Short Selling Ban and Intraday Dynamics

Julio A. Crego*
CEMFI
<jacrego@cemfi.edu.es>

Abstract

Since September 2008 regulators from different countries, motivated by suspicions regarding an increase in investors' aggressiveness, have implemented several temporary short selling restrictions. In this paper, I study the effect of such policies in the context of the 2012 Spanish short selling ban. The results of this paper highlight an important policy trade-off: on the one hand, I provide evidence that, in line with regulator beliefs, investor aggressiveness is extremely high prior to the ban and, it reverts just after the ban implementation. On the other hand, using a novel identification strategy, I find that this policy increases the bid-ask spread. The causal interpretation of these results is obtained using intraday data under the assumption that the exact time of the implementation is random. The results obtained under this methodology are much smaller than the ones found in previous literature.

Keywords: Short Selling, Market Quality, Hawkes Process, Aggressiveness

*I am indebted to my advisor Dante Amengual for his encouragment and useful comments. For helpful suggestions, I also thank Manuel Arellano, Stéphane Bonhomme, Julio Gálvez, Roberto Pascual, Rafael Repullo, and Enrique Sentana. I am extremely thankful to BME for providing the data. Errors and omissions are exclusively my own responsibility.

1 Introduction

On July 23, 2012, at 14:28 (CET), the Spanish analogue to the US SEC (Comisión Nacional del Mercado de Valores, CNMV) implemented a short selling ban. From that precise moment up until the January 31, 2013, investors could not increase their short positions in Spanish stocks. This action was taken in coordination with the Italian counterpart with the belief that some speculators were performing aggressive short selling in Spanish and Italian stocks, regardless of the fundamentals.

In this paper, I analyze the effect of this policy on aggressiveness and market quality. First, I document that, in line with regulators' beliefs, the level of aggressiveness was abnormally high at trading hours before to the ban. Second, I find suggestive evidence that the policy was effective, since investor aggressiveness reverted to usual levels just after the implementation. Finally, regarding market quality, I use a novel identification strategy to show that even if the policy increases the bid-ask spread,thus, an indication of a reduction in market quality, this increase is much lower than the one found by previous literature.

While numerous studies have analyzed the effect of short selling bans on different measures of market quality, the effect of this type of policy on aggressiveness is still not well understood.

Closest to my analysis is the work by Boehmer et al. (2013) they use intraday data form NASDAQ, NYSE and BATS to compute daily mean effective spreads paid by the short sellers. This quantity reflects how much short sellers are willing to pay for immediacy. Thus, it can be considered as a measure of aggressiveness. The authors show that aggressiveness was stable before the 2008 US short selling halt; however, at the day of the ban they find, that aggressiveness increased to a level that was maintained up to the end of the ban. They acknowledge, however, that this effect is probably driven by the composition of short sellers. While all investors were short selling before the ban, only market-makers were allowed to short sell once the policy was implemented. Since market makers might be more aggressive in aggregate, overall aggressiveness should increase, in spite of the policy not affecting individual investors' aggressiveness.

This paper differs from Boehmer et al. (2013) along two main dimensions: the measurement of aggressiveness and the context. Regarding the first dimension, I propose a new measure based on the classification in Biais et al. (1995), the intensity of sell-initiated transactions.[1] Intuitively, this quantity represent the likelihood of observing an investor who pays the bid-ask spread to sell some shares. As this measure does not depend on the bid-ask spread, I can differentiate between the effect of the ban on aggressiveness and its effect on liquidity. To model the arrival of sell-initiated transactions I use a continuous-time intensity model -Hawkes process- that allows capturing most of the intraday effects. In particular it allows me to disentangle an increase in aggressiveness caused by the ban from the usual inertia present in order submissions buy (sell) market orders tend to be followed by buy (sell) market orders.

In terms of the context, I focus my analysis on the 2012 Spanish short selling ban because of its unique features. First, the Spanish stock market in 2012 was extremely

[1]During the paper I will use the term sell-initiated transaction or just sales to indicate those transactions in which the liquidity maker is the buyer.

concentrated as more than 90% of the trading took place using the main stock platform. Therefore, results are not likely to be affected by any sample selection bias. Second, the Spanish short selling ban affected all investors equally, including market makers; consequently, the effect of the ban on aggressiveness is not affected by investor composition. Finally, the ban was implemented in the middle of the trading day which was an unexpected event. I use this latter feature to identify a causal effect of the ban on market quality.

While these characteristics make the 2012 Spanish ban special, the results and techniques presented in this paper are relevant outside the Spanish context, since similar convictions about aggressiveness led regulators of several countries to implement similar policies between 2008 and 2011. For instance, in 2011, the ESMA (European Securities and Markets Authority) forbade short selling in most European countries, arguing on their public statement that short positions were not related to fundamentals:

> *"Given these recent market developments, ESMA wants to emphasize the requirements in the Market Abuse Directive referring to the prohibition of the dissemination of information which gives, or is likely to give, false or misleading signals as to financial instruments, including the dissemination of rumours and false or misleading news".[...] "Today some authorities have decided to impose or extend existing short selling bans in their respective countries. They have done so either to restrict the benefits that can be achieved from spreading false rumours"*
>
> European Securities and Markets Authority (ESMA)
> Public Statement August 11, 2011

Another example is the 2008 US short selling halt enforced by the US SEC that directly targets aggressive short selling, as stated in the official press release:

> *"This decisive SEC action calls a time-out to aggressive short selling in financial institution stocks, because of the essential link between their stock price and confidence in the institution. The Commission will continue to consider measures to address short selling concerns in other publicly traded companies."*
>
> US Securities and Exchange Commission
> Press Release September 19, 2008

This paper makes the following contributions. First, I provide an economic interpretation of the Hawkes process which allows to disentangle aggressiveness by speculators or informed traders (baseline aggressiveness), from the one caused by uninformed who learn about the fundamental or investors who split orders to reduce price impact (inertia). Second, using this methodology, I show that inertia was stable before and after the ban while the baseline aggressiveness decreased when the policy is implemented. These findings are consistent with an increase in speculators' aggressiveness before the ban, as the regulator suspected.[2]

[2]The theoretical relationship between an increase in aggressiveness and speculative trading when market volatility is high is presented in Goettler et al. (2009).

Yet the reduction in aggressiveness is just one side of the coin. The other side of the coin is the effect on market quality. Previous literature has shown that short selling restrictions are detrimental for market quality. In particular, Beber and Pagano (2013), using data from different countries, conclude that the 2008 bans harmed liquidity. In the UK context, Marsh and Payne (2012) find similar results as Beber and Pagano (2013). Prior to 2008, Bris et al. (2007), using double listed firms and panel data, found that short selling restrictions are associated with less negative skewness and a slower incorporation of bad news into prices.[3]

A common thread of most of these papers is that the results hinge on two main assumptions, first, control and treatment groups are comparable and, second, the control group is not affected by the ban. However, both assumptions are unlikely to hold unless we rely on an experimental design. In the first assumption, the treatment group is usually defined by the regulator (for instance, financial firms) and thus it is not random. Furthermore, if the regulator's objective is to improve market quality, she would treat the firms whose market quality is at risk of deterioration. This selection into treatment will bias the results toward a negative effect of the policy on market quality. With respect to the second assumption, the treatment group might lose liquidity or volume in favor of the control group. Consequently, the estimates will be biased toward a bigger effect of the ban.

To my knowledge, the only paper that provides evidence from a randomized experiment is Kaplan et al. (2013). In this study, the authors generate an exogenous reduction to the cost of short selling by increasing the supply of lending shares in some random firms. They find almost no significant effect in most of the variables usually considered in the literature except for the bid-ask spread. This measure decreases when the experiment is finished (when the cost of short selling increases). While Kaplan et al. (2013) do not specifically study the effects of short selling bans, their results suggest that these might not be of first-order importance.

In this paper, I analyze the effect of the ban using a novel identification strategy that relies on the assumption that the exact time of the policy implementation is random. The availability of high-frequency data, coupled with the fact that the ban took place in the middle of the trading session, enables this assumption to hold. Using an approach similar to regression discontinuity in the time dimension, I estimate an increase of the percentage bid-ask spread of 0.0006 percentage points. This estimated effect is much lower than the effect found in previous literature. For instance, Beber and Pagano (2013) find an effect of 1.98 points using their whole panel and around 0.8 percentage points in the Spanish case.[4]

Finally, I address some additional methodological issues that arise when using Hawkes processes with high frequency data. These auxiliary results, which are gathered in the appendices, include the simulation algorithm of a multivariate Hawkes process and a Monte Carlo experiment to test the estimation using a new methodology proposed by Halpin and De Boeck (2013). This methodology improves the stability of the maximum likelihood estimator and also provides new insights about the Hawkes process. I also

[3] Numerous other papers have analyzed the effect of short selling bans in different contexts. For instance, Saffi and Sigurdsson (2010), Billingsley et al. (2011), Grundy et al. (2012), Kolasinski et al. (2013).

[4] The authors consider the bans implemented in 2008, therefore the sample is not the same

compute the Hessian of the multivariate Hawkes process analytically, using a formulation that makes it straightforward to implement in a computer program. Lastly, I propose new methods to deal with rounded time stamps, that is, the problem that most observations seem to take place at the same time. This feature, when the data is rounded, leads to biased results, as it yields oversize jumps in the intensity.

The paper is structured as follows. Section 2 describes the data used while section 3 introduces the econometric model and the analysis of rounded data. Section 4 presents the results, and in section 5 I conduct an analysis of the key market microstructure variables. Finally, Section 6 concludes.

2 Data

I use Trade and Quote data from July 16, 2012 to July, 27.[5] The trades can be unambiguously merged to the current state of the book by the use of the annotation number. Therefore the classification into seller-initiated and buyer-initiated transactions is perfect. Moreover, the BME electronic system (SIBE) accounts for more than 90% of the operations related to Spanish Stocks, which almost completely eliminates any possible selection bias arising from observing only operations done by specific types of traders.[6]

Nevertheless, the data have one important limitation: The time stamp available is rounded to the nearest second. In Appendix C I analyze the different methods proposed in the literature to adjust the time stamp and I show that they generate relevant biases. Furthermore, I develop, in the context of the Hawkes process, two new methods that reduce or even eliminate the bias. These methods can be easily extended to other intensity models. Even if this data issue is not present in other markets, the bundling of signals might create similar problems as Filimonov and Sornette (forthcoming) point out. They use the term bundling to refer to the fact that market platforms group several orders in one message and thus, the time between orders is no longer informative. Moreover, Filimonov and Sornette (forthcoming) point out that the mostly used protocol have precise time stamps rounded to the nearest seconds but any division inside the second is affected by bundling.

The main drawback of the Spanish market is the lack of liquidity. Since my analysis is based on high frequency data, I only consider the six largest firms in terms of market capitalization. These are: Banco Santander (SAN) and BBVA, two of the biggest European banks; Telefónica (TEF), the most important Spanish telecommunications firm; Iberdrola (IBE) and Repsol (REP) two important companies in the energy sector; and Inditex (ITX) whose business is focused on clothing manufacturing and distribution.

Regarding the different trading periods, I use only data in the general trading period since they constitute the information that is really observed by market participants, even if the data includes non-general trading periods (i.e. initial and final auctions). The Spanish stock exchange during this period is a purely limit order book market with total disclosure of best bid and ask prices, where algorithmic trading is allowed.

[5]All the data have been provided by BME (Bolsas y Mercados Españoles).
[6]Data from Bats Chi-X Europe, July 2012.

3 Model

In this section, I describe the econometric model that I use to characterize the intensity of sales. This measure is, loosely, the probability of observing a sale instead of not observing operations or observing a different operation (e.g. a buy). Therefore the random variable of interest is similar to a Bernouilli random variable that takes the value 1 if there is a sale or the value 0 if there is not a sale in a given period of time. In this paper, due to the high frequency of data, I use a continuous intensity process, thus each period length would tend to zero.

