How Duration Between Trades of Underlying Securities Affects Option Prices[*]

Álvaro Cartea and Thilo Meyer-Brandis

Forthcoming: Review of Finance

May 21, 2009

Abstract

We propose a model for stock price dynamics that explicitly incorporates random waiting times between trades, also known as duration, and show how option prices can be calculated using this model. We use ultra-high-frequency data for blue-chip companies to motivate a particular choice of waiting-time distribution and then calibrate risk-neutral parameters from options data. We also show that the convexity commonly observed in implied volatilities may be explained by the presence of duration between trades. Furthermore, we find that, ceteris paribus, implied volatility decreases in the presence of longer durations, a result consistent with the findings of Engle (2000) and Dufour and Engle (2000) which demonstrates the relationship between levels of activity and volatility for stock prices. Finally, by directly employing information given by time-stamps of trades, our approach provides a direct link between the literature on stochastic time changes and business time (see Clark (1973)) and, at the same time, highlights the link between number and time of arrival of transactions with implied volatility and stochastic volatility models.

Keywords: Duration between trades, waiting-times, stochastic volatility, operational clock, transaction time, high frequency data.

JEL Classifications: G12, G13

Most financial models assume that securities are continuously traded. However, in equity markets for example, trading happens discretely at random times. In the literature there have been several approaches to directly model the times between trades also known as duration. Early models that capture the impact of duration between trades include Diamond and Verrechia (1987) and Easley and O'Hara (1992). The work of Easley and O'Hara establishes the link between the existence of information, the timing of trades and the dynamics of security prices. One of their main contributions is to show that duration between trades affects the behavior of security prices and consequently that transaction prices are not a Markov process, as is currently assumed in many financial models.

Using ultra-high-frequency equity data, Engle (2000) studies the consequences of stochastic trade arrival times (see also Engle and Russell (1998)). This empirical study finds evidence that both stock returns and variances are found to be negatively influenced by long durations between trades. The study of Dufour and Engle (2000) shows that the stochastic component of duration can explain the relationship between short time durations, i.e. high trading activity, and both larger quote revisions and stronger positive autocorrelations of trades.

Recent work by Aït-Sahalia and Mykland (2003) focuses on the estimation of continuous-time models and its consequences, in particular the fact that high-frequency financial data are discretely sampled in time and that the time separating successive observations is often random. One of the main messages emerging from their findings is that for empirical purposes, researchers using randomly spaced data, "... should pay as much attention, if not more, to sampling randomness as they do to sampling discreteness".

When it comes to derivative pricing, most financial literature on discrete time models assumes that the distribution of the waiting-time $\tau_n = T_n - T_{n-1}$ between the nth and $(n-1)$th trades, occurring at times T_n and T_{n-1} respectively, is either constant (tree models) or exponentially distributed (compound Poisson process models). This prompts two questions. Firstly, to what extent are these assumptions deviating from the 'true' distribution of durations? Secondly, how will this deviation from the 'true' empirical distribution impact derivative prices? The first question is not a new line of research in the literature, but the second, despite its importance in asset pricing, has received very little attention.

1

When looking at data that involves the random arrival of events, trades in our case, it is customary to look at what is known as the survival function, which represents the probability that the waiting-time between two consecutive trades is greater than t. This function is given by

$$\Upsilon(t) = 1 - \int_0^t \upsilon(u)du, \tag{1}$$

where $\upsilon(t)$ denotes the probability density function (pdf) of the waiting times.

