ESSAYS IN DYNAMIC DURATION AND COUNT MODELLING

Lorenzo Braccini

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Francis X. Diebold
Professor of Economics

Graduate Group Chairperson

George Mailath
Professor of Economics

Dissertation Committee

Francis X. Diebold, Professor of Economics
Francis J. Di Traglia, Professor of Economics
Frank Schorfheide, Professor of Economics

ESSAYS IN DYNAMIC DURATION AND COUNT MODELLING

A Marta,

ai miei genitori, Stefano e Carla,

ed ai miei nonni.

ACKNOWLEDGEMENTS

The process of completing graduate studies is certainly a demanding one. Research requires hard work and sometimes it repays you with the satisfaction of an improved knowledge. However, sometimes it daunts you, it overwhelms you with the complexity of reality–where often the answers to the questions you pose just sound terribly alike other questions. Guidance and support are fundamental ingredients to successfully complete such process.

During my stay at the University of Pennsylvania, I had the honor of working with a wonderful advisor and an even better scholar and econometrician, economist and statistician. Prof. Diebold continued guidance, patience and wisdom greatly contributed to shape my view of the discipline and of the profession, and to train me as an econometrician and as a person. It would have been impossible to have achieved my goals without his help and I will always be extremely grateful to him.

I am also grateful to the econometrics group of the Department of Economics and, in particular, to my committee members Prof. Di Traglia and Prof. Schorfheide. Their guidance, support and advice have been an invaluable source of inspiration and improvement.

Thanks to Prof. Cheng, Ross Askanazi, Laura Liu, Minchul Shin, Jacob Warren, Molin Zhong and all of the Friday Research Group for sharing their knowledge and comments during seminars and private conversations.

The Department of Economics staff has been very supportive, kind and timely throughout my studies and special thanks go to Kelly Quinn and Charmaine Thomas.

During these years I also have been fortunate enough to encounter many great new friends with whom I shared the difficult moments and the good times, and that never let me down. Thanks to Alberto, Alessandro, Daniel, Devin, Ekim, Molin and Murat.

Alessandro and Molin, you have both been great roommates. Thanks for all the interesting discussions about econometrics, statistics, economics and probability (not to mention history, politics, music, movies etc. etc.) that we had but, most importantly, thanks for all the time we spent together, it has really been a great one.

A final thought to my wife and family. Marta, il tuo supporto, la tua dedizione e la tua pazienza sono stati impagabili. Mi hai sopportato nei momenti difficili e mi sei stata sempre accanto. Non potró mai ripagarti per quello che mi hai dato in questi anni. Stefano e Carla, grazie di avermi sempre incoraggiato, consigliato ed appoggiato nelle mie scelte, ma, soprattutto, grazie per tutto l'affetto che mi avete dato.

Lorenzo Braccini

Philadelphia, PA

April 22, 2015

ABSTRACT

ESSAYS IN DYNAMIC DURATION AND COUNT MODELLING

Lorenzo Braccini

Francis X. Diebold

In this dissertation we study the dynamic and static probabilistic structure of the distribution of equity transaction times on financial markets. We propose dynamic, non-linear, non-Gaussian state space models to investigate both the structure of the associated inter-trade durations, and the properties of the number of transactions over a mesh of fixed length. The economic motivation of the study lies in the relationship between the properties of transaction times and those of the time-varying volatility of equity returns and of market liquidity measures such as bid-ask spreads. We use high-frequency data extracted from the Trade and Quotes database to recover transaction time-stamps recorded down to the second or millisecond time scale depending on the sample of analysis. We focus our attention to a randomly selected sub-sample of the S&P100 index traded on U.S. financial markets. Starting from the work of Chen et al. (2013), we propose a dynamic duration model that is able to capture the salient features of the empirical distribution of inter-trade durations for the most recent samples, namely, over-dispersion, long-memory, transaction clustering and *simultaneous trading*. We employ this model to study the structural change in the properties of the transaction process by assessing its ability of fitting the data and its forecasting accuracy over a long span of time (1993-2013). As an alternative tool for the analysis of the transaction times process, and motivated by the necessity of reducing the computational burdens induced by the appearance of data-sets of unprecedented size, we propose a dynamic, long-memory model for the number of transactions over a mesh of fixed length, based on the Markov Switching Multifractal model proposed by Calvet and Fisher (2008). We perform goodness-of-fit and forecasting accuracy comparisons against competing models and find that the proposed model provides a superior performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Introduction

Equity transaction times are a very informative object for a wide variety of quantities associated with measures of risk and liquidity of the considered asset. Their irregular behaviour captures the speed of the information flow and uncover the ability of the market to incorporate such new information into the equilibrium price. A relatively large body of literature confirms this view both from a theoretical and an empirical perspective. Most notably, O'Hara (1995) proposes a model of strategic interaction where the pace of the trading process signals the presence of privately informed traders. More recently, starting from the subordinated change of time intuition of Clark (1973), Ané and Geman (2000) show that the empirical distribution of calendar time returns recovers normality when the cumulative number of transaction is taken as the stochastic clock of the price process. As an immediate implication, such observation implies that the time-varying characteristics of returns volatility are solely determined by the properties of the underlying distribution of transaction times. More explicitly, Deo et al. (2009) demonstrate that the memory parameter of the duration process propagates to that of the counting process and up, completely determining the memory parameter of the associated realized volatility process in a wide varieties of settings.

Recent events such as the Flash Crash and the late 2000s financial crisis testify to our relatively poor understanding of and ability to predict the behavior of financial markets, their price and liquidity dynamics, in particular at high frequencies. The rapid increase in availability of data sampled at ultra high-frequencies–second, millisenceond and nanosecond level–however, allows us to study the main features of equity transaction processes at incredibly fine time scales. Most importantly, high-frequency sampling enables us to study the primitive object of interest that drives the process of dissemination of new information into the market, i.e. transaction times.

The natural framework for the study of equity transaction times is provided by the mathematical theory of point processes and, in particular, by the theory of Cox point processes. In such a framework the central object of interest is the intensity function (or intensity field), $\lambda(t)$. This function completely determines the dynamic and static properties of the event times distribution. The most natural approach to the problem suggests to formulate an Exponential model for the duration process associated with the event times, conditional on the intensity function, where the duration between consecutive events is defined as the time difference between their times of occurrence. An alternative and common approach is to study the counting process associated with the event times, i.e. the number of events in a given time interval. In this case, time is divided into an evenly spaced mesh of size Δt, and the distribution of the number of events in each interval is assumed to follow a Poisson distribution, again conditional on the intensity function.

From a statistical perspective, we would like to exploit all the information at our disposal. It is then immediate to notice that the duration process includes all the information contained in the counting process, and therefore it is the preferred object of analysis. However, if we allow for the possibility of *simultaneous* events, i.e. events happening at the exact same time, the same definition of the duration process breaks down, since we are not able to assign a meaningful ordering to the point process itself anymore. In such a case, the only viable possibility is the study of the counting process. Moreover, the appearance of data-sets of unprecedented size, especially in recent years, has greatly increased the computational burdens induced by the analysis of the point process in its entirety, through its duration representation. Therefore, the study of the counting process, rather then the study of the duration process, can sometimes be a more convenient alternative in order to recover the fundamental object of interest, the intensity function $\lambda(t)$.

In this dissertation we propose new, dynamic, non-linear, non-Gaussian state space models to study both duration processes and time series of counts over a mesh of fixed length in time. Both models are able to capture the well-known and salient features of the empirical

distribution of their quantities of interest, namely, over-dispersion, long-memory and transaction clustering, and are based on the previous work of Chen et al. (2013) and Calvet and Fisher (2008).

We employ the proposed inter-trade duration model to analyze the structural change in the properties of the transaction process of 25 stocks randomly selected from the S&P100 index over a long span of time (1993-2003), and in a coherent framework. In particular, we reconcile the approach of the existing literature in inter-trade durations modeling with the empirical observation of a clustering of inter-trade durations at a value of zero (i.e. *simultaneous trading*), especially in the recent samples. We do so by constructing a model of interval censoring that explains the observed simultaneity as a result of the granular measurement of time in terms of a minimum unit of measure. This interpretation is consistent with the analyzed data for most of the considered assets and over most of the considered years.

We employ the proposed model for time series of counts per unit of time to study the number of transactions over an evenly spaced mesh of different lengths of the same 25 stocks during February 2010. We compare the proposed model to the benchmark model introduced by Rydberg and Shephard (2003) in terms of in-sample goodness-of-fit and out-of-sample forecasting ability. The proposed model dominates the benchmark in terms of in-sample goodness-of-fit, and it is at least as good as the benchmark in terms of out-of-sample forecasting accuracy.

As already outlined, the two different approaches have both advantages and imperfections. The proposed models improve over the state-of-the-art modeling perspective of financial inter-trade durations and number of transactions per unit time. However, our analysis also proves that more work is needed to further enhance our ability of fitting and forecasting the transaction times process, and hence better understanding the main characteristics of the dynamic behaviour of financial markets at high frequencies.

Overview of the Subsequent Chapters

In Chapter 2 we introduce the proposed model for the analysis, a censored and zero-augmented version of the Markov Switching Multifractal Duration model introduced by Chen et al. (2013), and illustrate its properties. The proposed model, unlike the original version, is able to adjust to the most recent characteristics of the empirical distributions of inter-trade durations. We describe the data employed in the analysis, trying to highlight its most important features from a dynamic and static perspective. We then discuss the estimation procedure, comment on the estimated parameters and investigate the structural change of our estimates over five years intervals from 1993 to 2013.

In Chapter 3 we introduce the Markov Switching Multifractal Conditional Poisson Model and discuss its properties. We then describe the data employed in the analysis trying to compare the empirical regularities in the number of transactions per unit time to the properties of the proposed model. We discuss the estimation procedure, and comment on the estimated parameters. We finally compare the proposed model to the benchmark model introduced by Rydberg and Shephard (2003) in terms of in-sample goodness-of-fit and out-of-sample forecasting accuracy and assess its superior performance.

CHAPTER 2 : Financial Trading Over The Years, A Multifractal Intensity Perspective

High-frequency transaction-level data on financial markets has long been exploited in the literature to study many objects of economic interest. Major attention has been directed towards measures of risk (e.g. price volatility) and liquidity (e.g. bid-ask spreads and transaction volumes) in the hope of shedding light upon the micro-structure properties of financial markets and of achieving a better quantitative understanding of their dynamics. While such high-frequency sampling enhances the value of statistical inference through the collection of vast datasets of unprecedented size–what is often referred to as "Big Data"–it also typically imposes upon the econometrician the burden of studying irregularly spaced observations. The quantities of interest, such transaction prices and volumes or quote revisions, are observed at (almost) the exact time they occur, which raises the question of whether the size of the temporal displacement between events conditions their probabilistic structure. Empirical and theoretical arguments, starting from Clark (1973) and Easley and O'Hara (1992), support the view that the amount of time elapsing between consecutive transactions constitutes information about the economically relevant variables of interest (such as prices or volumes).

On one hand plain time deformation arguments, starting from the simple observation that a price can only change when a trade actually occurs, directly imply that the number of transactions executed during a given time interval drives (time-varying) calendar time volatility of stock prices over the same interval. Note that, even if trade-by-trade price changes are assumed to have constant variance, the resulting calendar time price changes can display time-varying volatility. To see this, consider a simple random walk model for trade-by-trade (log) prices, $\log P_i = \log P_{i-1} + \epsilon_i$, where $\epsilon_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. The log return between t and t' will then be given by $\log P(t') - \log P(t) = \sum_{\{i:t_i \in (t,t']\}} [\log P_i - \log P_{i-1}]$, where t_i is the calendar transaction time. Thus $\log P(t') - \log P(t) \sim \mathcal{N}(0, N_{t,t'})$, where $N_{t,t'}$

represents the number of transactions between time t and t'. In this very simple case, then, stochastic volatility in (log) returns is *completely* determined by the number of transactions in the considered time interval. However, if stochastic volatility in stock prices depend–at least in part–on the underlying number of transactions realized in the considered time span, ultimately, the dynamic properties of inter-trade durations will likely correspond to the dynamic properties of price volatility, as discussed by Deo et al. (2009).

On the other hand financial market micro-structure models with strategic interactions–such as those in O'Hara (1995)–suggest that the pace of the transaction process, and thereby the waiting times between trades, informs the market observer about the presence of information-driven, speculative traders, who are trying to exploit their information advantage to maximize profits. In such a scenario, shorter inter-trade durations would induce the specialist to revise her quotes, widening bid-ask spreads to hedge against the losses deriving from information-driven trades. Moreover, signal extraction from information-driven trades will be faster in calendar time if the inter-trade durations are shorter, everything else equal, implying a quicker adjustment of transaction prices and quotes to the new equilibrium levels. Finally, shorter durations will affect market depth, as information-driven trades are associated with higher order volumes and the direction of trading is unambiguously determined depending on whether the signal about the equity value is good or bad. However, the precise effect would depend on the specialist's constraints (for example, whether or not she has inventory constraints).

Many studies, e.g. Engle and Lunde (2003) and Easley et al. (2008), have successfully documented the empirical relevance of inter-trade durations for fitting and forecasting quote and transaction dynamics in attempts to validate theretofore theoretical results. Inter-trade durations feature prominently as an essential component of the dynamic behaviour of financial markets' liquidity in its four dimensions, as introduced by Harris et al. (1991)–namely the bid-ask spread, market depth, immediacy and speed of price adjustment–and price volatility. Therefore, understanding the dynamic properties of inter-trade durations is

fundamental to characterizing the dynamic structure of price volatility and market liquidity.

In this work I study the evolution over the last two decades of the dynamic probabilistic structure of inter-trade durations. I consider the transaction activity of twenty-five equities traded on U.S. stock markets during the month of February for the years 1993, 1998, 2003, 2008 and 2013. All data is obtained from the Trade And Quote (TAQ hereafter) database, and transaction times are measured down to the second until 2003 and on a millisecond scale thereafter. Inter-trade durations are shown to display the well-known characteristics of over-dispersion (i.e a coefficient of variation bigger than one) and long memory (i.e. hyperbolic decay of the autocorrelation function as opposed to the exponential decay typical of short memory processes). As expected, in such a long period of time, even simple descriptive statistics suggest that the probabilistic structure of inter-trade durations has drastically changed. In particular, the mean rate of the processes, as measured by average duration, has collapsed from a bundle average of around 35 seconds in 1993 to an average of 0.67 seconds in 2013, and over-dispersion increased drastically in the 2008 and 2013 samples. Furthermore, a substantial fraction of transactions (more than 25% of the total observations in the last three samples considered) occur at exactly the same measured time. From a probabilistic point of view, this feature implies that measured durations should display a point mass at zero. Moreover, I document that the occurrence of zero durations is strongly autocorrelated. In order to incorporate zero-duration observations to the econometric analysis of inter-trade durations, I construct a model of censoring that takes into account the discretization mechanism imposed by the minimum recording time scale. The censoring model, in particular, allows for simultaneous transactions clustering by dictating that any inter-trade duration smaller than the measurement unit threshold (i.e. one second or one millisecond) is recorded as zero. The model thus justifies the clustering of inter-trade durations at zero in a way that is neutral with respect to the underlying properties of the point process. The zero observations result from the constraints of the recording mechanism and their autocorrelation structure reflects the dynamic properties of the original duration series. Lastly, the censoring model has the advantage that it does not

require the estimation of any additional parameter *vis à vis* the original model chosen to represent the duration process, while simultaneously allowing for the use of much bigger datasets which include the zero-duration observations.

Several inter-trade duration models have been proposed in the literature, starting from the seminal contribution of Engle and Russell (1998), introducing an observation-driven short memory model to analyze inter-trade durations (the Autoregressive Conditional Duration model, ACD hereafter). Jasiak (1999) tried to enhance long memory in the same framework by constructing a fractionally integrated version of the ACD model. Bauwens and Veredas (2004) propose a parameter-driven short memory model to analyze inter-trade durations (the Stochastic Conditional Duration model, SCD hereafter). Similarly, Deo et al. (2010) allow for long memory in the form of fractional integration in the same context. Finally, Chen et al. (2013), develop a parameter-driven, long memory model featuring multifractal scaling and multi-frequency components, the Markov Switching Multifractal Duration model (MSMD hereafter). In particular, as also shown by Zikes et al. (2012), the MSMD model accounts for most of the relevant properties displayed by the data and remains unbeaten in terms of forecasting ability and quality of fit, and will therefore be taken as the starting point for the subsequent analysis.

Nevertheless, this and all of the other models proposed to analyze inter-trade durations crucially rely on continuous distributions, and are hence unable to account for the empirical observation that observed durations cluster at zero. The choice of continuous and, in particular, exponential distributions conditional on a stochastic arrival rate, stems from the natural representation of financial inter-trade durations as realizations of a Poisson Process with time-varying intensity. However, in reality, all available data is necessarily measured in a discrete fashion (on a second or millisecond scale in this case). If a process is amply larger than the scale of its minimum unit of measure with very large probability, a continuous representation is likely to suffice. If the process instead lies in the scale of the minimum unit of measure with a relevant frequency, the model employed in the analysis

should directly account for the discretization mechanism.

Throughout, instead, the most common approach to operationalize these models has rather been to "thin" the original series by removing all zero observations. In older samples (e.g. 1993), where zero durations are more scarce, such an approach could result in negligible approximation errors. In more recent years, though, where the fraction of zero observations exceeds 25% even with a concomitant refinement of the measurement scale, thinning should not be expected to be neutral. In such an environment, estimation on the reduced sample will deliver inconsistent estimates. Intuitively, in fact, the excluded observations represent the smaller realizations of the process and the residual observations' ordering is altered.

The paper is organized as follows. In Section 1 I describe the data employed in the analysis and I compare the full and the thinned samples' main characteristics. Section 2 sketches the structure of the MSMD model of Chen et al. (2013), introduces the discretization mechanism imposed by the minimum recording scale and describes the Censoring MSMD model and its properties. Section 3, 4 and 5 discuss the estimation strategy and presents estimation results for the censoring model applied to the full sample as compared to the results produced by the MSMD model fitted on the thinned sample. In-sample goodness of fit and (pseudo-) out-of-sample forecasting accuracy comparison are performed. Finally, Section 6 concludes by relating the findings for inter-trade durations to market liquidity and price volatility.

2.1. Description of Data, A Long Retrospective

In this paper I will be focusing on high-frequency financial data of equity transactions. In particular, I will employ time series of inter-trade durations for 25 equities traded on U.S. stock exchanges throughout February in the years 1993, 1998, 2003, 2008 and 2013. I consider trading days starting at 10:00 AM and ending at 4:00 PM to avoid dealing with the micro-structure effects brought about by exchange opening procedures. The raw data sources are the Second TAQ Database for the years 1993, 1998 and 2003, and the Millisecond TAQ Database for the years 2008 and 2013. These databases provide information about trades and quotes of financial assets traded on all U.S. exchanges. Each trade (or quote) is time-stamped with second or millisecond level precision, depending on the considered year. Table 4 displays the list of selected equities, which directly follows from Chen et al. (2013). The original bundle was picked to be representative of the S&P100 market index in 1993. However, three equities among those considered have been delisted since 2012[1]. Regardless,

Table 1: Selected U.S. Equities, Ticker Symbols and Company Names

Symbol	Company Name	Symbol	Company Name
AA	ALCOA	INTC	Intel
ABT	Abbott Laboratories	JNJ	Johnson & Johnson
AXP	American Express	KO	Coca-Cola
BA	Boeing	MCD	McDonald's
BAC	Bank of America	MRK	Merck
C	Citigroup	MSFT	Microsoft
CSCO	Cisco Systems	QCOM	Qualcomm
DELL	Dell	T	AT&T
DOW	Dow Chemical	TXN	Texas Instruments
F	Ford Motor	WFC	Wells Fargo
GE	General Electric	WMT	Wal-Mart
HD	Home Depot	XRX	Xerox
IBM	IBM		

I have chosen to retain these equities for comparison consistency rather than to re-sample a new set of symbols from the time-constant S&P100 universe. On similar grounds, the

[1]Such equities are ALCOA, Dell and Xerox.

choice of the month, trading hours and sample length follow directly from the analogous values proposed and justified by Chen et al. (2013). I choose to consider samples every five years starting from 1993. This choice is arbitrary and has been made mainly for the sake of parsimony, given the growing sample sizes throughout the period.

As notted by Hautsch (2012) the raw data produced by the TAQ database must be cleaned to remove erroneously reported trades due to wrong and/or delayed recording of the trade itself. To this end, I follow the procedures reported in Barndorff-Nielsen et al. (2011) and Hautsch (2012) and check for: *i)* trades directly indicated as mis-recorded (CORR≠00) and *ii)* trades associated with zero values for the transaction price or the exchanged quantities.

Table 2: Descriptive Statistics, Bundle Averages, Full Sample, Seconds

Year	Mean	Med.	Max.	St. Dev.	Kurt.	CV	Trades	Zeros %
1993	42.1	20.0	850.9	61.1	25.68	1.41	19,343	5%
1998	18.7	10.1	341.8	24.4	23.83	1.35	57,232	11%
2003	5.21	2.96	113.1	6.57	25.85	1.47	225,910	23%
2008	0.86	0.14	59.37	2.10	54.00	2.60	718,310	48%
2013	0.90	0.01	75.98	2.56	62.03	2.94	609,780	28%

In this section I describe the preeminent properties of the inter-trade duration data year-by-year and across firms. Since I estimate the original MSMD model on a selected sub-sample, I aim to highlight the main differences between the thinned and the full sample, the former excluding and the latter including the zero observations.

Table 5 reports year-by-year bundle averages for simple descriptive statistics for the inter-trade duration series on the full sample while Table 3 reproduces the same quantities for the thinned sample. As expected, there are bigger proportional differences, in the considered quantities, for those years in which the presence of simultaneous trade is more relevant, that is 2003, 2008 and 2013.

First consider Table 5. The most striking feature of the full sample is the fact that the percentage of simultaneous trades, i.e. of inter-trade durations equal to zero, is quite large

Table 3: Descriptive Statistics, Bundle Averages, Thinned Sample, Seconds

Year	Mean	Med.	St. Dev.	Kurt.	CV	Trades
1993	43.4	21.3	61.6	25.08	1.35	17,788
1998	19.7	11.0	24.6	22.23	1.21	42,825
2003	5.96	3.56	6.75	21.10	1.09	100,450
2008	1.57	0.55	2.64	31.29	1.71	339,270
2013	1.22	0.05	2.93	45.23	2.43	429,140

in the 2003, 2008 and 2013 samples. As I already mentioned, all current state of the art models employed in the analysis of inter-trade durations fail to account for such clustering and the econometrician has to discard all the zero observations from the dataset prior to analysis. As an example, in 2008, the thinning procedure excludes almost 50% of the available observations. In this context, can such a substantial portion of the data be discarded safely, simply because it can't be explained by the available models? We should then include it in the analysis and explain its properties in a coherent and practical way.

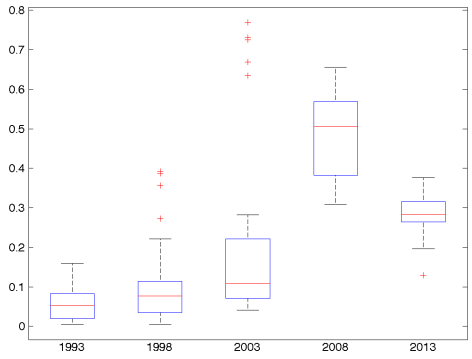
The average number of transactions sizably increased, peaking in 2008 and slightly decreasing in 2013, mirroring the behavior of the average mean duration which considerably dropped while staying constant over the last two periods considered. The average median duration, by contrast, continued to drop considerably also in 2013. Inter-trade durations' average standard deviation decreased, but proportionally less then the average mean, inducing very high coefficients of variation for the 2008 and 2013 samples. Also, average kurtosis exploded in the later samples.

Looking at the thinned sample in Table 3 we can see that the main qualitative features of the previous characterization remain unchanged. The most striking differences are that the number of transactions increases from 2008 to 2013 when ignoring zero-length durations, that the boost in kurtosis is relatively less pronounced, and that over-dispersion levels are smaller.

To give a more precise sense of the location shifts and the dispersion of some of the reported

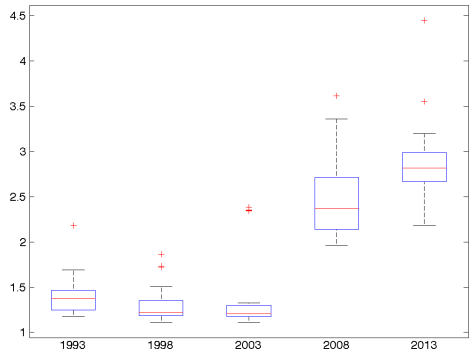
Figure 1: Descriptive Statistics Over The Years, Full Sample

(a) Proportion of Zero Durations



Year

(b) Coefficient of Variation



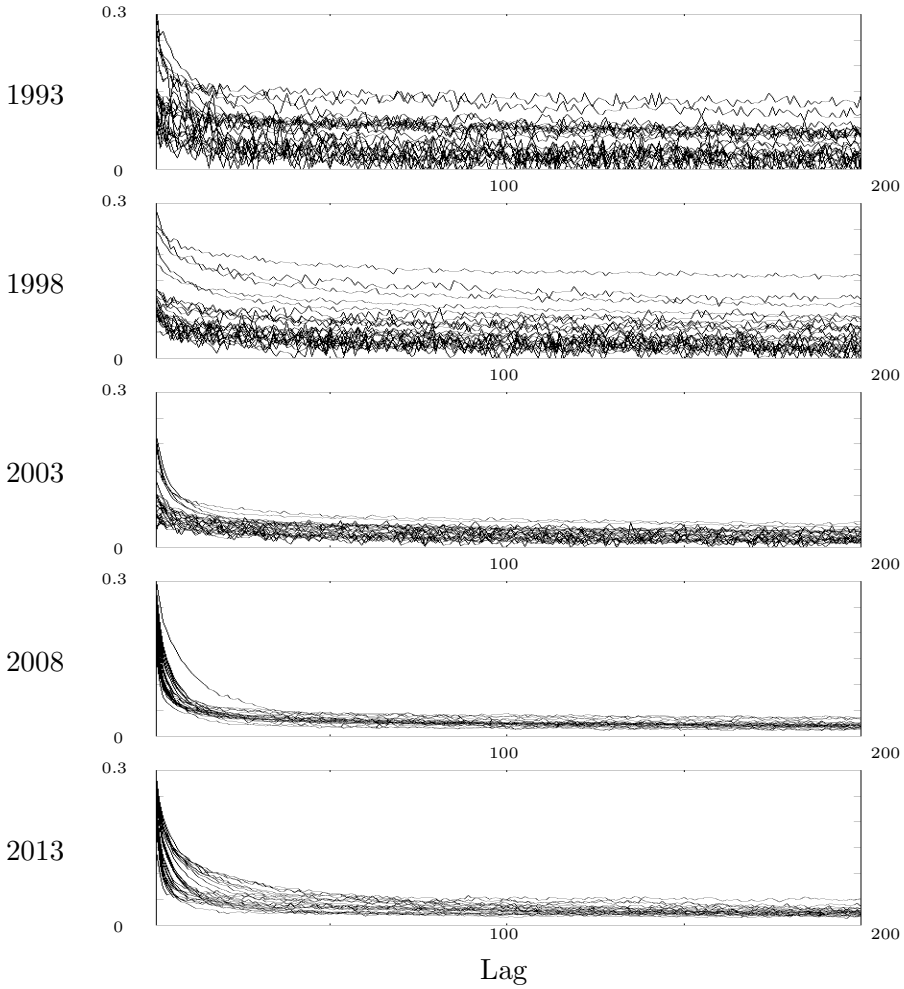
Year

descriptive statistics over the bundle of equities, Figure 27 depicts box and whisker plots over the considered years. Two statistics are considered–the coefficient of variation (right panel) and the proportion of simultaneous trades (left panel). Red crosses represent outliers. For the coefficients of variation, bundle distribution dispersion decreased until 2003 and then increased in 2008 and 2013. For the proportion of simultaneous trades it is interesting to note the presence of several outliers in the 2003 sample that display very high values.

As mentioned above another major characteristic of inter-trade durations is long-memory. However, such abrupt change in distributional characteristics should not be surprising. Financial markets and the trading process have recently benefited from many structural innovations, ranging from technological improvements in the speed of information transmission and computing power, to the development of new trading strategies enabled by the advent of high-frequency trading.

Figure 28 displays sample autocorrelation functions of the equity bundles for each year considered. The slow and hyperbolic decay rate proper of long memory processes is easily detectable from Figure 28; Figure 30 (left panel), moreover shows the corresponding Geweke-Porter-Hudak estimator of fractional integration which clearly indicates the presence of long-memory in the considered series. It is interesting to note how, over the years, the equity

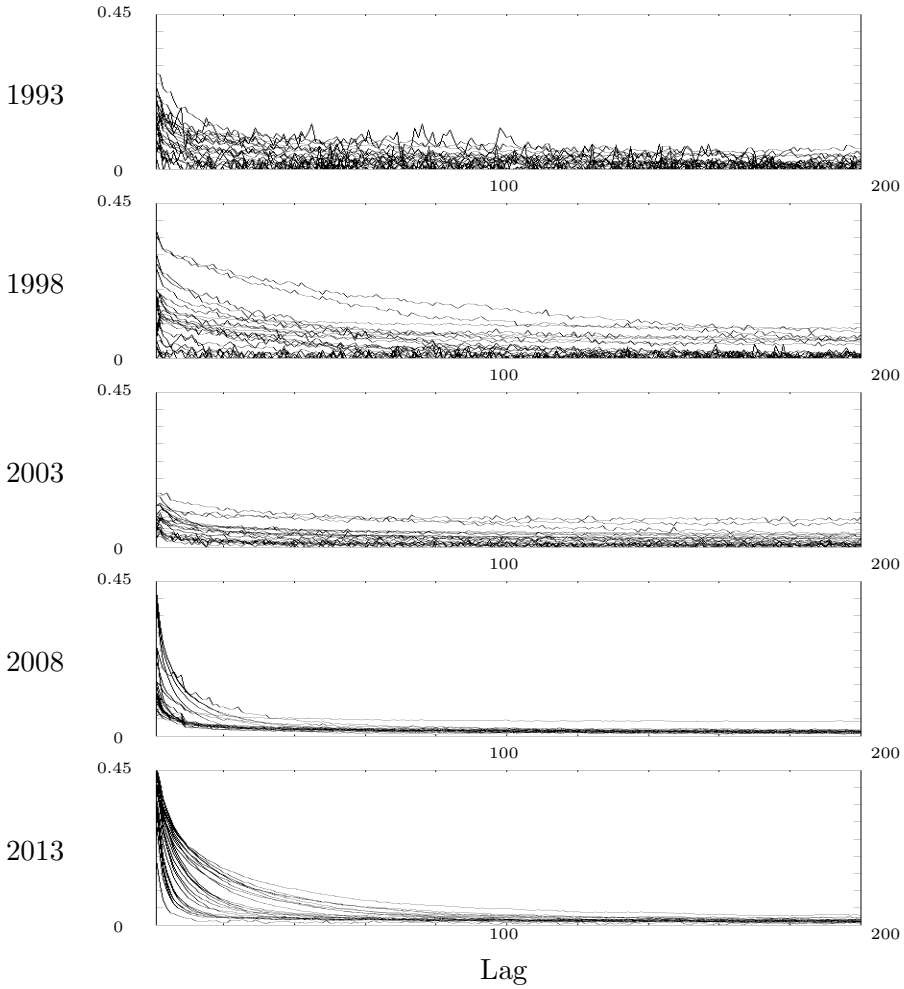
Figure 2: Inter-Trade Durations, Sample ACFs, Bundle, Full Sample



dynamic behavior has become more and more similar and that the level of dependence has tended to drop, especially in 2008.

In addition to the well known properties of unconditional over-dispersion and long memory displayed by inter-trade durations, I indicated above that clustering at zero is a very important feature of the data, especially in later samples. In Figure 29, I offer some more evidence of the dynamic properties of such clustering. In order to do concisely so, I construct a simultaneous trade indicator defined by $\mathbb{1}\{d_i = 0\}$ where d_i is the $i-th$ inter-trade duration. Figure 29 shows sample autocorrelation functions of the simultaneous trade indicator for the equity bundles over the considered years. AWith the exception of 1993, the simultaneous

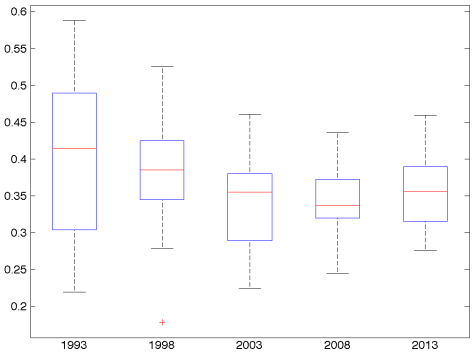
Figure 3: Simultaneous Trade Indicator, Sample ACFs, Bundle



trade indicator seems to be very strongly correlated. Furthermore, it also seems to display the slow decay typical of long memory processes. Figure 30 (right panel) depicts a box and whisker plot of the GPH estimator of fractional integration for the simultaneous trade indicator. For most of the stocks there seems to be long memory in simultaneous trades starting from 1998. If the cause of inter-trade duration clustering at zero is somehow related to the measurement discretization mechanism imposed by the minimum unit of measure in a given dataset, then the dynamic features of simultaneous trades would be directly determined by the underlying inter-trade durations dynamic properties. In particular, in such a case, given that inter-trade durations display long-memory dynamics, one would expect

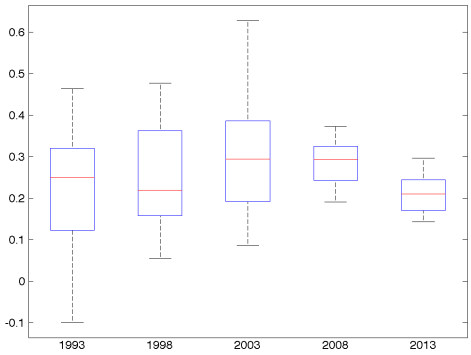
Figure 4: Geweke-Porter-Hudak Estimator for Fractional Integration, Full Sample

(a) Inter-Trade Durations



Years

(b) Simultaneous Trade Indicator



Years

the simultaneous trades to display them too.

To conclude, the data I use is characterized by heavy over-dispersion and fat tails, long memory and clustering at zero. All these features have become more and more pronounced over time. As such, in order to confidently statistically analyze the dynamic behavior of inter-trade durations, a candidate model should be able to take into account all the afore-mentioned characteristics. Now I will proceed to describe Chen et al. (2013) MSMD Model and sketch its key properties. I then build a censoring mechanism into the MSMD model that is constructed based on the measurement discretization scheme induced by the minimum unit of measure. The censoring mechanism is able to capture the clustering of durations at zero, including its underling dynamic structure.

2.2. Introducing A Censored And Zero-Augmented Markov Switching Multifractal Duration Model

2.2.1. The MSMD Model

Let $\{t_i\}_{i\in\mathbb{N}}$ be an infinite sequence of (strictly positive) trading times. Consider then the sequence of non-negative, real valued random variables $\{d_i\}_{i\in\mathbb{N}}$ with $d_i = t_i - t_{i-1}$, which we refer to as inter-trade durations. Let each d_i be a conditionally independent and exponentially distributed random variable with rate λ_i:

$$d_i \,|\lambda_i \overset{\text{iid}}{\sim} \text{Exp}\,(\lambda_i)$$

or, alternatively:

$$d_i = \frac{\epsilon_i}{\lambda_i}, \quad \epsilon_i \overset{\text{iid}}{\sim} \text{Exp}\,(1) \tag{2.1}$$

Where the second definition is justified by the mixture of exponential representation for a simple point process. The exponential rate, λ_i, is itself a stationary time series governed by a Markov-switching multi-fractal process, multiplicatively composed by independent Markov-switching chains $M_{k,i}$, for $k \in \left\{1, 2, \ldots, \bar{k}\right\}$, in the spirit of Calvet and Fisher (2004):

$$\lambda_i = \lambda \left(\prod_{k=1}^{\bar{k}} M_{k,i} \right), \qquad \lambda > 0 \tag{2.2}$$

Each Markov-switching chain $\{M_{k,i}\}_{i\in\mathbb{N}}$ is characterized by the following structure:

$$M_{k,i} = \begin{cases} M & \text{w.p} \quad \gamma_k \\ M_{k,i-1} & \text{w.p} \quad 1-\gamma_k \end{cases} \tag{2.3}$$

where M is a non-negative and independent random variable with cumulative distribution

function $F(M)$ such that $\mathbb{E}(M) = 1$. Following Chen et al. (2013), we will take M to follow a discrete distribution on support $\mathcal{M} = \{m_0, 2 - m_0\}$ with $m_0 \in (0, 2]$ and $f(m) = 1/2$ for both $m \in \mathcal{M}$. Therefore, each $M_{k,t}$ is an independent, two state Markov-switching process with transition matrix given by:

$$\mathbb{P}\mathrm{r}\left(M_{k,i} \,|M_{k,i-1}\right) = \mathcal{P}_k = \begin{bmatrix} 1 - \gamma_k/2 & \gamma_k/2 \\ \gamma_k/2 & 1 - \gamma_k/2 \end{bmatrix} \tag{2.4}$$

and states $\{s_1, s_2\} = \{m_0, 2 - m_0\}$. Note that, the largest eigenvalue of the transition matrix $\mathcal{P}_k$ is equal to one. The implied stationary distribution is independent of k since $\mathbb{P}\mathrm{r}(M = m) = 1/2$ for all $m \in \mathcal{M}$. The second eigenvalue, on the other hand, is instead given by $1 - \gamma_k$ and governs the persistence of the chain itself. In the remaining of the work, for the sake of parsimony, we impose the following restriction on the sequence of γ_k, $k \in \left\{1, 2, \ldots, \bar{k}\right\}$, that is:

$$\gamma_k = 1 - (1 - \gamma^*)^{b^{k-\bar{k}}} \tag{2.5}$$

with $b \in (1, \infty)$ and $\gamma^* \in (0, 1)$. In this specification, as $k \uparrow \bar{k}$ we see that $1 - \gamma_k \downarrow 1 - \gamma^*$. Moreover, as $\bar{k} \uparrow \infty$, $1 - \gamma_k$ ranges in the interval $[1, 1 - \gamma^*]$ producing an infinite variety of frequencies at which the different chains evolve over time. Also, note that γ^* bounds the highest (least persistent) frequency at which the multifractal is operational.

