deletions were executed. Clearly the deletion of zero price movement was the most aggressive filter where, on average, 31% of the original quote data was discarded. This means that, on average, 31% of the events in the order book were for the second and third level of the order book, which was not relevant information to this research.

The second most aggressive filter was the deletions of short intermarket durations, which amounted to nearly 10% of the original data (except for the first bond, which is only traded in the local platform). These durations were labelled as parallel dual market quoting and, from what we can see in Panel B, it seems that this is not an uncommon practice in the market place. Understandably, a market maker is better prepared to provide competitive liquidity if he can, given fast changing market conditions, quickly update his quotes in the local and European platforms. So, the investment in trading software is justified and this presence of algorithmic quoting is not a surprise.

Another interesting feature of Panel B is that the number of negative spreads is mostly high for long term bonds. Remember that the bonds with longer maturity are the ones with higher sensitivity to interest rate unexpected news. So, this high volatility combined with the different liquidities of both platforms (European and local) can result in market makers in one platform quickly updating their prices with respect to unexpected news, while the other platform lags its updating process. This can result in a temporary negative spread in the

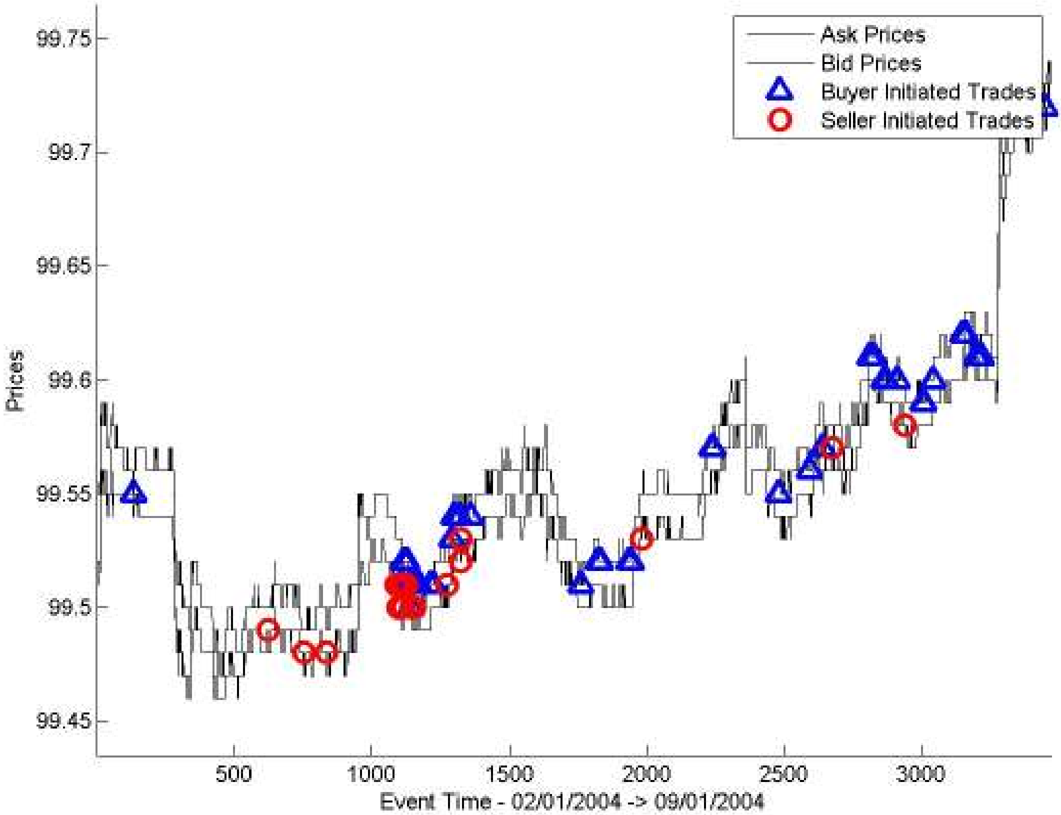
[19] When trading with informed traders, the market makers are likely to lose wealth. These losses are recovered by increasing the spread and trading with uninformed liquidity traders. See (Glosten, et al., 1985) for details.

market, implying the existence of risk free arbitrage profits. Note though that the number of negative spreads is small compared to the number of quote changes. While this is an interesting feature of the data, it is directly not related to our research so we decide to only report it, without any sort of adjustment.

From the final columns of Panel B, Table 2, we see that the data are mostly composed of quote updates. On average, across all bonds, there are approximately 199 quote updates for each trade. The Italian bonds (the first 4 entries in the table) are the most traded, with an average number of quotes per trade equal to approximately 31, which is significantly lower than for the rest of the countries (283 quote changes for each trade).

Next, we show a time series plot of a small subset of the data, particularly trades and quotes for the 1 year German Bond (DE0001137024) between the dates of 02/01/2004 and 09/01/2004.

Figure 3 – Quote Prices and Trades for DE0001137024



The sample plotted in Figure 3 was obtained after the pre-processing of the data and the scale of the horizontal axis is in event time. The red (blue) dots are the trades which were buyer (seller) initiated[20]. Not surprisingly, all buyer initiated trades were traded at the ask prices, while all sell initiated trades were executed at the bid price. Interestingly, from Figure 3 we

[20] These were not estimated from the data. The original database already contains the information regarding the side of the aggressor. These are identified according to the trade originator, that this, the entity which has hit the ask or bid prices.

see a significantly higher number of quote price variations than trades, which is also corroborated within the information in Table 2, Panel B. As an example, from event number (x axis) 0 to 500 only one sell trade is observed, while the quotes prices are showing a great deal of variation. This is also true for other intervals in the figure. These "unexplained" price changes can be caused by fundamental changes in the efficient price (e.g. news regarding interest rates).

The Models

The models used in this study are based on the theoretical implications for a microstructure model with spillover effects defined in the previous section of the paper. In total, we estimated three different models. We start with the benchmark representation (Hasbrouck, 1991) and we further extend it with the spillover variables in order to test our hypothesis, which is that the trading process in other assets affects the trading process of each instrument.

Model 1 – Hasbrouck (1991)

The econometric model of (Hasbrouck, 1991) is defined as:

$$r_{i,t} = \sum_{k=1}^{p} \beta^{r}_{i,k} r_{i,t-k} + \sum_{k=1}^{p} \gamma^{r}_{i,k} Q_{i,t-k} + \varepsilon^{r}_{i,t} \tag{11}$$

$$Q_{i,t} = \sum_{k=1}^{p} \beta^{Q}_{k} Q_{i,t-k} + \sum_{k=1}^{p} \gamma^{Q}_{k} r_{i,t-k} + \varepsilon^{Q}_{i,t} \tag{12}$$

Where:

$r_{i,t}$ − log return of mid quote for bond *i*, time *t*

$Q_{i,t}$ − Signed volume for bond *i*, time *t*

For this model, Equations (11) to (12), the idea is that trades and quotes evolve over time in a joint process[21]. The theoretical justification for the multivariate model was given previously in the paper. The $\beta^{r}_{i,k}$ parameter in Equation (11) measures the autoregressive part of the return equation. It is usually negative given the bid ask bounce effect[22], that is, a positive mid quote change is usually followed by a negative change. The parameter $\gamma^{r}_{i,k}$ measures the impact that the signed volume has on the mid quote changes, i.e. the impact of a trade. We

[21] Note that the original Hasbrouck model also had a contemporaneous trade variable in the return equation. For our case, given the way that we structured the data in event time, such a term is meaningless since at the time of a quote change, there are, by definition, no trades.