If we assume that the waiting-time between trades possesses an exponential distribution with parameter λ, then $\upsilon(t) = \lambda e^{-\lambda t}$ and $\Upsilon(t) = e^{-\lambda t}$. Employing General Motors (GM) consolidated trades (over the period April-June 2005) in Figure 1, as an example we show a log-log plot of empirical and fitted exponential survival functions.[1] We used 419,264 trades from all exchanges with a resulting average duration between consecutive trades of $\tau_o^e = 5.26$ seconds. The Figure also shows that the fitted exponential survival function with parameter $\lambda = 1/\tau_o^e$, (the dashed line), is a very poor fit when compared to empirical data (circles).[2]
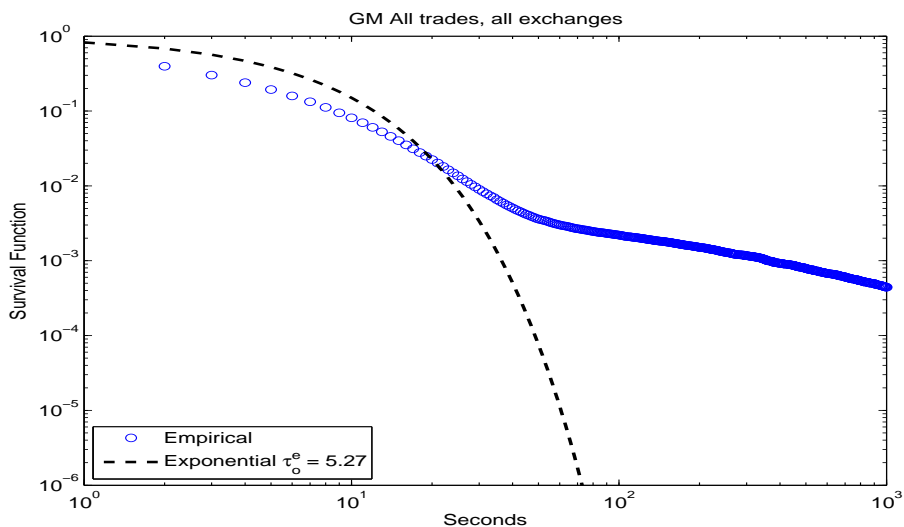
Figure 1. General Motors waiting-times: empirical and exponential.

Intuitively, the rationale for rejecting the exponential survival function as a possible candidate to model durations is its inability to capture the long durations between consecutive trades, see for example Engle (2000), Engle and Russell (1998) and Dufour and Engle (2000). Furthermore, assuming that the duration between consecutive trades is exponentially distributed is equivalent to assuming that the

2

number of trades follows a Poisson counting process or that the intensity of arrival of trades is constant. If this were the case, then the mean and variance of the data should be the same, a property known as 'equidispersion'. In fact, what is commonly observed in ultra-high-frequency models is 'overdispersion', i.e. where the variance is greater than the mean of the data, see Cameron and Trivedi (1996). For example, in the case of GM the variance of waiting times for trades is $3.4575 * 10^3$, while the mean is 5.27.

In general, assuming a specific distribution for the duration between trades is equivalent to assuming a particular trade arrival intensity. The seminal work of Clark (1973) argues that markets operate at different trading rates over different periods of time. This is due to the fact that information becomes available to traders at a varying rate. For instance, during periods when there is very little information trading is slow, while on days when new relevant information becomes available trading pace picks up and the price process evolves more quickly. To put it another way, markets operate on a 'business time', as opposed to a calendar time, basis.[3]

In Clark's work, it is argued that the trading volume is positively related to the number of intraday transactions and hence that the trading volume is related to the variability of price change. This positive relationship between price changes and trading volume, without controlling for the number of transactions, is further investigated in Epps and Epps (1976), Tauchen and Pitts (1983), Gallant, Rossi, and Tauchen (1992) and Karpoff (1987). However, in Jones, Kaul, and Lipson (1994) it is shown that the positive relationship between volume and volatility, measured as absolute or squared price changes, actually reflects the positive relationship between volatility and the number of transactions. In Jones et al it is further argued that, on average, the size of trades has virtually no incremental information content; any information about the trading behavior of agents is almost entirely contained in the frequency of trades during a particular interval. Finally, the finding that it is the number of trades, rather than the volume, that generates price volatility is further supported by Ross (1989). Ross's study shows that, under certain assumptions, the variance of price changes is directly proportional to the variance of the flow of information, a quantity that is better proxied by the number of transactions than by the volume of trades (see also Andersen (1996)).

Although our approach presents several new features, it retains a strong connection to the more mature and established literature that examines the relationship between price returns, volume and the

number of transactions. In this article, we concentrate on the question of how derivatives prices are calculated when durations possess a distribution function that better reflects the observed empirical behavior. Our contribution is threefold. Firstly, we propose a general model that explicitly incorporates waiting times as one of the building blocks of stock price dynamics under the physical measure. In particular, one of the key elements of our approach is to exploit the idea of time-changes by using transaction times, given by the high-frequency time-stamps of trades, to describe the "business time". Secondly, we show how option prices are calculated by choosing a risk-adjusted measure. Thirdly, based on empirical waiting-time data from blue-chip companies, we investigate a particular distribution for duration and we employ it to calibrate risk-neutral parameters to IBM options data.