Specifically, I define the counting process $J(t)$ as a random process that counts how many sales have taken place in the market up to time t, then the intensity (the probability of having a sale) is defined as: [7]

$$\lambda(t) = \lim_{\Delta \to 0} \frac{1}{\Delta} \mathcal{P}\left(J\left(t+\Delta\right) - J\left(t\right) | \mathcal{F}\left(t\right)\right).$$

I make use of the Hawkes process (Hawkes, 1971) to study the dynamics of intraday operations. This process has been previously used in the literature because of its self-exciting structure, same type of operations (sales) are clustered in time, and its simple multivariate representation that allows for cross-exciting effects, different types of operations are allowed to be clustered in time. Intuitively, modeling large and small sales as different types allow past events to affect the probability of a new big or small sale taking place differently. The effect of each transaction in every future transaction is decreasing deterministically in time; specifically in this paper, as in the literature (e.g. Bowsher (2007)), the decay is assumed to be exponential. More formally, the multivariate Hawkes process is totally defined by its intensity which is given by:

$$\lambda^m(t) = \lambda^m_{0,d} + \sum_{n=1}^{N} \int_0^t \alpha_{mn} e^{-\beta_{mn}(t-s)} dJ^n_s \quad \forall m \quad \lambda^m_{0,d} > 0 \; \alpha_{mn} \geq 0 \; \beta_{mn} \geq 0 \; \forall n, m \tag{1}$$

where $J_s \equiv (J^1_s, ..., J^N_s)$ is a càdlàg multivariate simple counting process at s. In this specification, $m = \{1, 2\}$ and $n = \{1, 2\}$ will indicate big and small sales and the baseline parameter $\lambda^m_{0,d}$, is allowed to be different on each day d, and also to change at the moment of the ban. The parameters α_{ij}'s correspond to the size of the jump on the intensity of the event i caused by the event j, while the parameters β_{ij}'s measure the reversion speed. A high β_{ij} would imply that the effect of an operation of type j on the intensity of type i does not last for long.

The multivariate Hawkes process was introduced by Bowsher (2007) in order to describe the buy/sell dynamics. Moreover, he shows the properties of its Maximum Likelihood Estimator. In this paper the estimation is done using an EM algorithm based on the branching structure of the process, which behaves better in terms of numerical error. I provide a detailed explanation of the algorithm as well as a Monte Carlo study of the estimates in Appendix A. Standard errors are obtained using the Hessian of the log-likelihood, whose expression is presented in Appendix B.

[7] $\mathcal{F}(t)$ is the filtration generated by the process $J(t)$, that is how much we know after having observed $J(t)$ from 0 to t.

Previous literature has used different models to describe the order dynamics. These models can be divided in two groups (see Hautsch, 2011): duration and intensity models. The Hawkes process belongs to the second group. Since this group is defined in the time domain, it is more suited to include time related effects such as a dummy at the moment of the ban. In contrast, duration models are defined in terms of the time between two consecutive observations, therefore around the ban implementation the estimate will capture mixed effects. Moreover in this type of models the cross exciting effects do not carry the same interpretation.

Among intensity models, the Hawkes process is one of the simplest models that allows for auto-dependence of the intensity process. This is very important for the analysis of the effect of some event at the middle of the day; otherwise, the effect of the auto-dependence might be attributed to the event such as the ban. Furthermore, the branching structure of the Hawkes process allows us to interpret economically the parameters of the model.

The branching structure representation of the Hawkes process, loosely, is a way of representing the same process by dividing it into different independent Poisson processes. In the case of the Hawkes process, events are divided in two classes: "immigrants", driven by the baseline parameter (λ_0^m); and "offspring", caused by prior events. Specifically, each order can have a daughter with intensity $\alpha_{mn}e^{-\beta_{mn}(t-t_p)}$, where t_p is the time when the "parent" event occurs.

In the context of market operations, this representation provides more insight about the properties of the Hawkes process. When the market starts, the probability of an investor demanding liquidity in the bid side of the limit order book is equal to λ_0^m.[8] Once an investor has done a seller-initiated operation at time t_p, the probability of a similar investor demanding the same type of liquidity remains λ_0^m. In addition, some uninformed investors might consider the first operation as a signal and demand liquidity at the same side of the book with probability $\alpha_{mn}e^{-\beta_{mn}(t-t_p)}$. Since investors' type is not observable, the operations of these uninformed investors might also be taken as a signal by the other uninformed investors in the market.

Regarding market microstructure models, like models à la Kyle (1985), the baseline could represent informed (or speculative) investors or investors with liquidity needs; while the offspring could represent all uninformed investors who infer the stock value from the signals created by the sales in the market, like high frequency traders.[9]

One common concern when dealing with the arrival of orders as a dependent variable is the fact that traders usually split a big potential order into several small orders to reduce market impact. Economically, all these small orders should be taken as one; however, in the data they appear as several orders. Since there is no information about the trader I cannot disentangle an split order from several traders that submit small orders. Nevertheless, the Hawkes process filters the order splitting from the baseline parameter. The filtering process can be illustrated by the branching representation: if an informed trader splits the order, the first order contributes to the baseline component

[8] I allowed for a different baseline parameter at the beginning of the day $\lambda_0+\delta e^{-\gamma(t-t_0)}$ and the estimate of δ was very close to zero and created a lot of instability in the numerical maximization, therefore is not considered. Note that if $\delta = 0$ then γ is not identified.

[9] I assume along the paper that the number of operations driven by liquidity needs is constant over the sample period, such that I attribute any increase or decrease in λ_0 to informed/speculative investors.

while following orders just affect its offspring.[10] Therefore the baseline parameter becomes economically significant at the cost of the inertia component.

4 Results

Since the focus in this paper is to study the behavior of the sales before and after the short selling ban, to explain changes in that behavior, the analysis is restricted to large and small sales. I classify operations according to their aggressiveness following Biais et al. (1995):

1. *Large Sale (LS)*: Sell marketable order whose volume exceeds the volume at the best bid.
2. *Small Sale (SS)*: Sell marketable order whose volume *does not* exceeds the volume at the best bid.

I use the term sell marketable order to include sell market orders as well as sell limit orders whose target price is lower or equal than the bid. If this price is exactly equal to the bid, I classify the operation as a small sale independently of the volume.

The Hawkes model is estimated using data from July 16, 2012 to July 27, 2012, consequently, the estimation window includes five days without ban, four with the ban and the day where the ban is implemented. To focus on the inference about the baseline parameter, all other parameters are assumed to change just at the moment of the ban while the baseline parameter is allowed to change every day and at the moment when the ban took place.[11]

Unfortunately, the amount of big sales is not enough to obtain stable estimates. Hence, I focus the analysis on small sales. Nevertheless, in the context of the Hawkes process the parameters which drive the different types are independent, as a consequence I can esimate the effect of the big sales on the small sale intensity.

In Table 1 I present the inertia parameter estimates obtained for Iberdrola and BBVA while in Figures 1 and 2, I plot the different baseline parameters for these firms. The figures show evidence that aggressiveness is doubled just before the ban and it reverts to the same levels just after the policy is implemented. Regarding the inertia, estimates are extremely similar before and after the ban which suggests that the effect of the policy is not important for information acquisition or order splitting, unless both effects cancel each other. Additionally, this stability in the parameters can be interpreted as evidence in favor of he Hawkes process as a model for order auto-dependence.

The point estimates obtained for the other firms are also stable before and after the ban. However, the standard error suffers from high numerical error in less liquid firms like Repsol or Inditex because the amount of large sales is very small. Therefore, in order to estimate the parameter standard errors and ensure a robust estimation for less liquid firms, I restrict the model such that large sales do not affect small sales. The results obtained for liquid firms are extremely similar as those obtained taking into account large

[10] Assuming some stability in the splitting behavior.

[11] Specifications where the baseline could be time dependent do not reject λ_0 constant. Moreover they increase the numerical error and might exacerbate the problem of the rounded time-stamp

sales. This statement can be visually checked by comparing the baseline estimates for Iberdrola and BBVA in Figures 1 and 2 with the same firms in Figure 3.

[Table 1]

[Figures 1 and 2]

[Figure 3]

In Figure 3, I extend the analysis to the six biggest Spanish firms in terms of market value. Every firm present the same pattern in terms of the baseline parameter: an increase just before the ban and a decrease afterwards. To interpret the estimate, I present the estimate normalized such that the first day is equal to one. The results indicate that aggressiveness is doubled before the ban in every firm. Moreover, the estimate corresponding to the last day is insignificantly different from the first day suggesting a reversion. This evidence is in line with a speculative movement stopped by the implementation of the short selling ban.

My strategy is limited by the fact that any factor with a similar pattern affecting all firms might cause the same effect. Since I use data one week before and one week after the short selling ban, it does not seem likely that a factor could display this pattern and affect the seller-initiated transactions of all the Spanish firms considered. Another concern is the possible endogeneity of the implementation time. Since the decrease in aggressiveness is observed just after the ban, this concern is only relevant if the exact hour of the ban is endogenous.

As an aside, in Figure 3 I show that Telefónica (TEF) presents an increase in the baseline parameter on July 26, 2012. On this day, Telefónica cut a revenue forecast and announced that the firm will not deliver dividends in 2012 and only half of the expected dividends in 2013. Thus, this peak reinforces the interpretation of the baseline parameter as insider trading, since there is evidence on an increase in informed trading just before the release of a dividend related news (see e.g. John and Lang (1991)).

5 Market Quality Measures

In this section I exploit the high frequency nature of the data and the special time of the implementation to make a causal inference about the effects of the short selling ban. My approach identifies the "instantaneous" effects of the policy under the assumption that there were no confounding factors affecting market quality around the time at which the ban was set. As measures of market quality I consider two main variables: the bid-ask spread, a measure of informed trading and inventory costs (see e.g. Kyle (1985), Glosten and Harris (1988)), and the order flow imbalance (OFI), a proxy for depth and the size of price changes (see Cont et al. (2013) among others). The validity of the instantaneous effect assumption is discussed below in the context of the bid-ask spread, in order to profit from the strong theoretical background present in previous literature.

The seminal paper of Diamond and Verrecchia (1987) considers the implementation of a short selling ban, concluding that it will decrease the speed at which bad news are incorporated in prices and the bid-ask spread converges to its minimum level. Therefore, unless the arrival of information about the fundamental has very low frequency, and the bid-ask spread is in the "steady state", we should observe a higher mean bid-ask spread right after the ban.

The main two assumptions in Diamond and Verrecchia (1987) are that both informed and uninformed investors are equally affected by the restriction and the proportion of both types of traders in the market is unaffected by the ban. However, the investors affected by the restriction were those that want to maintain a short position and thus we would expect they are not liquidity (uninformed) traders, but informed investors.[12] Thus, if the agents decide to acquire information about the fundamental as in Goettler et al. (2009), we expect that with a ban in place less agents become informed, reducing the proportion of informed in the market. The theoretical literature (for instance Kyle (1985), Glosten and Milgrom (1985)) predicts a reduction in the bid-ask spread as soon as the proportion of informed traders in the market decreases. The rationale is as follows: the market maker is a competitive agent with zero profits; therefore, she uses the bid-ask spread to earn from the uninformed the same amount (in expectation) that she loses from the informed. As the proportion of these latter agents is reduced, she does not need to extract so much surplus to the uninformed and hence, competition causes a decrease in the bid-ask spread.

While these two opposite channels might have an instantaneous impact on market quality, the policy might, nevertheless, affect market quality through other channels in a longer run. Since these effects are hard to isolate from other confounding factors, they are outside the scope of this paper.

Regarding the instantaneous effect, it does not only provide evidence on the effect of the ban on measures of market quality at the very short run that should be taken into account for policy perspectives but it can also give insight about the peak in aggressiveness: If this peak is caused by informed investors the bid-ask spread should decrease after the policy is implemented, since informed investors are leaving the market. However if the peak is caused by speculators, the only effect on the bid-ask might come through the Diamond and Verrecchia (1987) mechanism and thus there should be an increase in the bid-ask spread.