[22] When applied to trade prices, the bid/ask bounce is the result of impatient liquidity traders coming to the market and trading on the bid/ask quotes. This negative autocorrelation in traded prices changes is a function of fixed transaction costs such as the spread. Formal proofs can be seen in (Jong, et al., 2009) and (Hasbrouck, 2007). For mid quote values, this negative autocorrelation can be explained by the fact that the mean reversion of prices over a particular level (see Eq. (7)) results in a negative correlation for the price differences.

expect this parameter to be positive, that is, the incoming of buy/sell orders will increase/decrease the mid quote price.

For the trade equation of Hasbrouck model, (12), the parameter $\beta_{i,k}^{Q}$ measures the order flow of the trades, that is, the autoregressive pattern of the signed volume. While for the autoregressive representation given in (5) the sign of this parameter should be negative since c, b and α in (5) are positive, empirically we expect it to be positive. That is, a buy/sell order is most likely to be followed by a buy/sell order. One of the explanations is that this is the effect of traders splitting large trades across time. Another possible explanation would be that an informed trader in need of liquidity executes a large order and is followed by momentum liquidity traders. Both events would result in consecutive trades in the same direction. The coefficient $\gamma_{i,k}^{Q}$ relates to the impact that a mid quote change has on the trades. It is usually negative meaning that a drop (rise) in the mid quote price is most likely to be followed by a buy (sell) order. The usual explanation is the inventory problem. A market maker in need of reducing his exposure will improve quotes in order to induce trades.

Model 2 – Spillover Effects

For the extension of (11) - (12), which is the basis of our research, we start with the inclusion of two variables, the aggregated signed volume and the aggregated mid quote change. Again, this has foundations in the autoregressive representation of the theoretical spillover microstructure model defined earlier in the paper. This extension results in the next equations.

$$r_{i,t} = \sum_{k=1}^{p} \beta_{i,k}^{r} r_{i,t-k} + \sum_{k=1}^{p} \gamma_{i,k}^{r} Q_{i,t-k} + \sum_{k=0}^{p-1} \lambda_{1,i,k}^{r} r_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{r} Q_{i,t-k}^{AGG} + \varepsilon_{i,t}^{r} \tag{13}$$

$$Q_{i,t} = \sum_{k=1}^{p} \beta_{i,k}^{Q} Q_{i,t-k} + \sum_{k=1}^{p} \gamma_{i,k}^{Q} r_{i,t-k} + \sum_{k=0}^{p-1} \lambda_{1,i,k}^{Q} r_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{Q} Q_{i,t-k}^{AGG} + \varepsilon_{i,t}^{Q} \tag{14}$$

Where:

$r_{i,t}^{AGG}$ - Aggregated return for bond portfolio, betwenn t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio for time between t-1 and t.

$r_{i,t}$ − Log return of mid quote for bond i.

$Q_{i,t}$ − Signed volume of trade for bond i.

The aggregated variables are built as:

$$r_{i,t}^{AGG} = nSBonds^{-1} \sum_{j=1, j \neq i}^{nSBonds} \left[\exp\left(\sum_{z_j=1}^{n_j} \log \mathrm{Re}t_{j,z_j} \right) - 1 \right] \tag{15}$$

$$Q_{i,t}^{AGG} = \sum_{j=1, j \neq i}^{nSBonds} \sum_{z_j=1}^{n_j} P_{j,z_j} Vol_{j,z_j} Q_{j,z_j} \tag{16}$$

Before explaining the intuition behind the variables' creation, we need to explain how the data were aggregated. First we need to set how to define whether a bond is similar to another. For the paper we use the country as the criterion[23]. Exemplifying, for the one year Italian bond we will build a portfolio with the two, five and ten year Italian bonds. We choose to aggregate the bonds in a portfolio for reasons of simplicity[24]. This bond portfolio is then used to build the return and order flow variables in (15) and (16).

The idea behind the creation of $r_{i,t}^{AGG}$ and $Q_{i,t}^{AGG}$ is to measure the aggregate amounts of mid quote change and signed volume for the reference time *t*. For the first formula, (15), we are simply building the return of an equally weighted portfolio made of similar bonds. The term $\sum_{z_j=1}^{n_j} \log \mathrm{Re}t_{j,z_j}$ is the sum of log mid price changes for the same bond *j*, at time z_i, which gives the total logarithmic return for a bond in the portfolio for the period between *t* and *t*-1. Then, these returns are converted to the arithmetic formula and weighted by the number of bonds in question. For the outer sum in (15), the term *nSBonds* is the number of bonds which are similar to bond *i.* For our case, the value of *nSBonds* is constant and equal to three[25]. In the creation of these variables, when there is no trade or quote change for bond *j,* they are set to zero. For the creation of $Q_{i,t}^{AGG}$, Equation (16), we are interested in measuring the order flow for a portfolio of bonds. Each element $P_{j,z_j} Vol_{j,z_j} Q_{j,z_j}$ of the sum in (16) will give the money order flow, that is the amount of cash the position involved (price times volume), signed by the trade direction Q_{j,z_j}. When this value is summed across the different bonds, it gives the aggregated order flow of the bond portfolio in money terms.

Model 3 – Extended Hasbrouck's Model with intensity effects

The intuition behind this model is comparable to model 2, that is, we include the trading intensity variables following similar arguments in (Dufour, et al., 2000). The tweaking here is that the number of quotes and trades per unit of time is interacting with the mid price log returns of the bond itself[26]. That is, we are investigating whether the trading intensity in the similar bonds is increasing or decreasing the impact of explanatory variables in the benchmark model. The formal specification is given by:

$$
\begin{aligned}
r_{i,t} = \sum_{k=0}^{p-1} \lambda_{1,i,k}^{r} r_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{r} Q_{i,t-k}^{AGG} + \sum_{k=1}^{p} \left[\beta_{i,k}^{r} + \phi_{1,k}^{r} n_{i,t-k+1}^{T\&Q} \right] r_{i,t-k} + \\
\sum_{k=1}^{p} \left[\gamma_{i,k}^{r} + \phi_{2,k}^{r} n_{i,t-k+1}^{T\&Q} \right] Q_{i,t-k} + \varepsilon_{i,t}^{r}
\end{aligned}
\tag{17}
$$

[23] We also tested the robustness of the results by using a maturity criterion (instead of country). The main results of the research are very comparable. But we do find weaker evidence for a spill over effects across the bonds when the aggregation rule is based on the maturity instead of the country.

[24] Similar procedure was used in (Chordia, et al., 2000).

[25] Each bond has three counterparties from the same country.

[26] For the research, we also estimated a second version of model three where the intensity variables relates to the impact of $r_{i,t}^{AGG}$ and $Q_{i,t}^{AGG}$ over $r_{i,t}$ and $Q_{i,t}$. We do not find any surprising results, and therefore we do not formally report it. The estimated parameters can be found in appendix 4.

$$Q_{i,t} = \sum_{k=0}^{p-1} \lambda_{1,i,k}^{Q} Q_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{Q} r_{i,t-k}^{AGG} + \sum_{k=1}^{p} \left[\beta_{i,k}^{Q} + \phi_{1,k}^{Q} n_{i,t-k+1}^{T\&Q} \right] Q_{i,t-k} + \sum_{k=1}^{p} \left[\gamma_{i,k}^{Q} + \phi_{2,k}^{Q} n_{i,t-k+1}^{T\&Q} \right] r_{i,t-k} + \varepsilon_{i,t}^{Q} \quad (18)$$

Where:

$n_{i,t}^{T\&Q} = \left(nTrades_{i,t}^{AGG} + nQuotes_{i,t}^{AGG} \right) \left(\Delta t \right)^{-1}$ - Number of trades and quote changes for the bond portfolio related to bond i, for each unit of time.