Under the risk-adjusted measure we propose the use of a survival function that can capture long waits between trades and that nests, as a particular case, the exponential survival function. We then calibrate our model to IBM options data and find that in the vast majority of the cases the risk-neutral parameters of the stock dynamics responsible for modeling the duration between trades, indicate that the risk-neutral distribution of waiting times is not exponential.

As another illustration of our model, we chose to isolate the effect of the waits by calculating option prices with a survival function that allows for the possibility of long waits and the distribution of stock price revisions can be either Gaussian or CGMY (see Carr, Geman, Madan, and Yor (2002)). We observe that for different maturities the inclusion of waiting-times that are not exponentially distributed contribute to the implied volatility observed in financial markets. In particular, when we assume that price revisions are Gaussian, which asymptotically behaves like the classical Black-Scholes framework, the inclusion of non-exponential waiting-times is solely responsible for the emergence of the convexity in the volatility 'smile'. We also observe that, ceteris paribus, implied volatility decreases when waiting times are 'longer', a finding in line with those of Engle (2000) and Dufour and Engle (2000) which links the relationship of levels of activity and volatility for stock prices.

The rest of this article is organized as follows. Section 1 proposes a general model for stock prices, under the statistical measure, where duration between trades is random. Section 2 focuses on the pricing of instruments such as European-style options. Section 3 justifies the selection of particular waiting-time distributions and shows how European-style option prices may be calculated by employing widespread techniques such as those in Carr and Madan (1999). Section 4 calibrates risk-neutral

4

parameters for one of our models, using IBM options data. Section 5 produces numerical examples of how duration affects the shape and level of implied volatility. Section 6 concludes.

1. The Model: spot dynamics with duration

In this section, we propose a model which needs to satisfy three requirements. Firstly, every time a trade occurs stock prices must undergo a stochastic price revision. Secondly, the model must be able to explicitly incorporate the (random) duration between trades. Thirdly, the model must be capable of pricing basic financial instruments such as European-style options and one must be able to calibrate its risk-neutral parameters to the market.

Before presenting the model we need two more definitions: a counting process; and the hazard function. We denote the time of the nth trade by T_n and the duration between trades by $T_n - T_{n-1} = \tau_n$ with continuous pdf $\upsilon(t)$. Hence we can write

$$T_n = T_0 + \sum_{i=1}^{n} \tau_i, \qquad T_n - T_{n-1} = \tau_n, \qquad n = 1, 2, 3, \cdots.$$

The counting process, which represents the number of trades over the interval $[0, t]$, is defined by

$$N_t = \max\{n \geq 0 | T_n \leq t\}.$$

Further, the hazard function $u(t)$ is defined as

$$u(t) = -\frac{d}{dt} \ln \Upsilon(t), \qquad t \in \mathbb{R}^+, \tag{2}$$

where the survival function $\Upsilon(t)$ is that given above in equation (1). Intuitively, the hazard function represents the probability that a trade will happen in the next small time interval divided by the length of that time interval; i.e. the hazard function is the instantaneous intensity of a trade occurrence. Here we assume that $u(t)$ is strictly positive and continuous.

Stock price revisions

To model the stock price revisions, we assume that every time there is a trade, i.e. the counting process N_t increases by one unit, the price revision of the logarithm of the stock price $X(t) = \ln S(t)$ moves by i.i.d. Y. More precisely, we assume that the dynamics of the observed tick-by-tick microstructure of $X(t)$, under the physical measure $\mathbb{P}$, are described by

$$X(t) = X(0) + (r - D)t + \sum_{i=1}^{N_t} Y_i, \tag{3}$$

where the constants r and D denote the risk-free rate and the dividend yield. Note that for technical convenience, we consider a continuously compounded risk-free behavior with rate $(r - D)$ instead of capturing this deterministic trend in the jump price revisions $\sum_{i=1}^{N_t} Y_i$. At jump times (i.e. when there is a trade) there is no price difference between these two alternatives. However, with the continuous rate technicalities are simplified when it comes to derivatives pricing in section 2 below. We assume that the i.i.d. spacial shocks Y, which are independent of the waiting times, possess an infinitely divisible distribution. Given the above, the log-characteristic function of Y is given by the Lévy-Khintchine representation

$$\ln \mathbb{E}\left[e^{i\xi Y}\right] \equiv \Psi(\xi) = ai\xi - \frac{1}{2}\sigma^2\xi^2 + \int_{\mathbb{R}\setminus\{0\}} \left(e^{i\xi l} - 1 - i\xi l 1_{|l|<1}\right) W(dl). \tag{4}$$