The first approach I use to identify the causal effect is based on the high frequency of the data, that allows, by the use of a saturated model, to isolate the effect of the ban. Specifically, I estimate the following regression equation:

$$Y_\tau = \sum_{j=1}^{R} \rho_j Y_{\tau-j} + \eta_\tau \tag{2}$$

$$\eta_\tau = \alpha + \sum_{i=1}^{P} \gamma_i \cdot H_i(t) + \beta \cdot EVENT_\tau + \varepsilon_\tau \tag{3}$$

where $R = 1$ and $P = 10$ in the baseline regression but results remain similar as long as $R > 0$. t is clock time, τ is tick time, H_i is a i^{th} degree time polynomial and Y is the market quality variable and it is defined, depending on the specification, as:

[12] Spanish ban affected short positions, thus hedging strategies such as shorting stocks to cover a long position in calls were allowed.

$$Y = \frac{\text{Best Ask - Best Bid}}{\text{Opening price}} \qquad \text{or} \qquad Y = log\left(\frac{\text{Volume at the Best Ask}}{\text{Volume at the Best Bid}}\right)$$

The coefficient of interest is β , if the policy has an effect on the mean level of this variables this coefficient should be different from 0. This approach is similar to a regression discontinuity design in the time dimension where the assumption is that the mean of the market quality measures around the event time is only affected by the ban. To be able to interpret the results, I divide the bid-ask spread by the first day opening price. This normalization allows me to interpret this variable as a percentage effective spread, while it does not dependent on prices. Thus, I avoid the possible mechanical effect that a change in prices might cause due to the existence of a minimum tick.

The first two columns of Table 2 show the estimate of β in the case of the bid-ask spread and the OFI. From the first column, we observe that the spread increases in almost every firm considered, nevertheless, this effect is smaller than the one found in previous literature as discussed in the introduction. Regarding the OFI, the mean estimated coefficient is negative and significant which suggests that the policy reduces liquidity in the ask side of the book compared with the bid side.

Even if I only consider one week before and one week after the ban, some unobserved factors might affect the estimates. To explore this possibility, in Figure 4 and Figure 5 I present $\hat{\eta_\tau}$ for the six different firms in the case of the bid-ask spread and the OFI, respectively. This residual summarizes the time evolution of the variable Y in a clear and smooth way. The most striking fact of these figures is how similar the time polynomials are for different firms, for instance on July 27, 2012 most firms present a peak in the case of the Bid Ask spread and a drop in terms of the OFI. Unfortunately, due to the amount of news it is difficult to identify which event generates that patter, however, it could be driving the results toward a bigger effect if the time polynomial is not enough as a control.

[Figures 4 and 5]

In order to isolate the effect of the ban from these other effects, I consider a 1-day estimation window given the special feature of the Spanish short selling ban: its implementation in the middle of the trading day. This approach provides a better identification strategy but raises a trade-off. On the one hand, we would want to control for seasonal effects or confounding factors with the time polynomial as before, but, on the other hand, we would want to allow for a less instantaneous effect since the time dimension is small. In order to tackle this trade-off, I estimate Equation 4, where I do not include time trends, but instead of relying on standard errors I estimate the model each day separately and compare the estimates of β for each day. If the ban has an effect the estimate of β corresponding to the day of the ban should be the highest.

$$Y_\tau = \alpha + \sum_{j=1}^{R} \rho_j Y_{\tau-j} + \beta \cdot \mathbf{1}\{t > 14{:}28{:}00\} + \varepsilon_\tau \tag{4}$$

Figures 6 and 7 show the estimates of β each day and their 95% confidence intervals using as a dependent variable the bid-ask spread and the OFI respectively. To obtain a quantitative measure of the effect, I subtract from each estimate the average estimate obtained in that specific firm excluding the day of the ban. I present the results in the third and fourth column of Table 2. We observe that, even if the signs are maintained, the estimated effects are much smaller. These results are evidence that the time polynomial was not enough as a control.

[Figures 6 and 7]

These placebo tests are useful to control for seasonal trends inside a day but there is still the possibility that there was a confounding factor affecting market quality on the day of the ban which can bias the results. In order to reduce this concern, I conduct additional placebo tests estimating Equation 2 but changing the cut-off point. If the effect of the ban is causal, then the estimate of beta has to decline just after the time of the implementation. Figure 8 plots the estimate of β considering different cut-off points and, indeed, the estimate decreases next minutes after the ban in most firms with the exception of Iberdrola.

[Figure 8]

The identification strategy proposed in this paper presents two main concerns. Firstly, investors might have expected the ban since they knew Spain was in distress. Nevertheless, since previous similar restrictions were set when the market was closed, it seems very plausible that investors do not expect the ban at that specific time. The second concern is related to the effect of the announcement of the ban. If the announcement has an effect on investors but the ban does not, i.e a psychological effect, this effect would carry different policy implications. Even if this issue cannot be totally addressed using my proposed methodology, it would increase the estimated effect.

6 Conclusion

This paper analyzes the 2012 Spanish short selling ban, which is characterized by two main features: First, it was set under rumors of aggressive trading against Spanish stocks unrelated to fundamentals; and second, the ban was implemented in the middle of the trading day as opposed to previous short selling bans which were introduced outside of the trading period.

The first feature of the ban is not unique to the 2012 Spanish short selling ban, it was also present in similar policies such as the ones set in 2011 in Spain, Italy and Greece. However, the veracity of these rumors has not been tested. In this paper I present evidence supporting the rumors by documenting a peak in trade aggressiveness before the policy implementation. Moreover, using a novel interpretation of the branching structure of the Hawkes process I attribute this increase in aggressiveness to speculators or informed investors as opposed to uninformed investors who learn from the market. Additionally, I show that these investors reduce their aggressiveness right after the ban is implemented.

The second feature of the 2012 Spanish short selling ban, in conjunction with the availability of high frequency data, provides an opportunity to test the causal effects of this type of policies on market quality under a mild identification assumption. To identify a causal effect, I require that the exact instant at which the policy is implemented is random. Exploiting this unique feature, I find empirical evidence of a detrimental effect of the ban on market quality measured by the bid-ask spread. However, the magnitude of the effect is small compared o previous empirical results. This discrepancy might be a result of the context but it might also be due to a potential endogeneity in the diff-in-diff methodology used in previous literature, since the treatment is not usually randomly assigned.

While the ban's unique features permit the stud of this effect on speculation and market quality, he external validity of the results might be an issue, as the analysis focuses on Spain. Nevertheless, the Spanish stock market is very similar to other stock markets in continental Europe in terms of regulation, especially after the creation of ESMA in 2010.

In light of these results, policymakers than face a clear trade-off between the positive effect of stopping speculation and the negative effect of increasing the bid-ask spread; the net effect of these two opposite forces might be an interesting avenue for future research.

References

Alessandro Beber and Marco Pagano. Short-selling bans around the world: Evidence from the 2007−09 crisis. *The Journal of Finance*, 68(1):343–381, 2013.

Bruno Biais, Pierre Hillion, and Chester Spatt. An empirical analysis of the limit order book and the order flow in the paris bourse. *Journal of Finance*, 50(5):1655–89, December 1995.

Randall S Billingsley, Tunde Kovacs, et al. The 2008 short sale ban: Liquidity, dispersion of opinion, and the cross-section of returns of us financial stocks. *Journal of Banking & Finance*, 35(9):2252–2266, 2011.

Ekkehart Boehmer, Charles M. Jones, and Xiaoyan Zhang. Shackling short sellers: The 2008 shorting ban. *Review of Financial Studies*, 26(6):1363–1400, 2013.

Clive G. Bowsher. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876 – 912, 2007.

Arturo Bris, William N. Goetzmann, and Ning Zhu. Efficiency and the bear: Short sales and markets around the world. *The Journal of Finance*, 62(3):1029–1079, 2007.

Rama Cont, Arseniy Kukanov, and Sasha Stoikov. The price impact of order book events. *Journal of Financial Econometrics*, 12(1):47–88, 2013.

Daryl J. Daley and D. Vere-Jones. *An introduction to the theory of point processes*. Springer, 2nd edition, November 2003.

Douglas W Diamond and Robert E Verrecchia. Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics*, 18(2):277–311, 1987.

Vladimir Filimonov and Didier Sornette. Apparent criticality and calibration issues in the hawkes self-excited point process model: application to high-frequency financial data. *arXiv*, forthcoming.

Lawrence R Glosten and Lawrence E Harris. Estimating the components of the bid/ask spread. *Journal of Financial Economics*, 21(1):123 – 142, 1988.

Lawrence R Glosten and Paul R Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, 14(1): 71–100, 1985.

Ronald L. Goettler, Christine A. Parlour, and Uday Rajan. Informed traders and limit order markets. *Journal of Financial Economics*, 93(1):67 – 87, 2009.

Bruce D Grundy, Bryan Lim, and Patrick Verwijmeren. Do option markets undo restrictions on short sales? evidence from the 2008 short-sale ban. *Journal of Financial Economics*, 106(2):331–348, 2012.

Peter F. Halpin. An em algorithm for hawkes process. *Working Paper*, w.p.

Peter F. Halpin and Paul De Boeck. Modelling dyadic interaction with hawkes process. *Psycometrika*, 2013.

Nikolaus Hautsch. *Econometrics of financial high-frequency data*. Springer, 2011.

Alan G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

Alan G. Hawkes. A Cluster Process Representation of a Self-Exciting Process. *Journal of Applied Probability*, 11(3):493–503, September 1974.

Craig W. Holden and Stacy Jacobsen. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *The Journal of Finance*, forthcoming.

Kose John and Larry H. P. Lang. Insider trading around dividend announcements: Theory and evidence. *The Journal of Finance*, 46(4):pp. 1361–1389, 1991.

Steven N. Kaplan, Tobias J. Moskowitz, and Berk A. Sensoy. The effects of stock lending on security prices: An experiment. *The Journal of Finance*, 68(5):1891–1936, 2013.

Adam C Kolasinski, Adam Reed, and Jacob R Thornock. Can short restrictions actually increase informed short selling? *Financial Management*, 42(1):155–181, 2013.

Albert P. Kyle. Continuous Auctions and Insider Trading. *Econometrica*, 53(6):1315–1335, 1985.

K. Lange. *Numerical Analysis for Statisticians*. Springer, 2010.

P. A. W Lewis and G. S. Shedler. Simulation of nonhomogeneous poisson processes by thinning. *Naval Research Logistics Quarterly*, 26(3):403–413, 1979.

F Lorenzen. Analysis of order clustering using high frequency data: A point process approach. *Working Paper*, 2012.

Ian W. Marsh and Richard Payne. Banning short sales and market quality: The uk's experience. *Journal of Banking & Finance*, 36(7):1975 – 1986, 2012.

Y. Ogata. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30:243–261, 1978.

Y. Ogata. On lewis' simulation method for point processes. *Information Theory, IEEE Transactions on*, 27(1):23–31, 1981.

T. Ozaki. Maximum likelihood estimation of hawkes' self exciting process. *Ann. Inst. Statist Math.*, 31:145–155, 1979.

Pedro AC Saffi and Kari Sigurdsson. Price efficiency and short selling. *Review of Financial Studies*, page hhq124, 2010.

Alejandro Veen and Frederic P Schoenberg. Estimation of space time branching process models in seismology using an em type algorithm. *Journal of the American Statistical Association*, 103(482):614–624, 2008.

A Monte Carlo simulation

This appendix presents the Monte Carlo experiment done in order to test the validity of the estimation code. The estimation was done using an EM algorithm allowing for a rounding error following Veen and Schoenberg (2008), Halpin and De Boeck (2013) and Halpin (w.p.). These papers already check the asymptotic consistency of the estimator, however it seems appropriate to summarize some details in this appendix. Nevertheless, in order to maintain the brevity of this appendix, the reader will be referred to different references for the technical details.

This appendix is divided in 3 different sections: first, I present the simulation process, second the estimation procedure and finally the results of the Monte Carlo experiment.

A.1 Simulation

The simulation algorithm is based on the fact that at each moment t the intensity of the Hawkes process is deterministic until a new event arrives and therefore is just an inhomogeneous Poisson process that can be simulated following the thinning algorithm presented in Lewis and Shedler (1979) and extended to multivariate processes in Ogata (1981). Moreover, the simulation algorithm is closely related with the branching structure of the Hawkes process. Therefore to provide more intuition about this representation of the intensity process, I discuss the algorithm for the simulation of a N-variate Hawkes process with intensity, below:

$$\lambda^m(t) = \lambda_0^m + \sum_{n=1}^{N} \int_0^t \alpha_{mn} e^{-\beta_{mn}(t-s)} dJ_s^n \ \ \forall m \quad \alpha_{mn} \geq 0\ , \beta_{mn} \geq 0 \ \ \forall n, m \tag{5}$$

where $J_s \equiv (J_s^1, ..., J_s^N)$ is a càdlàg multivariate simple counting process at s.