Finally, note that the MSMD Model is characterized by a $\bar{k}$-dimensional hidden-state vector, $M_t = \left(M_{1,t}, \ldots, M_{\bar{k},t}\right)$, in which every coordinate can only take on two values. This amounts to a total of $2^{\bar{k}}$ possible states for the intensity process. Furthermore, we can collect the remaining parameters in the vector $\theta = (\lambda, m_0, b, \gamma^*)$, whose dimension is independent of $\bar{k}$. This nice separability allows $\bar{k}$ to be viewed as model index rather than as a standard parameter.

Interpretation As a Continuous Time Process

The framework proposed by Chen et al. (2013) and described above–in particular, the mixture-of-exponential representation–can be justified by its natural interpretation as a point process. Consider the sequence of arrival times for the event of interest $\{t_i\}_{i\in\mathbb{N}}$, where t_0 is normalized to 0, with associated intensity $\lambda(t)$. Let the associated counting process be given by:

$$N(0,t] = N(t) = \sum_{i=1}^{\infty} \mathbb{1}\{t_i \leq t\}$$

and further assume that:

a) The intensity process is such that $\lambda(t) = \lambda_i$ for any $t \in [t_{i-1}, t_i)$, where λ_i is defined by equations (2), (3), (4) and (5);

b) The distribution of $N(t)$ conditional on the relevant filtration $\mathcal{F}_t$ has the following properties:

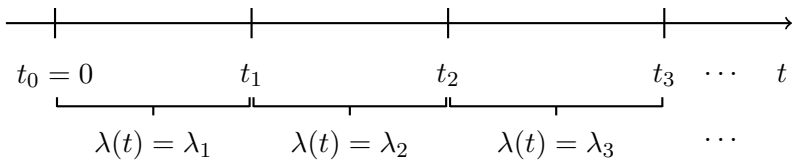
   i. $\Pr\left(N(t+dt) - N(t) = 1 \,|\mathcal{F}_t\right) = \lambda(t)dt + o(dt)$;

   ii. $\Pr\left(N(t+dt) - N(t) \geq 2 \,|\mathcal{F}_t\right) = o(dt)$;

   iii. $N(t') - N(t)$ is independent of $N(s') - N(s)$ for any disjoint intervals $(t, t']$ and $(s, s']$;

It is then straightforward to show that $d_i \,|\lambda_i$ are independent across i and distributed as exponential r.v. with rate parameter λ_i, as in the model presented above. Moreover, by the standard time deformation argument, we can find a (random) change of time that transforms the inhomogeneous process into a unit-rate Poisson process, thereby justifying the mixture-of-exponential representation highlighted above. Figure 32 gives a graphical representation

Figure 5: Example of the Considered Process



for the described process. The intensity is constant over each of the intervals $[t_{i-1}, t_i)$ at some

value λ_i and, therefore the inter-event duration is conditionally exponentially distributed with constant rate λ_i. Finally note that, in this setting, the probability of observing $t_i = t_{i+1}$ is equal to zero, i.e. the process is *simple*. This is inherent in the continuous nature of the exponential distribution, which has no point mass at zero, and hence assigns zero probability to simultaneous occurrences of different transactions.

Properties of the Base Model

In the following subsection I will briefly recall the primary properties of the MSMD model, with emphasis on the parallels to the main characteristics displayed by the data.

Stationarity, Ergodicity and Finitness of Moments

Provided that $\gamma_k > 0$, the transition matrices for the Markov chains imply that the processes $\{M_{k,i}\}_{k\in\{1,\dots,\bar{k}\}}$ are strictly stationary and ergodic. Moreover, their independence implies that the vector process $\xi_i = \left(M_{1,i}, \dots, M_{\bar{k},i}\right)$ is strictly stationary and ergodic as well. Seeing as how λ_i is a measurable function of ξ_i, it is straightforward to demonstrate that λ_i as well must be a strictly stationary and ergodic process. Similarly, d_i is a measurable function of λ_i, and hence the duration process, too, is ergodic and strictly stationary. Given the structure of the MSMD model it is easy to show that the r^{th} moment of d_i is given by $\mathbb{E}\left(\lambda_i^{-r}\right) r!$, and hence each moment exists if and only if the corresponding moment for the directing measure λ_i exists. Note that, for every $i \in \mathbb{N}$, λ_i can only take on a finite number of possible values which is bounded above, say, by $\bar{\lambda}$ and below by 0. Therefore all moments for λ_i exist, and hence all the moments for d_i exist as well.

Overdispersion

In Section 1 I showed that overdispersion is a salient characteristic of the considered data. This is built into the MSMD model; note that the process d_i in the model has mean $\mathbb{E}(d_i) = \mathbb{E}(1/\lambda_i)$ and variance $\mathbb{V}\text{ar}(d_i) = 2\mathbb{E}(1/\lambda_i^2) - \mathbb{E}(1/\lambda_i)^2) = \mathbb{E}(1/\lambda_i)^2 + 2\mathbb{V}\text{ar}(1/\lambda_i)$.

Therefore:

$$\frac{\sqrt{\mathbb{V}\text{ar}(d_i)}}{\mathbb{E}(d_i)} = \frac{\sqrt{\mathbb{E}(1/\lambda_i)^2 + 2\mathbb{V}\text{ar}(1/\lambda_i)}}{\mathbb{E}(1/\lambda_i)} \geq 1$$

where the inequality is strict if and only if $\mathbb{V}\text{ar}(1/\lambda_i) > 0$, i.e. $\mathbb{V}\text{ar}(M) > 0$. When $\mathbb{V}\text{ar}(\lambda_i) = 0$, in fact, the model boils down to a simple Poisson process with constant intensity, hence generating an equi-dispersed unconditional distribution.

Long Memory

Chen et al. (2013) show that the MSMD model produces long memory dynamics in the sense that:

$$\sup_{n \in \mathcal{I}_{\bar{k}}} \left| \frac{\log \rho(n)}{\log n^{-\delta}} - 1 \right| \longrightarrow 0 \qquad \text{as} \qquad \bar{k} \longrightarrow +\infty$$

where $\rho(n) = \mathbb{C}\text{orr}(d_i, d_{i+n})$, $\delta = \log_b \mathbb{E}\left(\widetilde{M}^2\right)$, $\widetilde{M} = M^{-1}/\mathbb{E}(M^{-1})$ and:

$$\mathcal{I}_{\bar{k}} = \left\{ k \in \mathbb{N} : \alpha_1 \log_b \left(b^{\bar{k}}\right) \leq \log_b k \leq \alpha_2 \log_b \left(b^{\bar{k}}\right) \right\}$$

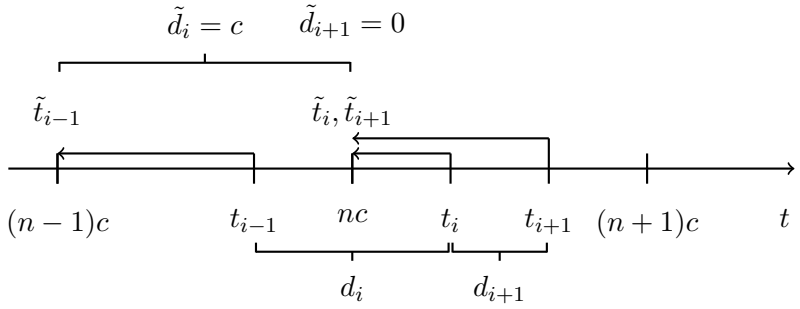
for any α_1, $\alpha_2 \in (0, 1)$ and $\alpha_2 > \alpha_1$. The hyperbolic decay of the autocorrelation function implied by the MSMD model is thus able to account for the long memory dynamics highlighted as one of the principal features of the data in Section 1.

2.2.2. Accounting For Simultaneous Events

As mentioned in Section 1, the analysis dataset is characterized by the relevant presence of simultaneously occurring events, i.e. observations for which $t_i = t_{i-1}$, and therefore $d_i = 0$. Because data is recorded in discrete time, the rationale behind such an observation can be twofold. On the one hand observing $d_i = 0$ could just reflect the discretization imposed by the measurement device on the underlying true continuous-time process, as shown in Figure 6. Suppose we are only able to measure time in multiples of an arbitrary fixed number c.

Then if an event occurs at some time t_i in between $t = nc$ and $t = (n+1)c$, it will be recorded

Figure 6: Censoring Scheme, A Graphical Example



as occurring at time $\tilde{t}_i = nc$. Therefore, if it is the case that $t_i, t_{i+1} \in [nc, (n+1)c)$, the recorded inter-event duration would be given by $\tilde{d}_{i+1} = \tilde{t}_{i+1} - \tilde{t}_i = 0$ even though the true duration d_i is strictly positive. It is possible, then, for the observed simultaneity present in the considered data to be a byproduct of censoring technologically imposed by the measurement mechanism. On the other hand the observed simultaneity could reflect true simultaneity in the true unobserved process, in which case even perfect observation of exact realizations of event times would not eliminate zero-durations (though blunt discretization could exacerbate the recorded presence of zero-duration events).

As a first step towards accounting for simultaneous events in a MSMD-like framework I will consider the censoring mechanism imposed by the discrete recording activity. Note that one can write the observed process as:

$$\tilde{t}_i = \sum_{j=0}^{\infty} jc \mathbb{1}\{t_i \in \mathcal{I}_j\}$$

with $\mathcal{I}_j = [jc, (j+1)c)$. Therefore, the implied observed duration process is given by:

$$\tilde{d}_i = \sum_{j=0}^{\infty} jc \mathbb{1}\{t_i \in \mathcal{I}_j\} - \sum_{j=0}^{\infty} jc \mathbb{1}\{t_{i-1} \in \mathcal{I}_j\} \tag{2.6}$$

Note that $\tilde{d}_i = 0$ if and only if, for some $j^* \in \mathbb{N}$, $t_i, t_{i-1} \in \mathcal{I}_{j^*}$. Hence the probability of observing $\tilde{d}_i = 0$ can be written as:

$$\Pr(\tilde{d}_i = 0) = \Pr(N_{j^*+1} - N_{j^*} \geq i - N_{j^*})$$

where $j^* = \tilde{t}_{i-1}/c$ and $N_j = N(cj)$. As one can immediately notice, the probability for the point mass at zero induced by the censored recording mechanism depends on the history of the intensity $\lambda(t)$ on the whole interval $\mathcal{I}_{j^*}$, since it reflects the probability distribution of the counting process on the same interval. This is true because the placement of t_{i-1} over the interval $\mathcal{I}_{j^*}$ directly affects the probability of t_i being in the same interval by bounding the maximum value that the associated duration d_i can take.

By exploiting the distributional characteristics of the counting process we can then express the probability of observing a censored duration as:

$$\Pr(\tilde{d}_i = 0) = 1 - \exp\left(-\Lambda_{j^*}^{j^*+1}\right) \sum_{r=0}^{i-1-N_{j^*}} \frac{(\Lambda_{j^*}^{j^*+1})^r}{r!} \tag{2.7}$$

where $\Lambda_j^{j'}$ is the compensator over the interval $[cj, cj')$ and is given by:

$$\Lambda_j^{j'} = \int_{cj}^{cj'} \lambda(s)\, \mathrm{ds}$$

Similarly, we can compute the full probability distribution for $\tilde{d}_i$, i.e. $\Pr(\tilde{d}_i = nc)$ for any $n \in \mathbb{N}$. To observe a value of nc for the censored duration, in fact, it is necessary and sufficient that, for some $j^* \in \mathbb{N}$, $t_i \in \mathcal{I}_{j^*+n}$, and $t_{i-1} \in \mathcal{I}_{j^*}$. Hence the probability of

observing $\tilde{d}_i = nc$ can be written as:

$$\begin{aligned}
\mathbb{P}\mathrm{r}(\tilde{d}_i = nc) =& \mathbb{P}\mathrm{r}\left(\{N_{j^*+1} - N_{j^*} = i - 1 - N_{j^*}\}, \{N_{j^*+n} - N_{j^*+1} = 0\},\right. \\
& \left.\{N_{j^*+n+1} - N_{j^*+n} \geq 1\}\right) \\
=& \mathbb{P}\mathrm{r}\left(N_{j^*+1} - N_{j^*} = i - 1 - N_{j^*}\right) \times \mathbb{P}\mathrm{r}\left(N_{j^*+n} - N_{j^*+1} = 0\right) \times \\
& \mathbb{P}\mathrm{r}\left(N_{j^*+n+1} - N_{j^*+n} \geq 1\right) \\
=& \frac{(\Lambda_{j^*}^{j^*+1})^{i-1-N_{j^*}}}{[i-1-N_{j^*}]!} \exp\left(-\Lambda_{j^*}^{j^*+1}\right) \exp\left(-\Lambda_{j^*+1}^{j^*+n}\right)[1 - \exp\left(-\Lambda_{j^*+n}^{j^*+n+1}\right)]
\end{aligned}$$

However, rather than focusing on the discreteness imposed by the censoring mechanism in its entirety, this analysis wants to justify the model proposed in the next paragraph as an approximation able to take into account the observed simultaneity deriving from the discretization imposed by the recording technology, only at the scale of the minimum unit of measure.

2.2.3. A Censoring MSMD Model

Continuing from the discussion above, I assume that the observed duration $\tilde{d}_i$, is strictly positive if and only if the true duration d_i is greater then the cutoff value c for the censored record:

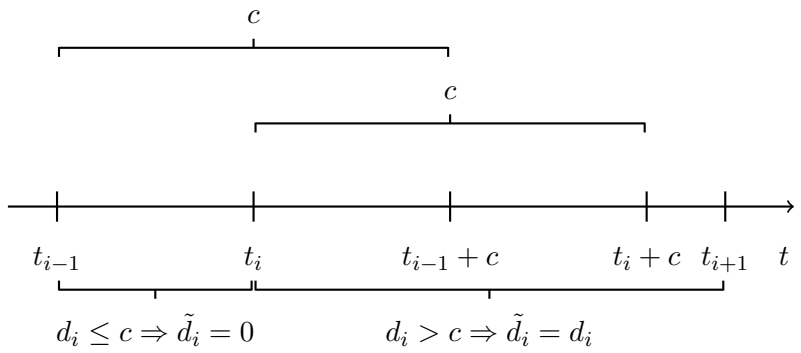
$$\tilde{d}_i = \mathbb{1}\{d_i > c\} d_i, \quad d_i = \frac{\epsilon_i}{\lambda_i}, \quad \epsilon_i \overset{\text{iid}}{\sim} \mathrm{Exp}\,(1) \tag{2.8}$$

The specification for the measurement equation implies that:

$$\mathbb{P}\mathrm{r}(\tilde{d}_i = 0) = \mathbb{P}\mathrm{r}\left(d_i \leq c\right) = 1 - \exp\left(-\lambda_i c\right)$$

which is an approximation of the expression in equation (7). Given that this representation is just an approximation , Figure 7 depicts the exact censoring mechanism that would have

Figure 7: Example of the Considered Process



given rise to the simplified specification in equation (8). As we can see, the described model is equivalent to a censoring mechanism that records strictly positive inter-event durations, $\tilde{d}_i > 0$, if and only if such events are far enough apart in time.

The approximation employed in equation (8) overstates the probability that simultaneous events occur consecutively. To see this, notice that as we observe a sequence of consecutive zero durations, say $\{\tilde{d}_{i+1} = 0, \ldots, \tilde{d}_{i+r-1} = 0\}$, with $\tilde{d}_i > 0$, the true probability of observing the $(i+r)$-th duration to be zero is decreasing in r, since, from the discussion above, it is exactly equal to:

$$\mathbb{P}\mathrm{r}(\tilde{d}_{i+r} = 0) = 1 - \exp\left(-\Lambda_{j^*}^{j^*+1}\right) \sum_{n=0}^{r} \frac{(\Lambda_{j^*}^{j^*+1})^n}{n!}$$

Therefore, the constructed censoring model inflates the probability of observing a streak of inter-trade durations equal to zero *vis à vis* the true implications from the recording mechanism.

Properties of The Model

As above, I trace out the properties of the censoring model, keeping in mind as reference both the empirical regularities present in inter-trade duration data and the properties of the original MSMD model.

Stationarity, Ergodicity and Finiteness of Moments: Because the censoring model yields a measurable function of the original duration process from the MSMD specification, it inherits its strict stationarity, ergodicity and moment finiteness properties.

Overdispersion: The censored model, like the MSMD model, exhibits unconditional over-dispersion. Moreover, as opposed to the uncensored specification, censoring induces unconditional over-dispersion even if the intensity process is non-stochastic–i.e., it implies *conditional* over-dispersion, too. These new results are summarized in the following Proposition:

Proposition 2.2.1. *Let the distribution of $\tilde{d}_i$ be described by equation (8). Then,*

i) if $c = 0$ then $c_v(\tilde{d}_i) \geq 1$, where equality holds if and only if λ_i is constant;

ii) if $c > 0$ then $c_v(\tilde{d}_i) > 1$;

Proof. Note that *i)* simply refers to the uncensored version of the model presented above and therefore the result goes through without modifications. As for *ii)*, first note that the following statements are equivalent:

$$c_v(\tilde{d}_i) > 1 \quad \Longleftrightarrow \quad c_v^2(\tilde{d}_i) = \frac{\mathbb{V}\mathrm{ar}(\tilde{d}_i)}{\mathbb{E}(\tilde{d}_i)^2} > 1 \quad \Longleftrightarrow \quad \frac{\mathbb{E}(\tilde{d}_i^2)}{\mathbb{E}(\tilde{d}_i)^2} > 2$$

Further realize that, given two random variables X, Y and a real number $k > 0$:

$$\frac{\mathbb{E}\left(X^2\,|Y\right)}{\mathbb{E}\left(X\,|Y\right)^2} \geq (>)k \quad \Longrightarrow \quad \frac{\mathbb{E}\left(X^2\right)}{\mathbb{E}\left(X\right)^2} \geq (>)k$$

By Jensen's inequality, in fact, we have that:

$$\mathbb{E}\left(\mathbb{E}\left(X\,|Y\right)^2\right) \geq \mathbb{E}\left(X\right)^2$$

Therefore:

$$\mathbb{E}\left(X^2\,|Y\right) \geq k\mathbb{E}\left(X\,|Y\right)^2 \quad \Longrightarrow \quad \mathbb{E}\left(\mathbb{E}\left(X^2\,|Y\right)\right) = \mathbb{E}\left(X^2\right) \geq k\mathbb{E}\left(\mathbb{E}\left(X\,|Y\right)^2\right) \geq k\mathbb{E}\left(X\right)^2$$

Finally let us consider the quantity:

$$\frac{\mathbb{E}(\tilde{d}_i^2\,|\lambda_i)}{\mathbb{E}(\tilde{d}_i\,|\lambda_i)^2} = \frac{\exp\left(-c\lambda_i\right)\left[\left(c+\lambda_i^{-1}\right)^2+\lambda_i^{-2}\right]}{\left(\exp\left(-c\lambda_i\right)\left(c+\lambda_i^{-1}\right)\right)^2} = \exp(c\lambda_i)\left[1+\left(1+c\lambda_i\right)^{-2}\right] = \psi\left(c\lambda_i\right)$$

Since, by definition, $\lambda_i > 0$, we have that $c\lambda_i > 0$ as long as $c > 0$. Let $x = c\lambda_i$. Then note that:

$$\psi(0) = 2$$

and

$$\frac{\partial\psi(x)}{\partial x} = \exp\left(x\right)\left[\frac{(1+x)^3+(1+x)-2}{(1+x)^3}\right] \geq (>)0 \quad \text{if} \quad x \geq (>)0$$

Therefore we can conclude that $\psi(x) > 2$ for $x > 0$, proving the desired result.

□

A Zero-Augmented Extension of the Censoring MSMD Model

As mentioned above, the observed simultaneity in the data could also be produced, in whole or in part, by simultaneity fundamental to the point process itself. In order to allow for this possibility, in the spirit of Hautsch et al. (2010), I build a Zero-Augmented MSMD model (ZA-MSMD), introducing a point mass at zero (and zero alone) independent of the

censoring mechanism. Specifically, the specification for the ZA-MSMD model is as follows:

$$\tilde{d}_i = \mathbb{1}\{d_i > c\} d_i, \quad d_i = \mathcal{S}_i \frac{\epsilon_i}{\lambda_i} \tag{2.9}$$

with

$$\epsilon_i \overset{\text{iid}}{\sim} \text{Exp}\,(1) \quad \text{and} \quad \mathcal{S}_i \overset{\text{iid}}{\sim} \text{Bernoulli}(1-p)$$

This specification for the measurement equation implies that:

$$\mathbb{P}\text{r}(\tilde{d}_i = 0) = \mathbb{P}\text{r}\,(d_i \leq c) = p + (1-p)\,[1 - \exp\,(-\lambda_i c)] = 1 - \exp\,(-\lambda_i c)\,(1-p)$$

As we can see, the probability of observing a zero duration is now additionally inflated by the term Bernoulli $(1-p)$. Finally, we can readily extend the overdispersion argument to the zero-augmented model:

Corollary 2.2.2. *Let the distribution of $\tilde{d}_i^{ZA}$ be described as in (9) and the distributions of $\tilde{d}_i$ be described as in (8). Then,*

$$c_v(\tilde{d}_i^{ZA}) > c_v(\tilde{d}_i) \geq 1$$

Proof. It is sufficient to note that:

$$\mathbb{E}(\tilde{d}_i^{\text{ZA}}) = p\mathbb{E}(\tilde{d}_i) \quad \text{and} \quad \mathbb{E}[(\tilde{d}_i^{\text{ZA}})^2] = p\mathbb{E}(\tilde{d}_i^2)$$

and therefore:

$$\frac{\mathbb{E}[(\tilde{d}_i^{\text{ZA}})^2]}{\mathbb{E}(\tilde{d}_i^{\text{ZA}})^2} = \frac{1}{p}\frac{\mathbb{E}(\tilde{d}_i^2)}{\mathbb{E}(\tilde{d}_i)^2} > \frac{\mathbb{E}(\tilde{d}_i^2)}{\mathbb{E}(\tilde{d}_i)^2} \geq 2$$

□

2.2.4. Intra-Day Calendar Effects

In high-frequency data, intra-day calendar (seasonal) effects deriving from the micro-structure of the considered market are quantitatively non-negligible and, hence, need to be accounted for in the most flexible and parsimonious way possible. The standard approach in this class of models would consist of specifying the inter-trade duration, d_i, as the product of a deterministic and a stochastic component–respectively the seasonal effect at time t_i and the seasonally adjusted duration at the same time index, respectively. In particular, (1) would then become:

$$d_i = \exp(s_i) d_i^{\mathrm{SA}} \tag{2.10}$$

where s_i represents the calendar effect for duration i. This choice is very convenient since, if there are no zero observations in the considered sample, the seasonal effect can be easily pre-estimated in logarithms through OLS:

$$\log(d_i) = \mathbb{E}\left[\log(d_i^{\mathrm{SA}})\right] + s_i + \xi_i$$

where $\xi_i = \log(d_i^{\mathrm{SA}}) - \mathbb{E}\left[\log(d_i^{\mathrm{SA}})\right]$. The seasonally adjusted series would then be constructed as follows:

$$\widehat{d}_i^{\mathrm{SA}} = d_i \exp(-\widehat{s}_i)$$

I follow this approach when estimating the MSMD model on the thinned sample.
However, when considering the cases of censoring or zero augmentation, such approach would naturally translate into the following specification:

$$\tilde{d}_i = \mathbb{1}\left\{d_i^{\mathrm{SA}} \exp(s_i) \geq c\right\} d_i^{\mathrm{SA}} \exp(s_i)$$

In this case pre-estimation by OLS would result in inconsistent estimates for the seasonal effects since the sub-sample selection to discard the zero observation would introduce correlation between the seasonal effects and the log-linear specification errors ξ_i. Therefore, I will estimate the seasonal effects and the MSMD parameters together, by employing Maximum Likelihood methods for the censoring model.

Following to Ghysels et al. (2004) and Chen et al. (2013) I divide the trading day into half-hour windows displaying constant calendar effects in the form:

$$s_i = \sum_{j=1}^{12} \alpha_j x_{j,i}$$

where $x_{j,t}$, $j = 1, 2, \ldots, 12$, are dummy variables taking the value one when t_i is in the j-th daily window. An immediate implication is that this specification requires the estimation of twelve coefficients and its flexibility directly depends on the chosen number of intra-day windows. Moreover, simple identification arguments demonstrate that it is impossible to disentangle the mean of the seasonally adjusted process when employing a full set of dummy variables. Since I am studying structural change over time for the inter-trade duration processes, it is fundamental to keep track of the levels around which they are located across the different years. In the subsequent analysis, therefore, I decide to measure seasonal effects relative to the first half-hour daily window and hence assume that $\alpha_1 = 0$.

2.3. Estimation: Methodology

In this section I first discuss the estimation methodology employed for the MSMD and the Censoring MSMD models and issues in their identification. I then report the estimation results for both the thinned and the full samples and proceed to compare estimated parameter values, goodness of fit and forecasting accuracy over the considered years. I conclude by a comparison of the results generated by the two procedures and a discussion of the most relevant changes in the inter-trade duration processes over the last two decades based on the estimated quantities.

2.3.1. Likelihood, Filtering and Identification

The MSMD Model can be estimated readily by Maximum Likelihood, by utilizing an appropriate filtering technique. Denote the set of parameters to be estimated with $\theta = \{\lambda, m_0, b, \gamma^*\}$. Given $\bar{k}$, the Likelihood function can be immediately factorized as:

$$\mathscr{L}_{\bar{k}}\left(d_{1:N}\,;\,\theta\right) = \prod_{i=1}^{N} \mathbb{P}\mathrm{r}\left(d_i \,\middle|d_{1:i-1}\,;\,\theta, \bar{k}\right) \tag{2.11}$$

Further recall that the intensity process, being multiplicatively composed of $\bar{k}$ independent , first order, Markov switching chains with transition matrix $\mathcal{P}_k$, is itself a first order Markov switching chain with finite support, $\Lambda_{\bar{k}}$, composed of $2^{\bar{k}}$ elements and transition matrix $\mathcal{P}^{\lambda}_{\bar{k}}$. The new support and transition matrix are simple to calculate. In particular, let $A_1 = [m_0, 2 - m_0]$ and $A_k = A_{k-1} \otimes A_1$ for any $1 < k \leq \bar{k}$. Then $\Lambda_{\bar{k}}$ consists of the elements in the vector $A_{\bar{k}}$. Similarly, letting $P_1 = \mathcal{P}_1$ and $P_k = P_{k-1} \otimes \mathcal{P}_k$ for any $1 < k \leq \bar{k}$, we will have that $\mathcal{P}^{\lambda}_{\bar{k}} = P_{\bar{k}}$. I can then employ a simple filter in the spirit of Hamilton (1989) as described in Durbin and Koopman (2001) to evaluate the Likelihood at any given parameter values, θ. For the ease of exposition, I drop the θ and $\bar{k}$ notation in what follows.

i. For every period $i \geq 1$, suppose we know $\mathbb{P}\mathrm{r}\left(\lambda_{i-1} \left| d_{1:i-1} \right.\right)$;

ii. We can then iterate λ_{i-1} forward according to $\mathcal{P}_k^{\lambda}$ to get the conditional joint distribution:

$$\mathbb{P}\mathrm{r}\left(\lambda_i, \lambda_{i-1} \left| d_{1:i-1} \right.\right) = \mathbb{P}\mathrm{r}\left(\lambda_i \left| \lambda_{i-1} \right.\right) \mathbb{P}\mathrm{r}\left(\lambda_{i-1} \left| d_{1:i-1} \right.\right)$$

By integrating out λ_{i-1} we can get the predicted distribution for the intensity given the information up trough duration $i-1$:

$$\mathbb{P}\mathrm{r}\left(\lambda_i \left| d_{1:i-1} \right.\right) = \sum_{\lambda_{i-1}} \mathbb{P}\mathrm{r}\left(\lambda_i, \lambda_{i-1} \left| d_{1:i-1} \right.\right)$$

iii. We can then compute the joint distribution of the state and the measurement:

$$f\left(\lambda_i, d_i \left| d_{1:i-1} \right.\right) = f\left(d_i \left| \lambda_i \right.\right) \mathbb{P}\mathrm{r}\left(\lambda_i \left| d_{1:i-1} \right.\right)$$

where $f\left(d_i \left| \lambda_i \right.\right)$ the p.d.f. of the relevant distribution for the model at hand. In particular:

- For the MSMD model:

$$f\left(d_i^{\mathrm{SA}} \left| \lambda_i \right.\right) = \lambda_i \exp(-\lambda_i d_i^{\mathrm{SA}})$$

- For the Censoring MSMD model:

$$\begin{aligned} f(\tilde{d}_i \left| \lambda_i \right.) = {} & \mathbb{1}\{\tilde{d}_i = 0\} \left\{1 - \exp[-c\lambda_i \exp(-s_i)]\right\} \\ & + \mathbb{1}\{\tilde{d}_i > c\} \lambda_i \exp(-s_i) \exp[-\lambda_i \exp(-s_i)\tilde{d}_i] \end{aligned}$$

- For the Censoring and Zero Augmented MSMD model:

$$f(\tilde{d}_i \,|\lambda_i\,) = \mathbb{1}\{\tilde{d}_i = 0\}\,\{1 - \exp[-c\lambda_i \exp(-s_i)](1-p)\} \\ + \mathbb{1}\{\tilde{d}_i > c\}(1-p)\lambda_i \exp(-s_i)\exp[-\lambda_i \exp(-s_i)\tilde{d}_i]$$

iv. Finally, integrating out λ_i from the previously calculated joint distribution, we can obtain the period Likelihood:

$$f(d_i \,|d_{1:i-1}\,) = \sum_{\lambda_i} f\,(\lambda_i, d_i \,|d_{1:i-1}\,)$$

and the updated distribution for the latent state:

$$\mathbb{P}\mathrm{r}\,(\lambda_i \,|d_{1:i}\,) = f\,(\lambda_i, d_i \,|d_{1:i-1}\,)\,/\,f\,(d_i \,|d_{1:i-1}\,)$$

Recursively repeating the procedure for every $i = 1, \ldots, N$, we can finally evaluate the Likelihood function at a point θ. The ML estimator for θ $(\widehat{\theta})$ will be then found by numerically maximizing the Likelihood function.

The MSMD model suffers from three potential sources of identification failure. Firstly, when $\gamma^* = 1$ or $\gamma^* = 0$, the sequence of renewal probabilities, $\{\gamma_k\}_{k=1}^{\bar{k}}$, becomes a sequence of constants regardless of b. In such a situation, the value of b would not affect the Likelihood function and hence the decay parameter could not be identified. Further, even if the exact cases are ruled out by the assumption that $\gamma^* \in (0, 1)$, the decay parameter b becomes weakly identified as the parameter value approaches the boundaries of the parameter space. The second, similar, case of identification failure is when the decay parameter b is equal to 1. Again, the sequence of renewal probabilities $\{\gamma_k\}_{k=1}^{\bar{k}}$ becomes a sequence of constants regardless of γ^* which ca not, therefore, be identified. Similarly, even as 1 is explicitly ruled out from the parameter space, the same weak identification problem arises for γ^* when b approaches its lower bound. Finally, in the unfortunate case that $m_0 = 1$, the intensity

components collapse to become identical across states[2] and therefore neither b nor γ^* can be identified. Furthermore, in the case of the censoring and zero-augmented specification, all the parameter of the MSMD model will be poorly identified if c and p are sufficiently large since most observations will be zeros.

The numerical implementation of the estimation procedure is performed mixing `Matlab` and `Fortran 90` codes. In particular, the filtering step relies on the faster language, `Fortran 90`, while the less computationally intensive part of the algorithm is written in `Matlab`. For the optimization algorithm, I choose the latest available version of Hansen's CMA-ES (Hansen, 2006), whose source code is available at `https://www.lri.fr/~hansen/cmaes_inmatlab.html`.

In the remainder of this section I will first present estimates for the MSMD model estimated on the thinned sample, with particular attention to the bundle distribution of each parameter throughout the considered years, performing in-sample goodness of fit and (pseudo) out-of-sample forecasting accuracy comparisons. Secondly, I will present the estimates for the Censoring and Zero Augmented MSMD model, again commenting on the bundle distribution for each parameter, performing in-sample goodness of fit and (pseudo) out-of-sample forecasting accuracy comparisons. Finally I will compare the two procedures pointing out their most striking differences and highlight the most important features of the structural change occurred in the last two decades for inter-trade duration processes.

[2]Recall that the case in which $m_0 = 1$ implies that $\mathbb{V}\text{ar}(\lambda_t) = 0$.

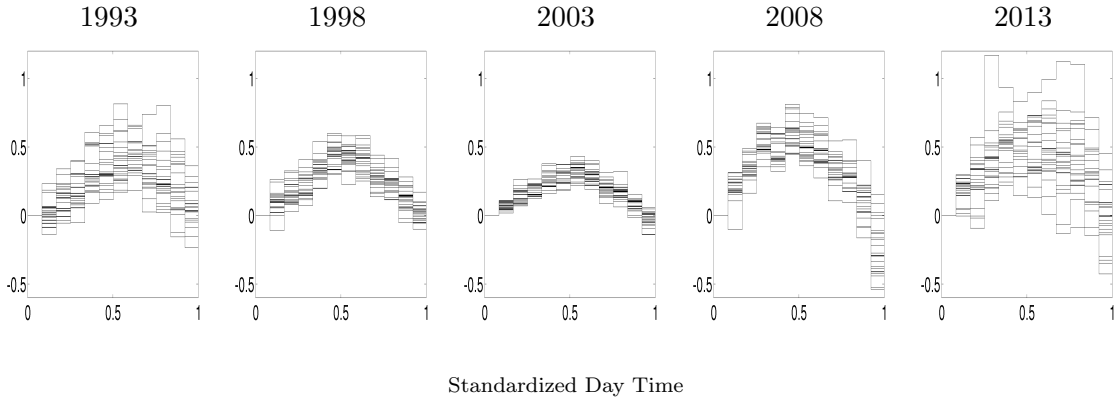
2.4. Estimation: Results For Thinned Sample

In this section I report estimation results for the basic MSMD model estimated by exploiting the thinned samples. I first comment on the estimated intra-day calendar effects, then proceed to the MSMD model parameters and finally perform in-sample goodness of fit and forecasting accuracy comparisons over the considered years.

Intra-Day Calendar Effects

Figure 8 plots bundles of the estimated intra-day calendar effects over the considered years.

Figure 8: Intra-Day Calendar Effects: $\widehat{\alpha}_j$, Bundle



For a more precise assessment, Figure 47 in Appendix B and Table 48 – 52 depicts and reports the estimated intra-day seasonal effects for each equity in the bundle over the considered years.

As expected, the estimated seasonal profiles are hump-shaped over the day, implying a faster trading pace at the beginning and at the end of the trading day; this is consistent with the preceding literature. Furthermore, the seasonal effects seem to decrease in magnitude from 1993 to 2003, after which such trend is reversed. Bundle heterogeneity follows a similar pattern, decreasing until 2003 and growing again afterwards.

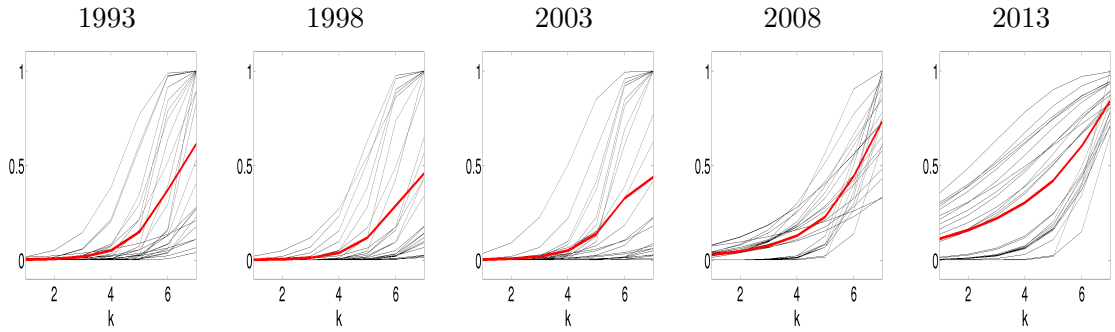
Model Parameters

Figures 9, 10, 11, 12 and 13 report time series histogram plots of estimated parameters (or

combinations thereof) over the considered years.