Here $a \in \mathbb{R}$, $\sigma \geq 0$, the truncation function $l 1_{|l|<1}$ ensures integrability around the origin, and $\Psi(\xi)$ is known as the characteristic exponent of the distribution with triplet (a, σ^2, W). For technical simplicity, we assume that the distribution of the spacial shocks Y is given by a continuous density $g(y) > 0$, $y \in \mathbb{R}$. Note that if we denote by $N(\omega, dt, dz) = N(dt, dz)$ the integer valued jump measure associated with the process $\sum_{i=1}^{N_t} Y_i$, we can rewrite the dynamics (3) as[4]

$$X(t) = X(0) + (r - D)t + \int_0^t \int_{\mathbb{R}_0} zN(dt, dz). \tag{5}$$

In the financial literature, the two most common models of the type described in equation (3) are: discrete time models (tree models) with deterministic, equally spaced, time steps τ_n; and compound Poisson models where the τ_n's are i.i.d. exponentially distributed, random variables. In the latter, $X(t)$

belongs to the class of Lévy processes which have been extensively studied and applied in finance over the recent years.

For example, a conditionally Gaussian model arises when it is assumed that price revisions in (3) arise from a Gaussian distribution, with $Y \sim N(\mu, \sigma^2)$, and that the counting process N_t is a homogeneous Poisson process, which is equivalent to assuming that the waiting-time distribution between trades is exponential. However, as is well known, the Gaussianity of price revisions is not supported by empirical studies, especially over short-time periods. Most efforts to improve these models have focused on the spacial shocks aspect, as opposed to the distribution of the waiting times τ, despite the crucial role that these waiting times play in the distributional properties of stock prices.

A major reason why people only reluctantly depart from exponentially distributed waiting times, is the loss of Markovianity (even if empirical studies confirm the non-Markovianity of prices). Indeed Markovianity is important for many issues, including derivatives pricing, where expectations conditioned on past market evolution have to be computed. With the exception of the exponential waiting-time distribution, the log-stock $X(t)$ is not Markovian for a general waiting-time distribution in model (3). Indeed, let $H(\omega, t) = H(t)$ denote the so-called backward recurrence time (i.e. the time elapsed since the last trade) defined by

$$H(t) = t - T_{N_t}, \tag{6}$$

where T_{N_t} represents the last trade time before t. Then it is well known that the intensity of the counting process N_t is given by $u(H(t))$, see Jacobsen (2006). Consequently, the predictable compensator of the jump measure $N(dt, dz)$ is the random measure

$$\nu(\omega, dt, dz) = \nu(dt, dz) := u(H(t))g(z)\, dt dz, \tag{7}$$

where $u(t)$ is the hazard function given in (2) and $g(z)$ the probability density of the shocks Y. From this it follows that the process is not Markovian as long as $u(t)$ is not constant. Intuitively, for general hazard functions $u(t)$, it is important to know the time elapsed since the last trade and thus the process is not memoryless. However, if we enlarge the state space with the backward recurrence time $H(t)$, then we have the following result.

Theorem 1 *The two-dimensional process $(X(t), H(t))$ is a time-homogeneous Markov process.*

This is an important property which we will use below to price options. For a proof see appendix A.

A special example is the well-known case resulting from the assumption that the waiting times τ are exponentially distributed with parameter λ. For this particular case, the survival function is given by $\Upsilon(t) = e^{-\lambda t}$ and the hazard function becomes $u(t) = \lambda$; note that the hazard function is independent of the backward recurrence time $H(t)$. In this case the compensating measure (7) becomes $\nu(dt, dz) = \lambda g(z) \, dt \, dz$, which is the compensating measure of the compound Poisson process $X(t)$, and it is not necessary to consider the two-dimensional process $(X(t), H(t))$ because $X(t)$ already is Markovian.