In order to initialize the algorithm, set 0 as the time of the first event and the type, m, can be randomized such that:

$$P(\{t_0^m = 0\}) = \frac{\lambda_0^m}{\sum_{n=1}^N \lambda_0^n}$$

Afterwards, following Lewis and Shedler (1979), create a maximum intensity $\lambda^*(v)$ and draw a realization of a Poiss($\lambda^*(t_{i-1})$):

$$s = s - \frac{1}{\lambda^*(t_{i-1})} * ln(U) \ \text{ where } \ U \sim U(0,1)$$

where due to the càdlàg property and the fact that $\alpha \geq 0$ we have that:

$$\lambda^*(v) \equiv \sum_{n=1}^{N} \lambda^n(v) + \sum_{n=1}^{N} \alpha_{mn} e^{v - t_{i-1}} \Rightarrow \lambda^*(t_{i-1}) \geq \lambda^m(s) \ \forall m, s \ \ s.t \ \ s \geq t$$

The following step is the thinning stage, where the time s will be taken as a new event with probability $\sum_{n=1}^N \lambda^n(s)/\lambda^*(t_{i-1})$; otherwise, a new potential event s should be drawn from a Poiss($\lambda^*(s)$)

Finally, the type of event is set as in the initialization part:

$$P\left(\{s \in \{t_0^m, ..., t_{N_m}^m\}\}\right) = \frac{\lambda^m(s)}{\sum_{n=1}^N \lambda^n(s)}$$

There is an important technical issue that should be addressed. In most of the contexts (e.g. firms defaults) we do not observe the first event, therefore when the simulation is done the first events should be eliminated (Burn-in period). This correction will generate edge effects (see Daley and Vere-Jones, 2003) as we see in the data. However, in order to be able to realize a big number of simulations in the Monte Carlo experiment I am restricted to small samples and therefore edge effects will bias the estimates. Moreover, it can be thought that in each trading day, a new Hawkes process starts; therefore, there is no need to eliminate those observations. As a robustness check, I repeat the Monte Carlo experiment for a larger sample with edge effects and less simulations obtaining similar results (Not provided).

The quality of the simulation can be checked using the Random time change theorem (see Daley and Vere-Jones, 2003, Theorem 7.4.I), that implies that:

$$\{\Lambda_i^m\}_{i=1}^{N_m} \sim exp(1) \;\; \forall m \;\; \text{where} \;\; \Lambda_i^m \equiv \int_{t_{i-1}^m}^{t_i^m} \lambda^m(s)ds$$

I have done 2 different simulations, the first one corresponds to the simulation that is going to be used in the Monte Carlo experiment ($\approx$ 1000 events) and the result is presented in Figure 9 where the distribution of the residuals, $\{\Lambda_i^m\}_{i=1}^{N_m}$, is compared with a standard exponential using QQ-plots. The second one corresponds to a bigger sample ($\approx$ 100,000 events) simulation, and the same results are plotted in Figure 10 .

[Figures 9 and 10]

A.2 Estimation

The estimation of the Hawkes process is done by Maximum Likelihood Estimation, which has been proved to deliver consistent and asymptotically normal estimates (Hawkes, 1971; Ogata, 1978). However, if the log-likelihood function is maximized using standard numerical algorithms, the estimates are unstable and the computation time is increasing quadratically with the size of the estimation window.[13]

To solve the stability problem, Veen and Schoenberg (2008); Halpin and De Boeck (2013) proposed the use of an EM algorithm using the branching structure representation of the Hawkes process, that is equivalent to the intensity representation (Hawkes, 1974).

For simplicity, let's consider a sample $\{t_1, t_2 ..., t_N\}$ from a univariate Hawkes process with intensity given by:

$$\lambda(t) = \lambda_0(t) + \sum_{t_k<t} \alpha e^{-\beta(t-t_k)}.$$

Then, the branching structure consists of dividing the events in two classes:

[13]Given the exponential parametrization, the likelihood can be computed in a more efficient manner using an iterative method (linearly increasing)

1. Immigrants: Operations that take place because of a process that is independent of the previous operations. This process will follow an inhomogeneous Poisson distribution with intensity $\lambda_0(t)$.

2. Offspring: Each operation can be a "child" of previous operations. That is each "parent" operation will trigger operations following an inhomogeneous Poisson process with intensity $\alpha e^{-\beta(t-t_k)}$.

Assuming that all these processes are independent, we get the intensity of the Hawkes process. In order to present the log-likelihood, some additional notation is needed:

- $\lambda_k(t) \equiv$ Intensity of having a "child" at t for a "parent" born at t_k .
- $\mathbb{1}\{Z_s = z_k\} \equiv 1$ if the event at t_s is a daughter of the event at t_k.
- $\mathbb{1}\{Z_s = z_0\} \equiv 1$ if the event at t_s is an immigrant.
- $\theta \equiv$ Vector of parameters.
- $\Theta \equiv [0, M]^3 \equiv$ Parameter space, where M is some big real number.
- $\Lambda_k(\underline{t}, \bar{t}) \equiv \int\limits_{\underline{t}}^{\bar{t}} (1 - \lambda_k(t))dt$
- $r \equiv$ Iteration index for the EM algorithm.
- $\mathcal{F} \equiv \sigma(\{t_1, ..., t_N\}) \equiv$ Sigma algebra generated by the sample.

If the branching structure is known (we can differentiate the "parent" of each "child"), the log-likelihood will be trivially given by:

$$\mathcal{L}(\theta|\mathcal{F}) = \sum_{k=0}^{N} \left[\int_0^T ln(\lambda_k(s))\mathbb{1}\{Z_s = z_k\}dN_s - \Lambda_k(0, T) \right]$$

Obviously, the branching structure is not known in most of the cases, but the estimation problem is very similar to the estimation of a mixture of probability distributions when the weights are not known, hence the EM algorithm is naturally applied. The reader is referred to Lange (2010) for the technical details of the EM algorithm and to Halpin and De Boeck (2013, Appendix) for the technical details about its use in this context. The estimation algorithm will be as follows:

- E-Step: The expectation of the log-likelihood is taken over the conditional distribution of Z computed using the parameters of the iteration before, that is:

$$\mathcal{Q}(\theta|\mathcal{F};\ \theta^{r-1}) = \sum_{k=0}^{N} \left[\int_0^T ln(\lambda_k(s))Pr(\{Z_s = z_k\}|\mathcal{F};\ \theta^{r-1})dN_s - \Lambda_k(0, T) \right]$$

  where:

$$Pr(\{Z_s = z_k\}|\mathcal{F};\ \theta^{r-1}) = \frac{\lambda_k^{r-1}(s)}{\sum_{k<s} \lambda_k^{r-1}(s)}.$$

- M-Step: Update the parameters setting:

$$\theta^r = \arg\max_{\theta\in\Theta} \mathcal{Q}(\theta|\mathcal{F};\ \theta^{r-1}).$$

Finally in order to solve the quadratic increase in the computing time, in this paper as proposed by Halpin (w.p.) the algorithm includes a rounding error $\omega = 10^{-10}$ such that the following approximation is done (as proposed by Halpin (w.p.)):

$$\lambda_k^{r-1}(s) < \omega \Rightarrow Pr(\{Z_s = z_k\}|\mathcal{F};\ \theta^{r-1}) \equiv 0.$$

Finally, I present the results of the Monte Carlo experiment that are complementaries to those of Halpin (w.p.).

A.3 Results

In this section I present the results of the Monte Carlo experiment, the experiment consist of 500 simulations without burn-in period of a bivariate Hawkes Process with intensity given by:[14]

$$\lambda^1(t) = 0.15 + \sum_{t_k<t} 0.5e^{1.2(t-t_k^1)} + 0.1e^{0.5(t-t_k^2)}. \tag{6}$$

$$\lambda^2(t) = 0.25 + \sum_{t_k<t} 0.2e^{0.7(t-t_k^1)} + 0.6e^{1.3(t-t_k^2)}. \tag{7}$$

In Table 5, some statistics are presented, concretely, the first columns present the true value, the second and third column corresponds to the mean and standard deviation of the estimates and finally the fourth column shows the inter-quartil range. Moreover, Figure 11 shows the densities of the estimates.

[Table 5 and Figure 11]

B Gradient and Hessian

This appendix just presents the Hessian of the ML in analytical form (For the univariate process the more simple analytical expression can be found in Ozaki (1979)). Lets consider a sample of M-variate Hawkes process between $[0, T]$ given by the sequences $\{t_1^m, ..., t_{N_m}\}$ $\forall m = \{1, ..., M\}$. Therefore the log-likelihood of each type of event is given by:

$$\mathcal{L}^m(\theta|\mathcal{F}) = \sum_{k=1}^{N_m} ln(\lambda(t_k^m)) - \int_0^T (1 - \lambda^m(t))dt$$

where $\lambda^m(t)$ is defined as in Equation 5.

[14]It is aimed to reflect the sample size of Small Sales (see Section 3 for a definition of SS in Iberdrola.

It is trivial to see that the parameters of different types are independent, therefore I show below the Hessian only for the parameters of the m type, $\theta^m \equiv \{\alpha_{m1}, ..., \alpha_{mM}, \beta_{m1}, ..., \beta_{mM}, \lambda_0^m\}$. For convenience, let introduce some notation:

- $\lambda_i^m \equiv \lambda(t_i^m)$.
- $A_n^i \equiv \sum_{t_k^n < t_i^m} e^{-\beta_{mn}(t_i^m - t_k^n)} \;\; \forall \;\; n = \{1, 2, ..., M\}$.
- $B_n^i \equiv \sum_{t_k^n < t_i^m} e^{-\beta_{mn}(t_i^m - t_k^n)}(t_i^m - t_k^n) \;\; \forall \;\; n = \{1, 2, ..., M\}$.
- $C_n^i \equiv \sum_{t_k^n < t_i^m} e^{-\beta_{mn}(t_i^m - t_k^n)}(t_i^m - t_k^n)^2 \;\; \forall \;\; n = \{1, 2, ..., M\}$.
- $\mathbf{A_i} \equiv \{A_1^i ... A_M^i\}' \;\; \mathbf{B_i} \equiv \{B_1^i ... B_M^i\}' \;\; \mathbf{C_i} \equiv \{C_1^i ... C_M^i\}'$.
- $\beta_m \equiv \{\beta_{m1} ... \beta_{mM}\}' \;\; \alpha_m \equiv \{\alpha_{m1} ... \alpha_{mM}\}'$.

The gradient is given by:

$$\frac{\delta \mathcal{L}^m}{\delta \alpha_{mn}} = \sum_{i=1}^{N_m} \frac{1}{\lambda_i^m} A_n^i + \sum_{k=1}^{N_n} \frac{1}{\beta_{mn}} \left(e^{-\beta_{mn}(T - t_k^n)} - 1 \right) \;\; \forall \;\; n = \{1, ..., M\}$$

$$\frac{\delta \mathcal{L}^m}{\delta \beta_{mn}} = -\alpha_{mn} \sum_{i=1}^{N_m} \frac{1}{\lambda_i^m} B_n^i + \alpha_{mn} \{\mathbf{D}\}_n \;\; \forall \;\; n = \{1, ..., M\}$$

where [15]

$$\{\mathbf{D}\}_n = -\sum_{k=1}^{N_n} \frac{1}{\beta_{mn}^2} \left(e^{-\beta_{mn}(T - t_k^n)} - 1 \right) - \frac{1}{\beta_{mn}} e^{-\beta_{mn}(T - t_k^n)} (T - t_k^n) \;\; \forall \;\; n = \{1, ..., M\} \tag{8}$$

$$\frac{\delta \mathcal{L}^m}{\delta \lambda_0^m} = -T + \sum_{i=1}^{N_m} \frac{1}{\lambda_i^m}$$

The Hessian can be easily computed by pieces as: [16]

$$\frac{\delta^2 \mathcal{L}^m}{\delta \alpha_m \delta \alpha_m'} = -\sum_{i=1}^{N_m} \frac{1}{\lambda_i^m} \mathbf{A_i} \mathbf{A_i}'$$

$$\frac{\delta^2 \mathcal{L}^m}{\delta \beta_m \delta \alpha_m'} = \sum_{i=1}^{N_m} \frac{1}{\lambda_i^m} \left(\frac{1}{\lambda_i^m} \mathbf{A_i} \left(\alpha_m \odot \mathbf{B_i} \right)' - diag(\mathbf{B_i}) \right) + diag(\mathbf{D})$$

$$\frac{\delta^2 \mathcal{L}^m}{\delta \alpha_m \delta \lambda_0^m} = -\sum_{i=1}^{N_m} \left(\frac{1}{\lambda_i^m} \right)^2 \mathbf{A_i}$$