In particular, Figure 9 depicts the estimated renewal probabilities for the Markov chains $M_{k,i}$, $\widehat{\gamma}_k$, as a function of k. As described in Section 2, the renewal probabilities are functions of b and γ^*. The former determines the decay of γ_k as a function of k. A smaller b implies a flatter profile. The latter, on the other hand, defines the maximum renewal probability among the chains, that is when $k = \bar{k}$. All else equal, a higher γ^* implies lower persistence for the individual chains. The most apparent feature of the change in the renewal probability

Figure 9: Time Series of Estimated Renewal Probability Profiles: $\widehat{\gamma}_k$, Bundle of Firms (Black) and Bundle Average (Red)



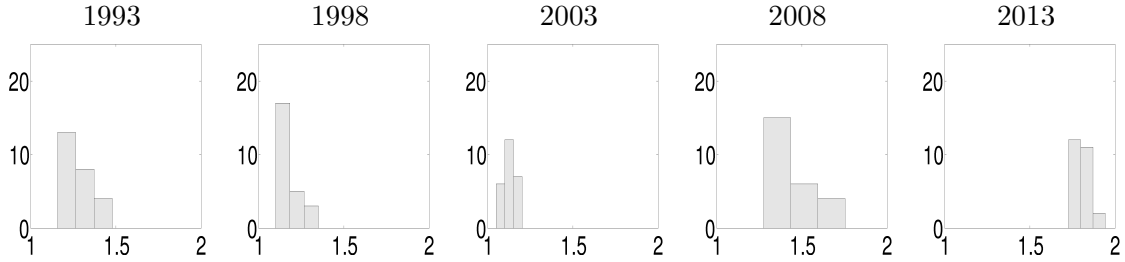
profiles across time is certainly the decreasing persistence of the least persistent chain and the progressive flattening of the profiles themselves. Both these facts are reflected in the decreasing pattern for $\widehat{b}$ and the increasing path for $\widehat{\gamma}^*$, as depicted in Figures 12 and 13. In a nutshell, the individual chains become less persistent and, as a consequence, the intensity functions and by proxy the inter-trade duration processes become, in turn, less strongly

Figure 10: Time Series of Histograms of Estimated Parameters: $\log \widehat{\lambda}$

correlated, confirming the empirical evidence provided in Section 1.

The preceding discussion focused on the dynamic aspects of the probabilistic structure of inter-trade durations. Figures 10 and 11 depict histograms of the estimates of the parameters that govern the properties of the unconditional distribution of inter-trade durations, $\widehat{\lambda}$ and $\widehat{m}_0$. In particular, $\widehat{\lambda}$ (roughly) represents the average intensity while $\widehat{m}_0$ determines

Figure 11: Time Series of Histograms of Estimated Parameters: $\widehat{m}_0$



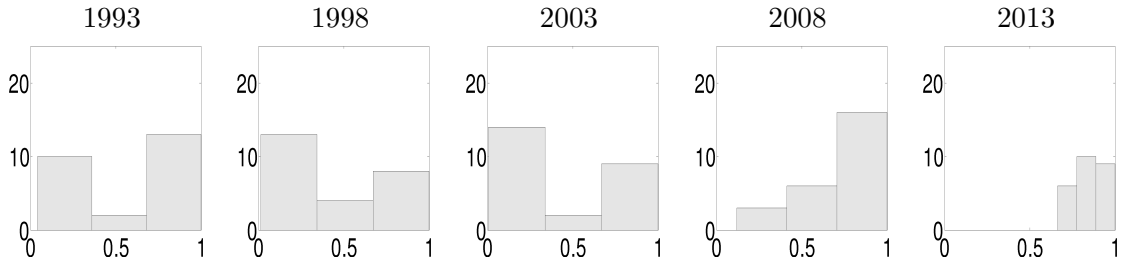
the variance of each Markov chain $M_{k,i}$. As expected from the reduced form evidence, the average intensity λ has sizably increased since 1993, accounting for the associated explosion of the number of transactions. Similarly m_0 is estimated to have grown closer to its upper bound of 2. As stressed before, this fact implies that the individual chains, and hence the intensity functions and the duration processes, become more volatile relative to their mean–i.e. they become more overdispersed.

Figure 12: Time Series of Histograms of Estimated Parameters: $\widehat{b}$

It is also interesting to note another pattern in the persistence parameter γ^*. In 1993, in fact, the considered bundle is clearly divided into two groups characterized by high and low persistence. Such bifurcation appears to have persisted until 2003. Starting with the 2008 sample, however, this bi-modality start to disappear and all considered equities converge

towards lower persistence levels, both in terms of least persistent chain and, worth to recall, in terms of flatness of the renewal probability profiles.

Figure 13: Time Series of Histograms of Estimated Parameters: $\widehat{\gamma}^*$



In-Sample Goodness of Fit and Pseudo-Out-of-Sample Forecasting Accuracy

In order to evaluate the fit of the MSMD model on the thinned sample I consider a specification test based on the degree of autocorrelation of the residual durations, that is, the ϵ_is. This approach evinces the ability of the MSMD model to account for the serial correlation displayed by the data and allows for comparisons across equities and years. I compute the residual durations, or filtered errors $\widehat{\epsilon}_i$, according to the following formula:

$$\widehat{\epsilon}_i = \mathbb{E}(\epsilon_i \,|d_{1:i}) = d_i \mathbb{E}(\lambda_i \,|d_{1:i}) = d_i \sum_{i=1}^{2^{\bar{k}}} \lambda_i \mathbb{P}\mathrm{r}(\lambda_i \,|d_{1:i})$$

where $\mathbb{P}\mathrm{r}(\lambda_i \,|d_{1:i})$ is the filtered probability distribution for the unobserved states produced by the Hamilton filter. I then calculate Ljung-Box statistics employing the first five hundred autocorrelations for the residual durations. This statistic is distributed as a χ^2_{500} random variable. Figure 14 depicts the aforementioned L-B statistics and the relevant 5% critical value. We can see that for all the considered samples, in a relevant fraction of cases, the test cannot be rejected, which suggests that the MSMD is able to account for the dynamic properties of the data.

Forecasting accuracy is measured by employing predictive R^2. This choice of metric has been made in order to allow for comparisons over samples whose distributions are characterized by a great degree of location heterogeneity. Point forecasts are computed as conditional

Figure 14: Time Series of Histograms of Ljung-Box Statistics Using 500 Acorrs. (Red Vertical Line is 5% Critical Value)

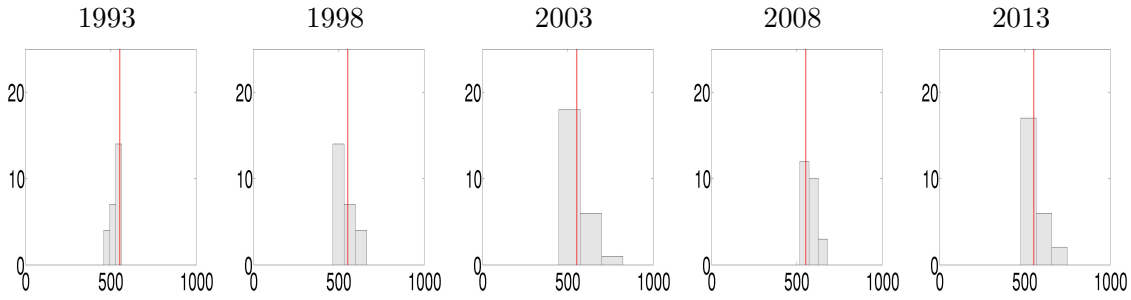
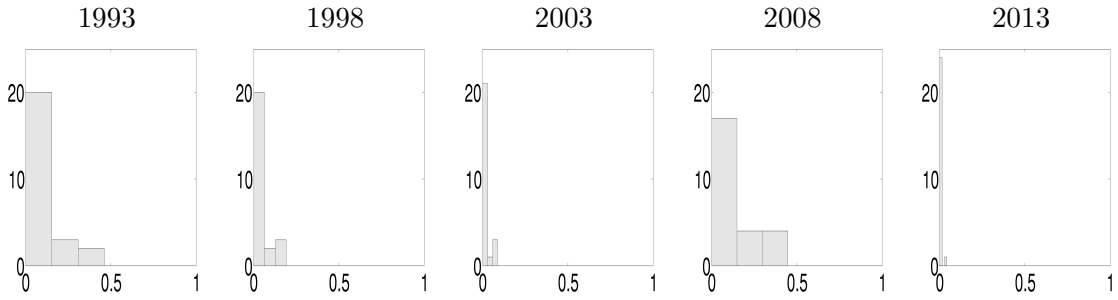


Figure 15: Time Series of Histograms of Predictive R^2's: One-Trade-Ahead Predictive Horizon



means according to the following formula:

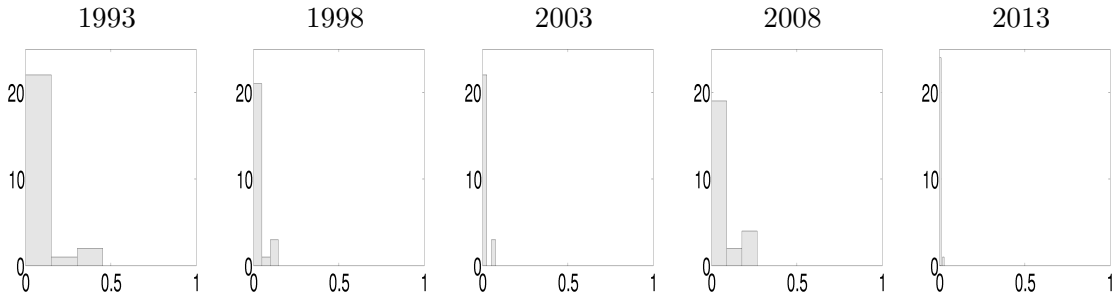
$$\mathbb{E}(d_{i+h}\,|d_{1:i}\,) = \mathbb{E}(\lambda_{i+h}^{-1}\,|d_{1:i}\,) = \sum_{i=1}^{2^{\bar{k}}} \lambda_i^{-1} \mathbb{P}\mathrm{r}(\lambda_{i+h}\,|d_{1:i}\,)$$

and

$$[\ldots, \mathbb{P}\mathrm{r}(\lambda_{i+h}\,|d_{1:i}\,), \ldots] = [\ldots, \mathbb{P}\mathrm{r}(\lambda_i\,|d_{1:i}\,), \ldots] \times (\mathcal{P}_{\bar{k}}^{\lambda})^h$$

where $\mathcal{P}_{\bar{k}}^{\lambda}$ is the transition matrix, and $\mathbb{P}\mathrm{r}(\lambda_i\,|d_{1:i}\,)$ is, again, the filtered probability distribution for the unobserved states produced by the Hamilton filter. Figures 15 and 16 reproduce time series of histogram plots for the predictive R^2 at two forecasting horizons, namely one and twenty trades ahead. It is striking to note how the predictability of inter-trade durations by the MSMD has collapsed over the years. In particular, in 2013, almost all considered equities are characterized by extremely low predictive R^2 at both horizons considered.

Figure 16: Time Series of Histograms of Predictive R^2's: Twenty-Trade-Ahead Predictive Horizon

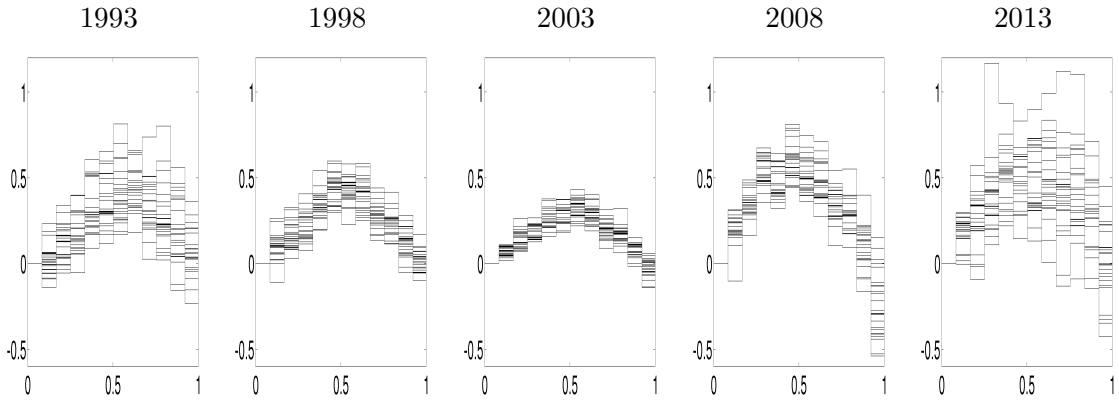


2.5. Estimation: Results For The Censoring Model

In this section I report the estimation results for the Censoring MSMD model produced by exploiting full samples, including the zero durations, or simultaneous trades. I first comment on the estimated intra-day calendar effects, then move on to the MSMD model parameters and finally perform in-sample goodness of fit and forecasting accuracy comparisons over the considered years.

Intra-Day Calendar Effects The estimated intra-day calendar effects are very similar to those of the thinned samples. Figure 17 depicts the seasonal effects as a function of standardized daily time. Again, I find the expected hump-shape characterizing the profiles

Figure 17: Intra-Day Calendar Effects: $\widehat{\alpha}_i$, Bundle



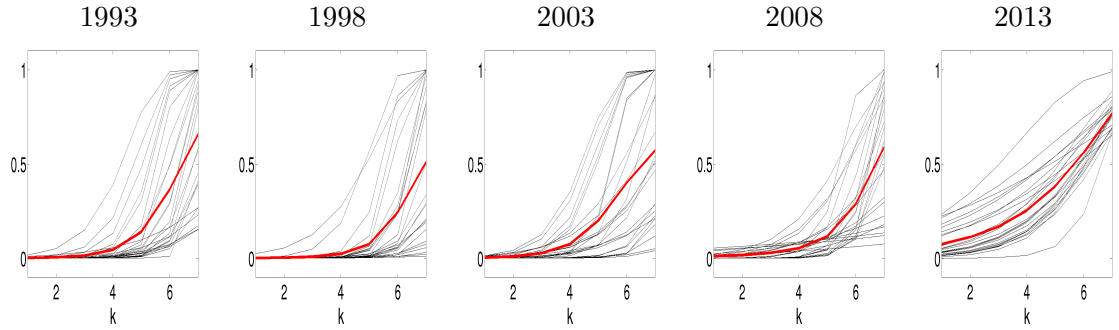
with higher trading intensity at the beginning and at the end of the trading day. As before, the seasonal effects seem to decrease in magnitude from 1993 to 2003, before reversing through 2013. Again, bundle heterogeneity follows a similar pattern, decreasing until 2003 and growing thereafter.

Model Parameters

In parallel to Figure 9, Figure 18 depicts the estimated renewal probabilities for the Markov chains $M_{k,i}$, $\widehat{\gamma}_k$. Recall that the renewal probabilities are functions of b and γ^* and that

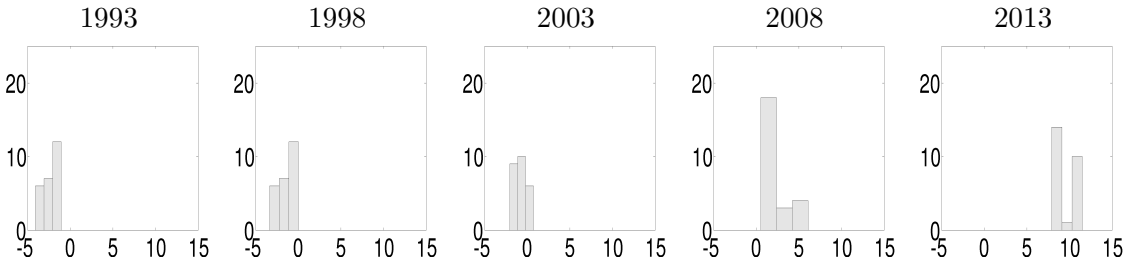
they determine, respectively the decay of γ_k as a function of k and the highest renewal probability among the same chains. The qualitative features of the renewal probability

Figure 18: Time Series of Estimated Renewal Probability Profiles: $\widehat{\gamma}_k$, Bundle of Firms (Black) and Bundle Average (Red)



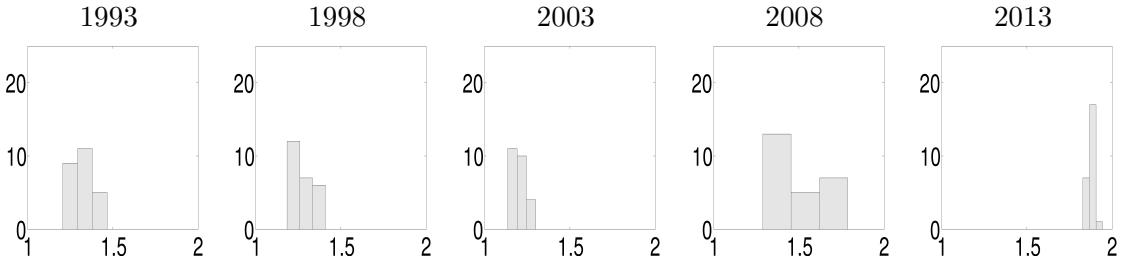
profiles are almost unchanged. The decreasing persistence of the least persistent chain and the progressive flattening of the profiles are still very evident from the figure. Just like before, both of these facts are reflected in the decreasing pattern for $\widehat{b}$ and the increasing path for $\widehat{\gamma}^*$, as depicted in Figure 21 and Figure 22. Quantitatively, the biggest differences are registered for the 2003, 2008 and 2013 samples–as expected, given the relative importance of zero durations–which point towards less persistence in the individual chains, and hence in the intensity and duration processes. Moving on to discuss the estimates of the parameters

Figure 19: Time Series of Histograms of Estimated Parameters: $\log \widehat{\lambda}$



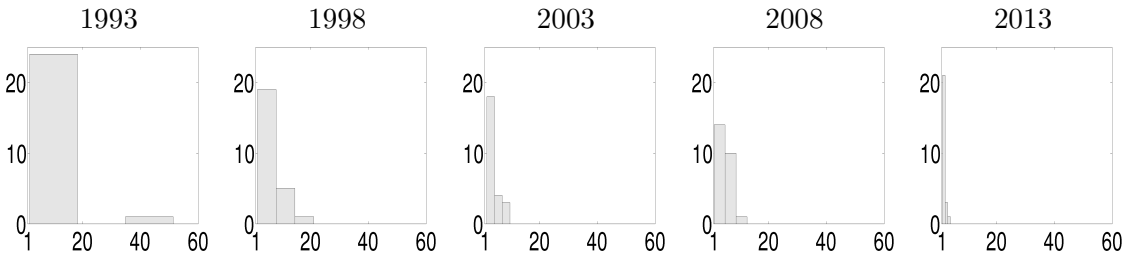
that govern the properties of the unconditional distribution of inter-trade durations, we turn to Figures 10 and 11, which show time series histogram plots for $\widehat{\lambda}$ and $\widehat{m}_0$. Recall that $\widehat{\lambda}$ represents the average intensity while $\widehat{m}_0$ determines the variance of each Markov chain $M_{k,i}$. Again, as in the thinning case, the average intensity λ sizably increases over the

Figure 20: Time Series of Histograms of Estimated Parameters: $\widehat{m}_0$



years, reflecting the increase of the number of transactions. However, as will be made more precise in the next subsection, the quantitative differences between the estimates provided by the two methods are quite substantial. Lastly, m_0 is, as above, estimated to grow closer to its upper bound of two, implying more overdispersed inter-trade duration distributions.

Figure 21: Time Series of Histograms of Estimated Parameters: $\widehat{b}$



As we can note from Figure 21, the collapse of $\widehat{b}$ towards one is even more stark in the context of the censoring model. Moreover, the bundle bifurcation between low and high persistence equities persists through the 2008 sample, differently from before, as highlighted by Figure 22.

Finally, estimates of p, the probability of *iid* simultaneous trade, are provided in Appendix B. It is worth mentioning that the values of p are statistically significant only for the 2008 sample.

In-Sample Goodness of Fit and Pseudo-Out-of-Sample Forecasting Accuracy

In order to evaluate the fit of the Censoring MSMD model estimated on the full sample I

Figure 22: Time Series of Histograms of Estimated Parameters: $\widehat{\gamma}^*$

compute the residual durations, or filtered errors $\widehat{\epsilon}_i$, which are now calculated as:

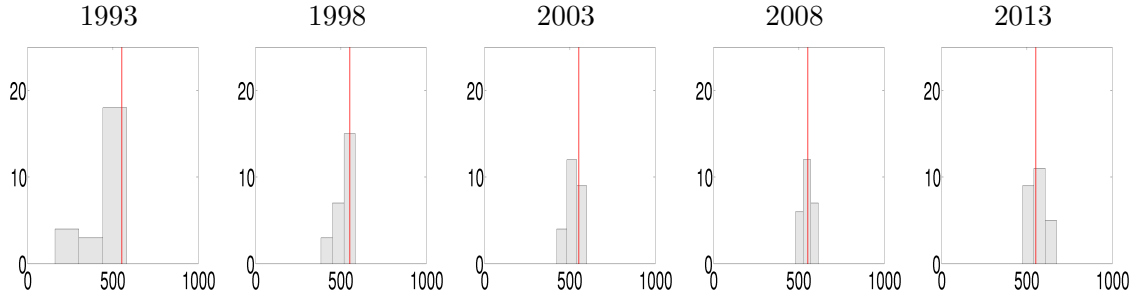
$$\begin{aligned}\widehat{\epsilon}_i = \mathbb{E}(\epsilon_i|\tilde{d}_{1:i}) =& \Pr(\mathcal{S}_i = 0|\tilde{d}_{1:i}) + \Pr(\mathcal{S}_i = 1|\tilde{d}_{1:i}) \left[\mathbb{1}\{\tilde{d}_i = 0\}\tilde{d}_i \exp(-s_i)\mathbb{E}(\lambda_i|\tilde{d}_{1:i}) \right. \\ &\left. + \mathbb{1}\{\tilde{d}_i = 0\}\mathbb{E}\left(\frac{1 - \exp(-c\lambda_i \exp(-s_i))(c\lambda_i \exp(-s_i) + 1)}{1 - \exp(-c\lambda_i \exp(-s_i))} \,\middle|\, \tilde{d}_{1:i}\right)\right]\end{aligned}$$

where

$$\Pr(\mathcal{S}_i = 0|\tilde{d}_{1:i}) = \mathbb{1}\{\tilde{d}_i = 0\}\frac{\widehat{p}}{f(\tilde{d}_i|\tilde{d}_{1:i-1})}$$

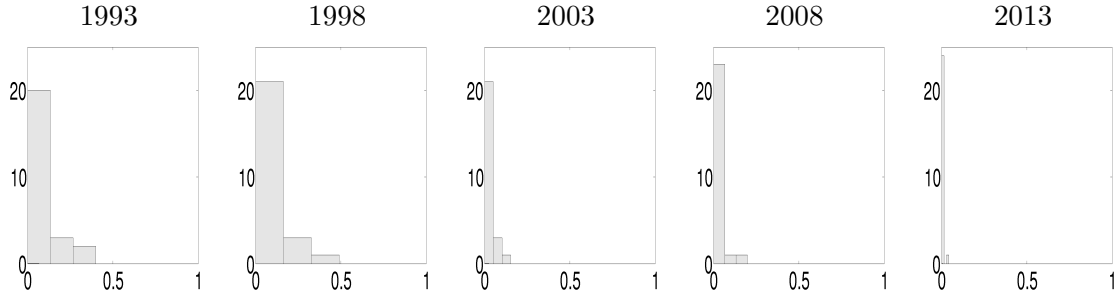
is the filtered probability of the simultaneous trade indicator $\mathcal{S}_i$ taking on the value of zero. The relevant expectations can be computed according to the filtered distribution of the latent state λ_i, provided by the Hamilton filter. In this case the filtering procedure is slightly more complicated, as the observation of a zero duration in conjunction with the latent state filtered location determines the relative likelihood of censoring and *iid* simultaneous trade. As before, I then calculate Ljung-Box statistics employing the first five

Figure 23: Time Series of Histograms of Ljung-Box Statistics Using 500 Acorrs.
(Red Vertical Line is 5% Critical Value)



hundred autocorrelations for the residual durations. Figure 23 depicts the afore-mentioned L-B statistics and the relevant 5% critical value for the Censoring model. Similarly as before, the test cannot be rejected in a relevant fraction of cases indicating that the Censoring MSMD is able to account for the dynamic properties of the full sample data, unlike the MSMD model.

Figure 24: Time Series of Histograms of Predictive R^2's: One-Trade-Ahead Predictive Horizon



Forecasting accuracy is once again measured by employing the predictive R^2. Point forecasts are computed as conditional means according to the following formula:

$$\mathbb{E}(\tilde{d}_{i+h}|\tilde{d}_{1:i}) = (1-\widehat{p})\mathbb{E}\left[\exp(-c\lambda_{i+h}\exp(-s_{i+h}))(\lambda_{i+h}^{-1}\exp(s_{i+h})+c)\,\middle|\,\tilde{d}_{1:i}\right]$$

Again, the relevant expectation can be computed according to the filtered distribution of the latent state λ_i, provided by the Hamilton filter and the transition matrix $\mathcal{P}_{\tilde{k}}^{\lambda}$.

Figure 24 and Figure 25 depict time series of histogram plots for predictive R^2 at two forecasting horizons, namely one and twenty trades ahead. Again, it is striking to note the extent to which the predictability of inter-trade durations has collapsed over the years. In particular, when the full sample is considered and the censoring mechanism is taken into account, almost all considered equities are characterized by extremely low predictive R^2 at both considered horizons in 2008 and 2013, which contrasts with the thinning approach. It is worth pointing out, though, that the 2008 sample is the only one characterized by a statistically significant presence of *iid* simultaneous trade, which is intrinsically unpredictable and was completely discarded from the analysis in the thinning approach.

Figure 25: Time Series of Histograms of Predictive R^2's: Twenty-Trade-Ahead Predictive Horizon

2.5.1. Comparisons and Structural Change

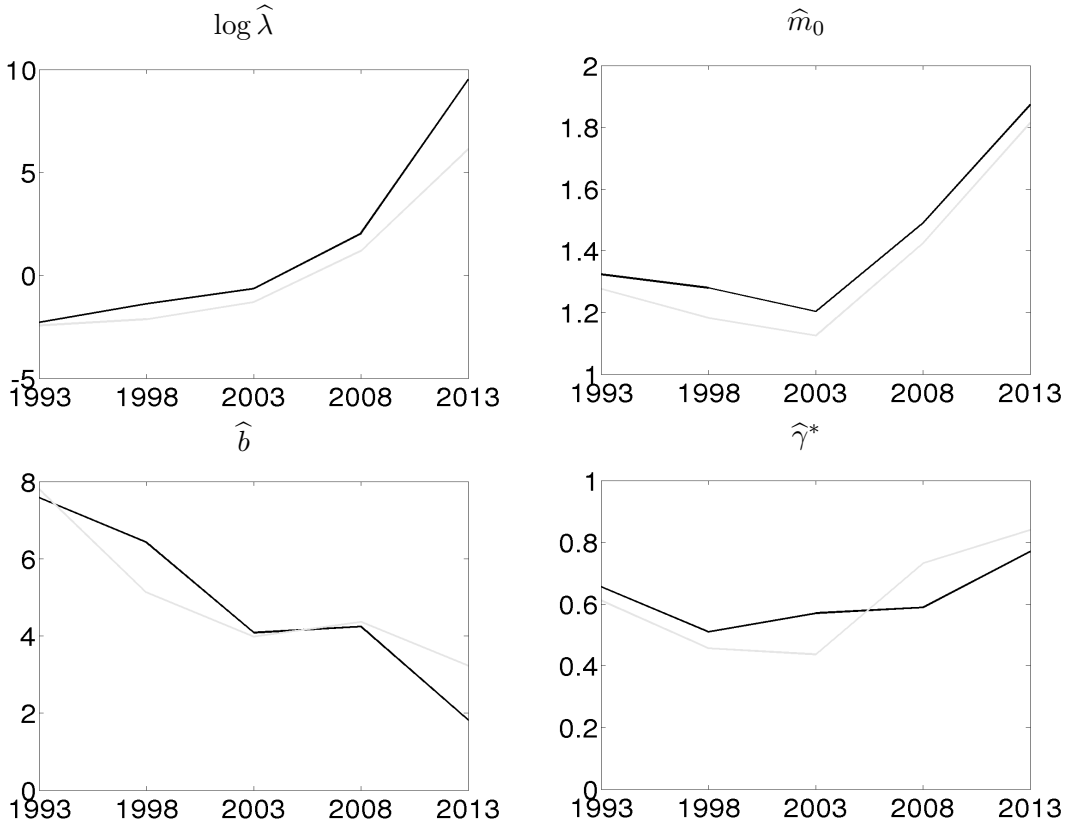
In the previous two Sections I reported estimates, in-sample fit and forecasting ability results for the two different approaches over the sample of considered years. I emphasized the structural change over time as filtered through the lens of each respective approach.

However, as mentioned in the introduction, the main theoretical difficulty with the thinning method is that it could produce severely biased and inconsistent estimators. This property is particularly salient when a hefty proportion of observations is censored. The analogy with the Tobit framework is natural. Excluding the censored observations or failing to account for the censoring mechanism produces biased and inconsistent estimators.

Therefore, beyond discussing differences in structural change through the lens of the different approaches, it is useful to quantitatively assess the inconsistency generated by the thinning procedure as compared to the censoring model.

Figure 26 shows time series plots of bundle averages of the estimated parameters over the different years. The black lines indicate the Censoring MSMD model estimates based on the full sample while the light gray lines refer to the MSMD model estimated on the thinned samples. As reported in Appendix B, individual equity differences between the two approaches are almost always statistically significant (with the exception of 1993). Quantitatively, the biggest differences are registered for $\widehat{\lambda}$, the average intensity, larger then its thinning method counterpart by up to a factor of ten in 2013. Also, the intensity's vari-

Figure 26: Time Series of Bundle Averages of Estimated Parameter Values: Censoring (Black) Vs. Thinning (Gray)



ance and overdispersion, as represented by $\widehat{m}_0$, is always bigger in the censoring approach. Instead, $\widehat{b}$ and $\widehat{\gamma}^*$ tend to be estimated as larger by the censoring model through 2003 and smaller in the two most recent samples, 2008 and 2013.

In a nutshell, ignoring censoring in the context of inter-trade durations, even when the data is measured up to the millisecond scale, produces significantly smaller parameter estimates: most notably for the average rate of the intensity process, but also for its variability and overdispersion. Results on the impact on the dynamic characterization of the probabilistic structure of inter-trade durations of censoring are quantitatively relevant, but, mixed depending on the year considered.

Finally, it is worthy of remark that, according to both methods, inter-trade durations have become more and more difficult to forecast both at short and at long horizons. The increased

unpredictability of inter-trade durations should be reflected in increased unpredictability of the calendar-time quantities of interest such as daily log-returns or transaction volumes, and therefore of price volatility and, ultimately, market liquidity.

2.6. Concluding Remarks

Inter-trade durations are one of the fundamental processes determining the behaviour of volatility and liquidity in financial markets. The ability to produce models capable of explaining the dynamic probabilistic structure of inter-trade durations and to generate accurate point and distributional forecasts translates into the ability to explain and forecast calendar time log-returns or transaction volumes at high frequencies. In turn, the ability to explain and forecast calendar time log-returns, volumes or bid-ask spreads impacts the ability of econometricians to understand and predict time-varying volatility or market liquidity.

In this work I propose an inter-trade duration model of censoring that accounts for the discretization imposed by the recording mechanism near the minimum unit of measure for transaction times, and is capable of including zero-valued durations to the econometric analysis.

Censoring has two main advantages. First, it avoids inconsistency in estimation by explicitly accounting for the censoring mechanism; prior work estimating duration models was forced to select a sub-sample of interest by discarding the simultaneous durations, which involves data reductions of increasingly substantial size. Secondly, unlike models of pure simultaneous trade, where transactions in general lose a concept of ordering, censoring maintains the meaning of this construct. Moreover, the underlying duration process can be–as in continuous time, uncensored settings–modelled as a *simple* point process, such as a Poisson process with time-varying intensity.

I use the Censoring MSMD model to study 25 U.S. equities from the S&P100 universe over a span of time ranging from 1993 to 2013. The Censoring MSMD model fits and forecasts the data as well as the MSMD model estimated on the thinned samples in all considered years, with the exception of 2008. Besides similar levels of in-sample goodness of fit and out-of-sample forecasting ability the censoring model has the appealing property that it avoids

the estimate inconsistencies mentioned above and achieving better estimate precision by enlarging the sample considerably. Furthermore, differences in estimated parameters are found to be statistically significant and, in some cases, relevantly large.

The most noteworthy results in terms of structural change highlighted by the analysis relate to the dynamic properties of inter-trade durations and their degree of predictability. Inter-trade durations have become more overdispersed and less persistent since the 1990s. As a result, their degree of predictability has plummeted over the last two decades, especially in 2008 and 2013, regardless of the model employed. Such statistical properties of inter-trade durations, as already discussed, translate into analogous properties of calendar time quantities such as prices and volumes, and hence can be used to characterize price volatility and market liquidity. Recent events such as the Flash Crash or the late 2000s financial crisis testify to our relatively poor understanding of and ability to predict recent behaviour of financial markets, and, in particular, of market liquidity, consistent with the findings in this work.

Overall, the transaction process has undergone massive structural change in both its static and its dynamic probabilistic structure since the early 1990s. Additional work is required in order to better understand the recent developments of financial market behaviour and, especially to strengthen our ability to predict inter-trade durations and, in turn, quantities such as price volatility, bid-ask spread dynamics and market depth.

CHAPTER 3 : A Markov Switching Multifractal Conditional Poisson Model

High frequency transaction level data on financial markets has been now long exploited in the literature to study several object of interests. In particular, it's widely recognized that high frequency sampling greatly benefit the estimation precision of price volatility[1]. Intrestingly enough, as already pointed out by Chen et al. (2013), time deformation arguments suggest that the serial correlation proper of calendar-time price volatility is strongly driven by the serial correlation in calendar-time trade counts. By taking one step further, one could reasonably argue that the serial correlation observed at the trade count level parallels that for transaction time arrivals (or inter trade durations) and, eventually, the serial correlation in the information flow charachterizing the trade process itself. Moreover, as consistently found in the stochastic volatility literature, such serial correlation features very high persistency, in particular, in the form of long memory. Also, not only inter-trade durations and trade counts are closely linked to price volatility, but they serve as natural measure for market liquidity and their variability translates into liquidity risk. In principle, therefore, the researcher would like to study the information flow itself and propagate, then, the gathered information to the higher levels. The lowest level of the trading process that can be analyzed employing easy to obtain data would then be the inter-trade duration process.

Chen et al. (2013) propose a parameter driven model to study inter-trade durations, featuring the long memory property for its implied dynamics. They perform a comparative study on a sample of equities randomly selected from the $S\&P100$ index, and traded during February 1993. They demonstrate that the proposed model outperforms the benchmark ACD(1,1) model of Engle and Russell (1998), both in terms of in-sample goodness of fit measures and of out-of-sample predictive ability. A potential pitfall of the proposed anal-

[1]The "realized volatility " literature, in fact, exploits convergence of the sample quadratic variation to its population analogue, as sampling frequency increases, to obtain less variable estimates of the price volatility. Cfr. with Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002).

ysis, though, is the fact that the trade data is dated back to 1993. Recent hardware and software developments, in fact, drastically impacted the trading process on the financial markets. In a nutshell, things may now look much different than in the past. Given the striking increase in trading volumes and transaction levels registered since the beginning of the new millennium, performing a similar analysis to Chen et al. (2013) would require the ability of handling a much bigger sample dimensionality and computing times.

Focusing on the trade counts would, instead, allow to fix the size of the sample regardless of the transaction activity level. Obviously, though, in a measure depending on the aggregation level - i.e. the time interval length over which we are counting the trades - part of the information embedded in the higher resolution inter-trade durations data would be discarded. However, as for inter-trade durations, trade counts are intimately - and directly - related with price volatility an market liquidity, and, therefore, a natural object of interest for the economic analysis. In this paper, I am thus going to construct a model for trade counts mutuated from the Markov Switching Multi Fractal Conditional Duration model of Chen et al. (2013).

After its introduction in Calvet and Fisher (2001), the Poisson Multifractal framework has been mainly employed in the stochastic volatility literature. Typical examples are Calvet and Fisher (2004) and Lux and Kaizoji (2007). However, in the counts literature such framework has never been exploited to build a parameter-driven model. As highlighted by Jung and Tremayne (2011), the counts literature is composed by a vast universe of models which can be mainly divided into two categories, parameter-driven and observation driven models. Models such as Zeger (1988) SAM or Shephard (1995) Generalized Linear ARMA models, can be included in the former group.Heinen (2003) ACP and - e.g. Jung and Tremayne (2006) - INAR models can be instead listed in the latter class. In their comparisons, moreover, Jung and Tremayne (2011) find that the ACP model(or variations thereof) tends to ouperform the remaining ones when brought to high frequency financial data.

Taking the last exposed fact in its right consideration, therefore, I will finally estimate and compare the Markov Switching Multi Fractal Conditional Poisson model to the ACP benchmark, in order to evaluate its relative performance.

The chapter develops as follows. Section 1 presents and describes the data employed throughout the analysis. Section 2 introduces the proposed model and illustrates its properties. Section 3 discuss the estimation procedure and Section 4 depicts the resulting estimates and compares the MSMCP model to the ACP benchmark in terms of in-sample fit and out-of sample forecasting ability. Finally, Section 5 concludes.

3.1. Description of Data

As anticipated in the section above, in this paper I will be focusing on high frequency financial data of equity transactions. In particular, I will employ time series of number of transaction per Δ-length time intervals, for 25 equities traded on the NASDAQ exchange[2] through February 2010. I consider trading days starting at 10:00 AM and ending at 4:00 PM to avoid dealing with the microstructure effects deriving from exchange opening procedures. The raw data source is the Millisecond TAQ Database[3], which provides information about trades and quotes of financial assets traded on the major U.S. exchanges. Each trade (or associated quote) is time-stamped at millisecond precision, allowing us to study the dynamics of the process in very high detail with respect to its time scale. Table 4 displays

Table 4: Selected U.S. Equities, Ticker Symbols and Company Names

Symbol	Company Name	Symbol	Company Name
AA	ALCOA	INTC	Intel
ABT	Abbott Laboratories	JNJ	Johnson & Johnson
AXP	American Express	KO	Coca-Cola
BA	Boeing	MCD	McDonald's
BAC	Bank of America	MRK	Merck
C	Citigroup	MSFT	Microsoft
CSCO	Cisco Systems	QCOM	Qualcomm
DELL	Dell	T	AT&T
DOW	Dow Chemical	TXN	Texas Instruments
F	Ford Motor	WFC	Wells Fargo
GE	General Electric	WMT	Wal-Mart
HD	Home Depot	XRX	Xerox
IBM	IBM		

the choice for the selected equities, which directly follows from Chen et al. (2013). The original bundle was picked to be representative for the market index S&P100 in 1993. However, two equities - amongst those considered - are not even listed in the index anymore as of February 2010[4]. Nonetheless, I rather choose to retain such equities for comparative

[2]Exchange Symbol 'B', 'T' and 'X'.