Δt - Number of seconds between event t-1 and t.

$r_{i,t}^{AGG}$ - Return for the bond portfolio, happening between t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t}$ − log return of mid quote for bond i.

$Q_{i,t}$ − Signed Volume of Trade for bond i.

Note that for last equation we added duration (time between events) indirectly by using Δt, which is in the denominator of $n_{i,t}^{T\&Q}$. Therefore, the model in (17) and (18) has (Dufour, et al., 2000) as a nested case[27]. All of the models exposed before were estimated by ordinary least squares and we use the Newey and West robust covariance matrix for the standard errors. We also did not allow for day spillover in the lagged part of the models. If observation t is using information from the previous day in the lagged regressors, we exclude it from the estimation. With this procedure, one can say that we are performing a weighted OLS estimation.

Results

Next, in the Table 3, we show the results from the estimation of the benchmark model, (Hasbrouck, 1991). For the analysis of results in the next section, we choose to only look at the sum of parameters in each dependent variable. That is, we assess the long term impact of the regressors over the explained variable. We compute the null hypothesis that this sum is equal to zero with a formal Wald test. This is the same approach used in (Hasbrouck, 1991)), (Dufour, et al., 2000), (Furfine, et al., 2005) and greatly facilitates the analysis of our large scale model.

[27] On a side note, we also estimated the models without the duration variables and the conclusions from the results were very similar and so the duration variable, solely, is not driving our results.

Table 3 – Parameters from Hasbrouck Model, Equations (11) – (12)

Panel A - Quote Equation for Hasbrouck Model, (11)

bondCode	Sum of Betas	Sum of Gammas	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	-0.02	0.01***	11,187.11***	311.67***	0.01
IT0001488102	-0.38***	0.05***	116,849.01***	67.85***	0.07
IT0003652077	0.06***	0.04***	9,727.53***	12.15**	0.01
IT0003472336	-0.17	0.06***	143,335,542.63***	172.8***	0.02
DE0001137024	-0.33***	0.04***	2,198,547.55***	79.74***	0.04
DE0001137057	-0.06	0.02***	89,178.91***	10.12*	0.01
DE0001141430	0.1***	0.04***	17,709.12***	2.54	0.01
DE0001135242	0.17***	0.09***	60,858.99***	3.42	0.02
FR0104756962	-0.08***	0.02***	9,615.32***	155.51***	0.01
FR0106589445	-0.13***	0.03***	17,151.05***	20***	0.01
FR0106589437	0.04	0.03***	9,313.99***	2.73	0.01
FR0010011130	0.17***	0.06***	91,796.95***	14.83**	0.02

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following (Newey, et al., 1987).

The Model is:

$$r_{i,t} = \sum_{k=1}^{p} \beta_{i,k} r_{i,t-k} + \sum_{k=1}^{p} \gamma_{i,k} Q_{i,t-k} + \varepsilon_{i,t}^{r}$$

$r_{i,t}$ – log return of mid quote for bond i, time t (times 10,000 (basis point scale)).

$Q_{i,t}$ – Signed volume of trade for bond i, time t (divided by 1,000,000 (minimum trading quantity))

Panel B - Trade Equation for Hasbrouck Model, (12)					
bondCode	Sum of Betas	Sum of Gammas	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	0.3***	-0.12**	120,126.82***	28.01***	0.04
IT0001488102	0.44***	-1.14***	65,807.52***	983.51***	0.08
IT0003652077	0.34***	-0.07***	137,617.94***	165.78***	0.05
IT0003472336	0.32***	-0.01*	324,542.59***	269.45***	0.05
DE0001137024	0.35***	-0.14***	366,281.29***	98.37***	0.06
DE0001137057	0.36***	-0.17**	551,812.74***	31.83***	0.06
DE0001141430	0.38***	-0.01*	1,735,692.13***	143.23***	0.06
DE0001135242	0.28***	0.01	1,065,429.35***	137.84***	0.04
FR0104756962	0.23***	-0.01	613,626.46***	174.6***	0.03
FR0106589445	0.34***	-0.12***	556,167.41***	155.98***	0.04
FR0106589437	0.45***	-0.01	1,381,104.58***	142.38***	0.10
FR0010011130	0.26***	-0.01	852,947.88***	-1.1	0.03

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following (Newey, et al., 1987).

The Model is:

$$Q_{i,t} = \sum_{k=1}^{p} \beta_k Q_{i,t-k} + \sum_{k=1}^{p} \gamma_k r_{i,t-k} + \varepsilon_{i,t}^{Q}$$

$r_{i,t}$ − log return of mid quote for bond i, time t (times 10,000).

$Q_{i,t}$ − Signed volume of trade for bond i, time t (divided by 1,000,000)

As one can see, the signs of most of the parameters are as expected. For Table 3, Panel A (the quote equation), the betas are negative for more than half of the cases and the gammas are all positive. This means that the bid ask bounce effect is present in the data, but it is not strongly defined. The gamma parameter in the quote equation is, for all bonds, positive and statistically significant at the 1% level. This indicates that a trade has an impact on the mid quote price. For example, on average across all bonds, the existence of five buy trades in the previous time periods, each with a volume of 1.000.000 bonds (the minimum), will result in a log return increase of approximately 0.034 basis points ($nBonds^{-1}\sum_{i=1}^{nBonds}\sum_{k=1}^{5}\gamma_{i,k} \cong 0.034$) in the mid-quote price. Another interesting piece of information from Panel A of Table 3 is that the value of the sum of gamma is, in a modest way, increasing within maturities of the bonds. This is intuitive since bonds with higher maturity have higher sensitivity to interest rate dynamics. Therefore, the information absorbed from the order flow regarding this risk factor should be priced more heavily for bonds with higher maturities.

For Panel B (trade equation), we see that the autoregressive pattern in the dependent variable (signed volume) is strongly defined, meaning that a buy (sell) order is most likely to be followed by another buy (sell) order. This is generally linked to traders splitting their whole order into small pieces over time, resulting in consecutive trades on the same side of the market. We also see from Table 3, Panel B, that the effect of a mid quote change on the trade signed volume is mostly negative, meaning that a decrease (increase) in the mid quote price is most likely to be followed by a buy (sell) order. An argument towards this effect is the inventory problem of the market makers. If a dealer has an excessive capital committed to a particular asset, then this risky exposure will force him to rapidly liquidate part of the position by submitting improvements in the bid/ask price. This results in a positive (negative) log return of the mid quote price being followed by a sell (buy) trade.

From Table 3, Panels A and B, its possible to check that serial correlation (Breush-Godfrey LM Test) is an issue for the majority of the estimated models[28]. We also see a great degree of heteroscedasticity, measured by the Breush-Pagan test. While the OLS parameters are unbiased and consistent in the presence of heterokesdasticity and serial correlation, the standard errors are not[29]. In order to overcome such an issue, we use a Newey and West robust covariance matrix[30] in all standard errors. These are then used as inputs in the Wald hypothesis tests.

Next, we follow the analysis for the second model, Equations (13)-(14).

[28] Note that one of the bonds has a negative value for Breush-Godfrey test which is, in general, unexpected. But, remember that the estimated trade equation doesn't include a constant so it is possible for the r squared of the auxiliary regression in the Breush-Godfrey test to be negative.

[29] These standard errors will be biased downwards, which is a dangerous feature for any hypothesis testing. See Maddala (1991) for further information.

[30] See (Newey, et al., 1987).