[15] The notation $\{X\}_n$ is use to represent the n^{th} entry of a column vector X.
[16] $diag(X)$ corresponds to a diagonal matrix whose diagonal is the vector X.

$$\frac{\delta^2 \mathcal{L}^m}{\delta \beta_m \delta \beta_m'} = \sum_{i=1}^{N_m} \frac{1}{\lambda_i^m} \left(diag\left(\alpha_m \odot \mathbf{C_i}\right) - \frac{1}{\lambda_i^m}\left(\alpha_m \odot \mathbf{B_i}\right)\left(\alpha_m \odot \mathbf{B_i}\right)' \right) + diag(\mathbf{F})$$

where

$$\begin{aligned}\{\mathbf{F}\}_n = \alpha_{mn} \sum_{k=1}^{N_n} \frac{1}{\beta_{mn}} e^{-\beta_{mn}(T-t_k^n)}(T-t_k^n)^2 + \frac{2}{\beta_{mn}^2} e^{-\beta_{mn}(T-t_k^n)}(T-t_k^n) \\ + \frac{2}{\beta_{mn}^3}\left(e^{-\beta_{mn}(T-t_k^n)} - 1\right) \quad \forall \ \ n = \{1, ..., M\}\end{aligned} \tag{9}$$

$$\frac{\delta^2 \mathcal{L}^m}{\delta \beta_m \delta \lambda_0^m} = \sum_{i=1}^{N_m} \left(\frac{1}{\lambda_i^m}\right)^2 \left(\alpha_m \odot \mathbf{B_i}\right)$$

$$\frac{\delta^2 \mathcal{L}^m}{\delta \lambda_0^{m^2}} = \sum_{i=1}^{N_m} - \left(\frac{1}{\lambda_i^m}\right)^2$$

C Dealing with rounded time stamp

From Equation 1, it can be seen that the most important term of the intensity, in terms of the estimation of self- or cross-exciting effect (α's and β's) is the $e^{-\beta(t-s)}$. The problem is that most of the observations will take place on the same second and therefore $(t-s)$ will be 0 if the data is rounded. Aside from technical issues, this will bias the results as it will oversize the jump in the intensity.[17]

In order to deal with this problem, two methods have been commonly used: the most naive method is to eliminate observations of the same type that take place at the same time and adding a uniform error which I denote as the *Eliminate* method. The second one which is widely used, e.g. Bowsher (2007) consists in adding a uniform distributed error to the observations respecting the order (if the order is available); I refer to this method as the *Uniform* method. Lorenzen (2012) studied the problems of these methods and concludes that the estimates will not be biased only if the observations are not too clustered.

I propose two different approaches: the first one, labeled *Metropolis*, tries to solve this problem without using additional information and therefore it can reduce but not completely eliminate the bias, a detailed explanation of this method is deferred to Appendix C.1 The second one exploits the annotation number and is presented in this section.

In order to test the asymptotic unbiasedness of the estimator, I conduct a Monte Carlo experiment with 250 simulations with similar size and concentration that we find in daily sales in the data. Concretely, the intensity components of the bivariate simulated process are given by:

$$\lambda^1(t) = 0.05 + \sum_{t_k<t} 0.5e^{1.2(t-t_k^1)} + 0.1e^{0.5(t-t_k^2)} \tag{10}$$

$$\lambda^2(t) = 0.025 + \sum_{t_k<t} 0.2e^{0.7(t-t_k^1)} + 0.6e^{1.3(t-t_k^2)} \tag{11}$$

The results of the Monte Carlo experiment can be seen in the Table 3, where in the first column I present the true parameters, in the second column the estimates of the MLE before rounding the time stamp are presented (Non-Feasible) and finally from the third to the fifth column the results of the different methodologies discussed earlier. As expected the two previous methods used in the literature deliver biased estimates, and also the *Metropolis* method proposed have a bias, however the bias is significantly reduced. In this paper, most of the conclusions depend on the baseline parameter, and it can be seen that this parameter is estimated without bias using *Metropolis* or even using *Uniform*.

[Table 3]

In order to eliminate the bias in all the parameters, as commented above, I use the annotation number as an "instrument" for time. The methodology is very simple, at every second I compute the maximum and the minimum annotation number in the market

[17] The main technical issue is that, by definition, the Hawkes process is a simple point process and then by assumption the probability of getting two observation at the same time is 0, see Daley and Vere-Jones (2003) for a technical definition of a simple process.

(considering any type of LOB movements) and I linearly interpolate each observation. The reader might realize that this procedure is very similar to applying the uniform methodology to every firm and every observation (including the movements of liquidity in the book) respecting the order. But obviously, the proposed procedure will reduce the computing time since in order to get the maximum and minimum of every second you do not need to store the whole dataset.

Holden and Jacobsen (forthcoming) propose a method based on time interpolation that is very similar to the use of the annotation number. Therefore the results in Appendix C.2, where I present some results comparing the Uniform and the "instrument" method are easily extended to the new method proposed in Holden and Jacobsen (forthcoming).

The interesting feature of this approach is that the highest the number of movements in one second the better is the approximation; if the number of movements, however, is very low, then the concentration is small and thus the rounded time stamp is not really a problem as pointed out by Lorenzen (2012).

This new method has the important advantage that it does not influence the residual.[18] This is not the case when we use the other methods, therefore measures of goodness of fit normally used in the literature that are based on the random time change theorem become invalid with previous methods but remain valid with this method.[19] In particular the Metropolis method will overestimate the goodness of fit by construction since the random time change theorem is used to separate the data inside a second. Additionally, this method, as opposed to the *Uniform* method does not create any additional dependence between durations or residuals.

In order to simulate the annotation number I simulate 35 firms at the same time, from a Hawkes process with 4 processes, the ones given by Equation 10 and Equation 11 and other 2 processes with more concentration in order to get realistic movements in the annotation number. Then I estimate only the same bivariate process as before. This exercise is done 300 times and each time the 35 firms are used as a simulation which gives a total of 10500 simulations.

The Monte Carlo experiment results are presented in Table 4. First and second columns are devoted to compare the estimates using the estimation before rounding the simulated data and the estimates done using the new methodology. Moreover, the Monte Carlo experiment is used to assess the accuracy of the standard error estimation and the results are presented in the last columns.

[Table 4]

C.1 Metropolis Method

This appendix presents an algorithm to deal with rounded data when information on the annotation number is not available. This algorithm will reduce significantly the bias in the parameters but as it can be seen in Table 3 the estimates will not be consistent. The algorithm consists of an iterative procedure given by 3 steps:

1. The parameters are estimated using the *Uniform* approach.

[18] Defined as the integral of the intensity between two events.

[19] Intuitively, they are based on the residuals distribution. See Daley and Vere-Jones (2003).

2. The original times are split following the joint distribution conditional on the number of events in a second given by [20]:

$$f(\{t_1, ..., t_m\}|M = m) \propto f(\{t_1, ..., t_m\}, M = m) = e^{-\int_{T_0}^{T_1} \lambda(t)dt} \prod_{i=1}^{m} \lambda(t_i) \tag{12}$$

where M is the number of events between $[T_0, T_1]$ and $\{t_1, ..., t_m\}$ is the time sequence of all observations in the interval $[T_0, T_1]$.

In order to simulate from (12) I use the Metropolis-Hastings algorithm using as proposal for $(t_i - T_0)$ a $Beta(1, \gamma)$ where γ is chosen in each iteration to match the mean of $t_i - T_0$ in the previous iterations. Different values of the first parameter of the proposal have been used and convergence is very similar.

3. Estimate the Hawkes parameters again and go to the second step until convergence.

C.2 Uniform vs Instrument

Let define $\Omega \equiv \{t_0, ..., t_n\}$ the set of times at which operations take place and $\{N_0, ..., N_{n-1}\}$ the corresponding annotation numbers. Let assume w.l.o.g that $N_0 = 0$ and $t_n \leq 1$.

Usually we are not interested in the whole sample but in a subset $\omega \equiv \{t_1^*, ..., t_m^*\} \subset \Omega$ that could be trades but not quotes or the operations of just one firm in the market. Henceforth, to ease notation, time will be indexed by 2 subscripts, such as $t_{s,j} = t_s^* = t_j$

The uniform method will consist in drawing from a uniform m numbers ϵ_i; and order them which is equivalent to draw the numbers from a beta distribution. That is, $\epsilon_i \sim B(i, m + 1 - i)$. The mean square error for the two methods is given by:

$$\mathbb{E}\left[\left(t_{s,j} - \bar{t}_{s,j}^{IV}\right)^2\right] = \mathbb{E}\left[t_{s,j}^2\right] - 2\frac{N_j}{n}\mathbb{E}\left[t_{s,j}\right] + \left(\frac{N_j}{n}\right)^2 \tag{13}$$

$$\mathbb{E}\left[\left(t_{s,j} - \bar{t}_{s,j}^{U}\right)^2\right] = \mathbb{E}\left[t_{s,j}^2\right] - 2\mathbb{E}\left[t_{s,j}\right] \cdot \frac{s}{(m+1)} + \frac{s(s+1)}{(m+1)(m+2)} \tag{14}$$

The non-random nature of the IV approach results in a MSE of 0 if the true times take place uniformly in Ω. On the other hand, in the case of the uniform approach,the MSE at the minimum $(t = \frac{s}{m+1})$ is $\frac{s(m+1-s)}{(m+1)^2(m+2)}$. Therefore even if the times are uniformely distributed the Uniform method does not have a MSE $= 0$ unless $m \to \infty$

Note that the MSE will depend on the ordinal of each operation. The sum over all operations inside a minute is given by:

$$\sum_{s=1}^{m} \mathbb{E}\left[\left(t_{s,j} - \bar{t}_{s,j}^{U}\right)^2\right] = \sum_{s=1}^{m} \mathbb{E}\left[t_{s,j}^2\right] - 2\sum_{s=1}^{m} \mathbb{E}\left[t_{s,j}\right] \frac{s}{(m+1)} + \left(\frac{m}{3}\right) \tag{15}$$

If we evaluate this expression at $t_{s,j} = \frac{s}{m+1}$ the MSE becomes:

[20] The formulas assume a univariate process given by $\lambda(t) = \lambda_0 + \sum_{t_k<t} \alpha e^{\beta(t-t_k)}$, the extension to the multivariate case is trivial but requires much more notation.

$$\sum_{s=1}^{m} \mathbb{E}\left[\left(t_{s,j} - \bar{t}^{U}_{s,j}\right)^2\right] = -\frac{m(2m+1)}{6(m+1)} + \left(\frac{m}{3}\right) \tag{16}$$

which goes to $\frac{1}{6}$ as $m \to \infty$.

In order to compare both methods I will use the difference in MSE that will depend on $\alpha \equiv \frac{N_j}{n}$, $t_{s,j}$, s and m. In the application considered in this paper $m \leq 10$ and the 10% quantile in the observed data, the number of annotations per second is $N > 1000$, therefore, in Figure 12 I present the difference in MSE for $m = 10$, $N = 1000$ and different values of $\alpha \equiv \frac{N_j}{n}$, t_s, j and s.[21]

In Figure 12, the 45^o line corresponds to the times being uniformly distributed over Ω, in this case the IV is more precise independently on the distribution over ω which is a direct implication of Equation 14 and Equation 13. Moreover, the uniform approach is more accurate only if the observations are uniformly distributed over ω but they are not over Ω. The self-exciting nature of the process used in this paper will imply that the distribution over ω is not uniform $(t = \frac{s}{m+1})$ and therefore the IV approach will behave better in terms of MSE unless annotations are extremely high concentrated.[22]

[Figure 12]

Finally it should be noticed that using the Uniform approach, we create another technical problem since the duration process will become serially dependent and thus the process will not be a poisson process as defined by Daley and Vere-Jones (2003). Therefore, most of the results from previous literature (goodness-of-fit, estimation, consistency...) would not necessarily hold.

[21] This parameter values correspond to the worst case scenario for the IV approach.

[22] Unless $\alpha = 0$, which according to the estimation is not a relevant case.