[3]Which can be accessed, for example, through the Wharton Research Data Services, at `https://wrds-web.wharton. upenn.edu/wrds/`.

[4]Such equities are ALCOA and Xerox.

reasons over resampling from the S&P100 universe a new set of symbols.

Similarly, the choice of the month, trading hours and sample length follow directly from those made in Chen et al. (2013) and are justified by comparative reasons. As for the year, I decided to pick the first one I could dispose of following the aftermath of the financial crisis.

As suggested by Hautsch (2012) the raw data produced by the TAQ database must be cleaned removing erroneously reported trades due to wrong and/or delayed recording of the trade itself. Performing this task I follow the procedures reported in Barndorff-Nielsen et al. (2011) and Hautsch (2012) and check for: *i)* trades directly indicated as misrecorded (CORR≠00) and *ii)* trades associated with zero price or zero quantity marks. I don't check instead for Intermarket Sweep Orders and regard them as separate transactions. Less than 0.1% of the trades are cut from each series and considered wrong according to this method.

After having cleaned the raw data, I create Δ-length bins over the considered time period[5] and count the number of trades laying in each bin.

Table 5: February 2010, Descriptive Statistics, Comparing WMT with the Bundle, $\Delta = 10$ seconds

	Mean	Med	Max	Min	St Dev	Skew	Kurt	OD	Trades
WMT	5.04	1	203	0	9.45	4.43	40.66	17.72	206,818
Average over Bundle	5.44	1	192	0	10.74	4.44	39.54	23.05	223,137

First of all, starting from now on, I will describe the collected data either referring to the bundle of equities in general, or looking at the particular case of Walmart[6]. I choose to follow the Walmart equity closely because it can be interpreted as the "typical" equity in the bundle, being its descriptive statistics and sample autocorrelations average in the bundle itself. This can be easily seen in Table 5 and Figure 31, respectively comparing Walmart

[5]For example, constructing 1 second bins over the 19 working days summing up to 6 hours of trade each, would yield a time series with 410,400 observations starting with the number of trades in the interval from 10:00.001 AM to 10:01.000 AM of February 1st, 2013 and concluding with the same quantity between 3:59.001 PM and 4:00.000 PM of February 28th, 2013.

[6]Symbol 'WMT'.

statistics to the bundle averages and displaying the bundle of sample autocorrelations.

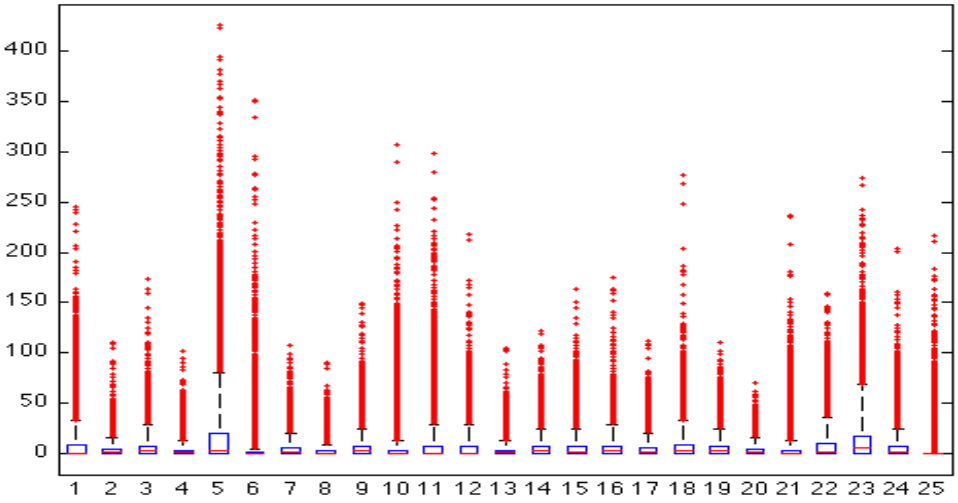
Figure 27 displays Box and Whisker Plots for the considered samples, respectively for $\Delta = 10$ seconds and $\Delta = 1$ minute. Detailed descriptive statistics can be found in Table 43, Table 44 and Table 45 in the Appendix. Equities are numbered in alphabetical order, as they are arranged in Table 4. From Figure 27, we can immediately realize that the sampling distributions are charchterized by very fat right tails and pronounced overdispersion, defined as the ratio of variance and mean. This first visual impression is confirmed by Table 43 and Table 44 in the Appendix. In particular, for $\Delta = 10$ seconds, overdispersion is at a minimum of 7.08 for Qualcomm and at a maximum of 66.83 for Citigroup, while kurtosis ranges between Texas Instruments minimum at 16.09, and Citigroup cap at 178.68. The qualitative picture for $\Delta = 1$ minute is not much different. Overdispersion minimum increases to Qualcomm 15.92 and the maximum is 114.8 for Bank of America. Kurtosis now ranges between Texas Instruments minimum at 10.15, and Citigroup maximum at 37.06. Looking at Table 45, also in the Appendix, reporting the descriptive statistics for $\Delta = 10$ minutes, we can see that overdispersion is even higher while kurtosis is, on average, smaller. It seems that, as the count interval increase, the unconditional distribution becomes more and more spread while the fatness of the right tail is reduced. This trend can also be recognized comparing panel (i) and panel (ii) of Figure 27. In particular, in fact, in panel (ii) the boxes and whiskers cover a wider section of the support of the unconditional distribution and the "outliers" presence[7] is reduced as compared to panel (i).This means that the range between the first and the third quantile is narrower using the 10 seconds bins, while the probability of extreme events[8] tends to be higher.

Looking at a Quantile-Quantile Plot of the sample distribution of Walmart against a Poisson distribution with intensity estimated by the sample average of the series (Figure 28), we can note in more detail their substantial differences in terms of fatness and length of the tails, confirming the above analysis.

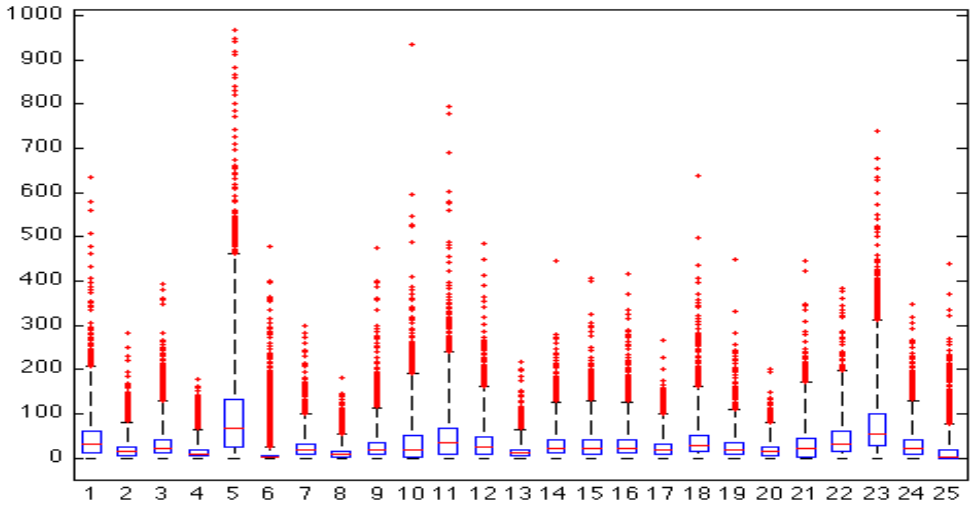
[7]Represented by the red crosses.
[8]Identified as those events exceeding the whisker.

Figure 27: Box and Whisker Plots - Number of Trades per Δ-Length Time Interval, Sample Distributions, Δ = 10 seconds

(a) Δ = 10 seconds



(b) Δ = 1 minute

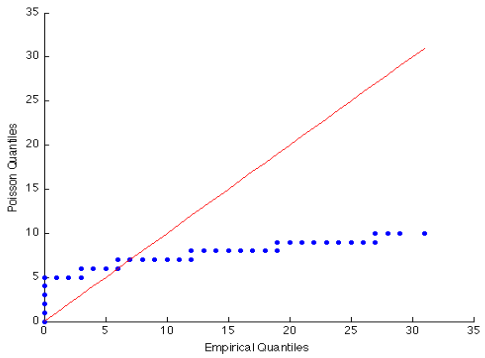


Equity (Ordered As in Table 4)

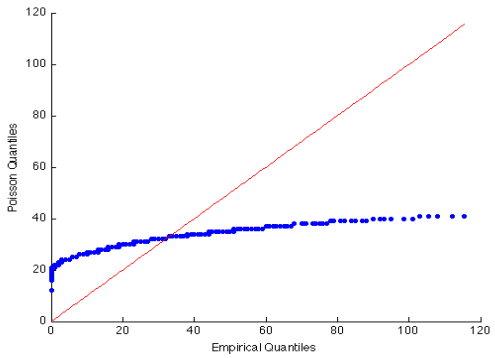
Note: Box top and bottom respectively represent the 75th and 25th percentile. The median is in between. Whiskers extend for three interquartile ranges above and below the box. The red crosses are those observations exceeding the value of the whiskers. Data has been collected from 10:00 to 16:00 every working day between 02/01/2010 and 02/28/2010 for a total of 19 days.

Figure 28: February 2010, Q-Q Plots, WMT, Sample vs Poisson

(a) Δ: 10 seconds



(b) Δ: 1 minute



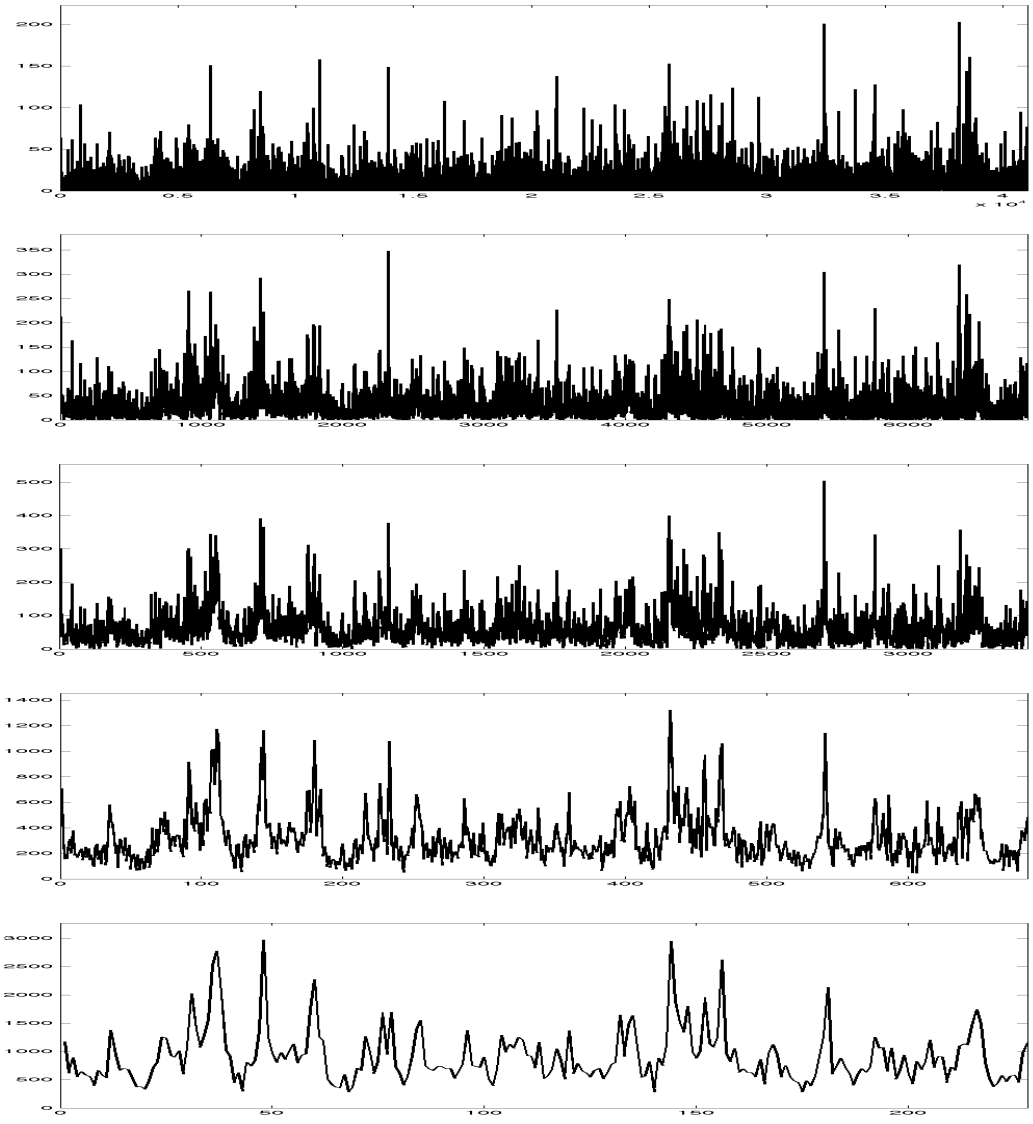
Note: The intensity parameter for the Poisson distribution is estimated by the sample average of the WMT series.

The sizeable kurtosis and overdispersion strikingly featuring the sample unconditional distribution of the number of transactions are generated by the time clustering displayed when looking at the time plot of the relevant series. Figure 29 shows the time series for the number of transactions in Δ-length intervals for the Walmart equity, respectively for $\Delta = 10$ seconds, 1 minute, 2 minutes, 10 minutes and 30 minutes. The time clustering is apparent, as well as it is its similarity across different interval lengths, or observational frequencies. Such scaling properties with respect to the observational frequency have long been emphasized in the financial asset returns literature[9] and are similar to those found in the February 1993 sample by Chen et al. (2013).

Another charachteristic worth noticing is given by the persistence that the series display. Figure 30 and Figure 31, respectively show the sample autocorrelations for Walmart and for the entire bundle of equities at various sampling frequencies. The persistence is apparent at all sampling frequencies, yielding to slowly decaing autocorrelations that remain relatively high after many lags. In particular, the sample autocorrelations display an hyperbolic - rather than exponential - decay. This is equivalent to say that the series of transaction counts feature long-memory. This charchteristic of the data coheres with the vast findings

[9]Especially by Mandelbrot, for example in Cootner et al. (1997) and Calvet et al. (1997).

Figure 29: February 2010, WMT, Number of Transactions per Δ-Length Time Interval

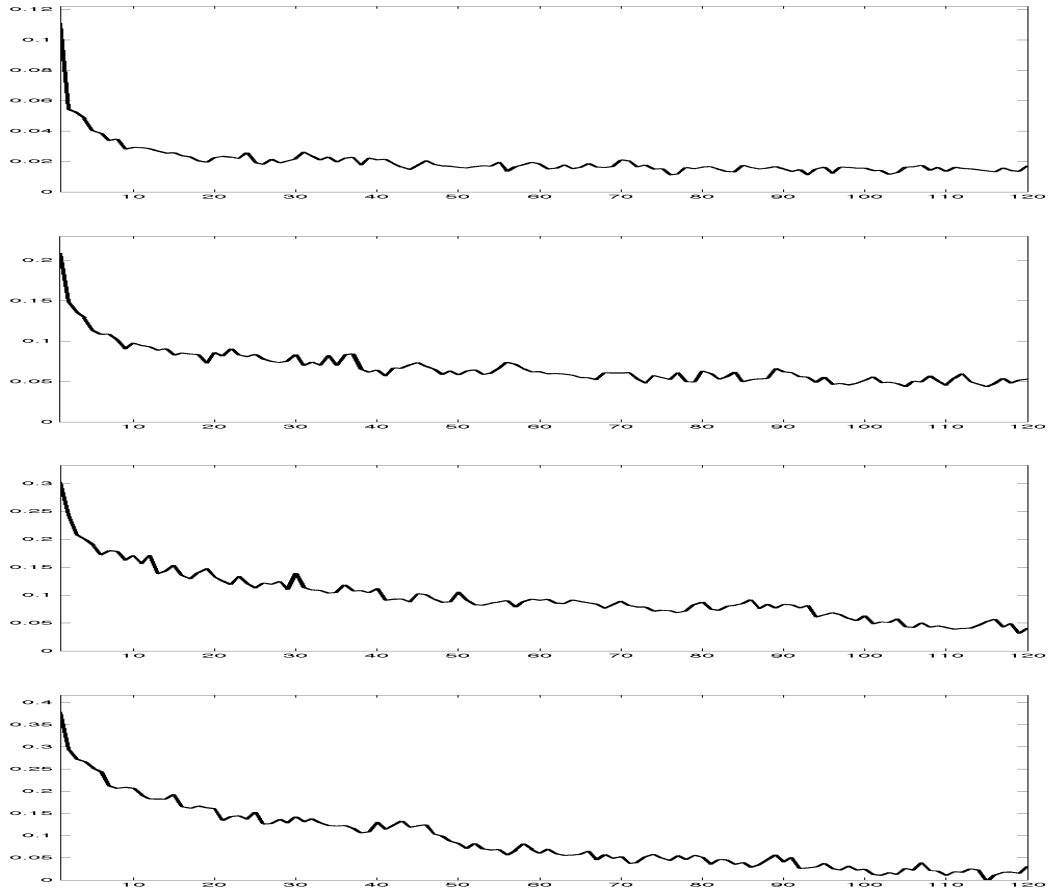


Note: From top to bottom panel, respectively: $\Delta = 10$ seconds, $\Delta = 1$ minute, $\Delta = 2$ minutes, $\Delta = 10$ minutes and $\Delta = 30$ minutes. Data has been collected from 10:00 to 16:00 every working day between 02/01/2010 and 02/28/2010 for a total of 19 days.

on long memory in the financial literature and, in particular, with the findings in Chen et al. (2013) for inter-trade durations data.

To conclude, in a nutshell, the data on the number of transactions per fixed length in terval

Figure 30: February 2010, WMT, Number of Transactions per Δ-Length Time Interval, Autocorrelations



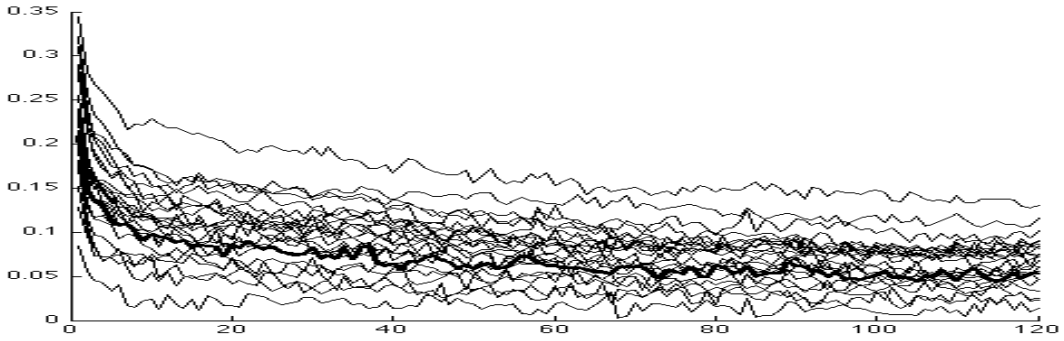
Note: From top to bottom panel, respectively: $\Delta = 1$ second, $\Delta = 10$ seconds, $\Delta = 30$ seconds and $\Delta = 1$ minute.

in time is charachterized by heavy overdispersion and fat tails, self-similar time clustering and long memory. All these charachteristics parallels those for inter-trade durations and requires an adequate treatment when the series are statistically analyzed.

In the next section I will introduce the Markov Switching Multi Fractal conditional Poisson model and sketch its properties in relation to the main charchteristics displayed by the data and analyzed above.

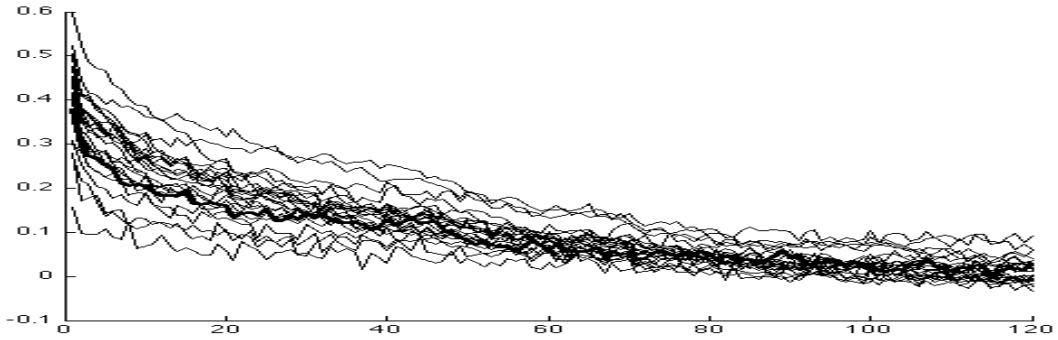
Figure 31: February 2010, Number of Transactions per Δ-Length Time Interval, Bundle of Sample Autocorrelations, Bold Profile is WMT

(a) Δ: 10 seconds



Lag

(b) Δ: 1 minute



Lag

3.2. The Model

Exploiting the framework proposed by Chen et al. (2013), I want to study the number of event occurrances in equally sized half open half closed intervals in time, $(t, t+\Delta]$, of length Δ. Consider the sequence of nonnegative, integer valued random variables $\{N_t\}_{t\in\mathbb{N}}$ with $N_t = N\left((t-1)\Delta, t\Delta\right]$ being the number of such occurrances. Let N_t be conditionally independent and Poisson distributed r.v. with intensity $\lambda_t\Delta$:

$$N_t \,|\lambda_t \sim \text{Poisson}\left(\lambda_t\Delta\right) \tag{3.1}$$

The Poisson intensity, λ_t, is itself a stationary time series governed by a markov-switching multi-fractal process, multiplicatively composed by independent markov-switching chains $M_{k,t}$, for $k \in \left\{1, 2, \ldots, \bar{k}\right\}$, in the spirit of Calvet and Fisher (2004):

$$\lambda_t = \lambda\left(\prod_{k=1}^{\bar{k}} M_{k,t}\right), \qquad \lambda > 0 \tag{3.2}$$

each markov-switching chain $\{M_{k,t}\}_{t\in\mathbb{N}}$ is charachterized by the following structure:

$$M_{k,t} = \begin{cases} M & \text{w.p} \quad \gamma_k \\ M_{k,t-1} & \text{w.p} \quad 1-\gamma_k \end{cases} \tag{3.3}$$

where M is a nonnegative and independent random variable with distribution function $f(M)$ such that $\mathbb{E}\left(M\right) = 1$. In this case we will take M to follow a discrete distribution on support $\mathcal{M} = \{m_0, 2-m_0\}$ with $m_0 \in (0, 2]$ and $f(m) = 1/2$ for all $m \in \mathcal{M}$. Therefore, each $M_{k,t}$ is an independent, two state markov-switching process with transition matrix given by:

$$\mathbb{P}\text{r}\left(M_{k,t}\,|M_{k,t-1}\right) = \mathcal{P}_k = \begin{bmatrix} 1-\gamma_k/2 & \gamma_k/2 \\ \gamma_k/2 & 1-\gamma_k/2 \end{bmatrix} \tag{3.4}$$

and states $\{s_1, s_2\} = \{m_0, 2-m_0\}$. As we can note, the largest eigenvalue of the transition matrix $\mathcal{P}_k$ is equal to one. The implied stationary distribution is independent of k since

$\mathbb{P}\mathrm{r}\,(M = m) = 1/2$ for all $m \in \mathcal{M}$. The second eigenvalue is instead given by $1 - \gamma_k$ and governs the persistence of the chain itself. In what follows, for the sake of parsimony, we impose the following restriction on the sequence of γ_k, $k \in \left\{1, 2, \ldots, \bar{k}\right\}$ that:

$$\gamma_k = 1 - (1 - \gamma^*)^{b^{k-\bar{k}}} \tag{3.5}$$

with $b \in (1, \infty)$ and $\gamma^* \in (0, 1)$. In this specification, as $k \uparrow \bar{k}$ we see that $1 - \gamma_k \downarrow 1 - \gamma^*$. Moreover, as $\bar{k} \uparrow \infty$, $1 - \gamma_k$ ranges in the interval $[1, 1 - \gamma^*]$ producing an infinite variety of frequencies at which the different chains are evolving over time. Also, note that γ^* bounds the highest (less persistent) frequency at which the multifractal is operational.

Finally, note that the MSMCP Model is charachterized by a $\bar{k}$-dimensional hidden-state vector, $M_t = \left(M_{1,t}, \ldots, M_{\bar{k},t}\right)$, in which every coordinate can only take on two values. This amounts to a total of $2^{\bar{k}}$ states for the intensity process. Furthermore, we can then collect the remaining parameters in the vector $\theta = (\lambda, m_0, b, \gamma^*)$ which dimension is independent of $\bar{k}$. This nice separability candidates $\bar{k}$ to be viewed as model index rather than as a standard parameter.

3.2.1. Interpretation As a Contnuous Time Process

Consider the point process of arrival times for the event of interest $\{t_i\}_{i \in \mathbb{N}}$, where t_1 is normalized to 0, with associated intensity $\lambda(t, \omega)$. Now I'll split the real line in half open and half closed intervals of the following form: $((j-1)\Delta, j\Delta]$ for any $j \in \mathbb{N}$. Let's now define the number of such events taking place in the jth interval as:

$$N_j = \sum_{i \geq 1} \mathbb{1}_{\{(j-1)\Delta < t_i \leq j\Delta\}}$$

In general, I will refer to $N(t)$ as the number of events occurring in the interval $(0, t]$. Further, assume that:

a) The intesity process is such that $\lambda(t, \omega) = \lambda_j(\omega)$ for any $(j-1)\Delta < t \leq j\Delta$, where $\lambda_j(\omega)$

is defined by equations (2), (3), (4) and (5);

b) The distribution of $N(t)$ conditional on λ_j has the following properties:

i. For any $(j-1)\Delta < t \leq j\Delta$:

$$\Pr\left(N(t+dt) - N(t) = 1 \,|\lambda_j\right) = \lambda_j dt + o(dt)$$

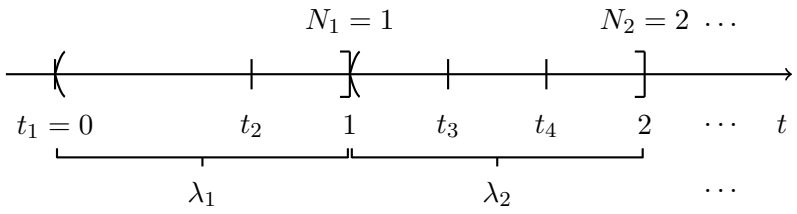
ii. For any $(j-1)\Delta < t \leq j\Delta$:

$$\Pr\left(N(t+dt) - N(t) \geq 2 \,|\lambda_j\right) = o(dt)$$

iii. Conditional on the relevant intensities, $N(t) - N(s)$ is independent of $N(t') - N(s')$ for any $t > s$, $t' > s'$ such that $(s, t]$ and $(s', t']$ don't overlap;

It is then straightforward to see that $N_j \,|\lambda_j$ are independent across js and distributed as Poisson with rate parameter $\lambda_j \Delta$, as in the model presented before.

Figure 32: Example of the Considered Process, $\Delta = 1$



The structure of the point process is explained in Figure 32 where Δ is normalized to be equal to one. As we can see the partition covers the entire real line except for the origin. The intensity is constant over each of the intervals $((j-1), j]$ at some value $\lambda_j(\omega)$ and, therefore the count over that interval is conditionally Poisson distributed with constant rate $\lambda_j(\omega)$.

3.2.2. Properties of The Model

In the following subsection I will sketch the properties of the MSMCP model, trying to draw a parallel with the main charachteristics displayed by the data.

Stationarity, Ergodicity and Finitness of Moments

Provided that $\gamma_k > 0$, the transition matrices for the markov chains imply that the processes $\{M_{k,t}\}_{k\in\{1,\dots,\bar{k}\}}$ are strictly stationary and ergodic. Moreover, their independence implies then that the vector process $\xi_t = \left(M_{1,t}, \dots, M_{\bar{k},t}\right)$ is strictly stationary and ergodic as well. Being λ_t a mesurable function of ξ_t, it is straigthforward to recognize that also λ_t must be a strictly stationary and ergodic process. Let's now turn our attention to the data count series N_t.

a) *Strict Stationarity:* pick a realization for the r.v. $N = (N_{t_1} = n_1, \dots, N_{t_r} = n_r)$ for any vector $(t_1, \dots, t_r) \in \mathbb{N}^r$ such that $t_1 < \cdots < t_r$. Let λ be analogously defined and let L be the lag operator so that $L^{-\tau}N = (N_{t_1+\tau} = n_1, \dots, N_{t_r+\tau} = n_r)$ for any $\tau \in \mathbb{N}$. Also let denote the support of the intensity process by $\mathcal{L} = \text{Supp}\,(\lambda)$[10]. By the Law of Total Probability we can then write:

$$\begin{aligned}
\mathbb{P}\text{r}\,(N) &= \sum_{\mathcal{L}} \left(\prod_{i=1}^{r} \mathbb{P}\text{r}\,(N_{t_i} \,|\lambda_{t_i}\,) \right) \mathbb{P}\text{r}\,(\lambda) \\
&= \sum_{L^{-\tau}\mathcal{L}} \left(\prod_{i=1}^{r} \mathbb{P}\text{r}\left(L^{-\tau}N_{t_i} \,\middle|L^{-\tau}\lambda_{t_i}\,\right) \right) \mathbb{P}\text{r}\left(L^{-\tau}\lambda\right) \\
&= \mathbb{P}\text{r}\left(L^{-\tau}N\right)
\end{aligned}$$

since λ_t is strictly stationary, implying that $\mathbb{P}\text{r}\,(\lambda) = \mathbb{P}\text{r}\,(L^{-\tau}\lambda)$ and $\mathcal{L} = L^{-\tau}\mathcal{L} = \text{Supp}\,(L^{-\tau}\lambda)$, and $\mathbb{P}\text{r}\,(N_{t_i}\,|\lambda_{t_i}\,)$ only depends on the value that the intesity is taking on and not, in any fashion, on its time index.

[10]Note that, fixed $\bar{k}$ and r, $\mathcal{L}$ is a finite set composed of $2^{r\bar{k}}$ elements.

b) *Ergodicity:* pick two realization of the r.v. $N = (N_t = n_1, \ldots, N_{t+k} = n_k)$ and $N' = (N_t = n'_1, \ldots, N_{t+k} = n'_k)$ for any $t, k \in \mathbb{N}$. Let L be the lag operator, then $L^{-m}N = (N_{t+m} = n_1, \ldots, N_{t+k+m} = n_k)$. Let define analogously λ and λ'. Again, by the Law of Total Probability we can write:

$$\begin{aligned}
&\Pr\left(N, L^{-m}N'\right) - \Pr\left(N\right)\Pr\left(L^{-m}N'\right) = \\
&= \sum_{\lambda\in\mathcal{L}}\sum_{\lambda'\in\mathcal{L}}\left(\prod_{s=t}^{t+k}\Pr\left(N_s\,|\lambda_s\right)\right)\left(\prod_{s'=t}^{t+k}\Pr\left(L^{-m}N'_{s'}\,\middle|L^{-m}\lambda'_{s'}\right)\right) \\
&\quad\times\left[\Pr\left(\lambda, L^{-m}\lambda'\right) - \Pr\left(\lambda\right)\Pr\left(L^{-m}\lambda'\right)\right] \\
&= \sum_{\lambda\in\mathcal{L}}\sum_{\lambda'\in\mathcal{L}}\left(\prod_{s=t}^{t+k}\Pr\left(N_s\,|\lambda_s\right)\right)\left(\prod_{s'=t}^{t+k}\Pr\left(N'_{s'}\,|\lambda'_{s'}\right)\right)\left[\Pr\left(\lambda, L^{-m}\lambda'\right) - \Pr\left(\lambda\right)\Pr\left(L^{-m}\lambda'\right)\right]
\end{aligned}$$

where $\mathcal{L}$ is defined as above as the support for the instensity vector λ and is a finite set. We can then directly write that:

$$\begin{aligned}
&\lim_{n\to+\infty}\frac{1}{n}\sum_{m=0}^{n}\left[\Pr\left(N, L^{-m}N'\right) - \Pr\left(N\right)\Pr\left(L^{-m}N'\right)\right] = \\
&= \sum_{\lambda\in\mathcal{L}}\sum_{\lambda'\in\mathcal{L}}\left(\prod_{s=t}^{t+k}\Pr\left(N_s\,|\lambda_s\right)\right)\left(\prod_{s'=t}^{t+k}\Pr\left(N'_{s'}\,|\lambda'_{s'}\right)\right) \\
&\quad\times\left\{\lim_{n\to+\infty}\frac{1}{n}\sum_{m=0}^{n}\left[\Pr\left(\lambda, L^{-m}\lambda'\right) - \Pr\left(\lambda\right)\Pr\left(L^{-m}\lambda'\right)\right]\right\} = 0
\end{aligned}$$

where the last equality holds since λ_t is ergodic and therefore:

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{m=0}^{n}\left[\Pr\left(\lambda, L^{-m}\lambda'\right) - \Pr\left(\lambda\right)\Pr\left(L^{-m}\lambda'\right)\right] = 0$$

c) Given the interpration of the MSMCP model as a Doubly Stochastic Poisson Process, the r^{th} moment of N_t exists if and only if the same moment for i ts directing measure, λ_t, exists [11]. Note that, for every $t \in \mathbb{N}$, λ_t can only take on a finite number of values bounded above, say, by $\bar{\lambda}$ and below by 0. Therefore all moments for λ_t exist, and hence

[11] This is a general result in the Theory of Point Processes. See, for example, Proposition 6.2.II of Daley and Vere-Jones (2008).

all the moments for N_t exist as well.

Overdispersion

This paragraph debates whether the MSMCP model shows overdispersion, that is, the property that $\mathbb{V}\text{ar}(N_t) > \mathbb{E}(N_t)$. Let's first derive expressions for both the quantities of interest:

$$\mathbb{E}(N_t) = \mathbb{E}[\mathbb{E}(N_t|\lambda_t)] = \mathbb{E}(\lambda_t) = \lambda\Delta\left(\prod_{k=1}^{\bar{k}}\mathbb{E}(M)\right) = \lambda\Delta = \Lambda$$

and

$$\begin{aligned}\mathbb{V}\text{ar}(N_t) &= \mathbb{E}\left[(N_t - \mathbb{E}(N_t))^2\right] = \mathbb{E}\left[(N_t - \Lambda)^2\right] = \mathbb{E}\left(N_t^2\right) - \Lambda^2 \\ &= \mathbb{E}\left[\mathbb{E}\left(N_t^2|\lambda_t\right)\right] - \Lambda^2 = \Lambda + \Lambda^2\mathbb{E}\left(M^2\right)^{\bar{k}} - \Lambda^2 \\ &= \Lambda\left(1 + \Lambda\left(\mathbb{E}\left(M^2\right)^{\bar{k}} - 1\right)\right)\end{aligned}$$

Therefore, the MSMCP model would display overdispersion if and only if $\mathbb{E}\left(M^2\right) > 1$, which is equivalent to the requirement that $\mathbb{V}\text{ar}(\lambda_t) > 0$. Overdispersion is, again, a standard result for Doubly Stochastic Poisson Processes as the MSMCP when $\mathbb{V}\text{ar}(\lambda_t) > 0$ (Daley and Vere-Jones, 2008). In the case that $\mathbb{V}\text{ar}(\lambda_t) = 0$, in fact, the model boils down to a simple Poisson Process with constant intensity Λ, hence generating an equidispersed unconditional distribution.

Long Memory

In the following paragraph, I want to show that the autocorrelation function generated by the MSMCP, $\rho(n) = \mathbb{C}\text{orr}(N_t, N_{t+n})$ displays an hyperbolic decay. This result is formally stated in the following proposition.

Theorem 3.2.1. *Suppose that the time series* $\{N_t\}_{t\in\mathbb{N}}$ *is characterized by equations (1)-(5). Pick* α_1, $\alpha_2 \in (0,1)$ *with* $\alpha_2 > \alpha_1$. *For every* $\bar{k} \in \mathbb{N}$, $\gamma^* \in (0,1)$, $b > 1$, $\lambda > 0$ *and*

$m_0 \in (0, 2]$ *define:*

$$\mathcal{I}_{\bar{k}} = \left\{ k \in \mathbb{N} : \alpha_1 \log_b \left(b^{\bar{k}} \right) \leq \log_b k \leq \alpha_2 \log_b \left(b^{\bar{k}} \right) \right\}$$

Then the autocorrelation function for the count series $\{N_t\}_{t \in \mathbb{N}}$ *satisfies:*

$$\sup_{n \in \mathcal{I}_{\bar{k}}} \left| \frac{\log \rho(n)}{\log n^{-\delta}} - 1 \right| \longrightarrow 0 \qquad as \qquad \bar{k} \longrightarrow +\infty$$

where $\delta = \log_b \mathbb{E}\left(M^2\right)$.

The proof for the proposition can be found in the Appendix.