Table 4 – Estimates for Spillover Effect Model, Equations (13) - (14)

Panel A - Quote Equation, Spill Over Model (13)

bondCode	Sum of Betas	Sum of Lambda 1	Sum of Gammas	Sum of Lambda 2	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	-0.0414	0.005***	0.0077***	0.0002**	137,348.9360***	373.4935***	0.03
IT0001488102	-0.6008***	0.0314***	0.0413***	0.0006	4,725,625.2142***	469.1931***	0.16
IT0003652077	-0.1528***	0.0865***	0.0283***	0.0039***	7,597,210.0498***	184.2412***	0.11
IT0003472336	-0.325**	0.1932***	0.0545***	0.0069***	157,879,155.4331***	223.2181***	0.06
DE0001137024	-0.489***	0.0171***	0.0304***	0.0005	2,765,667.6841***	187.6324***	0.11
DE0001137057	-0.3522***	0.0333***	0.0205***	0.0022	5,873,121.2147***	207.2242***	0.14
DE0001141430	-0.2554***	0.1252***	0.0299***	0.0001	34,099,966.7761***	147.7299***	0.27
DE0001135242	-0.0418	0.1788***	0.0723***	-0.0104	17,733,809.5304***	60.1146***	0.23
FR0104756962	-0.1294***	0.0043***	0.0107***	0.0007*	291,468.4461***	136.8616***	0.05
FR0106589445	-0.3821***	0.0334***	0.0211***	0.0027*	5,879,967.5077***	290.9236***	0.13
FR0106589437	-0.1277**	0.1029***	0.0197***	0.0031	4,867,753.2456***	42.6651***	0.12
FR0010011130	0.0048	0.1481***	0.0492***	0.0022	127,597,980.7001***	20.1695***	0.19

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following (Newey, et al., 1987).

The quote equation is given by:

$$r_{i,t} = \sum_{k=1}^{p} \beta_{i,k}^{r} r_{i,t-k} + \sum_{k=1}^{p} \gamma_{i,k}^{r} Q_{i,t-k} + \sum_{k=0}^{p-1} \lambda_{1,i,k}^{r} r_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{r} Q_{i,t-k}^{AGG} + \varepsilon_{i,t}^{r}$$

$r_{i,t}^{AGG}$ - Return for the bond portfolio, happening between t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t} -$ log return of mid quote for bond i.

$Q_{i,t} -$ Signed volume of trade for bond i.

Panel B - Trade Equation, Spill Over Model (14)							
bondCode	Sum of Betas	Sum of Lambda 1	Sum of Gammas	Sum of Lambda 2	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	0.2902***	0.002	-0.1586**	0.0051***	120,782.2264***	26.2391***	0.04
IT0001488102	0.433***	0.0029	-1.5539***	0.0444***	61,463.7266***	1097.9178***	0.08
IT0003652077	0.336***	0.0047	-0.1021***	0.0139***	137,775.3885***	170.0312***	0.05
IT0003472336	0.3119***	0.0029	-0.0144***	0.0052**	325,336.3705***	270.7336***	0.04
DE0001137024	0.3407***	-0.002	-0.1751***	0.0038***	367,038.8018***	100.4445***	0.05
DE0001137057	0.3534***	0.0053	-0.2467***	0.0081***	556,557.2159***	39.4061***	0.06
DE0001141430	0.3712***	0.0072	-0.0111**	0.0012*	1,735,874.6295***	135.0773***	0.06
DE0001135242	0.2703***	0.0076	0.0003	0.0012	1,069,067.1379***	139.1213***	0.03
FR0104756962	0.2205***	0.001	-0.0419	0.0022***	613,396.3199***	172.768***	0.02
FR0106589445	0.3307***	0.0037	-0.1739***	0.0064***	556,578.1721***	161.6024***	0.04
FR0106589437	0.4496***	0.0082*	-0.0003	-0.0002	1,380,621.4669***	142.1956***	0.10
FR0010011130	0.253***	0.0162	-0.0055	0.0012	887,408.9280***	2.3998	0.02

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following (Newey, et al., 1987).

The trade equation is given by:

$$Q_{i,t} = \sum_{k=1}^{p} \beta_{i,k}^{Q} Q_{i,t-k} + \sum_{k=1}^{p} \gamma_{i,k}^{Q} r_{i,t-k} + \sum_{k=0}^{p-1} \lambda_{1,i,k}^{Q} r_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{Q} Q_{i,t-k}^{AGG} + \varepsilon_{i,t}^{Q}$$

$r_{i,t}^{AGG}$ - Return for the bond portfolio, happening between t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t} -$ log return of mid quote for bond i.

$Q_{i,t} -$ Signed Volume of Trade for bond i.

For Table 4, Panel A, the coefficients of interest are the lambda values. These are measuring the effect of the quote changes and trades in the bond portfolio[31] towards the mid quote price change of the bond in question. For lambda 2, we find that only five (out of twelve) are statistically significant. But the sign of this sum is consistent across bonds, where eleven out of twelve are positive. Therefore, we argue for modest evidence of the importance of the order flow in the bond portfolio towards price changes in each bond.

For lambda 1, we find a strong dependency in the data. For this parameter, across all bonds, all the values are positive and statistically significant. This implies the existence of a positive correlation between the changes in the mid quote price of a bond portfolio and the changes of the mid quote price of the bonds. Looking at the values of adjusted R squared in Table 4, Panel A, and comparing to the ones obtained with the Hasbrouck model, Table 3, Panel A, we see a significant increase. The average difference in adjusted R squares in between the models (Table 4 and Table 3) is 0.11, with a maximum of 0.27 and a minimum of 0.02. We also see a clear "maturity" pattern for the values of the sum of lambda 1, where the lower the maturity of the bond, the lower the value of sum of lambda 1. This pattern is the result of the country wise construction of the portfolio. In general, the lower (higher) the maturity, the higher (lower) the relative variability of the portfolio price changes when comparing to the variability of the bonds in question. The lambda 1 parameters are then scaled given this difference in variability.

This positive correlation between $r_{i,t}$ and $r_{i,t}^{AGG}$ is not surprising. Similar results were found for US equities in Hasbrouck and Seppi (2001). The explanation is that, since the bonds have similar risk factors, then it can be expected that fundamental changes will make the prices of these instruments move in the same direction. But the interesting part of this effect is that these changes are first seen in the bond portfolio and then it follows through to the bonds. Remember that all models are estimated in event time so that all information in the independent variables is available at the time of the event in the dependent variable. This means that a significant part of the variation in the mid quote prices of the bonds is statistically predictable. In general, we can calculate how much the mid price would move at each arrival of information in the bond portfolio but, without any explicit parameterization of the time vector (e.g. ACD models), we cannot predict when such expected change in mid-quote would occur[32].

In Table 4, Panel B, we see the results from the estimation of the trade equation with the spillover extension. Again, the parameters of interest are the lambda values. For the sum of lambda 1, we see the pattern that the great majority of these values are positive. This means that positive (negative) changes in the bond portfolio will be followed by buy (sell) orders in the bond in question. We can explain this effect as a lagged update of the bonds prices with respect to the incoming of news to the market. If news about a common factor comes to the market and first impact the prices of other bonds, traders will realize such event and will trade on the un-updated bond until the equilibrium is once again reached. This effect can explain the positive correlation between $Q_{i,t}$ and the lags of $r_{i,t}^{AGG}$. But it should be clear that for only 8% (1/12) of the bonds this sum is statistically different than zero. We conclude that the effect of the returns on the bond portfolio quotes over trading volume is robust, but not particularly significant.