D Tables

	Iberdrola		BBVA	
	Pre-SSB	Post-SSB	Pre-SSB	Post-SSB
Jump Size Parameters				
α_{21}	0.006	0.005	0.003	0.002
	(0.001)	(0.001)	(0.000)	(0.000)
α_{22}	0.205	0.209	0.163	0.160
	(0.002)	(0.002)	(0.001)	(0.001)
Reversion Speed Parameters				
β_{21}	0.007	0.007	0.002	0.002
	(0.003)	(0.001)	(0.000)	(0.000)
β_{22}	0.283	0.278	0.226	0.203
	(0.004)	(0.005)	(0.004)	(0.004)

Table 1: Estimation: Large & Small Sales

Note: Estimates of a bivariate Hawkes process applied to big sales (labelled as 1) and small sales (labelled as 2) of Iberdrola and BBVA before and after the SSB. The parameters α_{ij} correspond to the size of the jump on the intensity of the event i caused by the event j; similarly the parameters β_{ij} measure the reversion speed.
Standard errors computed using the inverse information matrix are in parenthesis.

	Pool		Day by Day	
	Bid-Ask (b.p.)	OFI (%)	Bid-Ask (b.p.)	OFI (%)
SAN	0.464	-30.588	0.047	-0.283
	(0.075)	(3.587)	(0.006)	(0.228)
BBVA	-0.264	-18.730	0.047	-0.457
	(0.744)	(3.460)	(0.007)	(0.267)
TEF	0.199	-40.731	0.066	-0.748
	(0.123)	(4.904)	(0.007)	(0.244)
IBE	0.472	-4.788	0.116	0.120
	(0.228)	(6.195)	(0.016)	(0.411)
REP	0.104	25.189	0.006	1.627
	(0.030)	(4.879)	(0.002)	(0.318)
ITX	1.079	4.099	0.128	0.350
	(0.200)	(5.752)	(0.018)	(0.433)
Mean	*0.343*	*-10.920*	*0.068*	*0.113*
	(0.136)	*(2.000)*	*(0.004)*	*(1.487)*

Table 2: Add caption

Note: Estimates of the effect of the ban on he bid-ask spread and the order flow imbalance (OFI). The first two columns present the estimate obtained by regressing the variable of interest on its lag, a 10th degree time polynomial and a dummy that takes value one after the policy is implemented. To construct the estimates in the last two columns I regress, day by day, the dependent variable on its lag and a dummy that take value one after 14:28 p.m. (the time of implementation). In these columns I present the estimate for the day of the ban which I normalize subtracting the mean of the other estimates for each firm.

	True	Non-Feasible	Feasible		
		Non-Round	Metropolis	Uniform	Eliminate
Jump Size Parameters					
α_{11}	*0.5*	0.500	0.429	0.406	1.299
α_{21}	*0.2*	0.201	0.193	0.185	0.234
α_{12}	*0.1*	0.100	0.095	0.096	0.209
α_{22}	*0.6*	0.602	0.524	0.502	1.615
Reversion Speed Parameters					
β_{11}	*1.2*	1.200	0.999	0.944	4.735
β_{21}	*0.7*	0.704	0.680	0.653	0.672
β_{12}	*0.5*	0.504	0.495	0.504	0.968
β_{22}	*1.3*	1.302	1.120	1.070	4.9887
Baseline Parameters					
λ_0^1	*0.05*	0.050	0.049	0.049	0.0653
λ_0^2	*0.025*	0.025	0.024	0.024	0.033

Table 3: Monte Carlo: Rounding Data

Note: Estimates of the Hawkes process with intensity given by Equation 10 and Equation 11. The first column presents the true parameter values, the second column are estimates before rounding the simulated data, and the third to fifth columns present the estimates using three different methods to deal with rounded data. The *Metropolis* method is explained in the Appendix C, the *Uniform* method consists in summing a uniform random variable to the rounded time stamp and the *Eliminate* method is based on selecting one observation each second (and dropping the rest).

	Estimates		S.E. (Non-Feasible)		S.E. (Instrument)	
	Non-Feasible	Instrument	M.C.	Estimate	M.C.	Estimate
Jump Size Parameters						
α_{11}	0.501	0.500	0.024	0.015	0.019	0.015
α_{21}	0.200	0.202	0.013	0.008	0.010	0.008
α_{12}	0.101	0.102	0.010	0.006	0.007	0.006
α_{22}	0.601	0.599	0.029	0.018	0.024	0.018
Reversion Speed Parameters						
β_{11}	1.203	1.203	0.059	0.058	0.043	0.058
β_{21}	0.702	0.708	0.044	0.041	0.035	0.042
β_{12}	0.503	0.514	0.048	0.047	0.035	0.048
β_{22}	1.303	1.300	0.060	0.059	0.045	0.059
Baseline Parameters						
λ_0^1	0.050	0.050	0.000	0.002	0.000	0.002
λ_0^2	0.025	0.025	0.000	0.001	0.00	0.001

Table 4: Monte Carlo: Instrument

Note: Estimates of the Hawkes process with intensity given by Equation 10 and Equation 11. The first two column presents the parameter estimates before rounding and applying the *Instrument* method to the rounded time stamp. Third and fourth columns compare the Monte Carlo standard error of the parameters with its estimate using the asymptotic variance and data before rounding. Finally, fifth and sixth columns present the Monte Carlo standard error of the parameters with its estimate using the asymptotic variance and the *Instrument* method The columns labeled as M.C. (Monte Carlo) present the standard deviation of the estimates obtained in each simulation.

	True	Mean	Std	90% CI
Jump Size Parameters				
α_{11}	*0.5*	0.498	0.022	[0.462-0.535]
α_{21}	*0.2*	0.200	0.010	[0.185-0.216]
α_{12}	*0.1*	0.100	0.004	[0.093-0.107]
α_{22}	*0.6*	0.600	0.021	[0.569-0.634]
Reversion Speed Parameters				
β_{11}	*1.2*	1.199	0.031	[1.151-1.249]
β_{21}	*0.7*	0.700	0.017	[0.672-0.726]
β_{12}	*0.5*	0.497	0.013	[0.479-0.516]
β_{22}	*1.3*	1.302	0.025	[1.262-1.344]
Baseline Parameters				
λ_0^1	*0.15*	0.150	0.005	[0.141-0.158]
λ_0^2	*0.25*	0.250	0.007	[0.239-0.261]

Table 5: MC Descriptive Statistics

Note: Summary statistic of estimates obtained estimating 500 different simulated Hawkes Process with intensities given by Equation 6 and Equation 7.

E Figures

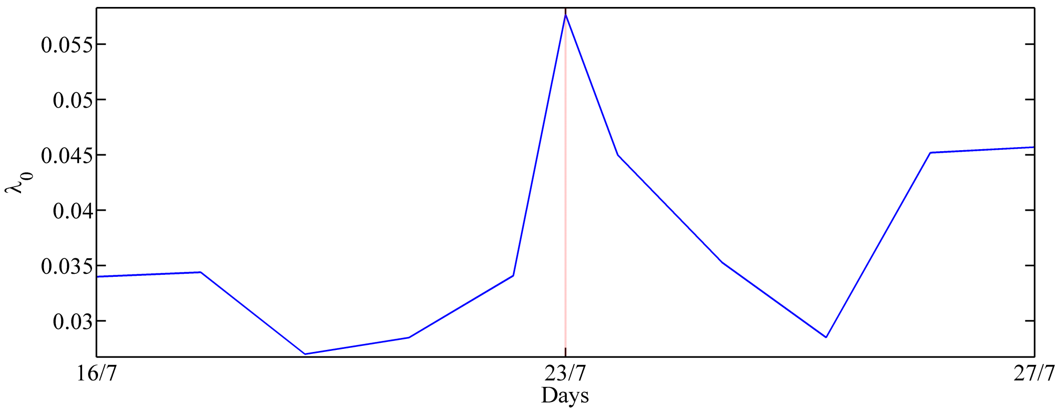

Figure 1: Baseline Intensity BBVA. This plot shows the baseline parameter for BBVA from July 16, 2012 to July 27, 2012, where the parameter is assumed to be constant during each day except for the case of the SSB date where it changes at the moment of the ban. Dashed lines represent the 95 % confidence interval. The vertical line corresponds to the SSB implementation.

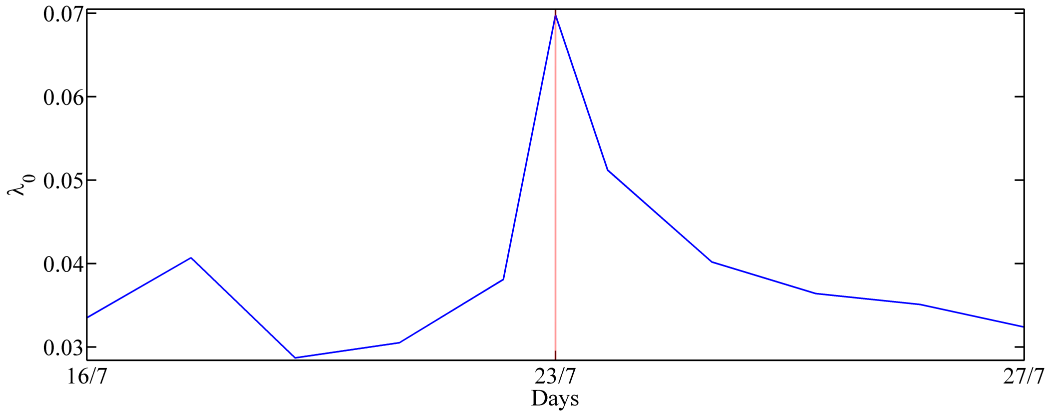

Figure 2: Baseline Intensity Iberdrola. This plot shows the baseline parameter for Iberdrola from July 16, 2012 to July 27, 2012, where the parameter is assumed to be constant during each day except for the case of the SSB date where it changes at the moment of the ban. Dashed lines represent the 95 % confidence interval. The vertical line corresponds to the SSB implementation.

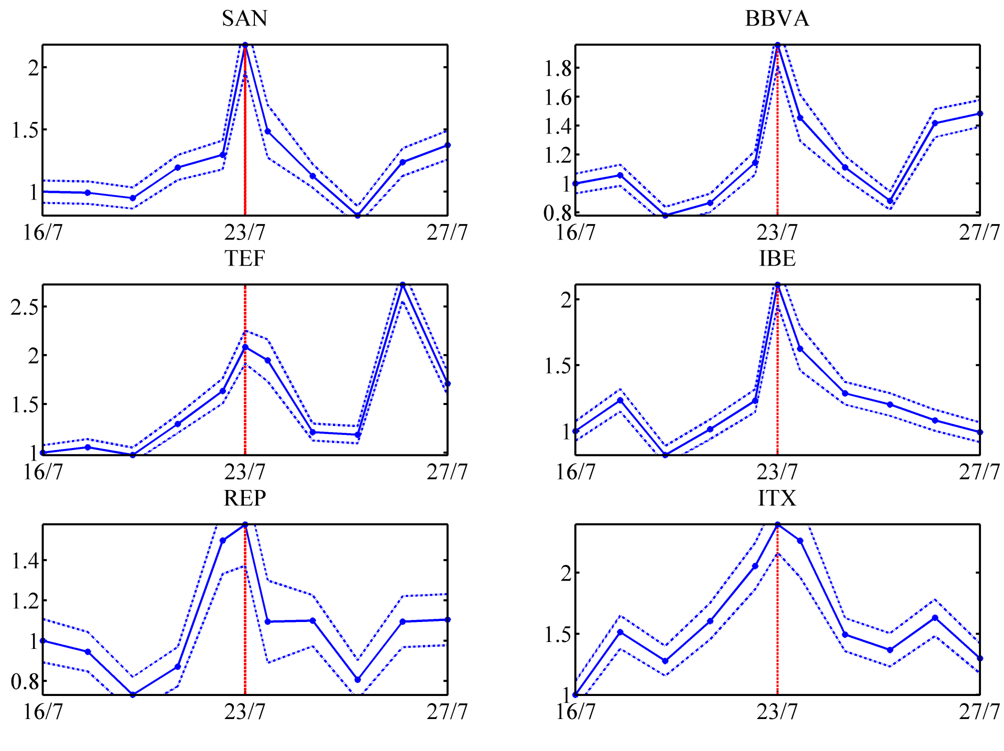

Figure 3: Baseline Intensity. This plot shows the baseline parameter for the six firms with highest market value from July 16, 2012 to July 27, 2012, where the parameter is assumed to be constant during each day except for the case of the SSB date where it changes at the moment of the ban. Dashed lines represent the 95 % confidence interval. Parameter of the first day is normalized to 1. The vertical line corresponds to the SSB implementation.