3.2.3. Intra-Day Microstructure and Calendar Effects

In high-frequency data, intra-day calendar (seasonal) effects deriving from the microstructure of the considered market are quantitatively non-negligible and, hence, need to be accounted for in the most flexible and parsimonious way. The standard approach in this class of models, would consist of specifying the conditional mean function, $\mathbb{E}(N_t \,|\lambda_t)$, as the product of a deterministic and a stochastic component, respectively the seasonal effect at time t and the intensity process at the same time index. In particular, (1) would then become:

$$N_t \,|\lambda_t \sim \text{Poisson}\left(\exp\left(s_t\right) \cdot \lambda_t \Delta\right) \tag{3.6}$$

where s_t represents the calendar effect at time t and λ_t is specified as before. Note that I follow a different approach in treating the calendar effects as compared to Chen et al. (2013). In particular, my strategy is different in two aspects.

Firstly, the seasonal effects are estimated toghether with the rest of the parameters. This choice stems from the fact that, not only a two-step estimation approach providing a deseasonalized series would yield to less efficient estimates, but, also, it would drastically affect

the non-negativity and the integer-valuedness nature of the series. On the other hand, this choice makes ML estimation more complicated, increasing the dimension of the optimization problem by the number of parameters needed to model the seasonal function.

The second difference is given by the choice of the parametric family for the seasonal effect deterministic function. Following Ghysels et al. (2004), Chen et al. (2013) divide the trading day into half-hour windows displaying constant calendar effects in the form:

$$s_t = \sum_{i=1}^{12} \alpha_j x_{j,t}$$

where $x_{j,t}$, $j = 1, 2, \ldots, 12$, are dummy variables taking the value one when t is in the j-th daily window. As it is immediate to notice, the afore-mentioned specification requires the estimation of twelve coefficients and its flexibility directly depends on the choosen number of intra-day windows. An alternative and common specification for the calendar effect function is given by the flexible Fourier series approximation proposed by Andersen and Bollerslev (1998) and exploited, amongst others, by Hautsch (2012):

$$s_t = \delta \bar{t} + \sum_{i=1}^{Q} \delta_{i,c} \cos(\bar{t} \cdot 2\pi i) + \delta_{i,s} \sin(\bar{t} \cdot 2\pi i) \tag{3.7}$$

Note that $\bar{t} \in [0, 1]$ is the normalized calendar time computed as the number of time units from the opening of the exchange until the calendar time t and divided by the length of the trading day measured in the same time unit. Also note that the number of additional parameters is a function of the order of the approximation, Q, and therefore the latter specification is more parsimonious then the former as long as $Q \leq 5$.

In what follows I will then employ the model described by equations (6) for the conditional distribution of the measurment, (7) for the seasonal effects and (2)-(5) for the intensity process.

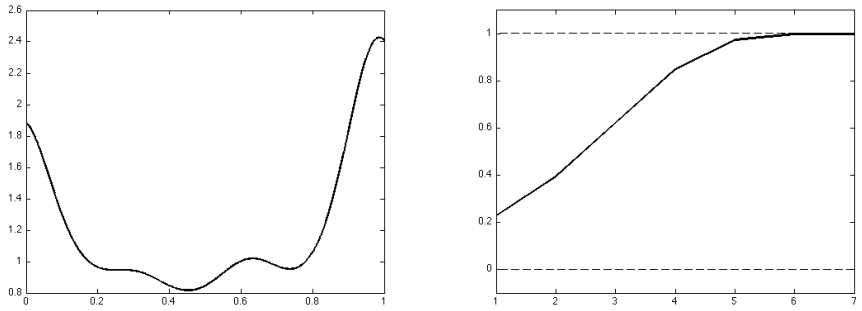
3.2.4. An Illustrative Simulation

For a preventive check of the model behavior against the stylized facts presented in Section 1, in what follows, I will show an illustrative simulation of the MSMCP model calibrated according to the latest estimates disposable for WMT. In particular, for the sake of brevity, I will only consider $\Delta = 10$ seconds. Further, $\bar{k}$ will be set equal to 7 and Q is choosen to be 3^{12}. The rest of the parameters are set as follows: $\lambda = 0.38$, $m_0 = 1.52$, $b = 1.94$ and $\gamma^* = 0.99$.

Figure 33 displays the estimated intra-day calendar effects for the stock WMT and the renewal probabilities for the markov chains, $\{\gamma_k\}_{k=1}^{\bar{k}}$. Intra-day calendar effects are non negligible and signal higher activity at the beginning and at the end of the trading day, which is a usual result in the literature[13]. The second panel of the figure instead, tells

Figure 33: WMT, Intra-Day-Calendar Effects and Estimated Renewal Probabilities

(a) $\exp(\widehat{s}_t)$ (b) $\widehat{\gamma}_k$



$\bar{t}$ k

us that that the markov chains are not very persistent. In fact, the most persistent chain ($k = 1$) has a rather high renewal probability of 0.22. Morevover, 4 out of 7 chains have renewal probability bigger than 0.90 and 3 of them higher than 0.99. Having this picture in mind, I'd rather expect very mild autocorrelations - though hyperbolically decaying - in the simulated sample.

[12] Again, in accordance with the "best" model selected through Likelihood comparisons.
[13] See Hautsch (2012) for example.

Figure 34: Simulated Sample, Number of Transactions per 10 Second Time Intervals, Calibrated on WMT

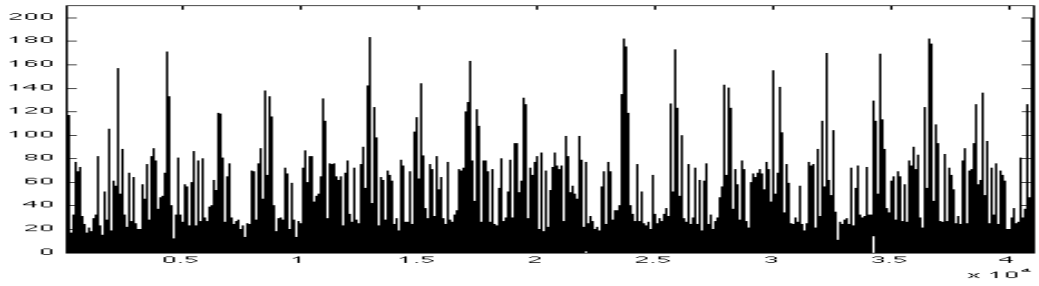


Figure 34 and Table 6 respectively plot the simulated time series and compare several descriptive statistics for the true and the simulated sample. The simulated time series plot seems to show a picture comparable with the true WMT counts sequence in terms of persistency, clustering and variation. This first visual impression is confirmed when looking at Table 6. As we can see the simulated sequence matches reasonably well the sample moments charachterizing the true WMT series. In particular, the model generates an

Table 6: February 2010, Descriptive Statistics, Comparing WMT with Simulated Sample, $\Delta = 10$ seconds, $\bar{k} = 7$, $Q = 3$

	Mean	Med	Max	Min	St Dev	Skew	Kurt	OD	Trades
WMT	5.04	1	203	0	9.45	4.43	40.66	17.72	206,818
Simulated Sample	4.61	1	200	0	10.37	6.33	64.88	23.30	189,355

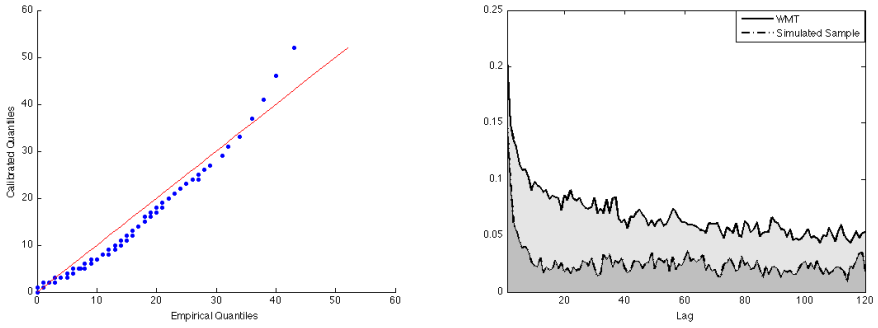
unconditional distribution with fairly robust overdispersion and very high kurtosis (actually, even too much).

Finally, Figure 35 is composed by two panels. The first one shows a quantile-quantile plot of the empirical distribution against the simulated ones. Compared to the static Poisson model of Figure 28, the improvement can be immediately noticed. However, we can see the excess kurtosis in the simulated distribution - already highlighted in Table6 - by noticing the right tail quantlies laying above the, red, 45 degree line. The second panel compares sample autocorrelations from the true and the simulated sequences. The model is able

to capture the slowly, hyperbolic, decay displayed by the true series autocorrelations as expected. it seems to fail though in matching the autocorrelations levels. In particular the autocorrelations from the simulated sample are consistently - and sizably - smaller then

Figure 35: WMT, Q-Q Plot Against Simulated Sample and Autocorrelations

(a) Q-Q Plot (b) Autocorrelations



those from the original series. An educated guessing would suggest that the model is unable to generate such fat tailed and overdispersed unconditional distributions without very high frequency movements in the intensity components, at least with $\bar{k} \leq 7$. The drastic mixing imposed by such high renewal probabilities induces then mild autocorrelations, as in this example. Worth to notice, though, not being able to fully capture the dynamic aspects of the series is certainly going to penalize the model forecasting ability.

3.3. Estimation: Methodology

The MSMCP Model can be easily estimated by Maximum Likelihood, employing an appropriate filtering technique. Denote the set of parameters to be estimated with $\theta = \{\lambda, m_0, b, \gamma^*\}$. Given $\bar{k}$, the Likelihood function can be immediately factorized as follows:

$$\mathcal{L}_{\bar{k}}\left(N_{1:T}\,;\,\theta\right) = \prod_{t=1}^{T} \mathbb{P}\mathrm{r}\left(N_t \left| N_{1:t-1}\,;\,\theta, \bar{k}\right.\right) \tag{3.8}$$

Further recall that the intensity process, being multiplicatively composed by $\bar{k}$ independent , first order, Markov switching chains with transition matrix $\mathcal{P}_k$, is itself a first order Markov switching chain with finite support, $\Lambda_{\bar{k}}$, composed by $2^{\bar{k}}$ elements and transition matrix $\mathcal{P}^{\lambda}_{\bar{k}}$. The new support and transition matrix can be easily calculated. In particular, let $A_1 = [m_0, 2-m_0]$ and $A_k = A_{k-1} \otimes A_1$ for any $1 < k \leq \bar{k}$, then $\Lambda_{\bar{k}}$ is just going to be composed by the elements in the vector $A_{\bar{k}}$. Similarly, let $P_1 = \mathcal{P}_1$ and $P_k = P_{k-1} \otimes \mathcal{P}_k$ for any $1 < k \leq \bar{k}$, then we will just have that $\mathcal{P}^{\lambda}_{\bar{k}} = P_{\bar{k}}$. I can then employ a simple filter in the spirit of Hamilton (1989) as described in Durbin and Koopman (2001) to evaluate the Likelihood at given parameter values, θ. For the sake of exposition, I drop the θ and $\bar{k}$ notation in what follows. For every period $t \geq 1$, suppose I know $\mathbb{P}\mathrm{r}\left(\lambda_{t-1} \left| N_{1:t-1}\right.\right)$. I can then iterate forward λ_{t-1} according to $\mathcal{P}^{\lambda}_{\bar{k}}$ and integrate out λ_{t-1} to get $\mathbb{P}\mathrm{r}\left(\lambda_t \left| N_{1:t-1}\right.\right)$, the predicted distribution for the intensity given information up to time $\Delta(t-1)$. I can then compute the joint distribution of the state and the measurement at time t, $\mathbb{P}\mathrm{r}\left(\lambda_t, N_t \left| N_{1:t-1}\right.\right)$, as $\mathbb{P}\mathrm{r}\left(N_t \left| \lambda_t\right.\right) \mathbb{P}\mathrm{r}\left(\lambda_t \left| N_{1:t-1}\right.\right)$, where $\mathbb{P}\mathrm{r}\left(N_t \left| \lambda_t\right.\right)$ is given by the p.m.f. of a Poisson distribution with rate λ_t. Finally, integrating out λ_t from the previously calculated joint distribution, one can obtain the period Likelihood, $\mathbb{P}\mathrm{r}\left(N_t \left| N_{1:t-1}\right.\right)$ and the updated ditribution for the latent state, $\mathbb{P}\mathrm{r}\left(\lambda_t \left| N_{1:t}\right.\right) = \mathbb{P}\mathrm{r}\left(\lambda_t, N_t \left| N_{1:t-1}\right.\right) / \mathbb{P}\mathrm{r}\left(N_t \left| N_{1:t-1}\right.\right)$. Recursively repeating the afore-described procedure for every $t = 1, \ldots, T$, we can finally evaluate the Likelihood function at point θ. The ML estimator for θ - $\widehat{\theta}$ - will be then found by numerically maximizing the Likelihood function.

As already discussed in the previous section, $\bar{k}$ will be used here as a model index rather than a parameter to be estimated. The notation MSMCP(κ) will be referring thus to a MSMCP model with $\bar{k} = \kappa$.

First recall that, as highlighted by Chen et al. (2013) and previously discussed, increasing $\bar{k}$ doesn't affect the dimensionality of the parameter vector, θ. Moreover, given any two models, MSMCP(j) and MSMCP(l) with $l > j$, they are nested only if either $b = 1$ and $\gamma^* = 1$, or $m_0 = 1$. Both cases are ruled out by the parameter constraints introduced in Section 2, implying that the models are non-nested. Therefore, model selection turns out to involve simple Likelihood comparison as adjusted criteria as BIC or AIC would produce exactly the same results.

Exactly as in Chen et al. (2013) and Calvet and Fisher (2004), the MSMCP model suffers from three potential sources of identification failure. Firstly, when $\gamma^* = 1$ or $\gamma^* = 0$, the sequence of renewal probabilities, $\{\gamma_k\}_{k=1}^{\bar{k}}$, becomes a sequence of constants regardless of b. In such a case then, the value of b would not affect the Likelihood function and hence the decay parameter couldn't be identified. Even if the exact cases are ruled ou by the assumption that $\gamma^* \in (0, 1)$, as the parameter value approaches its boundaries the decay parameter b becomes weakly identified. The second, and similar, case of identification failure is when the decay parameter b is equal to 1. Again, the sequence of renewal probabilities $\{\gamma_k\}_{k=1}^{\bar{k}}$ becomes a sequence of constants regardless of γ^* which can't, therefore, be identified. As 1 is explicitly ruled out from the parameter space, the same weak identification problem arises for γ^* when b approaches its lower bound. Finally, in the unfortunate case that $m_0 = 1$, the intensity components become identical across states[14] and therefore both b and γ^* can't be identified.

The numerical implementation of the estimation procedure is performed mixing `Matlab` and `Fortran 90` codes. In particular, the filtering step relies on the faster language, `Fortran 90`, while the less computationally intensive part of the algorithm is written in `Matlab`.

[14]Recall that the case in which $m_0 = 1$ implies that $\mathbb{V}\text{ar}(\lambda_t) = 0$.

As the optimization algorithm, I choose the latest available version of Hansen's CMA-ES (Hansen, 2006), which source code is available at `https://www.lri.fr/~hansen/cmaes_inmatlab.html`.

In what follows I will first present the estimates for the "best" model specification[15] for $\Delta = 10$ seconds and $\Delta = 1$ minute, commenting on the bundle distribution for each parameter and discussing the "best" model specification selection procedure. Secondly, I will compare the MSMCP model with the benchmark ACP$(1, 1)$ model of Heinen (2003) both in terms of in sample of goodness of fit measures and in terms of (pseudo) out of sample forecasting accuracy.

[15] In terms of selection of $\bar{k}$ and Q.

3.4. Estimation: Results

In this paragraph I will first discuss the selection of $\bar{k}$ and Q for the considered samples, then I will debate the parameter estimates over the bundle of equities for $\Delta = 10$ seconds and $\Delta = 1$ minute, trying to highlight the most relevant insights provided by the data. Detailed estimation results for every and each equity in the bundle for a variety of values of $\bar{k}$ can be found in the Appendix.

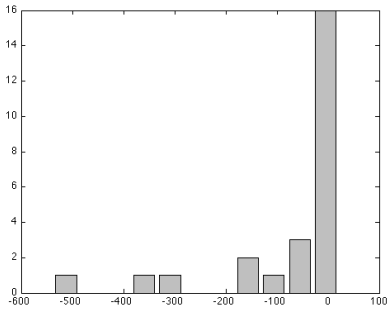
Figure 36, shows BIC[16] differential values for the MSMCP(7) with $Q = 3$ and $Q = 4$ for all equities in the bundle for the two different samples. The differentials are expressed as $BIC_4 - BIC_3$ so that positive values indicate that $Q = 3$ is preferred to $Q = 4$. As we can immediately note, for most equities and in both samples, the BIC differentials are very close to zero, but still somehow laying on the negative side, so that in the data a slight preference towards $Q = 4$ seems to be suggested. However, for the sake of parsimony, and given that no big improvements in the Likelihoods are registered from increasing the Fourier Form order to 4, I am going to focus, in what follows, on the case where $Q = 3$.

Figure 37 displays log-Likelihood mode differentials for the MSMCP model, given $Q = 3$, as a function of $\bar{k}$ for $\Delta = 10$ seconds. The reference value is given by the MSMCP(7) log-Likelihood mode. In particular, the differentials are framed in the form of $-\log \mathcal{L}_7 + \log \mathcal{L}_{\bar{k}}$, and negative (positive) values imply preference towards the MSMCP(7) (MSMCP($\bar{k}$)) specification of the model. As we can see from the picture, $\bar{k} = 7$ clearly dominates all other specifications and, in some cases, the Likelihood adjustments when increasing $\bar{k}$ are sizable, suggesting that further improvements can be achieved by increasing the number of markov chains above 7. However, as in Chen et al. (2013), I choose to set $\bar{k} = 7$ both for comparative and computational reasons. A completely similar evolution is found to be happening when focusing on the sample with $\Delta = 1$ minute. Also, looking at the tables in the Appendix can be registered as the estimates tend to remain stable independently of $\bar{k}$ for most of

[16] The Bayesian Information Criterion is computed according to the following formula: $BIC = -2 \log \mathcal{L} + k \log(n)$, where k indicates the number of estimated parameters and n the number of observations.

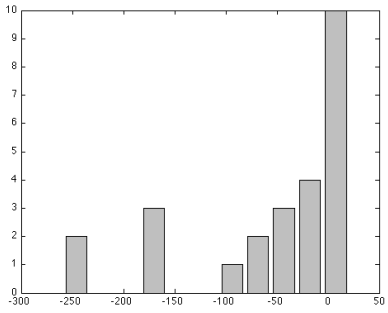
Figure 36: BIC Differences for $Q = 3$ and $Q = 4$, Bundle Distribution, $\bar{k} = 7$

(a) $\Delta = 10$ seconds



$\text{BIC}_4 - \text{BIC}_3$

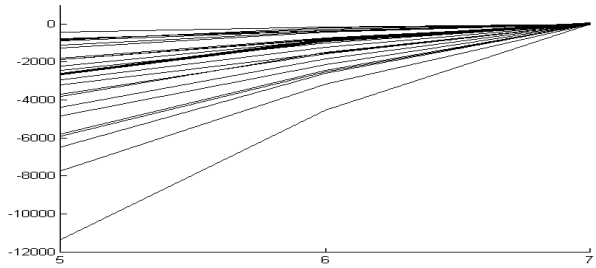
(b) $\Delta = 1$ minute



$\text{BIC}_4 - \text{BIC}_3$

the stocks. Therefore, in the subsequent exposition, I'm going to analyze the parameter estimates for the MSMCP model with $\bar{k} = 7$ and $Q = 3$, as previously justified.

Figure 37: Log-Likelihood Differentials, Profile Bundle, $\Delta = 10$ seconds, Bold is WMT



$\bar{k}$

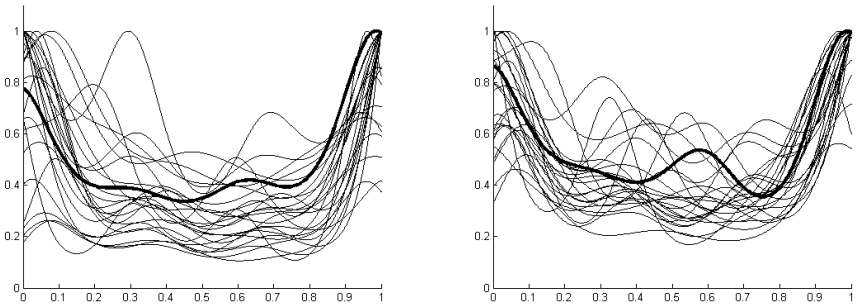
Figure 38 displays profile bundles for the estimated intra-day calendar effects - $\exp(\hat{s}_t)$ - over the standardized day time, $\bar{t}$, for $\Delta = 10$ seconds and $\Delta = 1$ minute. the calendar effects have been normalized to one at $\bar{t} = 1$ for comparison reasons. As expected[17], for almost all of the stocks, the activity peaks are registered at the beginning and at the end of the trading day, while transaction levels tend to be reduced in the middle of the day. The above dynamic is striking and consistent across equities, in both the considered samples, as shown by both panels.

Let's now turn our attention towards cross-equity ditributions of the estimated parameters,

[17]Cfr., for example, Hautsch (2012).

Figure 38: Estimated Intra-Day Calendar Effects, $\exp(\hat{s}_t)$, Profile Bundles, Bold is WMT

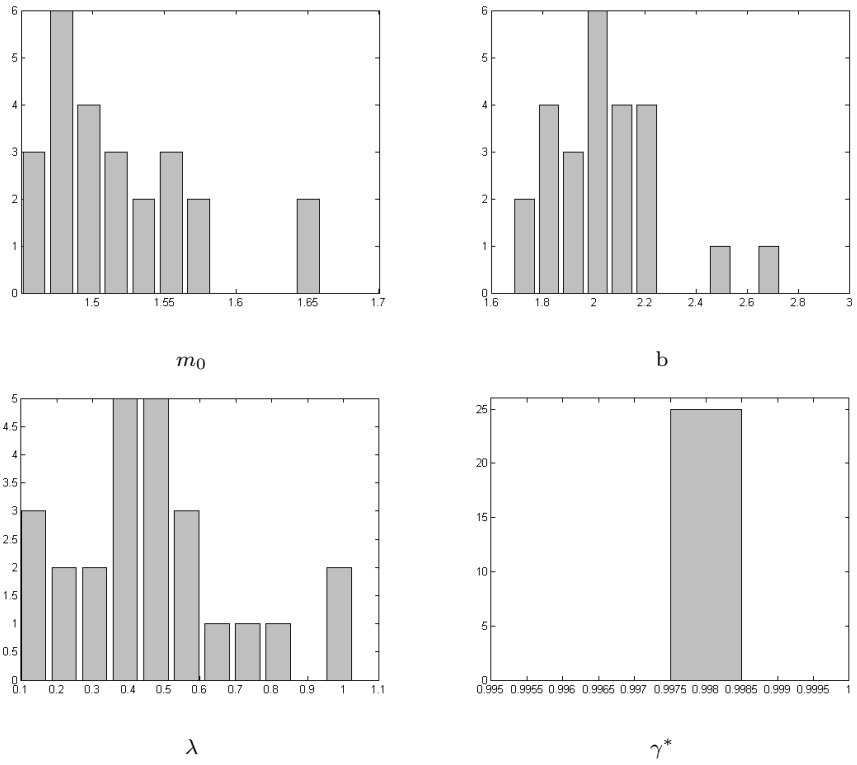
(a) $\Delta = 10$ seconds (b) $\Delta = 1$ minute



$\bar{t}$ $\bar{t}$

as shown in Figure 39 and Figure 40. The distribution of m_0 estimates tend to be slightly skewed to the right in both samples. In particular, for $\Delta = 10$ seconds, most of the estimates lie around roughly 1.5 but some are as high as 1.65. Instead, for $\Delta = 10$ minute, the estimates are centered around 1.3, with some climbing up to almost 1.5. The distribution for λ is not as skewed to the right as the one for m_0, at least for $\Delta = 10$ seconds. For most of the equities, in fact the estimated value for λ tend to be close to 0.5, and the right and left tails look very similar, almost simmetrically spreading the estimates between 0 and 1 around 0.5. For $\Delta = 1$ minute, instead, the λ estimates distribution looks more spread and skewed to the right, with values that, for some stocks, become as high as 1.3. The distribution for the estimates of the decay parameter, b, are, again, right skewed and centered around 2 for $\Delta = 10$ seconds. For $\Delta = 1$ minute instead, the estimated values seem to distribute much more evenly between 1.7 and 2.8. As compared to the results presented in Chen et al. (2013), we can note how the distribution is much less spread and concentrated towards much smaller values. Finally, the estimates for γ^* are very close to one for all stocks in both samples. This is a striking difference with respect to Chen et al. (2013), where the authors were able to divide the bundle of firms into two groups, one displaying high renewal probabilities and one displaying low renewal probabilities. The evidence coming from 2010 data suggests then such division is no longer taking place in the bundle of firms considered, at least, given the model specification.

Figure 39: Estimated Parameters, Histograms Over the Bundle, $\Delta = 10$ seconds



Being the latter charachteristic the most striking difference amongst the estimation results, as compared to the analysis in Chen et al. (2013), I display the profile bundle of estimated renewal probabilities for both samples, $\Delta = 10$ seconds and $\Delta = 1$ minute, in Figure 41[18]. As we can easily note, all the stocks, in both samples, display very high renewal probabilities for most of the components. In particular, for at least half of the considered equities, and in both samples, the lowest renewal probability is as high as 0.2. The low frequency components that are supposed to build up the long memory dependence in the model are somehow high frequency as well. This fact could explain the mismatch of the autocorrelation function registered in Section 2.4 when studying the simulated sample calibrated on WMT for $\Delta = 10$ seconds. Moreover, though, this feature of the estimated parameters could possibly signal the low adequacy of the choice of $\bar{k}$ being set to be equal to 7. In particular, allowing for a larger number of intensity components, the high frequency

[18]Recall that the renewal probabilities, γ_k, are computed according to equation (5).

Figure 40: Estimated Parameters, Histograms Over the Bundle, $\Delta = 1$ minute

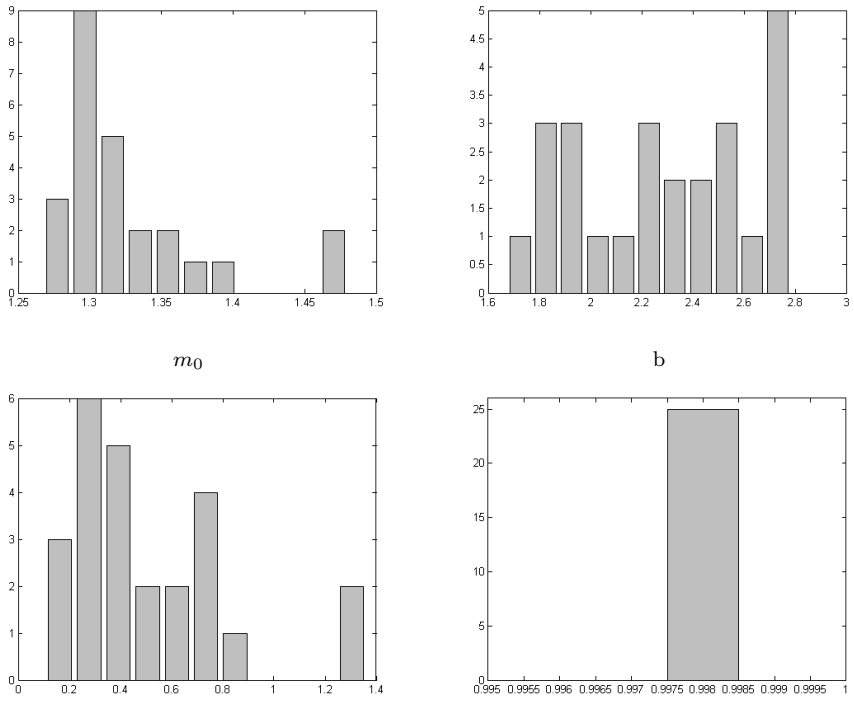
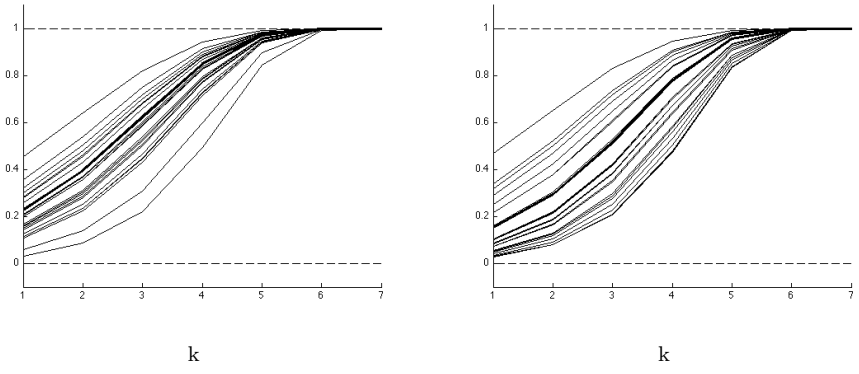


Figure 41: Estimated Renewal Probabilities, $\hat{\gamma}_k$, Profile Bundles, Bold is WMT

(a) $\Delta = 10$ seconds (b) $\Delta = 1$ minute



ones could still be retained while adding to the picture lower frequency dynamics able to match both the levels and the decay of the sample autocorrelations displayed by the equities in the bundle. If the number of components is, indeed, not sufficiently large to best capture the first order dynamics of the series, increasing $\bar{k}$ to values bigger than 7 could significantly

improve the forecasting performance of the model.

3.4.1. In-Sample Goodness of Fit Comparisons

As already discussed in at the beginning of the chapter, Heinen (2003) model is considered as the benchmark to study trade counts in financial markets. Also, it is appealing the fact that its structure is plainly mutuated from Engle and Russell (1998) ACD model, which is used as a benchmark for comparisons of the MSMCD model by Chen et al. (2013). Valuing simplicity, in what follows I will consider the basic version of the model, the ACP(1,1) specification. The model can be described by a measurement equation dictating a conditionally Poisson distribution for the trade count N_t:

$$N_t \,|\lambda_t \sim \text{Poisson}\left(\exp(s_t)\lambda_t\right)$$

and a transition equation for the latent intensity in the spirit of GARCH/ARCH modelling:

$$\lambda_t = \omega + \alpha \frac{N_{t-1}}{\exp(s_{t-1})} + \beta\lambda_{t-1} \tag{3.9}$$

Let collect all parameters to be estimated in $\theta_{ACP} = (\omega, \alpha, \beta, \lambda_0)$. Note that, provided λ_0, conditioning on the past observable information at time t, $\{N_s\}_{s=1}^{t-1}$, is equivalent to conditioning on the current intensity, λ_t. This fact can be ascertained by looking at the transition equation for the latent state, equation (9). For fixed parameter values, thus, the log-Likelihood function for the ACP(1,1) model is available in closed form and is given by:

$$\log \mathcal{L}\left(N_{1:T}\,;\,\theta_{ACP}\right) = \sum_{t=1}^{T} \left\{-\log(N_t!) + N_t \log(\exp(s_t)\lambda_t) - \exp(s_t)\lambda_t\right\}$$

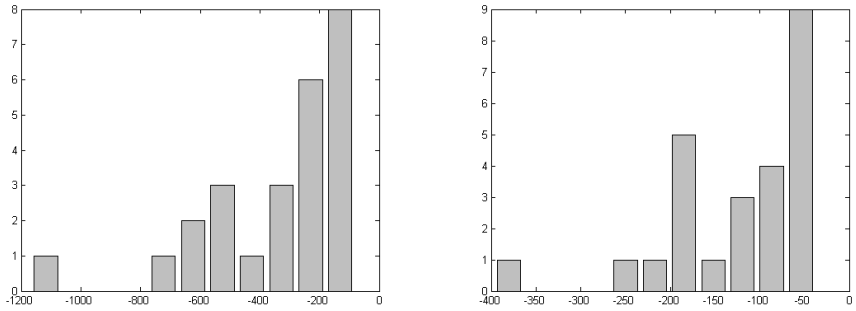
where, again, λ_t is computed recursively starting from λ_0, according to equation (9).

Figure 42 displays the distribution in BIC differences (expressed as $BIC_{MSMCP} - BIC_{ACP}$) for the bundle in both samples, $\Delta = 10$ seconds and $\Delta = 1$ minute. As can be easily seen, the MSMCP(7) dominates the ACP(1,1) for all equities and both samples, with differences

at least as negative as -100 for $\Delta = 10$ seconds and as -50 for $\Delta = 1$ minute. This

Figure 42: BIC Differences Between MSMCP(7) and ACP(1,1), Bundle Distribution, $Q = 3$

(a) $\Delta = 10$ seconds (b) $\Delta = 1$ minute



$\text{BIC}_{MSMCP} - \text{BIC}_{ACP}$ $\text{BIC}_{MSMCP} - \text{BIC}_{ACP}$

Note: BIC differentials are expressed interms of thousands of log-Likelihood points.

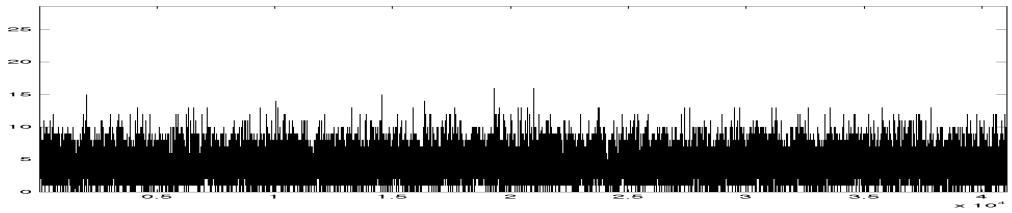
overwhelming evidence in favor of the MSMCP(7) model is supported when looking at the descriptive statistics from a simulated sample calibrated on WMT estimates for the ACP(1,1). These statistics are reported in Table 7 and the respective simulated sequence is displayed in Figure 43. As it can easily be seen, the ACP model produces sequences of counts which seem very different from the true series, at least from a first, visual, impression. Turning to Table 7, the model fails in matching almost all of the reported statistics, except for the mean. It seems, then, that the MSMCP model is able to generate much more skewed, fat tailed and overdispersed unconditional distributions then the ACP model and that this is the case in the considered samples. Lastly, consider the sample autocorrelations

Table 7: Descriptive Statistics, Comparing WMT with Simulated Sample, ACP(1,1), $\Delta =$ 10 seconds, $Q = 3$

	Mean	Med	Max	Min	St Dev	Skew	Kurt	OD	Trades
WMT	5.04	1	203	0	9.45	4.43	40.66	17.72	206,818
Simulated Sample	4.18	4	26	0	2.12	0.61	3.95	1.08	171,612

produced by the simulated ACP(1,1) model. They are shown in Figure 44 toghether with the sample autocorrelations for the WMT series, for $\Delta = 10$ seconds. As we can see, the

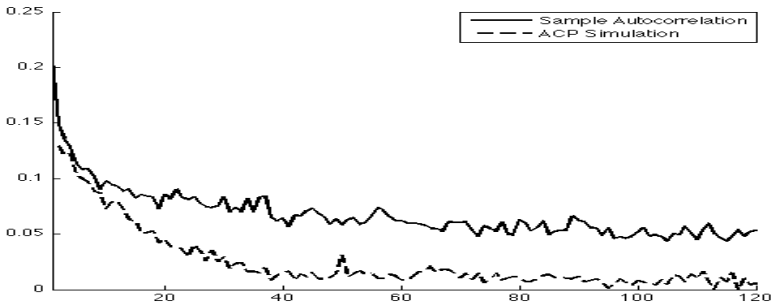
Figure 43: Simulated Sample, Number of Transactions per 10 Second Time Intervals, Calibrated on WMT



Observation

ACP model doesn't display that hyperbolic decay that is proper of the real series behaviour and which is produced by the MSMCP model. On the other hand though, at least for the very first autocorrelations the ACP model is much more able to match the true sample autocorrelations. Taking this fact into account, it could be the case that the forecasting performance of the ACP model would be superior to the one of the MSMCP model, at least for short horizons.

Figure 44: Simulated Sample, Autocorrelations Comparisons, $\Delta = 10$ seconds



Lag

3.4.2. Out-of-Sample Forecasting Accuracy Comparisons

In what follows I will perform a simple pseudo-out-of-sample forecasting experiment and directly compute RMSEs for several steps ahead forecasts. Given non-negligible computing times, I won't re-estimate the models for every forecasting origin. Instead, I use the parameter estimates delivered by the numerical optimization of the Likelihood on the full sample for each and every origin. The two samples have, obviously, different numerosity.

In particular, for $\Delta = 10$ seconds, I dispose of 41,040 observations while for $\Delta = 1$ minute I only have 6,840 of them. I then start with the first forecasting origin at observation 10,000 for $\Delta = 10$ seconds and 1,000 for $\Delta = 1$ minute and repeat the experiment untill the last usable observation of the considered sample. Also, for each forecasting origin I compute h-steps ahead forecasts with $h = 1, \ldots, 100$. The forecasting procedure plainly follows the one in Chen et al. (2013) and hence is not described here.