[31] The construction of this portfolio was obtained using aggregation rules over the maturity of the bonds.
[32] It would be interesting to see how a high frequency trading strategy based on this predictability would perform in an out of sample setting. Such a path of research was not followed as it diverges from our main objectives.

For lambda 2 in Panel B (Table 4) we see that the coefficient is predominately positive and statistically significant. Out of twelve cases, eleven are positive as expected and nine are statistically significant at 5%. This means that the order flow (sign and volume) in the bond portfolio is significantly (and positively) affecting the individual trades in each bond as expected. The derivations of the autoregressive form for the trade equation in the first part of the paper predicted this effect. Therefore, one could argue that the results from the econometric estimation modestly corroborate with the theoretical expectations from our microstructure model and the correlation between $Q_{i,t}^{AGG}$ and $Q_{i,t}$ is the result of a spillover effect. But this effect of correlation in between the trades could also be explained by informed traders splitting their orders across the bonds in different countries (but the same maturities). A trader with information (e.g. long term interest rate dynamics) may start trading the longer (more sensitive) maturities but will fill other legs of the order with shorter maturities according to his assessment of the risk of disclosing his private information through the fragmented operation. Another possible explanation is that this correlation across trades in different instruments is simply the incoming of news. Since the bonds are linked by similar risk factor, the incoming of news will generate trades in all instruments, resulting in cross correlations. From the results in Table 4, it is not possible to distinguish between these three alternative explanations.

Summarizing the results in Table 4, the main point in Panel A is the strong dependency of the mid quotes changes of the bonds with respect to the aggregated changes in the bond portfolio. This is assessed as a fundamental change in the common risk factor underlying the bonds. For Panel B, the main information is the positive correlation between trades in the bonds and trades in the bond portfolio. In our theoretical formulation, such a result was expected. In a microstructure sense, the idea is that mispricing in other instruments is motivating trades in the bond in question. But this positive correlation could also be explained by informed traders splitting large trades across different maturities or by news regarding a common risk factor. As said before, we were not able to distinguish such effects based solely on our results in Table 4.

The next part of our research is the analysis of the extension of the basic spillover effect model. Following similar arguments as in (Dufour, et al., 2000), we include a proxy for trading intensity in the parameters of the model.

Table 5 – Results from Extended Spillover Effect Model, Equations (17) - (18)

Panel A – Quote Equation for extended Spillover Model, (17)

bondCode	Sum of Lambda 1	Sum of Lambda 2	Sum of Betas	Sum of Phis 1	Sum of Gammas	Sum of Phis 2	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	0.0489***	0.0003***	-0.008	-0.0006**	0.0077***	-0.0001	147,199.1788***	218.1524***	0.03
IT0001488102	0.3118***	0.0006	-0.6038***	0.0003	0.0399***	0.0002**	4,732,028.0560***	482.8989***	0.16
IT0003652077	0.8621***	0.004***	-0.1337***	-0.0007*	0.0257***	0.0004*	7,573,803.3008***	180.621***	0.12
IT0003472336	1.928***	0.007***	-0.3152*	-0.0003	0.0326***	0.0028	408,153,082.6327***	116.1719***	0.07
DE0001137024	0.1699***	0.0005	-0.4833***	-0.0001	0.0313***	-0.0001***	2,699,144.7083***	182.4706***	0.11
DE0001137057	0.3312***	0.0024	-0.3962***	0.0003***	0.0198***	0.0001**	5,559,206.7241***	256.4371***	0.15
DE0001141430	1.2531***	0.0007	-0.2232***	-0.0007	0.0298***	0.0001	35,916,371.8527***	149.8814***	0.27
DE0001135242	1.7881***	-0.0098	-0.0272	-0.0005***	0.0656***	0.0007***	18,104,335.5121***	58.8948***	0.23
FR0104756962	0.0421***	0.0006*	-0.1245***	-0.0001	0.0114***	-0.0001**	286,037.7872***	11.5922**	0.05
FR0106589445	0.3331***	0.0026	-0.3603***	-0.0004**	0.0214***	-0.0001***	5,853,640.6427***	266.494***	0.13
FR0106589437	1.0223***	0.0032	-0.1189*	-0.0002	0.0151***	0.0008*	9,878,086.2961***	73.0228***	0.12
FR0010011130	1.4823***	0.0024	0.0071	-0.0001	0.0411***	0.001*	125,706,183.4365***	38.818***	0.19

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following (Newey, et al., 1987).

The estimated equation is given by:

$$r_{i,t} = \sum_{k=1}^{p} \lambda_{1,i,k}^{r} r_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{r} Q_{i,t-k}^{AGG} + \sum_{k=1}^{p} \left[\beta_{i,k}^{r} + \phi_{1,k}^{r} n_{i,t-k+1}^{T\&Q} \right] r_{i,t-k} + \sum_{k=1}^{p} \left[\gamma_{i,k}^{r} + \phi_{2,k}^{r} n_{i,t-k+1}^{T\&Q} \right] Q_{i,t-k} + \varepsilon_{i,t}^{r}$$

$n_{i,t}^{T\&Q} = \left(nTrades_{i,t}^{AGG} + nQuotes_{i,t}^{AGG} \right) (\Delta t)^{-1}$ - Number of trades and quote changes for the bond portfolio, for each unit of time.

Δt - Number of seconds between event t-1 and t.

$r_{i,t}^{AGG}$ - Return for the bond portfolio, happening between t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t} -$ log return of mid quote for bond i.

$Q_{i,t} -$ Signed volume of trade for bond i.

Panel B – Trade Equation for extended Spillover Model, (18)									
bondCode	Sum of Lambda 1	Sum of Lambda 2	Sum of Betas	Sum of Phis 1	Sum of Gammas	Sum of Phis 2	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	0.0503***	0.0019	0.2982***	-0.0006***	-0.195**	0.0005	122,117.3531***	27.3542***	0.04
IT0001488102	0.4346***	0.0031*	0.4696***	-0.0041***	-1.7022***	0.0041***	60,121.5153***	1078.4487***	0.08
IT0003652077	0.1455***	0.0044	0.3923***	-0.0071**	-0.1068***	0.0003	262,022.3505***	172.4459***	0.06
IT0003472336	0.0807***	0.0028	0.4306***	-0.0162***	-0.0137**	-0.0001	385,657.0244***	250.3709***	0.05
DE0001137024	0.0378***	-0.0016	0.3485***	-0.0008*	-0.1722***	-0.0001	368,539.7946***	85.1548***	0.05
DE0001137057	0.0805***	0.0055	0.3653***	-0.0008**	-0.2684***	0.0002*	553,493.7352***	37.407***	0.06
DE0001141430	0.0141**	0.0071	0.3982***	-0.0029**	-0.0079*	-0.0001*	1,674,969.1831***	125.6865***	0.06
DE0001135242	0.0099	0.0074	0.2807***	-0.001	-0.0015	0.0001	1,059,156.6595***	136.8517***	0.03
FR0104756962	0.0238***	0.001	0.2253***	-0.0001***	-0.0501	-0.0001	615,208.3230***	26.6177***	0.03
FR0106589445	0.0608***	0.0038	0.337***	-0.0007*	-0.1918***	0.0004*	552,014.3138***	180.5666***	0.04
FR0106589437	0.0076	0.0083*	0.5324***	-0.0138***	0.0018	-0.0001*	1,917,011.2135***	197.1651***	0.11
FR0010011130	0.0115	0.0145	0.3878***	-0.0166***	-0.0048	0.0001	845,149.1493***	-18.4761	0.03

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following (Newey, et al., 1987).