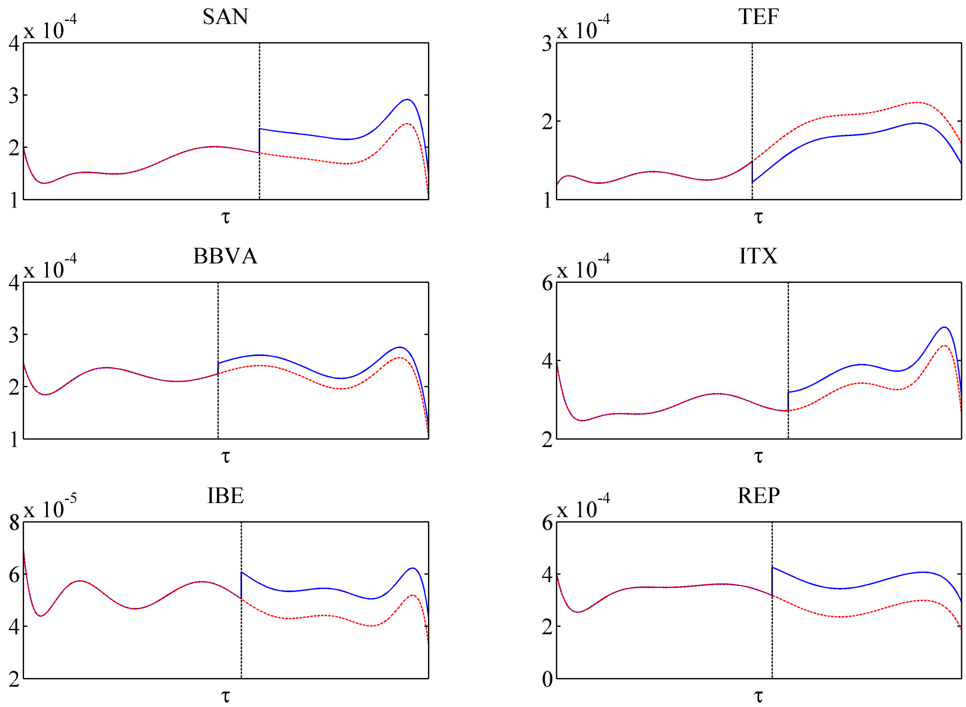

Figure 4: Y = bid-ask spread. The solid blue line is given by the residual, $\hat{\eta_\tau} = \hat{\alpha} + \sum_{i=1}^{p} \hat{\gamma_i} \cdot H_i(t) + \hat{\beta} \cdot EVENT_\tau$ and the dashed line is computed by assuming that $\beta = 0$. The vertical dash line represents the moment at which the SSB took place. The estimation period is from July 16, 2012 to July 27, 2012. The vertical line corresponds to the SSB implementation.

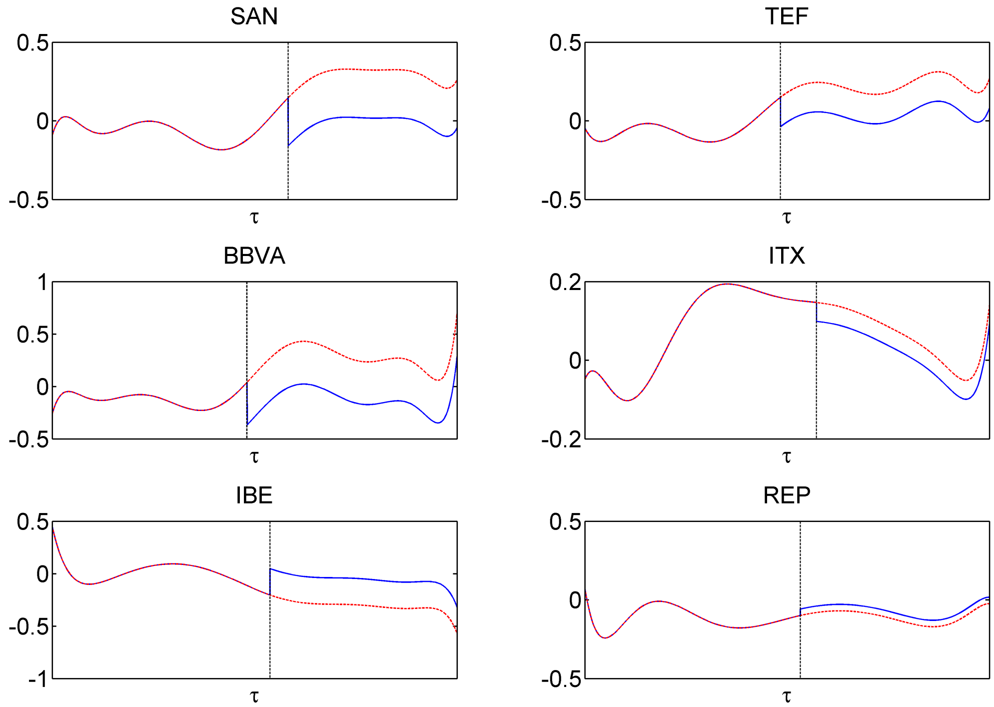

Figure 5: Y = order flow imbalance. The solid blue line is given by the residual, $\hat{\eta}_\tau = \hat{\alpha} + \sum_{i=1}^{p} \hat{\gamma}_i \cdot H_i(t) + \hat{\beta} \cdot EVENT_\tau$ and the dashed line is computed by assuming that $\beta = 0$, or equivalently, assuming that the ban has no effect. The vertical dash line represents the moment at which the SSB took place. The estimation period is from July 16, 2012 to July 27, 2012. The vertical line corresponds to the SSB implementation.

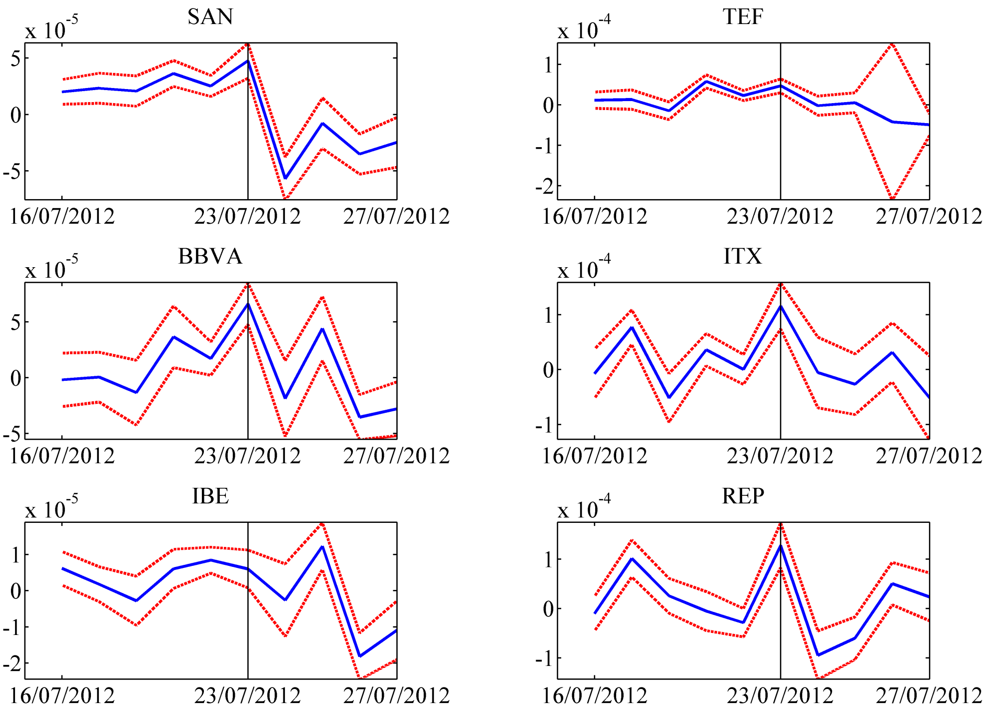

Figure 6: Y = bid-ask spread (Placebo tests). The continuous blue line represents the estimate of β in Equation 4 for each day in the sample: July 16, 2012 to July 27, 2012. The red dashed lines represent the 95 % confidence interval. The vertical line corresponds to the SSB implementation.

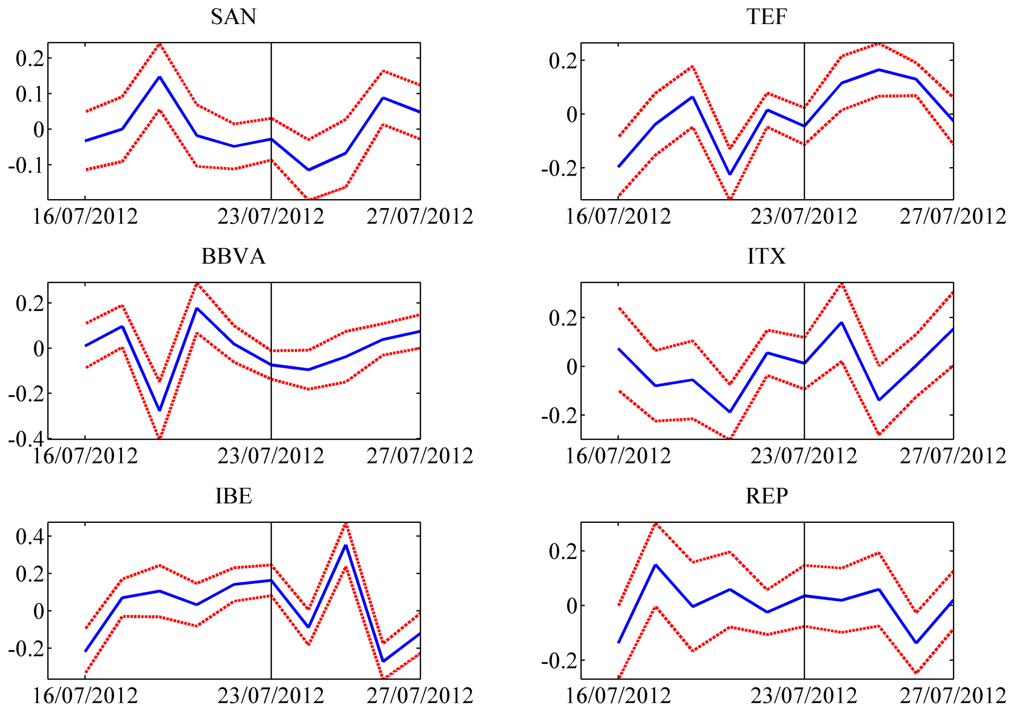

Figure 7: Y = order flow imbalance (Placebo tests). The continuous blue line represents the estimate of β in Equation 4 for each day in the sample: from July 16, 2012 to July 27, 2012. The red dashed lines represent the 95 % confidence interval. The vertical line corresponds to the SSB implementation.