Figure 45: RMSE Comparisons, Histograms Over the Bundle, $\Delta = 10$ seconds

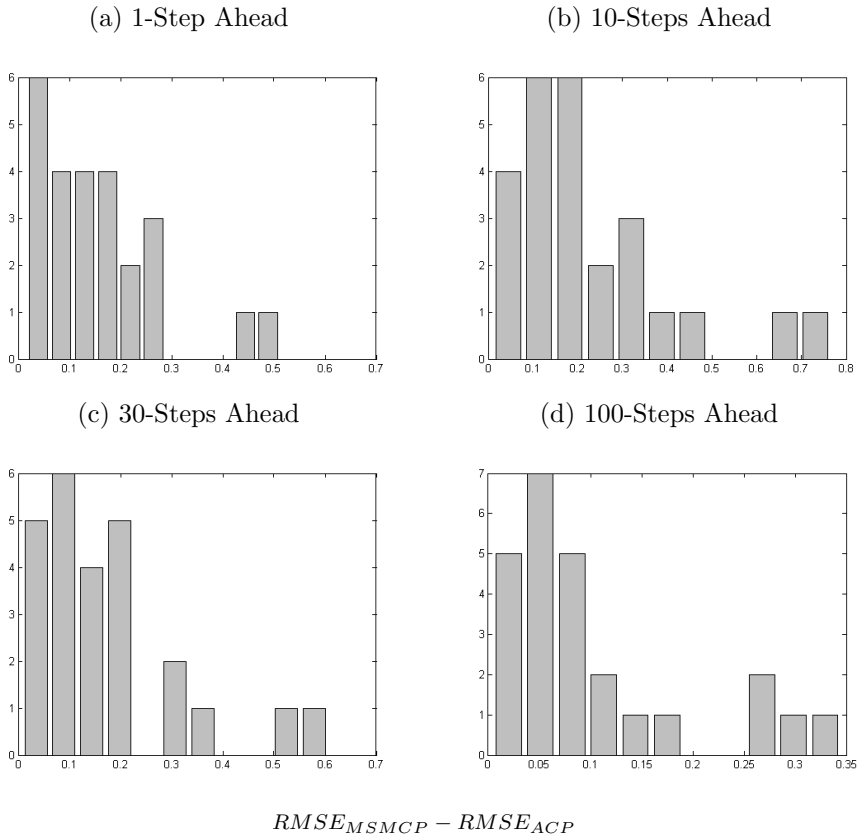
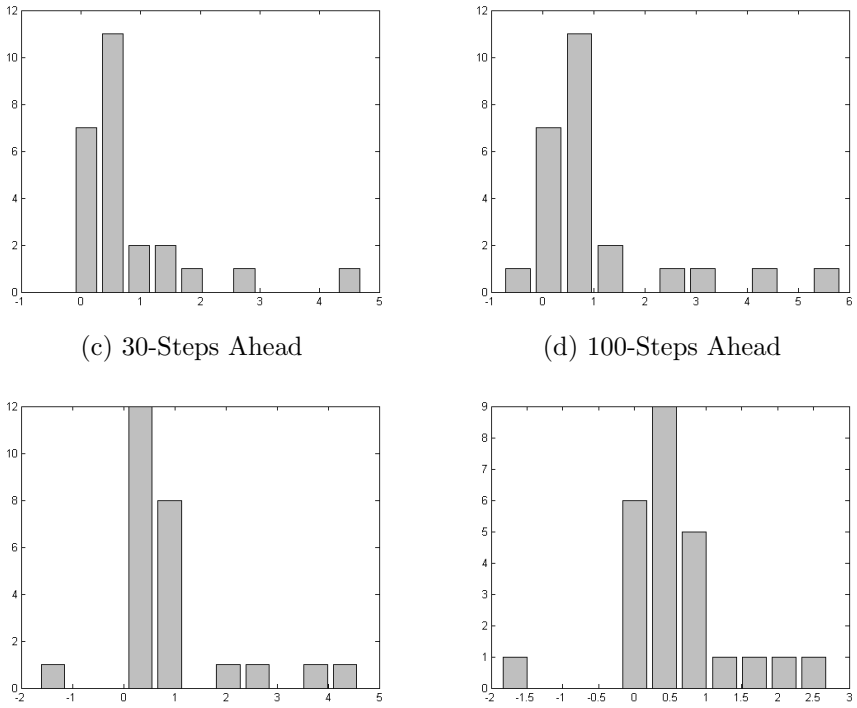


Figure 45 and Figure 46 respectively display the RMSE differential distributions for several steps ahead forecasts in the two samples. The differences are computed as $RMSE_{MSMCP} - RMSE_{ACP}$, so that negative values indicate that the MSMCP model outperforms the ACP model forecasting ability as measured by the RMSE. It is immediate to note that the differentials are very close to zero or slightly positive for both samples, suggesting that the ACP(1,1) is doing slightly better than the MSMCP in terms of forecasting ability.

Even if this result seems in contrast with the overwhelming evidence about the in-sample goodness of fit superiority of the MSMCP model, I already discussed how the MSMCP model fails in delivering a close approximation of the autocorrelation structure of the considered

Figure 46: RMSE Comparisons, Histograms Over the Bundle, $\Delta = 1$ minute

(a) 1-Step Ahead (b) 10-Steps Ahead



(c) 30-Steps Ahead (d) 100-Steps Ahead

$RMSE_{MSMCP} - RMSE_{ACP}$

series. Therefore, while being very effective in matching the moments of the unconditional distribution, yielding a very good in-sample fit, the MSMCP model with $\bar{k} = 7$ seems unable to capture the dynamic structure proper of the considered series and therefore poorly performs in terms of forecasting accuracy as compared to the ACP(1,1) model. Hopefully, increasing $\bar{k}$ to values greater than 7 would improve the forecasting capabilities of the model in the sample, adding additional low frequency components injecting more autocorrelation into the model and exploiting a wider range of operating frequencies.

3.5. Concluding Remarks

In this paper I construct a model for time series of counts and I use it to study the number of transaction in fixed time intervals for a selected bundle of US equities traded on NASDAQ OMX during February 2010.

The proposed model is able to outperform the benchmark ACP(1,1) model by Heinen (2003) in terms of in-sample fit but fails to improve on its forecasting precision. This seemingly puzzling result can be explained by the poor performance of the MSMCP model in matching the autocorrelation levels displayed by the data[19], which in turn can be explained by the inadequate number of intensity components exploited in the analysis, and dampening the autocorrelation structure.

However, given the nature of the model, further increases in the number of intensity components in order to improve its forecasting capabilities will translate into longer computing times and higher algorithm complexity. In this sense, there is a non-trivial trade off when coming to the choice of which model to use to study transaction count data and this paper doesn't provide a definitive answer to such question.

[19]Recall, in fact, that the MSMCP model is able to capture the slow decay in the autocorrelation function.

APPENDIX

A.1. Descriptive Statistics for Chapter 2

Table 8: 1993, Descriptive Statistics, Full Sample

Ticker	Company Name	Mean	Med.	Max.	St. Dev.	Skew.	Kurt.	C.V.	Trades	Zeros %
AA	ALCOA	135.7	60	2068	202.6	3.35	19.50	1.49	2,975	1.1
ABT	Abbott Labs	23.99	13	409	29.17	2.99	17.55	1.27	17,814	5.8
AXP	Am. Express	37.46	19	599	52.02	3.15	18.46	1.39	10,915	4.5
BA	Boeing	22.81	13	360	27.31	2.59	13.71	1.20	17,956	6.2
BAC	B. of America	51.21	29	811	65.43	3.10	18.14	1.28	7,980	1.6
C	Citigroup	16.75	8	339	25.04	3.82	25.22	1.49	24,474	7.9
CSCO	Cisco Systems	23.06	10	627	39.34	4.66	37.42	1.71	17,759	6.2
DELL	Dell	16.45	7	841	28.83	5.68	66.82	1.75	2,4915	8.3
DOW	Dow Chemical	58.69	32	1267	76.91	3.57	26.70	1.31	6,946	1.5
F	Ford Motor	25.18	11	714	38.59	3.81	27.63	1.53	16,253	5.9
GE	General Electric	27.39	14	388	35.94	2.83	14.99	1.31	14,941	2.2
HD	Home Depot	15.43	7	289	22.29	3.52	22.38	1.44	26,540	8.3
IBM	IBM	11.82	7	221	15.10	3.06	18.95	1.28	34,680	9.3
INTC	Intel	9.46	5	332	14.22	4.52	40.33	1.50	43,323	10.3
JNJ	Johnson & J.	16.19	9	355	20.28	3.35	23.60	1.25	25,311	5.3
KO	Coca-Cola	25.73	15	422	31.14	2.81	15.82	1.21	15,903	3.8
MCD	McDonald's	54.79	31	740	67.37	2.84	15.51	1.23	7,458	2.0
MRK	Merck	6.43	4	129	8.41	3.09	18.90	1.31	63,760	15.9
MSFT	Microsoft	13.72	6	438	21.81	4.51	38.44	1.59	29,855	9.5
QCOM	Qualcomm	94.67	35	2,171	164.88	4.21	30.59	1.74	4,271	4.2
T	AT&T	18.89	10	260	23.13	2.76	14.51	1.22	21,689	3.7
TXN	Texas Inst.	96.20	41	1,98	142.3	3.33	21.86	1.48	4,233	1.6
WFC	Wells Fargo	102.29	44	3,459	167.3	5.20	60.64	1.63	3,962	0.7
WMT	Wal-Mart	11.16	6	212	14.78	3.11	18.63	1.32	36,744	9.5
XRX	Xerox	138.91	64	1,845	194.3	2.95	15.77	1.40	2,909	0.5

Table 9: 1998, Descriptive Statistics, Full Sample

Ticker	Company Name	Mean	Med.	Max.	St. Dev.	Skew.	Kurt.	C.V.	Trades	Zeros %
AA	ALCOA	56.76	30	925	75.77	3.05	17.04	1.34	7,202	1.2
ABT	Abbott Labs	21.08	13	330	25.25	2.81	16.34	1.20	19,443	3.9
AXP	Am. Express	32.87	19	582	41.25	3.01	17.63	1.25	12,454	2.3
BA	Boeing	6.75	4	114	7.63	2.42	12.90	1.13	60,791	13.1
BAC	B. of America	26.15	15	563	33.06	3.13	19.96	1.26	15,673	1.7
C	Citigroup	19.67	11	298	25.21	2.88	15.94	1.28	20,830	7.1
CSCO	Cisco Systems	2.95	1	138	4.66	4.30	37.87	1.58	138,842	27.3
DELL	Dell	1.98	1	123	3.68	6.64	89.27	1.86	207,586	35.6
DOW	Dow Chemical	53.30	28	742	68.53	2.86	15.21	1.28	7,672	0.9
F	Ford Motor	20.00	11	312	25.69	2.76	14.79	1.28	20,480	5.4
GE	General Electric	8.77	5	164	10.15	2.70	15.02	1.16	46,745	7.7
HD	Home Depot	16.28	10	268	19.84	2.78	15.50	1.22	25,193	6.9
IBM	IBM	10.22	6	164	12.83	2.95	16.67	1.26	40,144	10.0
INTC	Intel	1.55	1	58	2.46	4.19	35.49	1.59	264,117	38.7
JNJ	Johnson & J.	14.07	8	299	17.08	3.00	19.47	1.21	29,147	5.9
KO	Coca-Cola	13.96	8	187	16.65	2.55	12.90	1.19	29,371	7.5
MCD	McDonald's	15.24	9	231	18.25	2.64	13.89	1.20	26,893	8.5
MRK	Merck	10.06	6	234	12.63	3.30	22.23	1.26	40,769	8.4
MSFT	Microsoft	1.73	1	82	3.08	5.06	49.91	1.78	237,284	39.2
QCOM	Qualcomm	7.87	2	347	14.99	4.52	36.79	1.90	52,128	22.1
T	AT&T	9.75	6	153	11.33	2.70	14.75	1.16	42,058	7.9
TXN	Texas Inst.	10.15	5	236	14.03	3.56	24.70	1.38	40,407	10.8
WFC	Wells Fargo	51.69	24	976	77.01	3.83	25.79	1.49	7,910	0.5
WMT	Wal-Mart	15.18	9	268	18.91	2.89	16.69	1.25	27,011	9.7
XRX	Xerox	38.42	20	752	49.78	3.012	19.00	1.29	10,663	1.9

Table 10: 2003, Descriptive Statistics, Full Sample

Ticker	Company Name	Mean	Med.	Max.	St. Dev.	Skew.	Kurt.	C.V.	Trades	Zeros %
AA	ALCOA	11.47	6	266	15.21	3.08	18.94	1.33	35,747	5.5
ABT	Abbott Labs	8.35	5	117	9.95	2.63	13.10	1.19	49,085	5.1
AXP	Am. Express	6.02	4	112	6.96	2.90	16.35	1.16	68,168	5.9
BA	Boeing	6.53	3	130	8.49	3.12	18.48	1.30	62,779	8.4
BAC	B. of America	4.97	3	152	6.45	3.39	24.21	1.30	82,519	10.8
C	Citigroup	2.90	2	68	3.58	3.00	18.70	1.23	141,672	20.0
CSCO	Cisco Systems	0.49	0	51	1.19	5.21	52.89	2.41	828,314	72.0
DELL	Dell	0.72	0	55	1.74	5.69	57.57	2.42	571,080	66.9
DOW	Dow Chemical	10.05	5	169	13.12	2.87	15.80	1.30	40,786	4.9
F	Ford Motor	6.38	4	116	7.59	2.76	15.43	1.19	64,308	10.8
GE	General Electric	1.86	1	36	2.22	2.53	14.28	1.19	220,055	28.3
HD	Home Depot	3.22	2	66	3.65	2.56	14.18	1.13	127,316	17.2
IBM	IBM	3.33	2	75	4.03	3.03	18.80	1.21	123,113	18.6
INTC	Intel	0.48	0	39	1.17	5.39	53.76	2.46	859,629	73.0
JNJ	Johnson & J.	4.92	3	93	5.74	2.91	17.55	1.17	83,343	11.2
KO	Coca-Cola	6.69	4	166	8.52	3.15	20.12	1.27	61,318	8.5
MCD	McDonald's	6.18	4	123	7.56	3.08	18.91	1.22	66,355	8.6
MRK	Merck	5.66	3	161	6.96	3.17	20.89	1.23	72,515	7.8
MSFT	Microsoft	0.32	0	29	0.77	4.93	48.78	2.40	1,277,142	77.0
QCOM	Qualcomm	0.81	0	68	1.99	6.73	84.83	2.45	505,437	63.4
T	AT&T	9.72	6	167	11.84	2.80	15.51	1.22	42,181	5.0
TXN	Texas Inst.	5.12	3	117	6.14	3.00	18.61	1.20	80,088	10.4
WFC	Wells Fargo	5.40	3	97	6.29	3.13	19.23	1.17	75,985	7.6
WMT	Wal-Mart	5.24	3	98	6.93	2.89	15.44	1.32	78,307	14.1
XRX	Xerox	13.45	8	256	16.08	2.59	13.84	1.19	30,487	4.1

Table 11: 2008, Descriptive Statistics, Full Sample

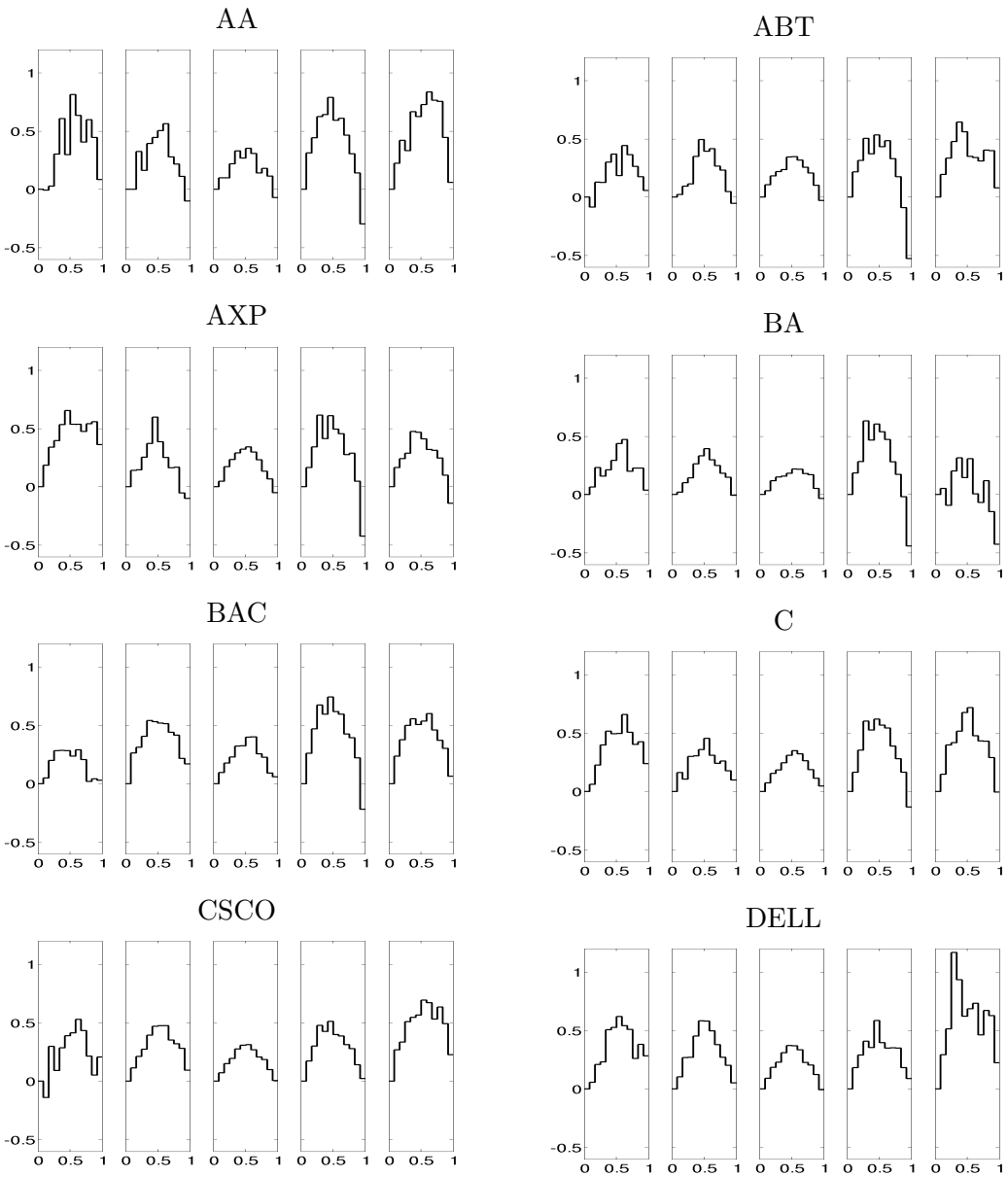
Ticker	Company Name	Mean	Med.	Max.	St. Dev.	Skew.	Kurt.	C.V.	Trades	Zeros %
AA	ALCOA	0.95	0.210	56.76	2.23	5.04	42.04	2.36	456,815	42.0
ABT	Abbott Labs	1.72	0.429	94.05	3.60	4.63	38.48	2.09	250,497	30.9
AXP	Am. Express	0.96	0.221	84.09	2.26	5.55	57.09	2.35	447,622	38.2
BA	Boeing	1.42	0.280	75.55	3.06	4.78	39.53	2.16	305,081	32.3
BAC	B. of America	0.37	0.000	28.59	0.93	5.86	59.80	2.51	1,161,278	54.3
C	Citigroup	0.24	0.000	25.94	0.70	6.68	76.98	2.91	1,800,842	65.6
CSCO	Cisco Systems	0.33	0.000	36.28	1.07	6.73	72.35	3.28	1,324,956	56.6
DELL	Dell	0.54	0.000	70.02	2.02	7.28	83.19	3.70	7.93	60.6
DOW	Dow Chemical	1.50	0.259	58.73	3.31	4.51	33.06	2.20	287,129	37.7
F	Ford Motor	1.06	0.000	81.85	2.94	5.62	52.98	2.76	405,362	57.6
GE	General Electric	0.44	0.000	27.64	1.13	5.31	45.86	2.58	986,023	55.8
HD	Home Depot	0.73	0.000	49.46	1.83	5.56	52.59	2.49	587,870	50.8
IBM	IBM	1.12	0.260	58.56	2.38	4.65	36.82	2.11	384,501	35.7
INTC	Intel	0.35	0.000	49.15	1.22	6.86	75.81	3.45	1,220,564	58.1
JNJ	Johnson & J.	1.02	0.219	50.64	2.27	4.58	34.58	2.23	423,358	41.4
KO	Coca-Cola	1.12	0.220	64.00	2.54	4.71	37.18	2.26	385,534	44.7
MCD	McDonald's	1.28	0.258	79.41	2.73	4.70	38.82	2.13	338,339	35.6
MRK	Merck	0.90	0.209	57.23	2.07	4.88	41.04	2.30	478,620	39.8
MSFT	Microsoft	0.24	0.000	38.40	0.83	7.40	90.57	3.44	1,799,810	59.2
QCOM	Qualcomm	0.63	0.014	64.12	1.92	6.39	67.40	3.04	685,436	47.5
T	AT&T	0.44	0.000	41.62	1.21	6.14	63.06	2.72	969,071	50.5
TXN	Texas Inst.	0.99	0.000	72.04	2.40	5.12	44.21	2.43	437,427	51.1
WFC	Wells Fargo	0.40	0.000	44.81	1.09	6.40	72.82	2.70	1,070,032	58.5
WMT	Wal-Mart	0.56	0.000	46.99	1.42	5.76	57.67	2.52	766,024	51.5
XRX	Xerox	2.24	0.263	128.4	5.12	4.61	35.86	2.29	193,127	38.3

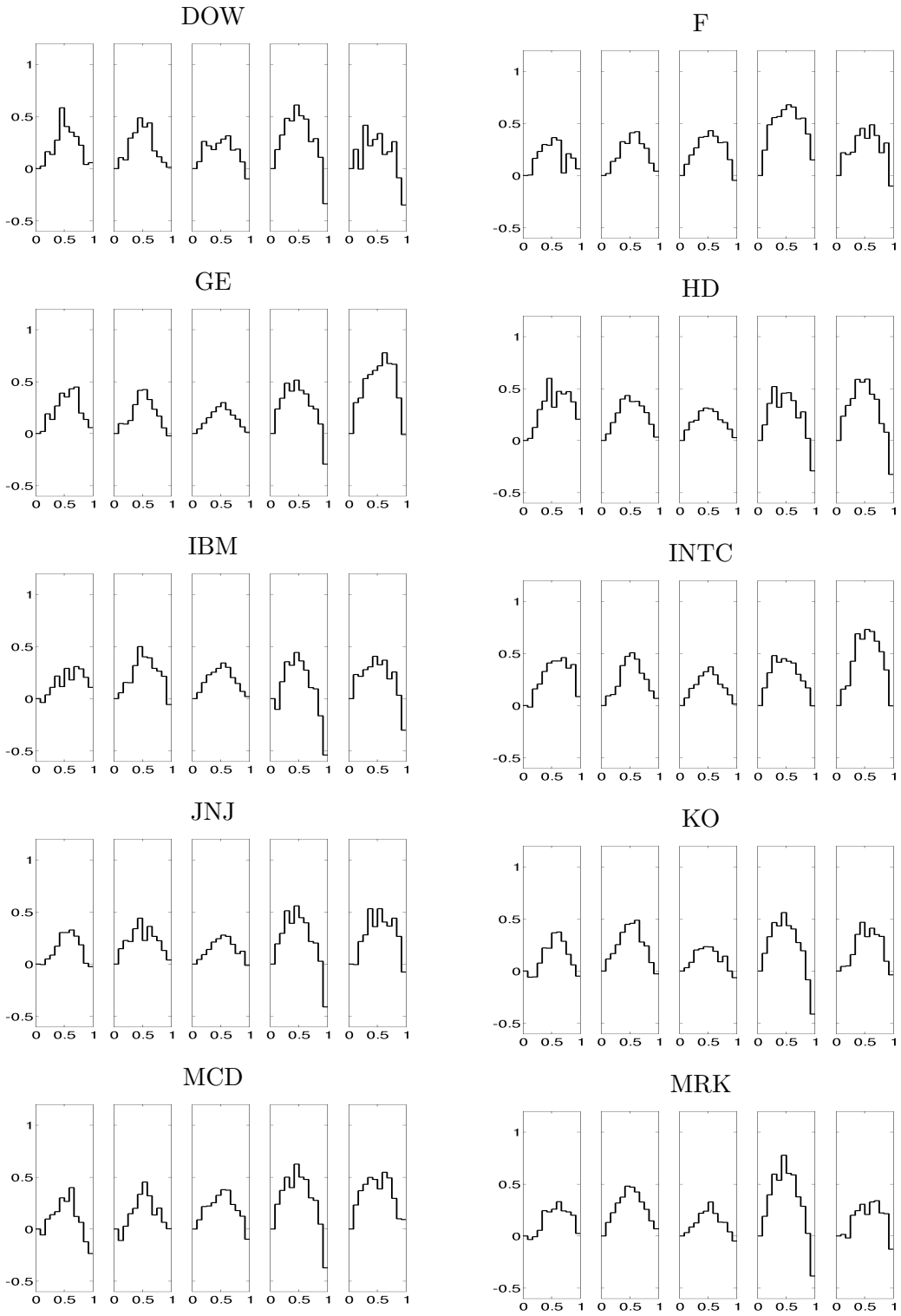
Table 12: 2013, Descriptive Statistics, Full Sample

Ticker	Company Name	Mean	Med.	Max.	St. Dev.	Skew.	Kurt.	C.V.	Trades	Zeros %
AA	ALCOA	0.90	0.005	87.33	2.63	5.95	59.33	2.92	456,067	26.7
ABT	Abbott Labs	0.97	0.005	94.36	2.67	5.66	55.04	2.76	424,783	28.5
AXP	Am. Express	1.45	0.009	102.2	4.24	5.77	53.55	2.92	282,210	24.9
BA	Boeing	1.58	0.017	106.2	4.27	5.48	49.06	2.70	259,421	19.6
BAC	B. of America	0.22	0.003	25.18	0.65	6.41	71.18	2.99	1,878,537	28.2
C	Citigroup	0.45	0.003	66.37	1.48	7.01	82.09	3.31	913,627	29.6
CSCO	Cisco Systems	0.48	0.002	52.26	1.51	6.25	65.12	3.14	855,253	36.2
DELL	Dell	0.61	0.001	155.3	2.68	11.38	229.3	4.39	671,228	37.7
DOW	Dow Chemical	1.28	0.005	105.1	3.87	5.94	56.45	3.02	320,166	28.1
F	Ford Motor	0.54	0.004	49.12	1.45	5.34	48.77	2.65	752,642	27.3
GE	General Electric	0.48	0.004	40.08	1.30	5.24	45.64	2.71	855,128	28.7
HD	Home Depot	1.32	0.015	83.70	3.47	5.21	44.73	2.63	311,307	21.8
IBM	IBM	2.41	0.083	115.0	5.48	4.35	31.40	2.27	170,185	13.0
INTC	Intel	0.43	0.002	30.77	1.24	5.41	47.15	2.87	947,197	32.9
JNJ	Johnson & J.	1.01	0.006	69.17	2.81	5.24	44.83	2.77	404,524	26.7
KO	Coca-Cola	0.65	0.003	63.67	1.87	5.52	50.06	2.89	634,012	31.9
MCD	McDonald's	1.43	0.015	79.05	3.59	4.73	36.41	2.51	286,556	22.7
MRK	Merck	0.63	0.005	55.66	1.79	5.48	49.57	2.83	650,137	26.7
MSFT	Microsoft	0.41	0.002	36.92	1.22	5.52	49.51	2.96	998,773	33.8
QCOM	Qualcomm	0.86	0.005	52.16	2.39	5.28	44.58	2.78	476,864	29.0
T	AT&T	0.57	0.003	45.74	1.68	5.42	46.77	2.94	718,819	30.2
TXN	Texas Inst.	0.98	0.002	110.7	3.52	7.17	82.71	3.57	415,672	35.8
WFC	Wells Fargo	0.52	0.002	54.46	1.67	6.46	69.12	3.23	792,310	31.4
WMT	Wal-Mart	0.87	0.006	82.98	2.64	6.17	61.50	3.01	469,485	25.7
XRX	Xerox	1.37	0.007	135.9	3.87	6.49	76.83	2.83	299,703	28.3

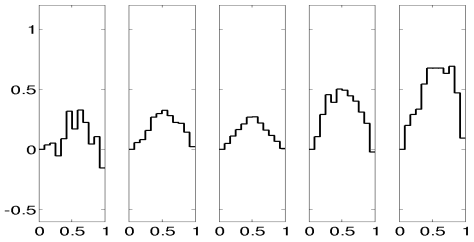
A.2. Estimation Results for Chapter 2

Figure 47: Estimated Intra-Day Seasonal Effects – $\widehat{\alpha}_i$ – as a Function of Standardized Day Time, All Years 1993-2013, Left to Right

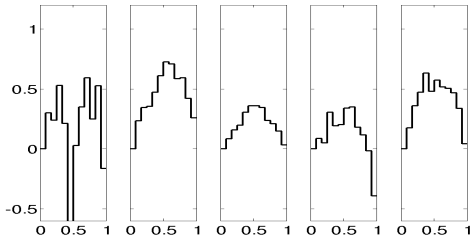



DOW
F
GE
HD
IBM
INTC
JNJ
KO
MCD
MRK

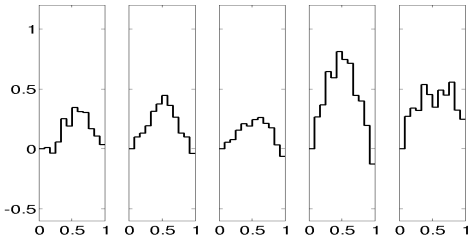
MSFT



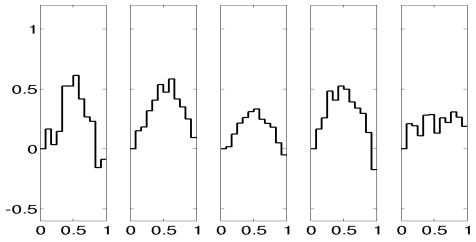
QCOM



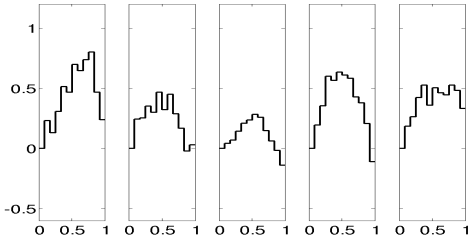
T



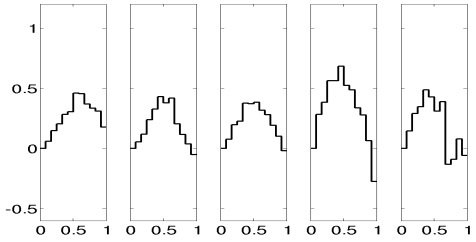
TXN



WFC



WMT



XRX

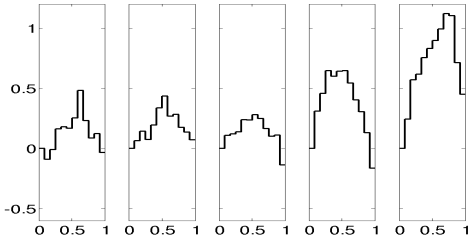


Table 13: 1993, Estimated Intra-Day Seasonal Effects

Ticker	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$	$\widehat{\alpha}_6$	$\widehat{\alpha}_7$	$\widehat{\alpha}_8$	$\widehat{\alpha}_9$	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$\widehat{\alpha}_{12}$
AA	−0.01	0.03	0.30	0.61	0.30	0.81	0.63	0.41	0.60	0.45	0.08
ABT	−0.09	0.13	0.13	0.30	0.37	0.19	0.44	0.36	0.26	0.18	0.06
AXP	0.18	0.34	0.40	0.53	0.65	0.53	0.53	0.48	0.54	0.56	0.36
BA	0.06	0.23	0.16	0.21	0.29	0.44	0.47	0.20	0.23	0.23	0.04
BAC	0.05	0.20	0.28	0.28	0.28	0.24	0.29	0.21	0.02	0.04	0.03
C	0.06	0.23	0.40	0.52	0.49	0.50	0.66	0.51	0.40	0.42	0.24
CSCO	−0.14	0.30	0.09	0.29	0.39	0.41	0.53	0.43	0.21	0.05	0.21
DELL	0.06	0.21	0.23	0.51	0.53	0.62	0.54	0.51	0.26	0.38	0.28
DOW	0.03	0.16	0.13	0.27	0.58	0.40	0.35	0.31	0.22	0.04	0.06
F	0.01	0.16	0.23	0.30	0.29	0.36	0.34	0.03	0.21	0.17	0.07
GE	0.02	0.19	0.13	0.27	0.39	0.35	0.43	0.45	0.20	0.14	0.06
HD	0.02	0.12	0.30	0.38	0.60	0.32	0.47	0.45	0.47	0.37	0.20
IBM	−0.04	0.04	0.11	0.22	0.12	0.29	0.18	0.31	0.29	0.20	0.11
INTC	−0.01	0.16	0.21	0.33	0.41	0.43	0.43	0.46	0.36	0.39	0.09
JNJ	−0.00	0.05	0.09	0.17	0.30	0.30	0.33	0.27	0.19	0.01	−0.03
KO	−0.06	−0.06	0.08	0.22	0.22	0.37	0.37	0.28	0.16	0.06	−0.05
MCD	−0.06	0.10	0.14	0.17	0.30	0.27	0.40	0.12	0.06	−0.12	−0.23
MRK	−0.04	−0.01	0.05	0.24	0.23	0.26	0.33	0.24	0.23	0.20	0.02
MSFT	0.04	0.05	−0.05	0.09	0.32	0.17	0.33	0.22	0.04	0.11	−0.15
QCOM	0.30	0.24	0.53	0.21	−1.10	0.03	0.35	0.59	0.25	0.53	−0.16
T	0.01	−0.04	0.06	0.25	0.19	0.34	0.31	0.30	0.17	0.11	0.03
TXN	0.16	0.04	0.15	0.52	0.52	0.61	0.42	0.27	0.23	−0.16	−0.09
WFC	0.23	0.13	0.31	0.51	0.47	0.70	0.65	0.74	0.80	0.47	0.24
WMT	0.06	0.15	0.20	0.28	0.31	0.46	0.46	0.37	0.33	0.31	0.18
XRX	−0.09	−0.01	0.16	0.18	0.17	0.25	0.48	0.23	0.09	0.12	−0.04

Table 14: 1998, Estimated Intra-Day Seasonal Effects

Ticker	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$	$\widehat{\alpha}_6$	$\widehat{\alpha}_7$	$\widehat{\alpha}_8$	$\widehat{\alpha}_9$	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$\widehat{\alpha}_{12}$
AA	0.00	0.33	0.16	0.39	0.45	0.50	0.56	0.28	0.22	0.11	−0.10
ABT	0.02	0.09	0.11	0.35	0.49	0.39	0.41	0.27	0.23	0.05	−0.05
AXP	0.14	0.14	0.25	0.37	0.60	0.39	0.25	0.16	0.17	−0.05	−0.10
BA	0.02	0.10	0.14	0.26	0.33	0.39	0.30	0.25	0.18	0.15	−0.01
BAC	0.26	0.31	0.41	0.54	0.53	0.52	0.51	0.44	0.41	0.22	0.17
C	0.16	0.10	0.30	0.30	0.36	0.45	0.31	0.24	0.26	0.18	0.10
CSCO	0.11	0.21	0.27	0.39	0.47	0.48	0.47	0.35	0.32	0.28	0.09
DELL	0.10	0.27	0.27	0.45	0.58	0.58	0.50	0.38	0.27	0.20	0.05
DOW	0.11	0.08	0.29	0.34	0.49	0.40	0.44	0.17	0.11	0.06	0.01
F	0.02	0.14	0.17	0.32	0.31	0.41	0.42	0.30	0.26	0.12	0.04
GE	0.10	0.09	0.13	0.28	0.42	0.42	0.33	0.23	0.17	0.05	−0.02
HD	0.06	0.17	0.27	0.40	0.43	0.37	0.38	0.33	0.27	0.16	0.03
IBM	0.06	0.15	0.15	0.32	0.50	0.40	0.39	0.29	0.26	0.22	−0.06
INTC	0.09	0.10	0.18	0.38	0.47	0.51	0.45	0.31	0.25	0.14	0.07
JNJ	0.15	0.23	0.22	0.34	0.44	0.23	0.36	0.26	0.22	0.13	0.04
KO	0.12	0.17	0.25	0.37	0.45	0.46	0.49	0.28	0.24	0.08	−0.03
MCD	−0.11	0.03	0.15	0.20	0.33	0.45	0.32	0.14	0.20	0.07	0.00
MRK	0.13	0.22	0.32	0.37	0.48	0.47	0.42	0.33	0.26	0.15	0.07
MSFT	0.06	0.08	0.16	0.27	0.30	0.33	0.28	0.22	0.22	0.14	0.02
QCOM	0.23	0.34	0.35	0.47	0.61	0.72	0.71	0.58	0.59	0.42	0.26
T	0.10	0.13	0.19	0.31	0.37	0.45	0.36	0.26	0.13	0.10	−0.04
TXN	0.15	0.18	0.32	0.41	0.54	0.47	0.58	0.42	0.35	0.25	0.09
WFC	0.24	0.25	0.35	0.30	0.47	0.32	0.45	0.29	0.17	−0.02	0.03
WMT	0.05	0.12	0.24	0.33	0.43	0.38	0.42	0.20	0.12	0.04	−0.05
XRX	0.06	0.14	0.07	0.20	0.34	0.43	0.27	0.28	0.18	0.14	0.07