The estimated equation is given by:

$$Q_{i,t} = \sum_{k=0}^{p-1} \lambda_{1,i,k}^{Q} Q_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \lambda_{2,i,k}^{Q} r_{i,t-k}^{AGG} + \sum_{k=1}^{p} \left[\beta_{i,k}^{Q} + \phi_{1,k}^{Q} n_{i,t-k+1}^{T\&Q} \right] Q_{i,t-k} + \sum_{k=1}^{p} \left[\gamma_{i,k}^{Q} + \phi_{2,k}^{Q} n_{i,t-k+1}^{T} \right] r_{i,t-k} + \varepsilon_{i,t}^{Q}$$

$n_{i,t}^{T\&Q} = \left(nTrades_{i,t}^{AGG} + nQuotes_{i,t}^{AGG} \right) \left(\Delta t \right)^{-1}$ - Number of trades and quote changes for the bond portfolio, for each unit of time.

Δt - Number of seconds between event t-1 and t.

$r_{i,t}^{AGG}$ - Return for the bond portfolio, happening between t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t} -$ log return of mid quote for bond i.

$Q_{i,t} -$ Signed volume of trade for bond i.

From Table 5, the parameter of interest is the sum of the phi coefficients. From Panel A, we see that the sum of the $\phi^{r}_{1,k}$ parameters is negative for 91% of the cases (ten out of twelve). But it is statistically different than zero at the 5% significance level for only six cases. In a modest way, this result implies that on average an increase in the quoting activity in the bond portfolio will strengthen the negative autocorrelation process for mid quote changes. A possible reasoning for this effect is that the increase in quoting activity in the bond portfolio relates to the presence of news in the market, which increases the volatility of the efficient price, increasing the spread and therefore making the bounce between bid and ask prices higher in absolute value. This would result in a stronger negative autocorrelation for mid quote changes. But, while we find consistency in the signs of the sum, the *p* values tell a different story. For only five cases we do find positive and statistically significant sums. Therefore this effect is weakly defined across different bonds.

When looking at the sums of $\phi^{r}_{2,k}$, Panel A, Table 5, we also find modest results. Among the bonds, the sum of this parameter is positive (and statistically significant) for six out of eight cases. While this is also not a strong result, the signs of the parameters are intuitive. A higher quoting activity in the bond portfolio indicates the presence of fundamental news in the market. Therefore, it can be expected that the market maker would be statistically sensitive to the activity in the market when pricing the incoming order flow. This results in a higher price impact when there is higher trading and quoting activity in the bond portfolio.

In Panel B, Table 5, we have the trade equation results of the extended spillover model. The sums of $\phi^{Q}_{1,k}$ show a consistent negative effect of trading and quoting activity towards the autocorrelation process of signed volume. For all bonds, the sum of this parameter is statistically significant and negative. This means that the higher the trading and quoting activity in the bond portfolio, the lower the chance that a buy (sell) order is followed by another buy (sell) order in the respective bond. The explanation we find for the negative signs of the sums of $\phi^{Q}_{1,k}$ is the inventory problem faced by the primary dealers (market makers). If the market is trading intensively, then it indicates the presence of news (or informed traders trading before news). If the impact of this news is uncertain, then there would be more motivation for the market maker to decrease his inventory limits given the higher risk profile of the situation. Therefore, while in a normal market situation the market maker would accommodate successive trades in the same direction, in a higher intensity trading situation the market maker would have more incentive to tighten the inventory control by improving quotes and motivating the trades in the opposite direction. This improvement of quotes corroborates with the result for the sum of parameters $\phi^{r}_{1,k}$ in the quote equation of the extended model (Table 5, Panel A) where an increase in quoting activity in the bond portfolio increases the autocorrelation process of the mid quote changes.

Conclusions

The main idea of the paper was to introduce spillover effects in a general microstructure model. We build the structural equations of which we draw our ideas and showed that it had a particular autoregressive representation for changes in mid quote prices and trades. Based on this result, we estimate different econometric specifications in order to test the consistency and robustness of the expected relationships in twelve fixed income instruments for the European bond market over the time period of one year.

The result of this exercise was positive. For the quote equation we see a significant increase in the adjusted R squared when comparing to the benchmark model in (Hasbrouck, 1991). This indicates the high explanatory power of some of the new variables. A strong positive correlation between the mid quote changes in the respective bonds and the bond portfolio was uncovered. This was expected as the bonds have similar risk factors and this correlation is the simple result of common fundamental changes in prices.

In the microstructure model built in the first part of the paper it was possible to show that the trade equation had an autoregressive representation which incorporated the trades in other instruments. This result was also seen empirically in the estimated econometric models, where a trade in the bond portfolio is positively related to trades in the respective bonds. But, as pointed out in the paper, the result was not strong and we were not able to distinguish the spillover effect of our microstructure model from the simple case of informed traders splitting orders across different maturities in the market. Therefore, we argue that the results of the study with respect to the robustness of the proposed microstructure model were positive but inconclusive.

In a second part of the paper, we follow the same arguments as (Dufour, et al., 2000) and investigate the effect of trading intensities on the parameters of our model. We find that these parameters do change according to the intensity with which the bond portfolio was being traded and quoted. First, we got the result that positive changes in the quoting activity of the bond portfolio strengthen the negative autocorrelation for mid quote changes. Second, we see that the trade and quote activity decreases the autocorrelation in the trade equation. Both results can be explained by the inventory problem, where the trading (and quoting) activity indicates news in the market, increasing the risk of market maker's exposures, and therefore forcing him to improve quotes in the opposite direction in order to motivate trades. This would cause stronger negative autocorrelation in the mid quote price changes and weaker positive autocorrelation in the trade equation for cases where the market is trading intensively.

The contributions of this paper to the existing literature rely on showing empirical evidence that trades and quotes of instruments with similar risk factors are not independent. In the document, we argued in favour of a simple microstructure model where the spillover effect had a transitory effect on prices. Even though we could not argue in favour of empirical evidence for our theoretical model given a higher complexity in the trading data, we do uncover a significant amount of evidence towards a spillover effect in the trading process of European bonds. The most reliable result we have is for the spillover of trading intensities in the trade equation. Further steps in the research would be to study empirically the effect of the spillover variables on the spread and volatility of the bonds, checking whether they have similar impacts as with the mid quote change and order flow.

References

Brandt M. and Kavajecz K. Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve [Journal] // ational Bureau of Economic Research, Working Paper Series. - 2003. - Vol. 9529.

Brennan M. and Subrahmanyam A. Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns [Journal] // Journal of Financial Economics. - 1996. - Vol. 41. - pp. 441-464.

Cheung Y., Jong F. and Rindi B. Trading European Sovereign Bonds - The Microstructure of the MTS Trading Platform [Journal] // European Central Bank, Working Paper. - 2005. - Vol. 432.

Chordia T., Roll R. and Subrahmanyam A. Commonality in Liquidity [Journal] // Journal of Financial Economics. - 2000. - 1 : Vol. 56.

Chordia T., Roll R. and Subrahmanyam A. Market Liquidity and Trading Activity [Journal] // Journal of Finance. - 2001. - 2 : Vol. 56.

Dufour A. and Engle R. Time and the Price Impact of a Trade [Journal] // Journal of Finance. - 2000. - Vol. 55. - pp. 2467-2498.

Dufour A. and Skinner F. MTS Time Series - Market and Data Description for the European Bond Market and Repo Database [Journal] // ICMA Centre Working Paper. - 2004. - Vol. 3.

Easley D. and O'Hara M. Price, trade and information in securities markets [Journal] // Journal of Financial Economics. - 1995. - Vol. 19. - pp. 69-90.