Another important characteristic of our model for the price dynamics (3) is that it captures many of the important features that relate the volume of transactions to (stochastic) volatility of returns. Note that a direct interpretation of our model is that the market operates on an operational clock where the instrument measuring the evolution of business time is the number of trades N_t (see Clark (1973)). Moreover, our model also reconciles the feature that volatility is not constant over identical time intervals; it is in fact stochastic. To see this last point, note that the conditional expected quadratic variation of the log-return between time s and t is given by

$$\mathbb{E}\left[\left(\int_s^t \int_{\mathbb{R}_0} z N(dr, dz)\right)^2 | \mathcal{F}_s\right] = \mathbb{E}\left[\int_s^t \int_{\mathbb{R}_0} z^2 u(H(r)) g(z) \, dr \, dz | \mathcal{F}_s\right] = f(H(s)) \int_{\mathbb{R}_0} z^2 g(z) \, dz$$

where $f(H(s))$ is defined by

$$f(H(s)) := \mathbb{E}\left[\int_s^t u(H(r)) \, dr | \mathcal{F}_s\right] = \mathbb{E}\left[\int_0^{t-s} u(H^h(r)) \, dr\right]_{h=H(s)}.$$

Here, we have used the Markov property of the process H (see Appendix A), and the notation H^h denotes that the initial value of H at time 0 is h. In this sense, unlike models based on Lévy processes, the conditional expected quadratic variations of log-returns are stochastic and are adapted to the backward recurrence time $H(s)$ (i.e. the time elapsed since the last trade). Also, the autocorrelation in the process H then leads to autocorrelation in squared log returns.[5]

2. Derivatives Pricing

One of the key requirements we have imposed on our model for stock price dynamics is that we can price financial instruments, such as European-style options written on the underlying stock $S(t)$. Therefore, in the first part of this section, we discuss the possible risk-neutral dynamics exhibited by $S(t)$ when we assume that, under the physical measure $\mathbb{P}$, the stock price follows (3). In the second part we then proceed to discuss derivatives pricing and derive an integro-pde characterization for the price process of European-style options in our framework. Further, under the assumption that a trade just has happened, we derive a second price description based on Fourier transform techniques which is much more efficient in practice both to price, and more importantly, to calibrate risk-neutral parameters.

On our stochastic basis $(\Omega, \mathcal{F}, \mathbb{P})$, let $\mathcal{F}_t$ be the filtration generated by the stock price $S(t)$; note that the same filtration is generated by the two-dimensional process $(X(t), H(t))$. Since $S(t)$ is obviously a semimartingale, theory tells us that we must specify an equivalent martingale measure (EMM) $\mathbb{Q}$, under which risk-neutral pricing of financial instruments, written on $S(t)$, can be performed. Generally, one of the consequences of employing asset models that capture jumps in prices is that the market is incomplete and that there is no unique EMM under which pricing of derivatives is achieved. In fact, there are many (possibly an infinite number of) EMMs and it is the market that selects the one EMM, under which pricing of instruments is performed and under which arbitrage opportunities are precluded.

The vast majority of models assume that calendar time and operational time coincide, i.e. when the calendar clock 'evolves' there is always a trade or the possibility to transact. In our model this is not the case, transaction time does not always coincide with calendar time and this element alone is an important source of market incompleteness. The work of Bossaerts, Ghysels, and Gouriéroux (1997) analyzes, in a discrete-time setting, the consequences of modeling price dynamics on two clocks: calendar and transaction time. The authors show that some of the simple no-arbitrage restrictions, such as put-call-parity for European options, might fail due to the fact that portfolio rebalancing can only occur in transaction time. This is at odds with the usual no-arbitrage results that stem from the assumption that calendar and transaction times are the same. Moreover, Bossaerts et al also discuss the conditions under which it is possible to price derivatives by arbitrage in highly incomplete markets (caused by the impossibility to hedge in continuous time).

In our model, we seek to specify a family of potential EMMs by allowing flexibility in the choice of the pricing measure and by then employing market data to identify which EMM the market is selecting. Importantly, when proposing the set of EMMs, we do not assume that under the pricing measure, transaction and calendar time are the same. Therefore, focusing on the waiting-time element of the model, the family of EMMs that we propose assumes that the market operates on a different clock from calendar time. Thus, one way to proceed is to assume that the stock dynamics under the risk-adjusted measure have the same structure as under the physical measure. In particular, we assume that the number of trades will be independent from price revisions, but we allow the distribution of the number of trades under the risk-neutral measure to differ from the distribution of the number of trades under the physical measure.[6] In addition, we also allow the distribution of the spacial shocks under the risk-neutral measure to differ from the equivalent distribution under the physical measure.