Table 15: 2003, Estimated Intra-Day Seasonal Effects

Ticker	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$	$\widehat{\alpha}_6$	$\widehat{\alpha}_7$	$\widehat{\alpha}_8$	$\widehat{\alpha}_9$	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$\widehat{\alpha}_{12}$
AA	0.10	0.10	0.22	0.33	0.27	0.35	0.31	0.14	0.18	0.11	−0.07
ABT	0.11	0.18	0.22	0.24	0.35	0.35	0.32	0.26	0.21	0.10	−0.03
AXP	0.05	0.17	0.23	0.29	0.32	0.34	0.30	0.23	0.13	0.07	−0.05
BA	0.03	0.12	0.15	0.16	0.18	0.22	0.22	0.18	0.17	0.05	−0.03
BAC	0.10	0.18	0.23	0.32	0.32	0.40	0.40	0.26	0.23	0.09	0.06
C	0.07	0.15	0.19	0.25	0.31	0.35	0.32	0.26	0.19	0.11	0.05
CSCO	0.07	0.15	0.19	0.28	0.31	0.31	0.27	0.21	0.18	0.10	0.01
DELL	0.09	0.19	0.23	0.31	0.37	0.37	0.33	0.23	0.21	0.12	−0.01
DOW	0.07	0.26	0.22	0.18	0.24	0.28	0.32	0.18	0.19	0.06	−0.10
F	0.11	0.20	0.26	0.37	0.37	0.43	0.38	0.32	0.32	0.15	−0.05
GE	0.05	0.10	0.15	0.21	0.26	0.30	0.23	0.18	0.14	0.06	0.01
HD	0.10	0.18	0.20	0.29	0.31	0.30	0.28	0.20	0.17	0.11	0.03
IBM	0.06	0.16	0.23	0.25	0.29	0.34	0.30	0.20	0.15	0.07	0.02
INTC	0.07	0.17	0.20	0.28	0.32	0.37	0.30	0.21	0.17	0.10	0.02
JNJ	0.05	0.09	0.14	0.21	0.24	0.28	0.26	0.19	0.10	0.12	−0.01
KO	0.03	0.08	0.20	0.21	0.23	0.23	0.19	0.09	0.14	0.00	−0.06
MCD	0.09	0.22	0.22	0.25	0.32	0.38	0.37	0.24	0.18	0.12	−0.10
MRK	0.03	0.09	0.12	0.22	0.24	0.33	0.22	0.13	0.13	0.04	−0.05
MSFT	0.05	0.11	0.17	0.21	0.27	0.27	0.22	0.16	0.11	0.07	0.01
QCOM	0.08	0.16	0.20	0.31	0.36	0.36	0.34	0.24	0.21	0.15	0.03
T	0.05	0.08	0.16	0.21	0.19	0.24	0.26	0.21	0.17	0.03	−0.06
TXN	0.02	0.12	0.21	0.26	0.31	0.33	0.25	0.21	0.18	0.05	−0.05
WFC	0.04	0.07	0.14	0.21	0.24	0.28	0.26	0.15	0.06	−0.02	−0.14
WMT	0.08	0.20	0.23	0.38	0.37	0.38	0.32	0.28	0.19	0.10	−0.02
XRX	0.11	0.12	0.14	0.24	0.24	0.28	0.25	0.16	0.10	0.11	−0.14

Table 16: 2008, Estimated Intra-Day Seasonal Effects

Ticker	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$	$\widehat{\alpha}_6$	$\widehat{\alpha}_7$	$\widehat{\alpha}_8$	$\widehat{\alpha}_9$	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$\widehat{\alpha}_{12}$
AA	0.31	0.44	0.62	0.64	0.79	0.59	0.61	0.46	0.31	0.14	−0.30
ABT	0.22	0.31	0.50	0.37	0.53	0.43	0.48	0.33	0.17	−0.09	−0.53
AXP	0.17	0.34	0.61	0.41	0.61	0.49	0.46	0.27	0.29	0.05	−0.42
BA	0.18	0.28	0.63	0.47	0.60	0.54	0.47	0.28	0.17	−0.02	−0.44
BAC	0.26	0.47	0.67	0.60	0.74	0.62	0.59	0.43	0.39	0.22	−0.22
C	0.16	0.35	0.60	0.53	0.62	0.57	0.54	0.39	0.28	0.17	−0.13
CSCO	0.17	0.30	0.48	0.43	0.51	0.40	0.39	0.31	0.28	0.14	0.02
DELL	0.18	0.29	0.41	0.35	0.58	0.39	0.35	0.35	0.35	0.18	0.09
DOW	0.18	0.32	0.48	0.46	0.61	0.51	0.48	0.26	0.29	0.11	−0.34
F	0.24	0.49	0.55	0.56	0.63	0.68	0.66	0.54	0.55	0.40	0.15
GE	0.24	0.34	0.48	0.41	0.51	0.41	0.38	0.26	0.24	0.09	−0.30
HD	0.15	0.36	0.52	0.32	0.45	0.46	0.38	0.22	0.27	0.02	−0.29
IBM	−0.10	0.16	0.36	0.32	0.44	0.36	0.27	0.11	0.09	−0.16	−0.54
INTC	0.17	0.32	0.48	0.42	0.45	0.43	0.41	0.30	0.24	0.17	−0.00
JNJ	0.19	0.30	0.51	0.39	0.56	0.45	0.40	0.22	0.20	0.03	−0.41
KO	0.17	0.33	0.46	0.43	0.56	0.43	0.40	0.27	0.19	−0.08	−0.41
MCD	0.24	0.37	0.50	0.40	0.62	0.50	0.48	0.30	0.28	0.05	−0.37
MRK	0.19	0.40	0.60	0.54	0.78	0.60	0.59	0.38	0.28	0.03	−0.38
MSFT	0.11	0.29	0.46	0.39	0.50	0.49	0.45	0.40	0.31	0.22	−0.02
QCOM	0.09	0.05	0.30	0.19	0.20	0.34	0.35	0.18	0.11	−0.02	−0.39
T	0.27	0.37	0.64	0.59	0.81	0.74	0.71	0.45	0.40	0.19	−0.13
TXN	0.17	0.26	0.48	0.41	0.52	0.50	0.39	0.34	0.30	0.14	−0.17
WFC	0.20	0.36	0.60	0.56	0.63	0.61	0.58	0.43	0.38	0.21	−0.11
WMT	0.28	0.38	0.56	0.56	0.68	0.52	0.49	0.34	0.28	0.06	−0.28
XRX	0.31	0.46	0.65	0.60	0.64	0.65	0.54	0.40	0.30	0.13	−0.16

Table 17: 2013, Estimated Intra-Day Seasonal Effects

Ticker	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$	$\widehat{\alpha}_6$	$\widehat{\alpha}_7$	$\widehat{\alpha}_8$	$\widehat{\alpha}_9$	$\widehat{\alpha}_{10}$	$\widehat{\alpha}_{11}$	$\widehat{\alpha}_{12}$
AA	0.22	0.42	0.33	0.67	0.62	0.73	0.83	0.76	0.76	0.44	0.06
ABT	0.19	0.33	0.48	0.64	0.56	0.35	0.34	0.31	0.40	0.40	0.08
AXP	0.16	0.24	0.29	0.47	0.47	0.41	0.32	0.31	0.25	0.10	−0.14
BA	0.05	−0.09	0.20	0.32	0.14	0.31	0.01	−0.07	0.12	−0.15	−0.43
BAC	0.24	0.38	0.50	0.56	0.51	0.53	0.60	0.46	0.37	0.30	0.07
C	0.15	0.40	0.42	0.52	0.68	0.72	0.48	0.43	0.43	0.29	−0.00
CSCO	0.27	0.33	0.51	0.54	0.57	0.69	0.67	0.53	0.63	0.49	0.23
DELL	0.29	0.51	1.17	0.93	0.62	0.68	0.73	0.46	0.67	0.63	0.22
DOW	0.18	−0.00	0.41	0.22	0.28	0.34	0.13	0.16	0.26	−0.09	−0.35
F	0.22	0.20	0.22	0.38	0.45	0.36	0.49	0.38	0.22	0.31	−0.10
GE	0.30	0.34	0.53	0.57	0.61	0.65	0.78	0.67	0.66	0.34	−0.01
HD	0.23	0.34	0.40	0.59	0.56	0.59	0.45	0.40	0.16	0.08	−0.33
IBM	0.23	0.21	0.28	0.30	0.41	0.32	0.37	0.19	0.26	0.03	−0.30
INTC	0.16	0.19	0.42	0.69	0.64	0.73	0.71	0.62	0.52	0.34	−0.00
JNJ	−0.01	0.22	0.28	0.53	0.36	0.53	0.40	0.36	0.44	0.26	−0.08
KO	0.05	0.05	0.16	0.35	0.47	0.33	0.41	0.35	0.33	0.09	−0.04
MCD	0.23	0.37	0.43	0.50	0.48	0.39	0.55	0.49	0.29	0.10	0.09
MRK	0.02	−0.02	0.20	0.24	0.30	0.21	0.33	0.34	0.22	0.21	−0.13
MSFT	0.20	0.29	0.34	0.54	0.67	0.68	0.68	0.63	0.69	0.47	0.09
QCOM	0.18	0.36	0.47	0.63	0.48	0.57	0.52	0.50	0.47	0.34	0.04
T	0.27	0.34	0.32	0.54	0.45	0.35	0.49	0.45	0.55	0.32	0.25
TXN	0.21	0.19	0.11	0.28	0.28	0.13	0.26	0.22	0.31	0.26	0.19
WFC	0.18	0.26	0.42	0.53	0.36	0.50	0.46	0.45	0.53	0.48	0.33
WMT	0.15	0.29	0.35	0.49	0.43	0.31	0.39	−0.13	−0.09	0.08	−0.06
XRX	0.24	0.57	0.62	0.75	0.83	0.90	0.99	1.12	1.10	0.71	0.45

Table 18: Estimated Parameter Values, AA

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.05	0.02	0.05	0.03	0.15	0.22	2.11	2.51	194.95	4,861.46
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.42	1.51	1.32	1.34	1.20	1.27	1.35	1.36	1.86	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	8.95	23.00	11.44	13.93	7.41	7.36	1.87	2.25	3.00	2.26
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	1.00	1.00	1.00	1.00	0.85	0.85	0.87	0.83
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.42	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−16,169	−17,136	−33,938	−35,630	−111,154	−125,690	−258,201	−568,454	310,645	166,067

Table 19: Estimated Parameter Values, ABT

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.07	0.09	0.07	0.09	0.16	0.22	1.86	1.91	209.22	5,351.07
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.20	1.33	1.16	1.21	1.13	1.18	1.47	1.48	1.77	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	4.78	6.22	5.42	7.52	5.13	4.20	8.00	8.29	1.50	1.70
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.18	0.28	0.65	0.93	1.00	1.00	1.00	0.99	0.81	0.68
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.31	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−66,212	−71,629	−72,412	−77,627	−138,577	−156,600	−271,507	−426,057	182,431	43,394

Table 20: Estimated Parameter Values, AXP

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.09	0.11	0.06	0.16	0.20	0.28	1.97	2.26	201.51	4,875.09
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.26	1.32	1.20	1.28	1.11	1.16	1.34	1.35	1.79	1.89
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	7.63	6.68	4.74	7.73	2.27	2.16	2.20	2.51	1.41	1.44
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	1.00	0.96	0.23	0.47	0.91	0.90	0.94	0.80
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.38	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−43,313	−48,828	−52,359	−55,084	−171,553	−198,984	−285,791	−582,995	48,838	−54,104

Table 21: Estimated Parameter Values, BA

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.09	0.25	0.20	0.42	0.22	0.52	2.33	2.47	146.09	3,161.54
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.19	1.34	1.10	1.27	1.19	1.26	1.48	1.49	1.78	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	5.42	9.49	8.12	1.47	7.01	5.13	7.13	7.34	1.50	1.42
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.53	0.64	0.02	0.01	1.00	0.98	0.97	0.94	0.98	0.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.32	–	0.01
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−66,567	−72,582	−151,369	−173,702	−161,984	−188,326	−277,701	−469,372	−32,296	−118,162

Table 22: Estimated Parameter Values, BAC

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.04	0.06	0.06	0.07	0.25	0.39	3.10	3.64	171.37	2,923.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.23	1.27	1.18	1.19	1.13	1.19	1.28	1.29	1.81	1.83
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.93	4.67	4.36	4.23	2.44	2.07	1.70	1.88	3.69	2.41
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	0.97	1.00	1.00	0.62	0.67	0.43	0.30	0.79	0.80
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.54	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−37,283	−38,846	−60,094	−65,911	−183,393	−225,612	−169,054	−967,482	2,515,563	1,897,076

Table 23: Estimated Parameter Values, C

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.13	0.20	0.09	0.20	0.35	0.55	3.67	5.09	266.67	49,091.12
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.21	1.32	1.20	1.40	1.08	1.15	1.29	1.32	1.73	1.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.67	6.79	3.62	8.68	1.91	2.25	1.56	1.70	1.46	1.85
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.27	0.46	0.91	0.52	0.08	0.25	0.33	0.18	0.87	0.89
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.65	–	0.01
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−77,076	−88,382	−73,281	−80,549	−237,295	−293,174	−117,971	−1,272,901	897,174	572,542

Table 24: Estimated Parameter Values, CSCO

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.09	0.12	0.39	0.62	0.66	2.06	11.62	352.77	2,260.65	65,744.14
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.30	1.33	1.18	1.27	1.11	1.21	1.50	1.76	1.85	1.87
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	1.59	1.71	2.79	3.20	4.79	1.80	1.59	1.29	2.68	1.98
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.21	0.28	0.07	0.19	0.02	0.21	0.73	0.17	0.71	0.70
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.50	–	0.04
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−63,106	−68,707	−207,411	−271,495	−320,655	−561,263	40,484	−851,692	773,531	475,843

Table 25: Estimated Parameter Values, DELL

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.16	0.19	0.42	0.74	0.62	1.87	57.67	106.66	1,961.88	77,102.61
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.31	1.40	1.18	1.24	1.15	1.24	1.72	1.74	1.81	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	2.28	3.07	2.77	3.10	3.78	2.30	5.84	3.28	2.07	2.17
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.11	0.13	0.03	0.07	0.06	0.52	0.99	0.72	0.71	0.76
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.59	–	0.04
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−76,383	−87,126	−233,606	−326,160	−292,985	−482,826	−109,390	−633,790	736,729	497,557

Table 26: Estimated Parameter Values, DOW

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.03	0.03	0.03	0.04	0.14	0.30	1.76	13.26	173.87	73,451.38
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
m_0	1.20	1.22	1.22	1.32	1.19	1.29	1.46	1.53	1.79	1.89
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	5.75	5.56	2.49	11.62	2.73	7.78	8.13	12.61	1.52	1.81
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.87	0.89	0.84
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.38	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−33,039	−34,669	−36,056	−37,649	−123,146	−140,805	−264,094	−453,913	68,028	−44,468

Table 27: Estimated Parameter Values, F

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.11	0.20	0.08	0.21	0.20	0.48	2.51	62.13	167.83	45,185.28
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.36	1.39	1.21	1.35	1.11	1.30	1.44	1.71	1.84	1.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	16.53	6.62	3.73	9.58	3.12	9.78	1.00	1.26	2.97	2.34
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	1.00	1.00	0.22	0.21	0.12	0.08	0.83	0.84
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.56	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−60,955	−65,389	−73,885	−80,151	−156,389	−173,197	−203,294	−474,965	522,731	277,719

Table 28: Estimated Parameter Values, GE

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.07	0.09	0.13	0.18	0.44	1.10	2.94	4.34	183.75	15, 163.21
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.23	1.27	1.11	1.17	1.05	1.16	1.34	1.37	1.93	1.94
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.95	4.64	6.77	11.50	1.88	4.06	1.31	1.47	6.67	4.07
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	0.14	1.00	0.01	0.04	0.33	0.22	0.66	0.66
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.56	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−59, 438	−63, 182	−132, 366	−146, 837	−286, 442	−357, 757	−252, 870	−926, 556	761, 382	485, 542

Table 29: Estimated Parameter Values, HD

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.09	0.24	0.08	0.21	0.34	0.63	1.61	1.96	175.48	3,976.84
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.20	1.28	1.13	1.27	1.07	1.16	1.30	1.31	1.78	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	8.36	7.69	5.96	7.95	2.58	4.43	1.77	2.12	1.45	1.40
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.41	0.92	0.18	0.15	0.03	0.12	0.65	0.56	0.95	0.79
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.51	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−82,724	−93,786	−83,774	−93,073	−229,198	−270,681	−304,728	−711,561	27,115	−82,915

Table 30: Estimated Parameter Values, IBM

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.12	0.22	0.13	0.35	0.31	0.72	2.47	2.71	143.81	2,745.52
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.18	1.27	1.17	1.27	1.09	1.20	1.38	1.39	1.80	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	2.94	5.31	4.77	3.25	1.92	4.03	5.70	6.32	1.52	1.61
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.07	0.59	0.09	0.12	0.06	0.23	0.90	0.84	1.00	0.99
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.36	–	0.00
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−105,700	−116,523	−114,703	−128,907	−223,668	−272,675	−334,514	−584,824	−85,163	−133,098

Table 31: Estimated Parameter Values, INTC

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.17	0.22	0.45	0.89	0.69	2.05	9.85	436.89	1,956.79	49,262.67
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
m_0	1.22	1.27	1.13	1.21	1.08	1.20	1.67	1.79	1.83	1.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	1.79	2.09	3.07	3.12	1.88	2.20	3.97	1.17	2.90	2.13
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.11	0.17	0.02	0.06	0.03	0.28	0.76	0.14	0.78	0.78
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.49	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−116,191	−133,260	−268,478	−377,441	−319,759	−568,897	23,719	−777,217	838,222	513,593

Table 32: Estimated Parameter Values, JNJ

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.09	0.14	0.13	0.17	0.28	0.66	1.70	4.42	203.03	4,762.54
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.16	1.21	1.14	1.21	1.11	1.22	1.34	1.51	1.77	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	7.06	4.15	5.66	6.77	3.62	7.09	1.68	6.29	1.37	1.41
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.23	0.38	0.43	0.90	0.06	0.18	0.60	0.49	0.83	0.65
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.41	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−87,502	−93,599	−94,714	−104,323	−189,673	−220,602	−292,560	−579,427	150,387	5,189

Table 33: Estimated Parameter Values, KO

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.05	0.08	0.12	0.52	0.28	0.26	1.18	1.70	187.78	61,327.94
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.16	1.22	1.14	1.34	1.18	1.20	1.30	1.42	1.76	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.03	4.15	4.12	8.25	11.63	3.18	1.95	6.59	1.58	1.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.89	0.96	0.34	0.15	1.00	1.00	0.78	0.69	0.83	0.76
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.45	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−63,336	−66,737	−94,145	−104,260	−160,585	−188,013	−296,308	−561,227	369,501	157,753

Table 34: Estimated Parameter Values, MCD

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.04	0.04	0.09	0.24	0.20	0.43	1.50	1.87	195.21	5,165.59
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.29	1.25	1.13	1.33	1.12	1.21	1.33	1.44	1.79	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	58.30	7.23	3.84	1.38	2.62	5.34	2.11	5.76	1.41	1.30
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	0.18	0.01	0.42	0.87	0.86	0.79	0.93	0.71
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.36	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−35,890	−36,896	−89,546	−96,883	−164,634	−182,475	−278,916	−499,034	21,516	−79,656

Table 35: Estimated Parameter Values, MRK

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.19	0.62	0.12	0.28	0.20	0.29	2.75	3.27	174.01	3,489.47
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.17	1.29	1.13	1.23	1.12	1.17	1.39	1.40	1.76	1.87
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	5.96	6.78	5.04	7.70	2.97	2.39	2.04	2.25	1.61	1.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.04	0.43	0.10	0.42	0.77	0.86	0.95	0.93	0.84	0.75
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.40	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−152,683	−175,798	−115,710	−131,739	−179,060	−210,303	−266,622	−587,271	414,326	206,734

Table 36: Estimated Parameter Values, MSFT

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.75	0.26	0.45	0.67	0.82	2.07	7.22	369.46	2,033.01	52,596.93
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
m_0	1.41	1.37	1.15	1.21	1.06	1.14	1.62	1.76	1.84	1.86
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.53	2.95	3.22	3.15	1.61	2.63	4.34	1.18	2.61	1.94
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.27	0.19	0.01	0.03	0.01	0.05	0.73	0.12	0.71	0.69
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.52	–	0.04
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−100,062	−100,653	−255,203	−343,471	−350,879	−650,643	203,048	−991,788	988,102	641,410

Table 37: Estimated Parameter Values, QCOM

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.07	0.09	0.45	0.44	0.49	1.72	15.62	7.58	5,987.32	5,102.78
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
m_0	1.48	1.50	1.35	1.40	1.17	1.24	1.75	1.64	1.94	1.87
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	2.40	2.67	4.67	4.04	3.98	3.22	31.13	7.02	27.04	1.41
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.28	0.17	0.14	0.37	0.05	0.54	1.00	1.00	1.00	0.69
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.47	–	0.04
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−22,392	−21,557	−106,660	−136,247	−287,236	−453,137	−212,649	−683,770	231,811	40,332

Table 38: Estimated Parameter Values, T

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.20	0.18	0.15	0.35	0.13	0.18	6.71	21.48	2,131.69	4,638.66
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.27	1.26	1.13	1.21	1.16	1.19	1.42	1.55	1.85	1.87
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	6.32	6.63	6.67	4.52	8.18	3.87	2.78	5.06	2.21	1.84
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.89	0.91	0.12	0.24	1.00	1.00	1.00	0.99	0.74	0.69
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.50	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−81,751	−84,064	−122,276	−135,940	−127,494	−145,090	−129,365	−795,655	648,328	414,824

Table 39: Estimated Parameter Values, TXN

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.03	0.04	0.12	0.29	0.22	0.32	2.56	2.16	1,547.63	99,580.42
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.03)
m_0	1.36	1.47	1.17	1.31	1.11	1.16	1.45	1.43	1.82	1.90
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.06	12.79	4.87	6.57	1.81	2.17	6.41	6.17	1.66	1.69
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	0.57	0.82	0.18	0.51	0.62	0.55	0.90	0.81
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.51	–	0.03
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−22,273	−22,859	−111,188	−129,335	−185,043	−212,353	−282,730	−585,376	235,820	78,026

Table 40: Estimated Parameter Values, WFC

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.03	0.03	0.04	0.08	0.23	0.47	2.56	3.05	2,169.89	64,349.94
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.02)
m_0	1.32	1.31	1.35	1.36	1.13	1.19	1.28	1.29	1.85	1.87
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	3.88	3.85	10.58	10.22	5.21	4.91	1.66	1.86	2.69	1.88
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	1.00	1.00	0.09	0.29	0.48	0.32	0.78	0.76
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.58	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−20,305	−21,539	−36,457	−38,231	−185,720	−221,744	−227,805	−952,270	765,921	495,875

Table 41: Estimated Parameter Values, WMT

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.45	0.26	0.08	0.42	0.23	0.41	2.23	2.63	189.37	3,590.31
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.33	1.25	1.13	1.35	1.14	1.22	1.30	1.31	1.76	1.87
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	10.30	6.23	2.94	5.22	5.45	4.37	1.69	1.94	1.49	1.52
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	0.83	0.89	0.43	0.19	1.00	1.00	0.58	0.47	0.93	0.79
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.51	–	0.02
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−112,078	−120,793	−88,314	−97,252	−174,172	−211,134	−269,005	−798,618	205,125	37,695

Table 42: Estimated Parameter Values, XRX

	1993		1998		2003		2008		2013	
	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring	MSMD	Censoring
λ	0.03	0.03	0.05	0.10	0.10	0.13	1.38	2.18	3,269.11	7,367.30
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
m_0	1.46	1.46	1.26	1.34	1.15	1.20	1.41	1.64	1.87	1.89
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
b	13.70	12.51	6.50	16.83	3.57	3.37	1.58	8.45	2.71	2.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ^*	1.00	1.00	1.00	1.00	1.00	1.00	0.73	0.62	0.74	0.67
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
p	–	0.00	–	0.00	–	–	–	0.38	–	0.05
	(–)	(0.00)	(–)	(0.00)	(–)	(–)	(–)	(0.00)	(–)	(0.00)
log-Lik.	−16,723	−16,945	−46,439	−48,790	−102,074	−110,842	−201,266	−329,692	133,753	33,557

Table 43: 2010, Descriptive Statistics, Number of Trades per Δ-Interval, $\Delta = 10$ seconds

Ticker	Mean	Med	Max	Min	St Dev	Skew	Kurt	OD	N	Trades
AA	7.22	0	246	0	15.00	4.01	29.80	31.15	41,040	296,333
ABT	3.20	1	110	0	5.98	4.01	32.19	11.16	41,040	131,458
AXP	5.37	2	174	0	8.62	4.01	34.27	13.85	41,040	220,266
BA	2.50	1	101	0	4.69	4.70	44.05	8.81	41,040	102,725
BAC	16.33	2	426	0	31.27	3.50	21.97	59.87	41,040	670,351
C	2.21	0	352	0	12.15	11.29	178.68	66.83	41,040	90,673
CSCO	4.01	1	108	0	6.79	3.65	24.85	11.50	41,040	164,422
DELL	2.09	0	90	0	4.39	4.81	43.90	9.21	41,040	85,714
DOW	4.89	2	149	0	8.56	4.32	36.29	14.96	41,040	200,800
F	5.97	0	307	0	15.74	4.89	39.10	41.51	41,040	244,873
GE	8.04	0	298	0	17.69	4.05	28.93	38.92	41,040	330,098
HD	5.85	0	218	0	11.46	3.98	30.65	22.48	41,040	239,954
IBM	2.57	1	104	0	4.84	5.16	53.36	9.14	41,040	105,371
INTC	5.03	2	121	0	8.47	3.44	21.86	14.24	41,040	206,573
JNJ	4.97	1	164	0	8.96	4.05	31.59	16.16	41,040	204,145
KO	5.14	2	175	0	8.96	4.32	38.38	15.64	41,040	210,746
MCD	3.96	1	112	0	6.80	3.77	27.27	11.70	41,040	162,371
MRK	6.54	2	277	0	11.53	4.67	49.48	20.33	41,040	268,414
MSFT	4.35	2	110	0	7.15	3.61	24.54	11.75	41,040	178,620
QCOM	3.07	1	70	0	4.66	3.05	18.53	7.08	41,040	125,904
T	5.26	0	236	0	12.37	4.28	32.29	29.12	41,040	215,681
TXN	7.14	1	159	0	13.43	3.09	16.09	25.25	41,040	293,152
WFC	12.50	5	274	0	19.56	3.23	20.39	30.59	41,040	513,196
WMT	5.04	1	203	0	9.45	4.43	40.66	17.72	41,040	206,818
XRX	2.67	0	217	0	9.99	6.77	69.38	37.34	41,040	109,775

Table 44: 2010, Descriptive Statistics, Number of Trades per Δ-Interval, $\Delta = 1$ minute

Ticker	Mean	Med	Max	Min	St Dev	Skew	Kurt	OD	N	Trades
AA	43.32	32	635	0	46.98	3.05	21.80	50.94	6,840	296,333
ABT	19.22	14	284	0	20.21	2.99	20.30	21.26	6,840	131,458
AXP	32.20	23	394	0	31.02	3.06	20.69	29.88	6,840	220,266
BA	15.02	10	178	0	16.23	2.90	16.42	17.53	6,840	102,725
BAC	98.00	69	966	0	106.07	2.43	12.49	114.80	6,840	670,351
C	13.26	3	478	0	33.74	5.03	37.06	85.89	6,840	90,673
CSCO	24.04	17	299	0	24.49	3.04	19.42	24.96	6,840	164,422
DELL	12.53	8	183	0	14.92	3.19	19.09	17.77	6,840	85,714
DOW	29.36	20	476	0	32.96	4.04	31.88	37.01	6,840	200,800
F	35.80	18	936	0	50.33	3.67	32.18	70.76	6,840	244,873
GE	48.26	34	794	0	57.09	3.25	24.04	67.53	6,840	330,098
HD	35.08	25	484	0	38.06	2.93	19.05	41.29	6,840	239,954
IBM	15.41	11	217	0	16.54	3.24	21.90	17.76	6,840	105,371
INTC	30.20	23	446	0	29.31	2.80	18.57	28.45	6,840	206,573
JNJ	29.85	21	405	0	31.03	2.86	18.57	32.27	6,840	204,145
KO	30.81	21	415	0	32.17	3.03	19.16	33.59	6,840	210,746
MCD	23.74	17	265	0	23.04	2.25	11.32	22.37	6,840	162,371
MRK	39.24	29	637	0	40.13	3.31	24.64	41.04	6,840	268,414
MSFT	26.11	19	448	0	26.97	3.33	25.14	27.86	6,840	178,620
QCOM	18.41	14	201	0	17.12	2.22	11.76	15.92	6,840	125,904
T	31.53	21	446	0	37.82	2.49	13.97	45.37	6,840	215,681
TXN	42.86	33	384	0	40.37	2.03	10.15	38.03	6,840	293,152
WFC	75.03	54	738	0	71.36	2.39	12.29	67.87	6,840	513,196
WMT	30.24	22	348	0	31.30	2.61	14.29	32.40	6,840	206,818
XRX	16.05	2	438	0	30.83	3.88	26.10	59.23	6,840	109,775

Table 45: 2010, Descriptive Statistics, Number of Trades per Δ-Interval, $\Delta = 10$ minutes

Ticker	Mean	Med	Max	Min	St Dev	Skew	Kurt	OD	N	Trades
AA	433.24	364	1,932	45	263.65	1.78	7.93	160.45	684	296,333
ABT	192.19	155	993	37	131.63	2.43	11.55	90.15	684	131,458
AXP	322.03	269	1,331	62	204.79	2.00	8.36	130.24	684	220,266
BA	150.18	121	947	19	108.66	2.43	12.94	78.61	684	102,725
BAC	980.05	806	4,335	64	686.54	1.79	7.17	480.93	684	670,351
C	132.56	84	1,110	1	149.27	2.15	9.06	168.09	684	90,673
CSCO	240.38	198	1,475	34	165.78	2.65	14.07	114.33	684	164,422
DELL	125.31	100	929	11	96.13	2.80	15.59	73.74	684	85,714
DOW	293.57	224	2,490	56	249.12	4.08	28.07	211.40	684	200,800
F	358.00	290	3,240	8	297.67	2.98	20.40	247.51	684	244,873
GE	482.60	404	3,112	14	353.72	2.75	15.32	259.27	684	330,098
HD	350.81	290	2,262	7	233.63	2.56	14.50	155.59	684	239,954
IBM	154.05	126	852	13	103.69	2.12	9.93	69.79	684	105,371
INTC	302.01	265	1,894	39	186.81	2.65	16.17	115.56	684	206,573
JNJ	298.46	250	1,518	41	198.77	1.95	8.41	132.37	684	204,145
KO	308.11	233	1,384	57	218.37	1.96	7.35	154.77	684	210,746
MCD	237.38	195	1,033	31	150.24	1.69	6.61	95.08	684	162,371
MRK	392.42	319	3,301	29	270.06	3.35	26.00	185.85	684	268,414
MSFT	261.14	209	1,862	22	190.18	2.78	15.67	138.50	684	178,620
QCOM	184.07	154	801	26	119.15	1.78	7.14	77.13	684	125,904
T	315.32	275	1,403	19	197.44	1.73	7.84	123.62	684	215,681
TXN	428.58	371	1,388	91	233.56	1.50	5.63	127.28	684	293,152
WFC	750.29	604	2,991	145	503.09	1.75	6.39	337.33	684	513,196
WMT	302.37	257	1,323	44	188.14	1.94	8.25	117.06	684	206,818
XRX	160.49	118	1,362	0	159.84	2.74	14.82	159.19	684	109,775

A.4. Proofs for Chapter 3

Proof of Proposition 1. Normalize $\Delta = 1$, for the sake of tractability. For $n \geq 1$, define $\mathcal{K}(n) = \mathbb{E}(N_t N_{t+n})/\mathbb{E}\big(N_t^2\big)$. Then we have:

$$\begin{aligned}
\mathcal{K}(n) &= \frac{\mathbb{E}\left(N_t N_{t+n}\right)}{\mathbb{E}\left(N_t^2\right)} = \frac{\mathbb{E}\left[\mathbb{E}\left(N_t \,|\lambda_t\right)\mathbb{E}\left(N_{t+n}\,|\lambda_{t+n}\right)\right]}{\mathbb{E}\left[\mathbb{E}\left(N_t^2\,|\lambda_t\right)\right]} \\
&= \frac{\mathbb{E}\left(\lambda_t\lambda_{t+n}\right)}{\mathbb{E}\left(\lambda_t+\lambda_t^2\right)} = \frac{\mathbb{E}\left(\lambda_t\lambda_{t+n}\right)}{\lambda+\mathbb{E}\left(\lambda_t^2\right)} = \frac{\lambda^2\left[\mathbb{E}\left(\prod_{k=1}^{\bar{k}} M_{k,t}M_{k,t+n}\right)\right]}{\lambda+\lambda^2\left[\mathbb{E}\left(\prod_{k=1}^{\bar{k}} M_{k,t}^2\right)\right]} \\
&= \frac{\lambda^2\left(\prod_{k=1}^{\bar{k}}\mathbb{E}\left(M_{k,t}M_{k,t+n}\right)\right)}{\lambda+\lambda^2\left(\prod_{k=1}^{\bar{k}}\mathbb{E}\left(M_{k,t}^2\right)\right)} = \frac{\lambda^2\left(\prod_{k=1}^{\bar{k}}\left[1+a\left(1-\gamma_k\right)^n\right]\right)}{\lambda+\lambda^2\left(1+a\right)^{\bar{k}}} \\
&= \frac{\prod_{k=1}^{\bar{k}}\left(1+a\left(1-\gamma_k\right)^n\right)}{c_{\bar{k}}^{-1}\left(1+a\right)^{\bar{k}}}
\end{aligned}$$

where $a = \mathbb{E}\left(M^2\right) - 1$ and $c_{\bar{k}} = \left(1+\frac{1}{\lambda}(1+a)^{-\bar{k}}\right)^{-1}$. Also, one can easily recognize that $0 < \mathcal{K}(n) \leq 1$. Therefore:

$$\frac{\mathcal{K}(n)}{c_{\bar{k}}} = \prod_{k=1}^{\bar{k}}\left(\frac{1+a\left(1-\gamma^*\right)^{nb^{k-\bar{k}}}}{1+a}\right) > 0$$

and then,

$$\log\left(\frac{\mathcal{K}(n)}{c_{\bar{k}}}\right) = \sum_{k=1}^{\bar{k}}\log\left(\frac{1+a\left(1-\gamma^*\right)^{nb^{k-\bar{k}}}}{1+a}\right)$$

Having this expression for $\log\left(\mathcal{K}(n)/c_{\bar{k}}\right)$, we can exploit the analysis in Calvet and Fisher (2004), Appendix A.1 and find the following bound for $\left|\frac{\log \mathcal{K}(n)}{\log n^{-\delta}} - 1\right|$, for any $n \in \mathcal{I}_{\bar{k}}$:

$$\left|\frac{\log \mathcal{K}(n)}{\log n^{-\delta}} - 1\right| \leq \frac{\log c_{\bar{k}} + \xi + 2a \log_b \bar{k} + a\left|\log\left(1-\gamma^*\right)\right|}{\delta \alpha_1 \bar{k} \log b} = \eta_{\bar{k}}$$

where $\xi = \delta \log b + a/\left(1-(1-\gamma^*)b^{-1}\right)$. Note that, as $\bar{k} \uparrow \infty$, the upper bound of $\mathcal{I}_{\bar{k}}$ diverges to infinity as well as its length as an interval on the real line, while $\eta_{\bar{k}}$ collapses towards zero. Hence we can safely infer that:

$$\sup_{n \in \mathcal{I}_{\bar{k}}} \left|\frac{\log \mathcal{K}(n)}{\log n^{-\delta}} - 1\right| \longrightarrow 0 \qquad \text{as} \qquad \bar{k} \longrightarrow +\infty$$

Now, let's turn our attention to the ratio of $\mathcal{K}(n)$ to the autocorrelation function $\rho(n)$. By definition we can write:

$$\begin{aligned}
\frac{\mathcal{K}(n)}{\rho(n)} &= \frac{\mathbb{E}\left(N_t N_{t+n}\right)/\mathbb{E}\left(N_t^2\right)}{\mathbb{C}\mathrm{ov}\left(N_t, N_{t+n}\right)/\mathbb{V}\mathrm{ar}\left(N_t\right)} = \frac{\mathbb{E}\left(N_t N_{t+n}\right)/\mathbb{E}\left(N_t^2\right)}{\left(\mathbb{E}\left(N_t N_{t+n}\right) - \lambda^2\right)/\left(\mathbb{E}\left(N_t^2\right) - \lambda^2\right)} \\
&= \left(\frac{\prod_{k=1}^{\bar{k}}\left(1+a\left(1-\gamma_k\right)^n\right)}{\prod_{k=1}^{\bar{k}}\left(1+a\left(1-\gamma_k\right)^n\right) - 1}\right)\left(\frac{1-\lambda+\lambda(1+a)^{\bar{k}}}{1+\lambda(1+a)^{\bar{k}}}\right) \\
&= \left(\frac{1}{1 - c_{\bar{k}}\frac{(1+a)^{-\bar{k}}}{\mathcal{K}(n)}}\right)\left(\frac{c_{\bar{k}}^{-1} - \lambda\left(c_{\bar{k}}^{-1} - 1\right)}{c_{\bar{k}}^{-1}}\right) \\
&= \frac{1-\lambda\left(1-c_{\bar{k}}\right)}{1 - c_{\bar{k}}\frac{(1+a)^{-\bar{k}}}{\mathcal{K}(n)}}
\end{aligned}$$

From the second equality and the fact that $\mathcal{K}(n) \leq 1$ we can find that one is a lower bound

for the above ratio. Also, noting that $0 \leq c_{\bar{k}} \leq 1$ yields the complete expression:

$$1 \leq \frac{\mathcal{K}(n)}{\rho(n)} \leq \left(1 - c_{\bar{k}} \frac{(1+a)^{-\bar{k}}}{\mathcal{K}(n)}\right)^{-1}$$

Again, for every $n \in \mathcal{I}_{\bar{k}}$, we have that $\log_b(n) \leq \alpha_2 \bar{k}$. Therefore, for such n we can write that $\log_b \mathcal{K}(n) \geq -\delta(1+\eta_{\bar{k}})\alpha_2 \bar{k}$, which in turn implies the following relationship:

$$\log_b \left(\frac{\mathcal{K}(n)}{(1+a)^{-\bar{k}}}\right) \geq \bar{k}\delta\left(1 - \alpha_2 - \alpha_2 \eta_{\bar{k}}\right)$$

As $\bar{k} \uparrow +\infty$, then $\mathcal{K}(n)/(1+a)^{-\bar{k}} \uparrow +\infty$ and therefore $\mathcal{K}(n)/\rho(n) \downarrow 1$. Hence, we can finally conclude that:

$$\sup_{n \in \mathcal{I}_{\bar{k}}} \left| \frac{\log \rho(n)}{\log n^{-\delta}} - 1 \right| \longrightarrow 0 \qquad \text{as} \qquad \bar{k} \longrightarrow +\infty$$

which proves the claim made in the statement of the Proposition.