Easley D. and O'hara M. Time and the Process of Security Price Adjustment [Journal] // Journal of Finance. - 1987. - Vol. 47. - pp. 577-605.

Engle R. and Russel J. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data [Journal] // Econometrica. - 1998. - 5 : Vol. 66. - pp. 1127-1162.

Furfine C. and Remolona E. Price Discovery in a Market Under Stress: the U.S. Treasury Market in Fall 1998 [Journal] // Federal Reserve Bank of Chicago - Working Paper. - 2005. - Vol. 06.

Glosten L. and Milgrom P. Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents [Journal] // Journal of Financial Economics. - 1985. - Vol. 14. - pp. 71-100.

Hasbrouck J. and Seppi D. Common Factors in Prices, Order Flows and Liquidity [Journal] // Journal of Financial Economics. - 2001. - 3 : Vol. 59.

Hasbrouck J. Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading [Book]. - [s.l.] : Oxford University Press, 2007.

Hasbrouck J. Measuring the Information Content of Stock Trades [Journal] // Journal of Finance. - 1991. - Vol. 46. - pp. 179-207.

Hasbrouck J. Modeling Market Microstructure Time Series [Journal] // New York University - Stern School of Business - Working Paper Series. - 1995. - Vols. FIN-95-24.

Jong F. and Rindi M. The Microstructure of Financial Markets [Book]. - New York : Cambridge Univeristy Press, 2009.

Kyle A. Continuous Auctions and Insider Trading [Journal] // Econometrica. - 1985. - 6 : Vol. 53. - pp. 1315-1335.

Maddala G. Introduction to Econometrics [Book]. - [s.l.] : John Wiley & Sons, 1991. - 3.

Martellini L., Priaulet P. and Priaulet S. Fixed-Income Securities, Risk Management and Portfolio Strategies [Book]. - [s.l.] : Jonh Wiley & Sons, 2003.

MTS Group Bond Markets in Europe and Beyond [Report]. - [s.l.] : MTS Group, 2007.

Newey B. and West D. A Simple, Positive semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix [Journal] // Econometrica. - 1987. - Vol. 55. - pp. 347-370.

White H. Asymptotic Theory for Econometricians [Book]. - New York : Academic Press, 1984.

APPENDIX

Appendix 1 – Derivation of Autoregressive Form of Hasbrouck Model

The microstructure model in Hasbrouck (1991) is given by this set of recursive equations:

$$m_t = m_{t-1} + z\nu_t^x + \nu_t^m \tag{A.1.1}$$

$$q_t = m_t + \alpha\left(q_{t-1} - m_{t-1}\right) + bx_t \tag{A. 1.2}$$

$$x_t = -c\left(q_{t-1} - m_{t-1}\right) + \nu_t^x \tag{A. 1.3}$$

Our interest is in finding the autoregressive form of the mid quote price changes, $q_t - q_{t-1}$. We start by adding $-q_{t-1}$ in both sides of the equation, resulting in:

$$q_t - q_{t-1} = m_t + \alpha\left(q_{t-1} - m_{t-1}\right) + bx_t - q_{t-1} \tag{A. 1.4}$$

Substituting the value of m_t and ν_t^x we get:

$$q_t - q_{t-1} = \left(m_{t-1} - q_{t-1}\right)\left(1 - \alpha - zc\right) + \left(z + b\right)x_t + \nu_t^m \tag{A. 1.5}$$

Now, rearranging and substituting for $q_{i,t-1}$ we get:

$$q_t - q_{t-1} = \alpha\left(m_{t-2} - q_{t-2}\right)\left(1 - \alpha - zc\right) + \left(zbc - b\left(1 - \alpha\right)\right)x_{t-1} + \left(z + b\right)x_t + \nu_t^m \tag{A.1.6}$$

This is a similar equation as in formula (10) in Hasbrouck (1991). The author further expands the equation by substituting for q_{t-2}. But, the pattern is already clear, each expansion of q_{t-p} will release $\alpha^{p-1}\left(zbc - b\left(1-\alpha\right)\right)x_{t-p}$. Using this simplification and generalizing we get:

$$\Delta q_t = \left(z + b\right).x_t + \sum_{k=1}^{\infty} \alpha^{k-1}\left(zbc - b\left(1 - \alpha\right)\right)x_{t-k} + \nu_t^m \tag{A. 1.7}$$

Therefore, this autoregressive system has a decaying value of lag parameters as long as $\alpha < 1$
.
For the derivation of the autoregressive system for the trades, $x_{i,t}$, we start with the trade equation:

$$x_t = -c\left(q_{t-1} - m_{t-1}\right) + \nu_t^x \tag{A. 1.8}$$

Substitute for q_{t-1}:

$$x_t = -cb.x_{t-1} - \alpha c\left(q_{t-2} - m_{t-2}\right) + \nu_t^x \tag{A. 1.9}$$

Again substitute for q_{t-2}:

$$x_t = -cb.x_{t-1} - \alpha cb.x_{t-2} + \left(\alpha\left(q_{t-3} - m_{t-3}\right)\right) + \nu_t^x \tag{A. 1.10}$$

But, the pattern is already clear. The generalized equation for trades will be:

$$x_t = -\sum_{k=1}^{\infty} \alpha^{k-1} cbx_{t-k} + \nu_t^x \tag{A. 1.11}$$

Appendix 2 – Derivation of Autoregressive Form of Simplified Spillover Model
This model will follow this set of recursive equations:

$$m_{i,t} = m_{i,t-1} + z_i v_{i,t}^x + v_{i,t}^m \tag{A.2.1}$$

$$q_{i,t} = m_{i,t} + \alpha_i \left(q_{i,t-1} - m_{i,t-1} \right) + b_i . x_{i,t} \tag{A.2.2}$$

$$x_{i,t} = -c_i \left(q_{i,t-1} - m_{i,t-1} \right) + \sum_{j=1, j \neq i}^{M} f_{i,j} \left(q_{j,t-1} - m_{j,t-1} \right) + v_{i,t}^x \tag{A.2.3}$$

Again, the interest is in the mid quote price changes $\Delta q_{i,t} = q_{i,t} - q_{i,t-1}$. It is easy to see that the insertion of the extra term in the trade equation will not change the autoregressive representation for $\Delta q_{i,t}$. Therefore, the solution of mid quote changes in the spillover model is the same as in Hasbrouck:

$$\Delta q_{i,t} = \left(z_i + b_i \right) x_{i,t} + \sum_{k=2}^{\infty} \alpha_i^{k-1} \left(z_i b_i c_i - b_i \left(1 - \alpha_i \right) \right) x_{i,t-k} + v_{1,t} \tag{A.2.4}$$

For the autoregressive formula of the trade equation, we start with:

$$x_{i,t} = -c_i \left(q_{i,t-1} - m_{i,t-1} \right) + \sum_{j=1, j \neq i}^{M} f_{i,j} \left(q_{j,t-1} - m_{j,t-1} \right) + v_{i,t}^x \tag{A.2.5}$$

And recognise that, from Hasbrouck's solution:

$$c_i \left(q_{i,t-1} - m_{i,t-1} \right) = c_i \sum_{k=1}^{\infty} \alpha_i^{k-1} b_i . x_{i,t-k} \tag{A.2.6}$$

Therefore, its easy to see that

$$q_{j,t-1} - m_{j,t-1} = \sum_{k=1}^{\infty} \alpha_j^{k-1} b_j x_{j,t-k} \tag{A.2.7}$$