More precisely, we assume that the market chooses from a class of EMMs whose densities with respect to $\mathbb{P}$ is given by the following stochastic exponentials

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(\int_0^t \int_{\mathbb{R}_0} \ln(\phi(z)\alpha(\omega,t)) N(dt,dz) - \int_0^t \int_{\mathbb{R}_0} (\phi(z)\alpha(\omega,t) - 1) \nu(dt,dz) \right), \qquad (8)$$

where the function $\phi(z)$ and the predictable process $\alpha(\omega,t)$ are such that (8) is a well defined $\mathbb{P}$-martingale. Further, we assume that $g_Q(z) = \phi(z)g(z)$ is the density of an infinitely divisible distribution satisfying

$$\int_{\mathbb{R}} (e^z - 1)g_Q(z)\,dz = 0, \qquad (9)$$

and that $\alpha(\omega,t)u(H(t))$ takes the form $u_Q(H(t))$ for a strictly positive and continuous hazard function $u_Q(t)$. Using Girsanov's theorem for random measures (see Jacod and Shiryaev (2002)), the jump measure $N(dt,dz)$ has the $\mathbb{Q}$-predictable compensator

$$\nu_Q(\omega,dt,dz) = \nu_Q(dt,dz) := u_Q(H(t))g_Q(z)\,dt\,dz, \qquad (10)$$

which has the same structure as the predictable compensator (7) under the $\mathbb{P}$ measure. It is straightforward to see from the structure of the $\mathbb{Q}$-compensator (10) that the log-stock price

$$
\begin{aligned}
X(t) &= X(0) + (r-D)t + \sum_{i=1}^{N_t} Y_i \\
&= X(0) + (r-D)t + \int_0^t \int_{\mathbb{R}_0} z N(dt, dz)
\end{aligned}
$$

has the same renewal process structure under $\mathbb{Q}$, as it has under $\mathbb{P}$. The alteration is only a different, but equivalent infinitely divisible distribution for the spacial shocks Y given through the density $g_Q(z)$, which is such that $\mathbb{E}^{\mathbb{Q}}[e^Y - 1] = 0$, as well as a different hazard function $u_Q(t)$ characterizing the waiting times. Now, the discounted stock price $e^{-(r-D)t}S(t)$ is given by

$$
e^{-(r-D)t}S(t) = S(0) \exp \left(\int_0^t \int_{\mathbb{R}_0} z N(dt, dz) \right).
$$

Because of condition (9) we can rewrite $e^{-(r-D)t}S(t)$ as

$$
e^{-(r-D)t}S(t) = S(0) \exp \left(\int_0^t \int_{\mathbb{R}_0} z N(dt, dz) - \int_0^t \int_{\mathbb{R}_0} (e^z - 1) v_Q^2(dt, dz) \right), \tag{11}
$$

which is an exponential martingale under $\mathbb{Q}$. Consequently, under the above conditions, (8) determines indeed a class of EMM.

Having specified a pricing measure $\mathbb{Q}$ from the above defined class, we now consider pricing of instruments written on $S(t) = \exp(X(t))$. Let F be a pay-off function of a European option with maturity T written on $S(t)$. Then the price process of this option is given as

$$
V(t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[F(S(T))|\mathcal{F}_t] \qquad 0 \leq t \leq T.
$$

Note that considering a European option written on $S(t)$ is equivalent to considering a European option written on $X(t)$ with pay-off function $G = F(\exp(\cdot))$. Thus, the value process of the option above can be rewritten as

$$
V(t) = e^{-r(T-t)} \mathbb{E}^{\mathbb{Q}}[G(X(T))|\mathcal{F}_t].
$$

11

Now, because of the time-homogeneous Markov structure of $(X(t), H(t))$, we can write

$$V(t) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[G(X(T))|X(t), H(t)] = e^{-r(T-t)}\mathbb{E}^x_{\mathbb{Q}}[G(X^h(T-t))] \mid_{x=X(t), h=H(t)} . \qquad (12)$$