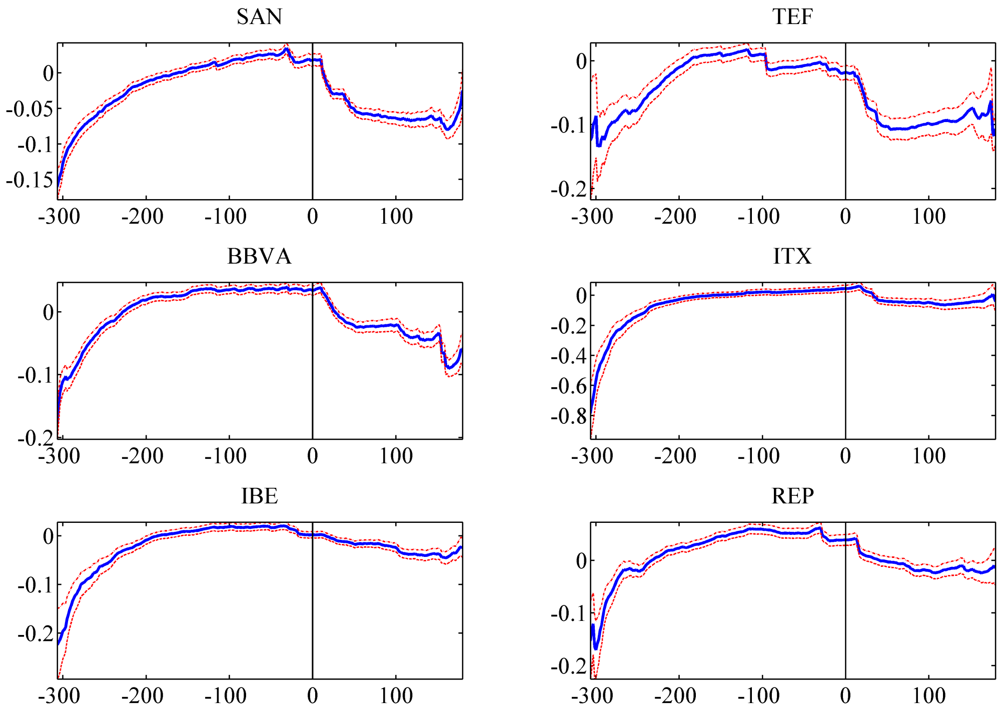

Figure 8: Y = bid-ask spread (Intra-day Placebo tests). The continuous blue line represents the estimate of β in Equation 4 changing the cut-off point. The x axis is measured in minutes from the time of the implementation. The red dashed lines represent the 95 % confidence interval. The vertical line corresponds to the SSB implementation.

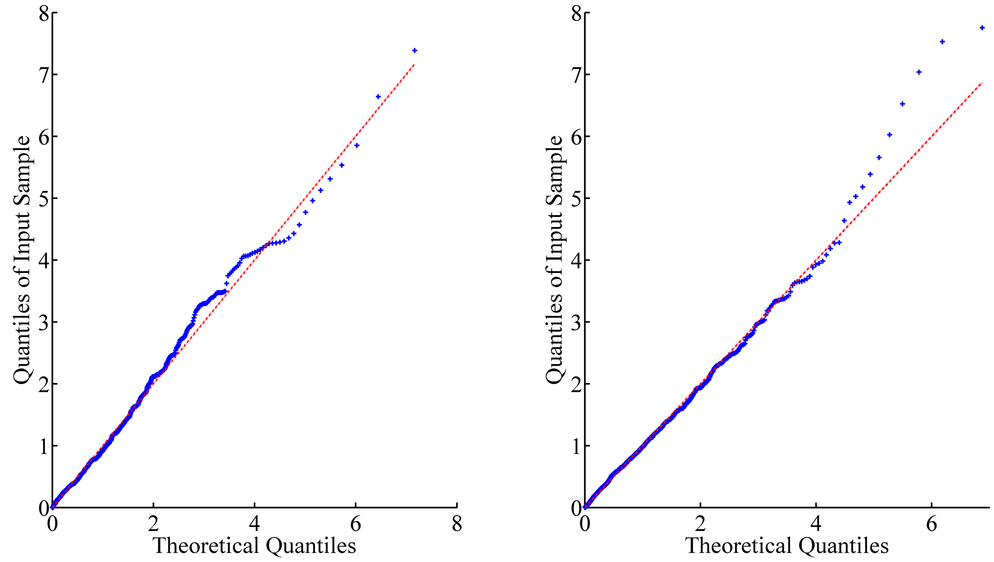

Figure 9: QQ-plots Residuals vs exp(1). QQ-plots of the residuals,$\{\Lambda_i^m\}_{i=1}^{N_m}$, of the simulated process against their theoretical distribution, exp(1). The graphs correspond, from left to right, to the first and second type of event of a bivariate process estimated in the interval [0,1000] (1189 events)with intensities given by: $\lambda^1 = 0.15 + \sum_{t_k<t} 0.5e^{1.2(t-t_k^1)} + 0.1e^{0.5(t-t_k^2)}$; $\lambda^2 = 0.25 + \sum_{t_k<t} 0.2e^{0.7(t-t_k^1)} + 0.6e^{1.3(t-t_k^2)}$.

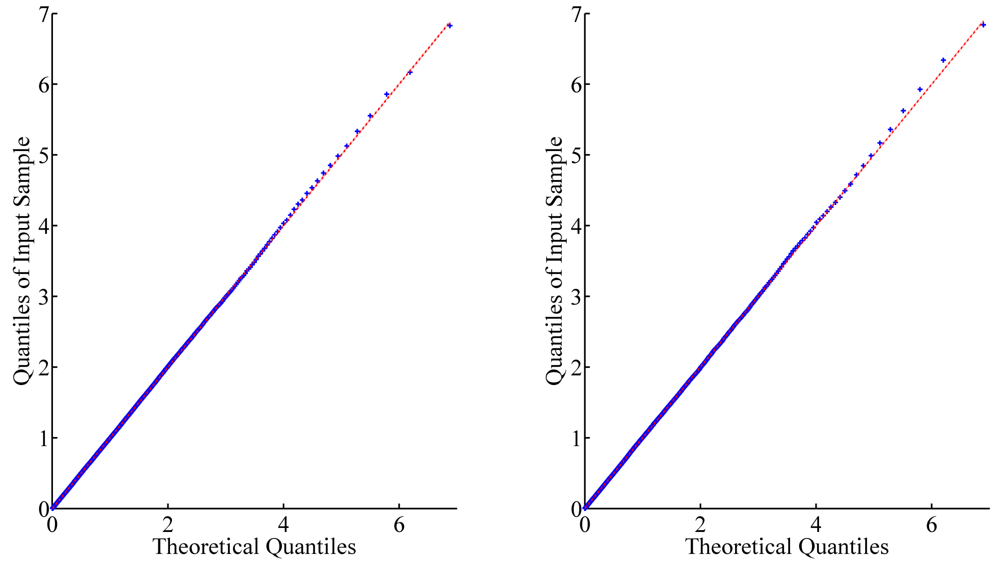

Figure 10: QQ-plots Residuals vs exp(1). QQ-plots of the residuals,$\{\Lambda_i^m\}_{i=1}^{N_m}$, of the simulated process against their theoretical distribution, exp(1). The graphs correspond, from left to right, to the first and second type of event of a bivariate process estimated in the interval [0,30600] (111650 events) with intensities given by: $\lambda^1 = 0.4 + \sum_{t_k<t} 0.6e^{0.8(t-t_k^1)} + 0.2e^{1.8(t-t_k^2)}$; $\lambda^2 = 0.5 + \sum_{t_k<t} 0.43e^{1.3(t-t_k^1)} + 0.2e^{1.9(t-t_k^2)}$.

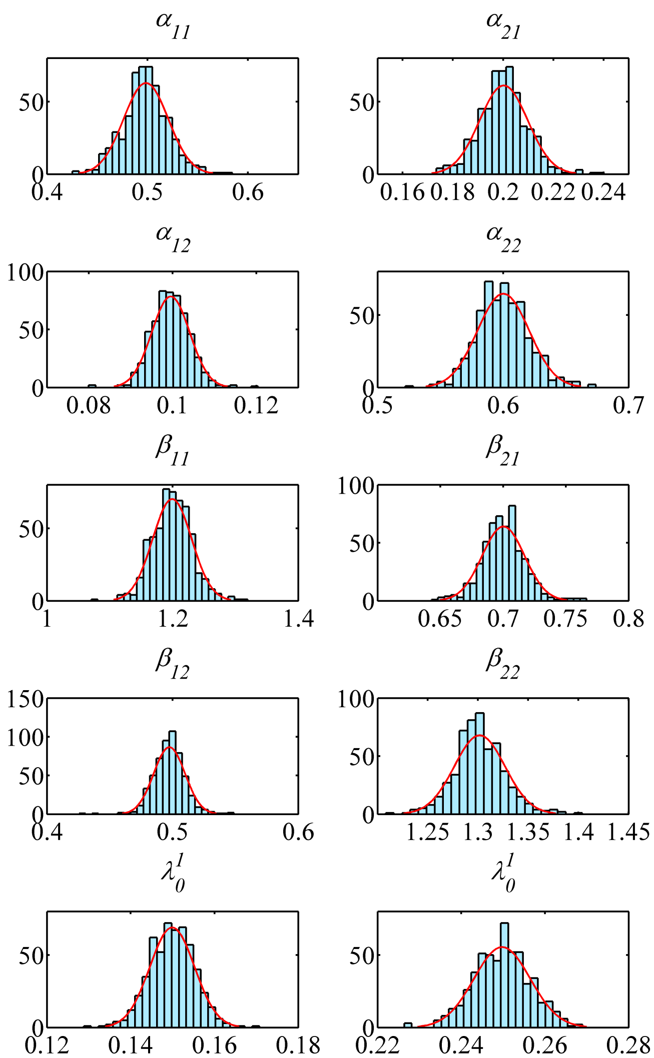

Figure 11: Monte Carlo Simulated Distributions. Summary statistic of estimates obtained estimating 500 different simulated Hawkes Process with intensities given by Equation 6 and Equation 7 which imply that the parameters in the figure notation are: $\alpha_{11} = 0.5$, $\alpha_{12} = 0.1$, $\alpha_{21} = 0.2$, $\alpha_{22} = 0.6$, $\beta_{11} = 1.2$, $\beta_{12} = 0.5$, $\beta_{21} = 0.7$, $\beta_{22} = 1.3$, $\lambda_0^1 = 0.15$ and $\lambda_0^2 = 0.25$. In red is plotted the closest normal distribution.

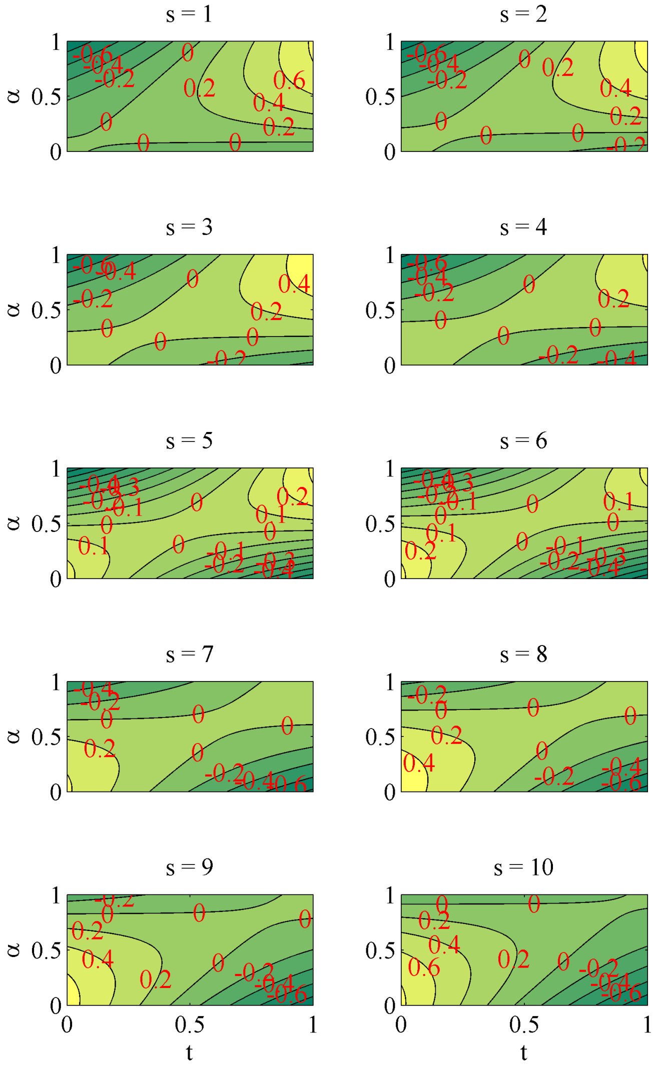

Figure 12: Difference in MSE: Uniform vs Instrument. Difference in MSE defined as: $\mathbb{E}\left[\left(t_{s,j}-\bar{t}^{U}_{s,j}\right)^2\right]-\mathbb{E}\left[\left(t_{s,j}-\bar{t}^{IV}_{s,j}\right)^2\right]$. This graph is done assuming there are 10 events inside a second, s indicates the ordinal position of each event, $\alpha \in [0,1]$ is the relative position of the operation relative to all the movements and t is the true time at which the order takes place.