Substituting back into A.2.6, the final representation is:

$$x_{i,t} = -\sum_{k=1}^{\infty} \alpha_i^{k-1} c_i b_i x_{t-k} + \sum_{j=1, j \neq i}^{M} f_{i,j} \left(\sum_{k=1}^{\infty} \alpha_j^{k-1} b_j x_{j,t-k} \right) + v_{i,t}^x \tag{A.2.8}$$

Appendix 4 – Results from extended Spillover Model, Version 2

Panel A – Quote Equation for extended Spillover Model

bondCode	Sum of Lambda 1	Sum of Lambda 2	Sum of Betas	Sum of Phis 1	Sum of Gammas	Sum of Phis 2	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	0.0407***	0.0001	-0.0418	-0.0108	0.0068***	0.0003	4,326.5062***	325.6475***	0.01
IT0001488102	0.1751***	-0.0003	-0.5916***	0.9315***	0.0414***	0.0044	2,232,619.5238***	527.9054***	0.17
IT0003652077	0.6744***	0.0038	-0.1639***	0.1893	0.0236***	0.0069	1,359,245.6350***	28.1477***	0.09
IT0003472336	1.7467***	0.0074*	-0.2**	-1.4827*	0.0411***	-0.0062	642,303.2565***	882.1091***	0.05
DE0001137024	0.2394***	0.0004	-0.4852***	-0.6792	0.027***	0.0201	1,864,336.1463***	152.7297***	0.16
DE0001137057	0.4204***	-0.0033	-0.2889***	-1.1148	0.0172***	0.0253	3,895,835.3121***	81.1789***	0.19
DE0001141430	1.2044***	-0.0026	-0.3029***	0.5658	0.019***	0.0134***	879,430.1457***	22.9774***	0.33
DE0001135242	1.4214***	-0.008	-0.0883***	0.3128	0.0764***	0.0564	570,904.3280***	10.4960*	0.15
FR0104756962	-0.0057	0.0008	-0.2057***	0.5218***	0.0101***	-0.0068	10,973.6441***	27.2206***	0.04
FR0106589445	0.2556***	0.0041	-0.5096***	0.8249***	0.0219***	-0.0184	36,910.6446***	106.7503***	0.13
FR0106589437	0.9334***	-0.0116***	-0.0872	0.0396	0.0119***	0.0909***	709,302.8176***	369.2947***	0.12
FR0010011130	3.0612***	-0.0116	-0.2577***	-3.7894	0.0469***	0.0823	3,177,288.5790***	29.3913***	0.38

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following Newey and West (1987).

$$r_{i,t} = \sum_{k=1}^{p} \beta^r_{i,k} r_{i,t-k} + \sum_{k=1}^{p} \gamma^r_{i,k} Q_{i,t-k} + \sum_{k=0}^{p-1} \left[\lambda^r_{1,i,k} + \phi^r_{1,k} n^{T\&Q}_{i,t-k} \right] r^{AGG}_{i,t-k} + \sum_{k=0}^{p-1} \left[\lambda^r_{2,i,k} + \phi^r_{2,k} n^{T\&Q}_{i,t-k} \right] Q^{AGG}_{i,t-k} + \varepsilon^r_{i,t}$$

$n^{T\&Q}_{i,t} = \left(nTrades^{AGG}_{i,t} + nQuotes^{AGG}_{i,t} \right) \left(\Delta t \right)^{-1}$ - Number of trades and quote changes for the bond portfolio, for each unit of time.

Δt - Number of seconds between event t-1 and t.

$r^{AGG}_{i,t}$ - Return for the bond portfolio, happening between t-1 and t.

$Q^{AGG}_{i,t}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t} -$ log return of mid quote for bond i.

$Q_{i,t} -$ Signed volume of trade for bond i.

Panel B – Trade Equation for extended Spillover Model									
bondCode	Sum of Lambda 1	Sum of Lambda 2	Sum of Betas	Sum of Phis 1	Sum of Gammas	Sum of Phis 2	Breush-Pagan Test	Breush-Godfrey LM Test	adj R2
IT0003248512	0.0929**	0.0041	0.3553***	-0.1323	-0.1235	-0.0193	140,849.6430***	67.0557***	0.07
IT0001488102	0.3975***	0.0077**	0.4331***	0.3234	-1.5658***	-0.022	60,936.4104***	1,094.8627***	0.08
IT0003652077	0.0761**	-0.016**	0.4037***	-0.0001	-0.0986***	0.0615***	36,788.3161***	4.8346	0.08
IT0003472336	0.0799	0.0002	0.3923***	-0.1871	-0.0234**	0.0398	73,929.8244***	9.3842*	0.08
DE0001137024	0.0531**	-0.0083	0.393***	-0.0402	-0.2516***	0.0349	131,361.9154***	124.8865***	0.07
DE0001137057	0.0233	0.0067	0.3977***	1.083	-0.3968***	-0.0701	172,388.8841***	89.8649***	0.09
DE0001141430	0.0022	0.0094	0.4839***	0.0781	-0.0083	-0.0123	318,734.8728***	133.0418***	0.16
DE0001135242	-0.0334	0.0122	0.3512***	0.3446	-0.0111	-0.0767*	177,675.1587***	102.2906***	0.07
FR0104756962	-0.0005	0.0056	0.2028***	0.4198	-0.0772	-0.0325	119,327.3701***	97.4394***	0.03
FR0106589445	0.0813**	-0.0226	0.3317***	0.0369	-0.2266***	0.2246	137,186.6040***	83.8415***	0.05
FR0106589437	0.0129	0.0094	0.5067***	-0.2792	-0.0058	-0.0567***	313,114.6273***	27.3610***	0.16
FR0010011130	-0.0426	-0.0115	0.3601***	0.8831	-0.014	0.1519	168,378.5677***	71.7770***	0.06

- All autoregressive parameters values are interpreted as a sum and we use a Wald test for testing the null hypothesis that this sum is equal to zero.
- *, ** and *** means rejection of the null hypothesis at the 10%, 5% and 1% levels, respectively.
- Breush-Pagan is a test for heteroscedasticity. The null hypothesis is of homocesdasticity (no heterokesdasticity). We use 5 lags in the test.
- The Breush-Godfrey is a test for serial correlation. The null hypothesis is of no serial correlation. We use 5 lags in the test.
- All standard errors are computed following Newey and West (1987).

$$Q_{i,t} = \sum_{k=1}^{p} \beta_{i,k}^{Q} Q_{i,t-k} + \sum_{k=1}^{p} \gamma_{i,k}^{Q} r_{i,t-k} + \sum_{k=0}^{p-1} \left[\lambda_{1,i,k}^{Q} + \phi_{1,k}^{Q} n_{i,t-k}^{T\&Q} \right] Q_{i,t-k}^{AGG} + \sum_{k=0}^{p-1} \left[\lambda_{1,i,k}^{Q} + \phi_{2,k}^{Q} n_{i,t-k}^{T\&Q} \right] r_{i,t-k}^{AGG} + \varepsilon_{i,t}^{Q}$$

$n_{i,t}^{T\&Q} = \left(nTrades_{i,t}^{AGG} + nQuotes_{i,t}^{AGG} \right) \left(\Delta t \right)^{-1}$ - Number of trades and quote changes for the bond portfolio, for each unit of time.

Δt - Number of seconds between event t-1 and t.

$r_{i,t}^{AGG}$ - Return for the bond portfolio, happening between t-1 and t.

$Q_{i,t}^{AGG}$ - Aggregated order flow for bond portfolio, between t-1 and t.

$r_{i,t}$ – log return of mid quote for bond i.

$Q_{i,t}$ – Signed volume of trade for bond i.