Here, $X^h(t)$ is the h-delayed renewal process starting in x, induced by $X(t)$, i.e. the first waiting-time in (3) has the distribution of $(\tau_1 - h)$, given $\tau_1 > h$. Furthermore, from (A2) and (A3) it follows that the generator of the Markov process $(X(t), H(t))$ is given by the integro-differential operator O, defined as follows:

$$Of(x, h) = (r - D)\frac{\partial}{\partial x}f(x, h) + \frac{\partial}{\partial h}f(x, h) + \int_{\mathbb{R}_0} \{f(x+z, 0) - f(x, h)\} u_Q(h)g_Q(z)\,dz, \qquad (13)$$

for $f \in C_0^{1,1}(\mathbb{R}^2)$. Here, $C_0^{1,1}(\mathbb{R}^2)$ is the space of continuous functions, with compact support and continuous derivatives in x and h. Then, with the usual Feynman-Kac considerations, we obtain the following description of the price process $V(t)$.

Theorem 2 *Let $F(\cdot)$ be the pay-off function of a European option with maturity T written on the stock $S(t)$. Let the function $G(\cdot) := F(\exp(\cdot))$ be the composition of F and* exp*, and assume that there exists a bounded solution $v(t, x, h) \in C^{1,1,1}([0, T], \mathbb{R}, \mathbb{R}_+)$ of the integro-pde*

$$\begin{cases} 0 = \frac{\partial}{\partial t}v(t, x, h) + Ov(t, x, h) \\ \\ v(T, x, h) = G(x), \quad (t, x, h) \in [0, T] \times \mathbb{R} \times \mathbb{R}_+. \end{cases} \qquad (14)$$

Then, the price at time t of the European option with pay-off $F(\cdot)$, and maturity T, is given as

$$V(t) = e^{-r(T-t)}v(t, X(t), H(t)).$$

Note that in the special case of an exponential waiting time distribution with parameter λ, the generator (13) becomes

$$Of(x, h) = (r - D)\frac{\partial}{\partial x}f(x, h) + \frac{\partial}{\partial h}f(x, h) + \int_{\mathbb{R}_0} \{f(x+z, 0) - f(x, h)\} \lambda g_Q(z)\,dz.$$

Thus, if a function $v'(t,x) \in C^{1,1}([0,T],\mathbb{R})$ solves

$$\begin{cases} 0 = \frac{\partial}{\partial t}v'(t,x) + \mathcal{O}'v'(t,x) \\ \\ v'(T,x) = G(x), \ (t,x) \in [0,T] \times \mathbb{R}, \end{cases} \tag{15}$$

where the generator $\mathcal{O}'$ is defined as

$$\mathcal{O}'f(x) = (r-D)\frac{\partial}{\partial x}f(x) + \int_{\mathbb{R}_0} \{f(x+z) - f(x)\}\lambda g_Q(z)\,dz,$$

$f \in C_0^1(\mathbb{R})$, then $v(t,x,h) := v'(t,x)$ solves (14). Consequently, for exponentially distributed waiting times, we obtain the usual pricing integro-pde (15) for compound Poisson processes which is independent of h.

The integro-pde representation of the option price (14) provides a method for computing option prices in our model. However, an alternative way to calculate prices of European-style instruments is to use transform methods (Carr and Madan (1999), Carr and Wu (2003)). These methods are very efficient and powerful to calibrate risk-neutral parameters from market data. Here we present the general result which we employ below in subsection 3.1.1, when we choose a particular survival function, to calibrate parameters to IBM options data in Section 4.

Proposition 1 *Let $F(\cdot)$ be the pay-off function of a European option with maturity T written on the stock $S(t)$, and let $G(\cdot)$ be as in Theorem 2. Assume that $\hat{q}(\xi,t,T)$, defined by*

$$\hat{q}(\xi,t,T) := \mathbb{E}^Q\left[e^{i\xi\sum_{i=N_t+1}^{N_T} Y_i} \mid \mathcal{F}_t\right], \tag{16}$$

is analytic in ξ in a strip that intersects the strip where the (complex) Fourier transform of G exists. Let $\hat{\xi} \in \mathbb{R}$ be such that the line $[-\infty + i\hat{\xi}, \infty + i\hat{\xi}]$ is part of this intersection. Then the value at time t of the option is given by

$$V(t) = \frac{e^{-r(T-t)}}{2\pi} \int_{-\infty+i\hat{\xi}}^{\infty+i\hat{\xi}} e^{-i\xi \ln S(t)} e^{-i\xi(r-D)(T-t)} \hat{q}(-\xi,t,T)\hat{G}(\xi)d\xi. \tag{17}$$

where the notation $\hat{G}(\xi) = \mathcal{F}[G(x)] = \int_{-\infty}^{\infty} e^{ix\xi}G(x)dx$ denotes the Fourier transform of $G(\cdot)$.