□

A.5. Estimation Results for Chapter 3

Table 46: Estimated Intra-Day Seasonal Effects, $\Delta = 10$ seconds, $\bar{k} = 7$, $Q = 3$

Ticker	Company Name	δ	$\delta_{1,c}$	$\delta_{2,c}$	$\delta_{3,c}$	$\delta_{1,s}$	$\delta_{2,s}$	$\delta_{3,s}$
AA	ALCOA	-0.50	0.20	0.03	0.15	-0.14	-0.13	-0.06
ABT	Abbott Labs	-0.23	0.55	0.27	0.16	0.11	-0.17	-0.08
AXP	American Express	-0.16	0.53	0.171	0.12	0.01	-0.03	-0.02
BA	Boeing	0.91	0.45	0.18	0.31	0.28	0.07	0.02
BAC	Bank of America	0.49	0.39	0.075	0.03	0.33	0.26	-0.25
C	Citigroup	0.22	0.28	0.42	0.36	0.07	-0.21	-0.41
CSCO	Cisco Systems	1.73	0.49	0.251	0.08	0.65	0.07	0.14
DELL	Dell	1.59	0.52	0.16	0.13	0.42	0.08	0.10
DOW	Dow Chemical	0.12	0.41	0.30	0.15	0.09	-0.06	0.08
F	Ford Motor	-0.42	0.17	-0.18	0.20	0.03	-0.08	-0.06
GE	General Electric	0.40	0.39	0.23	0.16	0.19	0.13	0.26
HD	Home Depot	-0.98	0.46	0.28	0.18	-0.38	-0.09	-0.03
IBM	IBM	0.64	0.34	0.15	0.10	0.19	0.04	0.05
INTC	Intel	0.21	0.34	0.13	0.06	0.28	0.02	-0.05
JNJ	Johnson & Johnson	-0.02	0.36	0.07	-0.02	0.17	0.05	0.06
KO	Coca-Cola	1.38	0.41	0.16	0.14	0.67	0.32	0.24
MCD	McDonald's	-0.86	0.43	0.21	0.11	-0.05	0.06	0.11
MRK	Merck	0.42	0.32	0.25	0.28	0.15	0.05	0.00
MSFT	Microsoft	1.46	0.38	0.23	0.22	0.62	0.14	0.04
QCOM	Qualcomm	-0.35	0.40	0.10	0.06	-0.08	-0.04	0.03
T	AT&T	0.13	0.28	0.20	0.01	0.20	0.11	0.04
TXN	Texas Instruments	-0.53	0.41	0.04	0.19	0.08	-0.09	-0.02
WFC	Wells Fargo	0.09	0.60	0.20	0.18	0.18	0.04	0.18
WMT	Wal-Mart	0.24	0.36	0.17	0.10	-0.01	-0.01	-0.06
XRX	Xerox	-0.67	0.44	0.17	-0.01	0.03	0.02	-0.02

Table 47: Estimated Intra-Day Seasonal Effects, $\Delta = 1$ minute, $\bar{k} = 7$, $Q = 3$

Ticker	Company Name	δ	$\delta_{1,c}$	$\delta_{2,c}$	$\delta_{3,c}$	$\delta_{1,s}$	$\delta_{2,s}$	$\delta_{3,s}$
AA	ALCOA	0.02	0.20	0.25	0.02	0.25	0.03	0.01
ABT	Abbott Labs	0.26	0.39	0.29	-0.07	0.09	0.02	0.02
AXP	American Express	0.22	0.38	0.19	0.04	0.16	-0.17	-0.02
BA	Boeing	0.70	0.37	0.01	0.15	0.28	0.09	0.12
BAC	Bank of America	0.12	0.35	0.28	-0.16	0.11	0.16	-0.10
C	Citigroup	0.08	0.09	0.17	0.15	0.26	0.01	-0.10
CSCO	Cisco Systems	0.57	0.39	0.25	-0.02	0.07	-0.12	-0.06
DELL	Dell	1.11	0.41	0.09	0.06	0.43	0.14	0.17
DOW	Dow Chemical	0.49	0.38	0.24	0.10	0.34	0.06	0.09
F	Ford Motor	-0.54	0.45	-0.09	0.25	-0.08	-0.55	-0.06
GE	General Electric	0.22	0.26	0.33	0.03	0.02	-0.11	0.28
HD	Home Depot	-0.33	0.46	0.25	0.02	0.18	0.00	0.00
IBM	IBM	0.54	0.37	0.15	0.18	0.01	0.03	0.17
INTC	Intel	0.46	0.39	0.13	0.13	0.15	-0.05	0.03
JNJ	Johnson & Johnson	0.56	0.36	0.15	0.12	0.26	0.11	0.15
KO	Coca-Cola	0.72	0.16	0.02	0.04	0.17	0.03	0.17
MCD	McDonald's	0.11	0.37	0.32	0.11	-0.03	-0.01	-0.07
MRK	Merck	-0.14	0.32	0.15	0.20	0.23	0.12	0.02
MSFT	Microsoft	0.93	0.52	0.26	0.01	0.32	0.08	0.04
QCOM	Qualcomm	0.23	0.49	0.10	0.11	0.00	-0.12	0.05
T	AT&T	0.40	0.26	0.32	0.17	0.18	0.02	-0.01
TXN	Texas Instruments	-0.62	0.29	0.15	0.07	-0.01	-0.07	-0.03
WFC	Wells Fargo	-0.11	0.15	0.14	0.14	-0.02	-0.05	0.11
WMT	Wal-Mart	0.14	0.26	0.24	0.07	0.07	0.05	-0.11
XRX	Xerox	0.50	0.37	-0.04	0.18	0.04	0.08	0.04

Table 48: AA - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.61	1.58	1.56	1.35
	(0.00)	(0.00)	(0.00)	(0.00)
b	2.53	2.02	1.78	1.80
	(0.00)	(0.02)	(0.01)	(0.00)
λ	1.09	0.92	0.70	0.73
	(0.00)	(0.07)	(0.01)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-118,021.82	-114,147.69	-111,538.65	-36,197.31
BIC	236,160.48	228,412.22	223,194.14	72,491.75
1-Step RMSE	-	-	14.72	43.84
10-Step RMSE	-	-	14.88	45.08
30-Step RMSE	-	-	14.87	45.08
100-Step RMSE	-	-	14.86	45.13
ACP(1,1)				
log-Likelihood	-	-	-396,698.11	-130,607.91
BIC	-	-	793,513.06	261,312.96
1-Step RMSE	-	-	14.55	42.66
10-Step RMSE	-	-	14.69	44.12
30-Step RMSE	-	-	14.75	44.50
100-Step RMSE	-	-	14.77	44.67

Table 49: ABT - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.59	1.56	1.54	1.33
	(0.00)	(0.03)	(0.03)	(0.00)
b	2.82	2.22	2.00	2.53
	(0.03)	(1.04)	(0.23)	(0.02)
λ	0.39	0.32	0.39	0.34
	(0.01)	(0.09)	(0.07)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-88,384.72	-87,561.79	-87,138.77	-26,943.17
BIC	176,886.28	175,240.43	174,394.38	53,983.47
1-Step RMSE	-	-	5.82	17.17
10-Step RMSE	-	-	5.94	18.41
30-Step RMSE	-	-	5.96	19.11
100-Step RMSE	-	-	5.96	19.44
ACP(1,1)				
log-Likelihood	-	-	-164,004.49	-53,823.01
BIC	-	-	328,125.83	107,743.15
1-Step RMSE	-	-	5.63	16.93
10-Step RMSE	-	-	5.71	17.96
30-Step RMSE	-	-	5.74	18.35
100-Step RMSE	-	-	5.81	18.95

Table 50: AXP - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.52	1.49	1.46	1.28
	(0.02)	(0.03)	(0.04)	(0.00)
b	2.99	2.34	1.98	2.64
	(0.75)	(0.57)	(0.53)	(0.02)
λ	0.71	0.51	0.56	0.60
	(0.06)	(0.50)	(0.26)	(0.01)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-111,095.60	-110,023.29	-109,236.77	-29,988.54
BIC	222,308.05	220,163.43	218,590.38	60,074.22
1-Step RMSE	-	-	8.14	26.06
10-Step RMSE	-	-	8.36	27.76
30-Step RMSE	-	-	8.38	28.63
100-Step RMSE	-	-	8.38	29.19
ACP(1,1)				
log-Likelihood	-	-	-209,572.40	-68,048.61
BIC	-	-	419,261.64	136,194.35
1-Step RMSE	-	-	7.97	25.71
10-Step RMSE	-	-	8.21	27.55
30-Step RMSE	-	-	8.29	28.10
100-Step RMSE	-	-	8.33	28.36

Table 51: BA - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.57	1.53	1.50	1.31
	(0.02)	(0.03)	(0.04)	(0.00)
b	3.50	2.68	2.23	2.78
	(0.54)	(0.35)	(0.33)	(0.01)
λ	0.25	0.29	0.17	0.20
	(0.04)	(0.03)	(0.06)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-80,664.94	-80,114.27	-79,779.68	-24,478.32
BIC	161,446.73	160,345.38	159,676.21	49,053.77
1-Step RMSE	-	-	4.48	13.39
10-Step RMSE	-	-	4.56	14.23
30-Step RMSE	-	-	4.60	14.82
100-Step RMSE	-	-	4.61	15.11
ACP(1,1)				
log-Likelihood	-	-	-126,752.55	-43,255.88
BIC	-	-	253,621.95	86,608.90
1-Step RMSE	-	-	4.39	13.28
10-Step RMSE	-	-	4.45	13.88
30-Step RMSE	-	-	4.48	14.29
100-Step RMSE	-	-	4.53	14.73

Table 52: BAC - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.54	1.51	1.49	1.29
	(0.00)	(0.00)	(0.00)	(0.00)
b	2.66	2.11	1.86	1.91
	(0.01)	(0.09)	(0.00)	(0.04)
λ	1.70	1.26	0.97	1.36
	(0.01)	(0.24)	(0.01)	(0.01)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-166,385.24	-159,570.85	-155,049.38	-45,543.84
BIC	332,887.33	319,258.55	310,215.61	91,184.81
1-Step RMSE	-	-	29.85	93.07
10-Step RMSE	-	-	30.56	99.42
30-Step RMSE	-	-	30.61	99.69
100-Step RMSE	-	-	30.57	100.34
ACP(1,1)				
log-Likelihood	-	-	-737,991.60	-243,909.87
BIC	-	-	1,476,100.05	487,916.88
1-Step RMSE	-	-	29.39	88.35
10-Step RMSE	-	-	29.80	93.57
30-Step RMSE	-	-	30.00	96.16
100-Step RMSE	-	-	30.28	98.29

Table 53: C - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.76	1.70	1.66	1.48
	(0.02)	(0.04)	(0.03)	(0.01)
b	4.49	3.31	2.73	2.55
	(0.28)	(1.29)	(0.20)	(0.13)
λ	0.45	0.36	0.37	0.31
	(0.04)	(0.14)	(0.04)	(0.02)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-58,617.41	-56,943.47	-55,419.00	-22,852.53
BIC	117,351.67	114,003.79	110,954.84	45,802.20
1-Step RMSE	-	-	12.06	32.92
10-Step RMSE	-	-	12.16	32.98
30-Step RMSE	-	-	12.27	33.04
100-Step RMSE	-	-	12.46	33.21
ACP(1,1)				
log-Likelihood	-	-	-235,763.72	-118,234.50
BIC	-	-	471,644.29	236,566.14
1-Step RMSE	-	-	12.03	32.45
10-Step RMSE	-	-	12.06	32.84
30-Step RMSE	-	-	12.06	32.81
100-Step RMSE	-	-	12.11	32.98

Table 54: CSCO - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.54	1.50	1.48	1.29
	(0.03)	(0.03)	(0.03)	(0.05)
b	2.94	2.37	2.10	2.38
	(0.42)	(0.16)	(0.34)	(0.95)
λ	0.21	0.25	0.15	0.29
	(0.04)	(0.03)	(0.06)	(0.07)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-97,641.92	-96,977.63	-96,789.23	-28,073.37
BIC	195,400.68	194,072.10	193,695.30	56,243.87
1-Step RMSE	-	-	6.40	20.27
10-Step RMSE	-	-	6.64	22.21
30-Step RMSE	-	-	6.68	22.84
100-Step RMSE	-	-	6.70	23.10
ACP(1,1)				
log-Likelihood	-	-	-176,822.00	-58,246.53
BIC	-	-	353,760.85	116,590.20
1-Step RMSE	-	-	6.31	19.97
10-Step RMSE	-	-	6.45	21.60
30-Step RMSE	-	-	6.54	22.07
100-Step RMSE	-	-	6.69	22.44

Table 55: DELL - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.59	1.56	1.53	1.33
	(0.05)	(0.04)	(0.02)	(0.18)
b	2.83	2.36	2.05	2.31
	(0.38)	(0.52)	(0.14)	(1.49)
λ	0.12	0.14	0.09	0.11
	(0.03)	(0.05)	(0.03)	(0.04)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-72,658.14	-72,217.63	-71,873.13	-23,664.16
BIC	145,433.13	144,552.11	143,863.11	47,425.45
1-Step RMSE	-	-	4.17	12.64
10-Step RMSE	-	-	4.27	13.51
30-Step RMSE	-	-	4.29	13.73
100-Step RMSE	-	-	4.30	13.86
ACP(1,1)				
log-Likelihood	-	-	-120,929.92	-43,660.61
BIC	-	-	241,976.69	87,418.36
1-Step RMSE	-	-	4.12	12.54
10-Step RMSE	-	-	4.20	13.21
30-Step RMSE	-	-	4.22	13.44
100-Step RMSE	-	-	4.26	13.68

Table 56: DOW - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.54	1.51	1.48	1.29
	(0.06)	(0.10)	(0.09)	(0.03)
b	3.00	2.47	2.08	2.71
	(0.29)	(0.71)	(0.36)	(0.74)
λ	0.45	0.35	0.44	0.44
	(0.11)	(0.16)	(0.17)	(0.07)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-105,811.86	-104,732.56	-104,015.70	-29,188.44
BIC	211,740.56	209,581.97	208,148.24	58,474.02
1-Step RMSE	-	-	7.88	25.22
10-Step RMSE	-	-	8.28	28.65
30-Step RMSE	-	-	8.33	30.18
100-Step RMSE	-	-	8.33	31.01
ACP(1,1)				
log-Likelihood	-	-	-198,790.24	-64,051.67
BIC	-	-	397,697.32	128,200.48
1-Step RMSE	-	-	7.61	24.49
10-Step RMSE	-	-	7.82	27.65
30-Step RMSE	-	-	7.97	29.30
100-Step RMSE	-	-	8.21	30.71

Table 57: F - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.64	1.60	1.58	1.39
	(0.00)	(0.01)	(0.01)	(0.00)
b	2.91	2.25	1.97	2.07
	(0.03)	(0.12)	(0.03)	(0.01)
λ	0.86	0.59	0.52	0.71
	(0.02)	(0.05)	(0.01)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-102,109.90	-98,738.15	-96,302.60	-34,569.36
BIC	204,336.64	197,593.15	192,722.05	69,235.85
1-Step RMSE	-	-	15.29	46.39
10-Step RMSE	-	-	15.57	49.78
30-Step RMSE	-	-	15.61	50.21
100-Step RMSE	-	-	15.63	50.44
ACP(1,1)				
log-Likelihood	-	-	-392,976.79	-153,351.72
BIC	-	-	786,070.43	306,800.58
1-Step RMSE	-	-	15.09	44.81
10-Step RMSE	-	-	15.28	47.52
30-Step RMSE	-	-	15.39	48.20
100-Step RMSE	-	-	15.52	48.80

Table 58: GE - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.60	1.58	1.55	1.35
	(0.01)	(0.01)	(0.00)	(0.00)
b	2.69	2.16	1.82	1.97
	(0.14)	(0.09)	(0.00)	(0.00)
λ	1.10	0.47	0.52	0.74
	(0.07)	(0.02)	(0.00)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-121,861.48	-117,269.76	-114,125.60	-37,592.07
BIC	243,839.80	234,656.37	228,368.05	75,281.27
1-Step RMSE	-	-	17.21	52.23
10-Step RMSE	-	-	17.42	55.25
30-Step RMSE	-	-	17.45	55.52
100-Step RMSE	-	-	17.48	55.72
ACP(1,1)				
log-Likelihood	-	-	-454,702.39	-151,609.14
BIC	-	-	909,521.63	303,315.41
1-Step RMSE	-	-	16.97	49.69
10-Step RMSE	-	-	17.10	52.19
30-Step RMSE	-	-	17.24	52.76
100-Step RMSE	-	-	17.41	54.60

Table 59: HD - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.59	1.57	1.55	1.32
	(0.01)	(0.00)	(0.00)	(0.00)
b	2.45	2.07	1.84	1.86
	(0.05)	(0.05)	(0.01)	(0.00)
λ	0.47	0.47	0.78	0.64
	(0.01)	(0.01)	(0.02)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-109,264.26	-106,959.88	-105,469.14	-32,995.01
BIC	218,645.36	214,036.60	211,055.13	66,087.15
1-Step RMSE	-	-	11.13	34.31
10-Step RMSE	-	-	11.28	36.01
30-Step RMSE	-	-	11.29	36.05
100-Step RMSE	-	-	11.29	36.12
ACP(1,1)				
log-Likelihood	-	-	-297,570.89	-100,729.55
BIC	-	-	595,258.62	201,556.23
1-Step RMSE	-	-	10.91	33.29
10-Step RMSE	-	-	11.12	34.98
30-Step RMSE	-	-	11.19	35.32
100-Step RMSE	-	-	11.22	35.50

Table 60: IBM - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.56	1.52	1.48	1.32
	(0.09)	(0.03)	(0.07)	(0.03)
b	3.84	2.61	2.17	2.71
	(2.15)	(0.80)	(0.47)	(0.50)
λ	0.27	0.22	0.18	0.24
	(0.09)	(0.04)	(0.07)	(0.03)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-82,000.13	-81,443.08	-81,177.83	-24,735.06
BIC	164,117.10	163,003.00	162,472.51	49,567.26
1-Step RMSE	-	-	4.58	14.24
10-Step RMSE	-	-	4.67	14.72
30-Step RMSE	-	-	4.69	14.97
100-Step RMSE	-	-	4.69	15.14
ACP(1,1)				
log-Likelihood	-	-	-129,568.65	-45,529.32
BIC	-	-	259,254.14	91,155.78
1-Step RMSE	-	-	4.56	14.09
10-Step RMSE	-	-	4.65	14.64
30-Step RMSE	-	-	4.68	14.81
100-Step RMSE	-	-	4.68	14.92

Table 61: INTC - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.53	1.51	1.49	1.30
	(0.02)	(0.02)	(0.04)	(0.02)
b	2.72	2.26	1.99	2.25
	(0.15)	(0.07)	(0.30)	(0.29)
λ	0.26	0.48	0.43	0.42
	(0.04)	(0.04)	(0.09)	(0.05)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-107,135.18	-106,052.98	-105,260.74	-30,001.91
BIC	214,387.20	212,222.80	210,638.33	60,100.96
1-Step RMSE	-	-	8.05	25.15
10-Step RMSE	-	-	8.29	26.98
30-Step RMSE	-	-	8.31	27.48
100-Step RMSE	-	-	8.32	27.64
ACP(1,1)				
log-Likelihood	-	-	-218,002.49	-71,726.66
BIC	-	-	436,121.82	143,550.46
1-Step RMSE	-	-	7.94	24.71
10-Step RMSE	-	-	8.11	26.31
30-Step RMSE	-	-	8.18	26.89
100-Step RMSE	-	-	8.25	27.22

Table 62: JNJ - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.57	1.54	1.51	1.31
	(0.03)	(0.11)	(0.07)	(0.03)
b	2.78	2.28	1.99	2.23
	(0.62)	(2.36)	(0.28)	(0.30)
λ	0.54	0.49	0.49	0.43
	(0.11)	(1.18)	(0.08)	(0.03)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-105,655.90	-104,211.30	-103,206.91	-30,509.42
BIC	211,428.64	208,539.45	206,530.66	61,115.98
1-Step RMSE	-	-	8.50	26.59
10-Step RMSE	-	-	8.71	27.94
30-Step RMSE	-	-	8.73	28.29
100-Step RMSE	-	-	8.74	28.38
ACP(1,1)				
log-Likelihood	-	-	-225,752.11	-76,234.60
BIC	-	-	451,621.07	152,566.33
1-Step RMSE	-	-	8.38	26.17
10-Step RMSE	-	-	8.58	27.52
30-Step RMSE	-	-	8.63	27.99
100-Step RMSE	-	-	8.66	28.20

Table 63: KO - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.56	1.52	1.50	1.30
	(0.05)	(0.03)	(0.02)	(0.01)
b	3.14	2.44	2.09	2.78
	(0.21)	(0.33)	(0.12)	(0.10)
λ	0.57	0.50	0.27	0.47
	(0.05)	(0.06)	(0.01)	(0.02)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-108,061.23	-106,683.19	-105,831.04	-30,159.64
BIC	216,239.30	213,483.22	211,778.93	60,416.42
1-Step RMSE	-	-	8.48	26.04
10-Step RMSE	-	-	8.78	28.25
30-Step RMSE	-	-	8.87	29.64
100-Step RMSE	-	-	8.91	31.24
ACP(1,1)				
log-Likelihood	-	-	-216,595.58	-70,863.26
BIC	-	-	433,308.00	141,823.66
1-Step RMSE	-	-	8.22	25.72
10-Step RMSE	-	-	8.41	27.63
30-Step RMSE	-	-	8.54	28.88
100-Step RMSE	-	-	8.73	30.34

Table 64: MCD - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.56	1.52	1.50	1.30
	(0.07)	(0.03)	(0.09)	(0.04)
b	2.99	2.37	2.12	2.38
	(0.50)	(0.51)	(1.37)	(0.41)
λ	0.44	0.38	0.59	0.37
	(0.11)	(0.11)	(0.29)	(0.08)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-97,450.46	-96,589.19	-96,178.41	-28,264.82
BIC	195,017.77	193,295.22	192,473.66	56,626.78
1-Step RMSE	-	-	6.59	19.66
10-Step RMSE	-	-	6.84	21.18
30-Step RMSE	-	-	6.88	21.85
100-Step RMSE	-	-	6.89	22.06
ACP(1,1)				
log-Likelihood	-	-	-179,717.09	-59,496.36
BIC	-	-	359,551.03	119,089.86
1-Step RMSE	-	-	6.41	19.29
10-Step RMSE	-	-	6.52	20.49
30-Step RMSE	-	-	6.57	21.18
100-Step RMSE	-	-	6.64	21.57

Table 65: MRK - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.55	1.52	1.49	1.29
	(0.01)	(0.02)	(0.00)	(0.04)
b	2.63	2.15	1.94	2.25
	(0.10)	(0.18)	(0.03)	(0.74)
λ	0.75	0.65	0.50	0.78
	(0.04)	(0.08)	(0.01)	(0.15)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-118,286.25	-116,556.28	-115,329.70	-32,537.16
BIC	236,689.35	233,229.40	230,776.25	65,171.46
1-Step RMSE	-	-	10.92	34.08
10-Step RMSE	-	-	11.20	37.15
30-Step RMSE	-	-	11.22	37.76
100-Step RMSE	-	-	11.24	37.86
ACP(1,1)				
log-Likelihood	-	-	-275,459.32	-88,886.31
BIC	-	-	551,035.48	177,869.76
1-Step RMSE	-	-	10.77	33.40
10-Step RMSE	-	-	10.96	35.85
30-Step RMSE	-	-	11.05	36.69
100-Step RMSE	-	-	11.18	36.89

Table 66: MSFT - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.52	1.49	1.47	1.28
	(0.09)	(0.03)	(0.04)	(0.02)
b	3.16	2.49	2.20	2.50
	(1.44)	(0.56)	(0.54)	(0.40)
λ	0.35	0.31	0.19	0.28
	(0.11)	(0.09)	(0.09)	(0.03)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-101,034.00	-100,314.29	-99,938.00	-28,408.82
BIC	202,184.84	200,745.43	199,992.85	56,914.78
1-Step RMSE	-	-	6.55	21.07
10-Step RMSE	-	-	6.81	23.34
30-Step RMSE	-	-	6.88	24.15
100-Step RMSE	-	-	6.89	24.48
ACP(1,1)				
log-Likelihood	-	-	-176,968.25	-59,097.81
BIC	-	-	354,053.34	118,292.75
1-Step RMSE	-	-	6.44	20.90
10-Step RMSE	-	-	6.65	22.90
30-Step RMSE	-	-	6.72	23.41
100-Step RMSE	-	-	6.81	23.74

Table 67: QCOM - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.51	1.48	1.45	1.29
	(0.05)	(0.03)	(0.04)	(0.03)
b	3.72	2.98	2.48	2.77
	(0.77)	(0.45)	(0.56)	(1.21)
λ	0.25	0.24	0.36	0.30
	(0.08)	(0.10)	(0.11)	(0.10)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-88,583.41	-88,292.12	-88,144.62	-25,699.34
BIC	177,283.66	176,701.08	176,406.08	51,495.83
1-Step RMSE	-	-	4.26	13.67
10-Step RMSE	-	-	4.43	14.91
30-Step RMSE	-	-	4.49	15.38
100-Step RMSE	-	-	4.51	15.59
ACP(1,1)				
log-Likelihood	-	-	-129,062.91	-44,596.71
BIC	-	-	258,242.66	89,290.55
1-Step RMSE	-	-	4.24	13.79
10-Step RMSE	-	-	4.39	15.70
30-Step RMSE	-	-	4.43	17.04
100-Step RMSE	-	-	4.46	17.46

Table 68: T - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.63	1.60	1.57	1.37
	(0.01)	(0.01)	(0.00)	(0.00)
b	2.70	2.20	1.90	1.82
	(0.06)	(0.11)	(0.02)	(0.01)
λ	0.74	0.40	0.37	0.33
	(0.02)	(0.02)	(0.01)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-99,385.88	-96,688.01	-94,534.43	-33,031.13
BIC	198,888.60	193,492.87	189,185.70	66,159.40
1-Step RMSE	-	-	12.07	35.23
10-Step RMSE	-	-	12.23	36.26
30-Step RMSE	-	-	12.24	36.28
100-Step RMSE	-	-	12.25	36.29
ACP(1,1)				
log-Likelihood	-	-	-333,220.04	-124,384.34
BIC	-	-	666,556.92	248,865.82
1-Step RMSE	-	-	12.02	34.72
10-Step RMSE	-	-	12.14	35.95
30-Step RMSE	-	-	12.19	36.20
100-Step RMSE	-	-	12.22	36.20

Table 69: TXN - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.58	1.55	1.52	1.30
	(0.00)	(0.00)	(0.00)	(0.01)
b	2.04	1.86	1.68	1.67
	(0.00)	(0.00)	(0.00)	(0.08)
λ	1.34	0.82	0.63	0.87
	(0.00)	(0.00)	(0.00)	(0.02)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-118,774.36	-116,247.40	-114,403.81	-34,904.37
BIC	237,665.56	232,611.65	228,924.47	69,905.88
1-Step RMSE	-	-	13.15	37.32
10-Step RMSE	-	-	13.22	38.13
30-Step RMSE	-	-	13.22	38.14
100-Step RMSE	-	-	13.21	38.17
ACP(1,1)				
log-Likelihood	-	-	-365,176.40	-108,797.99
BIC	-	-	730,469.65	217,693.11
1-Step RMSE	-	-	13.03	36.05
10-Step RMSE	-	-	13.08	37.19
30-Step RMSE	-	-	13.12	37.68
100-Step RMSE	-	-	13.15	38.08

Table 70: WFC - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.51	1.49	1.46	1.27
	(0.01)	(0.03)	(0.01)	(0.01)
b	2.59	2.22	1.87	2.33
	(0.06)	(0.18)	(0.12)	(0.17)
λ	1.27	1.25	1.03	1.32
	(0.08)	(0.13)	(0.23)	(0.07)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-151,590.49	-148,220.51	-145,697.60	-38,560.33
BIC	303,297.83	296,557.87	291,512.04	77,217.79
1-Step RMSE	-	-	18.32	57.39
10-Step RMSE	-	-	18.91	64.04
30-Step RMSE	-	-	18.93	66.18
100-Step RMSE	-	-	18.94	66.60
ACP(1,1)				
log-Likelihood	-	-	-433,061.79	-132,312.14
BIC	-	-	866,240.43	264,721.43
1-Step RMSE	-	-	17.81	55.57
10-Step RMSE	-	-	18.24	59.69
30-Step RMSE	-	-	18.38	61.56
100-Step RMSE	-	-	18.68	63.88

Table 71: WMT - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.58	1.55	1.52	1.31
	(0.09)	(0.03)	(0.01)	(0.01)
b	2.97	2.25	1.94	2.09
	(0.56)	(0.18)	(0.09)	(0.20)
λ	0.70	0.63	0.38	0.51
	(0.10)	(0.11)	(0.02)	(0.02)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-105,919.96	-104,162.00	-103,282.11	-31,113.47
BIC	211,956.76	208,440.84	206,681.07	62,324.08
1-Step RMSE	-	-	9.11	27.98
10-Step RMSE	-	-	9.29	29.73
30-Step RMSE	-	-	9.30	29.96
100-Step RMSE	-	-	9.31	30.06
ACP(1,1)				
log-Likelihood	-	-	-242,369.12	-83,474.68
BIC	-	-	484,855.08	167,046.50
1-Step RMSE	-	-	9.01	27.62
10-Step RMSE	-	-	9.19	29.22
30-Step RMSE	-	-	9.27	29.74
100-Step RMSE	-	-	9.31	29.92

Table 72: XRX - Estimation Results - $Q = 3$

	$\Delta = 10$ seconds			$\Delta = 1$ minute
MSMCP	$\bar{k} = 5$	$\bar{k} = 6$	$\bar{k} = 7$	$\bar{k} = 7$
m_0	1.73	1.69	1.66	1.48
	(0.01)	(0.03)	(0.02)	(0.00)
b	3.46	2.61	2.20	1.96
	(0.15)	(0.88)	(0.14)	(0.02)
λ	0.27	0.16	0.29	0.17
	(0.02)	(0.04)	(0.06)	(0.00)
γ^*	0.99	0.99	0.99	0.99
	(0.00)	(0.00)	(0.00)	(0.00)
log-Likelihood	-62,256.46	-60,148.08	-58,572.05	-25,606.06
BIC	124,629.77	120,413.00	117,260.95	51,309.26
1-Step RMSE	-	-	9.79	29.39
10-Step RMSE	-	-	9.95	30.56
30-Step RMSE	-	-	9.97	30.61
100-Step RMSE	-	-	9.97	30.61
ACP(1,1)				
log-Likelihood	-	-	-238,056.56	-114,770.52
BIC	-	-	476,229.96	229,638.18
1-Step RMSE	-	-	9.76	28.91
10-Step RMSE	-	-	9.91	30.00
30-Step RMSE	-	-	9.95	30.42
100-Step RMSE	-	-	9.95	30.48

BIBLIOGRAPHY

Andersen, T. G., T. Bollerslev, F. X Diebold, and P. Labys (2001), "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

Andersen, Torben G. and Tim Bollerslev (1998), "Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies," *Journal of Finance*, 53, 219–265.

Ané, Thierry and Hélyette Geman (2000), "Order Flow, Transaction Clock, and Normality of Asset Returns," *The Journal of Finance*, 55, 2259–2284.

Barndorff-Nielsen, Ole E., Peter Reinhard Hansen, Asger Lunde, and Neil Shephard (2011), "Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading," *Journal of Econometrics*, 162, 149–169.

Barndorff-Nielsen, Ole E. and Shephard (2002), "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society Series B*, 64, 253–280.

Bauwens, Luc and David Veredas (2004), "The stochastic conditional duration model: a latent variable model for the analysis of financial durations," *Journal of Econometrics*, 119, 381–412.

Calvet, Laurent and Adlai Fisher (2001), "Forecasting multifractal volatility," *Journal of Econometrics*, 105, 27–58.

Calvet, Laurent, Adlai Fisher, and Benoit Mandelbrot (1997), "A Multifractal Model of Asset Returns," Cowles Foundation Discussion Papers 1164, Cowles Foundation for Research in Economics, Yale University.

Calvet, Laurent E. and Adlai. J. Fisher (2004), "How to Forecast Long-Run Volatility: Regime Switching and the Estimation of Multifractal Processes," *Journal of Financial Econometrics*, 2, 49–83.

Calvet, L.E. and A. Fisher (2008), *Multifractal Volatility: Theory, Forecasting, and Pricing*, Academic Press advanced finance series, Academic Press.

Chen, F., Francis X. Diebold, and Frank Schorfheide (2013), "A Markov-Switching Multi-Fractal Inter-Trade Duration Model, with Application to U.S. Equities," *Journal of Econometrics*.

Clark, Peter K (1973), "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41, 135–55.

Cootner, P.H., E.F. Fama, W.S. Morris, H.M. Taylor, B.B. Mandelbrot, and R.E. Gomory (1997), *Fractals and Scaling In Finance: Discontinuity, Concentration, Risk*, Mandelbrot, Benoit B, Springer.

Daley, D. J. and D. Vere-Jones (2008), *"An Introduction to the Theory of Point Processes"*, Probability and its Applications (New York), Springer, New York, 2nd ed.

Deo, Rohit, Mengchen Hsieh, and Clifford M. Hurvich (2010), "Long memory in intertrade durations, counts and realized volatility of NYSE stocks," *Journal of Statistical Planning and Inference*, 140, 3715 – 3733, special Issue in Honor of Emanuel Parzen on the Occasion of his 80th Birthday and Retirement from the Department of Statistics, Texas A&M University Emmanuel Parzen.

Deo, Rohit, Clifford M. Hurvich, Philippe Soulier, and Yi Wang (2009), "Conditions For The Propagation Of Memory Parameter From Durations To Counts And Realized Volatility," *Econometric Theory*, 25, 764–792.

Durbin, J. and S.S.J. Koopman (2001), *Time Series Analysis by State Space Models*, Oxford statistical sciences series, Oxford University Press.

Easley, David, Robert F. Engle, Maureen O'Hara, and Liuren Wu (2008), "Time-Varying Arrival Rates of Informed and Uninformed Trades," *Journal of Financial Econometrics*, 6, 171–207.

Easley, David and Maureen O'Hara (1992), " Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 576–605.

Engle, Robert F. and Asger Lunde (2003), "Trades and Quotes: A Bivariate Point Process," *Journal of Financial Econometrics*, 1, 159–188.

Engle, Robert F. and Jeffrey R. Russell (1998), "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127–1162.

Ghysels, Eric, Christian Gourieroux, and Joann Jasiak (2004), "Stochastic volatility duration models," *Journal of Econometrics*, 119, 413–433.

Hamilton, James D (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–84.

Hansen, N. (2006), "The CMA evolution strategy: a comparing review," in "Towards a new evolutionary computation. Advances on estimation of distribution algorithms," (edited by Lozano, J.A., P. Larranaga, I. Inza, and E. Bengoetxea), 75–102, Springer.

Harris, L.E., Salomon Brothers Center for the Study of Financial Institutions, and Leonard N. Stern School of Business (1991), *Liquidity, Trading Rules, and Electronic Trading Systems*, Monograph series in finance and economics, New York University Salomon Center.

Hautsch, N. (2012), *Econometrics of Financial High-frequency Data*, Springer Berlin Heidelberg.

Hautsch, Nikolaus, Peter Malec, and Melanie Schienle (2010), "Capturing the Zero: A New Class of Zero-Augmented Distributions and Multiplicative Error Processes," SFB 649 Discussion Papers SFB649DP2010-055, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

Heinen, Andras (2003), "Modelling Time Series Count Data: an Autoregressive Conditional Poisson Model," CORE Discussion Papers 2003062, Universit catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Jasiak, Joann (1999), "Persistence in Intertrade Durations," Working papers, York University, Department of Economics.

Jung, Robert and A. Tremayne (2011), "Useful Models for Time Series of Counts or Simply Wrong Ones?" *AStA Advances in Statistical Analysis*, 95, 59–91.

Jung, Robert C. and A.R. Tremayne (2006), "Binomial Thinning Models for Integer Time Series," *Stat. Model.*, 6, 81–96.

Lux, Thomas and Taisei Kaizoji (2007), "Forecasting volatility and volume in the Tokyo Stock Market: Long memory, fractality and regime switching," *Journal of Economic Dynamics and Control*, 31, 1808–1843.

O'Hara, Maureen. (1995), *Market Microstructure Theory*, Blackwell Publishers Cambridge, Mass.

Rydberg, Tina Hviid and Neil Shephard (2003), "Dynamics of Trade-by-Trade Price Movements: Decomposition and Models," *Journal of Financial Econometrics*, 1, 2–25.

Shephard, Neil (1995), "Generalized linear autoregressions," Economics Papers 8., Economics Group, Nuffield College, University of Oxford.

Zeger, Scott L (1988), "A Regression Model for Time Series of Counts," *Biometrika*, 75, 621–629.

Zikes, Filip, Jozef Barunik, and Nikhil Shenai (2012), "Modeling and Forecasting Persistent Financial Durations," Papers 1208.3087, arXiv.org.