For a proof see appendix A.

We note that, depending on the assumptions regarding the waiting-time distribution $v(t)$, and/or the counting process N_t, expression (16) can be calculated analytically and the evaluation of European-style option prices becomes a straightforward task.

2.1. Modeling the leverage effect

One assumption in our model is that waiting-times and spatial shocks Y are independent. However, studies indicate that this assumption might not be supported by empirical evidence. In the literature of time-changed Lévy process models, this phenomenon is often referred to as leverage effect, see Barndorff-Nielsen and Shephard (2001). For example, in Carr, Geman, Madan, and Yor (2003) the authors propose to capture the leverage effect by extending their model, based on time-changed Lévy processes, by adding the rate of time-change to the log-stock price. One way to include the leverage effect in our framework is to assume dependence between waiting times and spacial shocks. We assume that, at time t, the probability density function of the spacial shocks g depends on the backward recurrence time $H(t)$. More precisely, the predictable compensator of the jump measure $N(dt, dz)$ is assumed to be of the form

$$v(\omega, dt, dz) = v(dt, dz) := u(H(t))g(H(t); z)\, dt\, dz,$$

where $u(t)$ is the hazard function given in (2) and $g(H(t); z)$ is the probability density function of the shocks Y with an additional parametric dependence on the stochastic process $H(t)$. For example, for Gaussian spacial shocks we can assume the expected jump size to be a function $f(H(t))$ of the backward recurrence time, i.e.

$$g(H(t); z) := \psi(v, \sigma^2; z)\,|_{v=H(t)},$$

where $\psi(v, \sigma^2; z)$ denotes the Gaussian density with expectation v and variance σ^2.

Although in this version of the model with leverage effects it does not seem possible to calculate the characteristic function of the log-stock price we can still proceed as in Theorem 2 to calculate

derivatives prices by solving the associated integro-pde. In this case, the integro-differential operator O takes the form

$$Of(x,h) = (r-D)\frac{\partial}{\partial x}f(x,h) + \frac{\partial}{\partial h}f(x,h) + \int_{\mathbb{R}_0} \{f(x+z,0) - f(x,h)\}u_Q(h)g_Q(h;z)\,dz,$$

for $f \in C_0^{1,1}(\mathbb{R}^2)$.[7]

3. Empirical survival function

In this section we look at empirical waiting-times of 23 blue-chip companies during the period April-June 2005. Our sample of stocks includes those from Dufour and Engle (2000) that were still being traded in 2005. All data were obtained from the TAQ database made available via WRDS.

Before proposing a model that captures the main properties of the empirical survival functions we address the question of how to treat the relatively frequent occurrences of consecutive trades when the duration between them is reported in the system with zero. From a practical point of view, time-stamps for every trade are rounded to the nearest second. A direct consequence of this is that trades that occur within the same second are recorded as if they had taken place simultaneously. On the other hand, there are cases when one trade is broken into various batches and these too are recorded as simultaneous trades. A common approach adopted in the literature has been to delete these trades. For instance, in our data set of IBM trades there are 178,512 durations of zero seconds. Deleting these observations would amount to discarding more than 28% of the 631,586 waits between trades.

Ideally, if one could discern which zero-duration trades are part of a large trade broken into batches, then these could be deleted and the remaining zero-duration trades could be kept by assigning them a waiting-time strictly greater than zero. From a mathematical standpoint, if we view the question of modeling durations as modeling the number of trades occurring on a given interval, we know that counting processes such as Poisson will assign zero probability to events where two or more trades take place at the same time. Therefore the need to assign waiting times that occurred within a second, but recorded as simultaneous trades, a duration strictly greater than zero. Instead of discarding all zero-duration observations the alternative we propose is to remove only those data points where there

was a zero waiting-time but there was no change in the price of the trade. For example, of the 178,512 instances of zero-duration in the restricted IBM data, 103,391 could be eliminated because they were accompanied by no change in price. The remaining 75,121 data points where price changes were different from zero were retained and were assigned a duration strictly greater than zero.[8] In Table 1, we show, for each stock, the number of data points omitted due to zero waiting times and no price changes (column "Out") and those included through assignment of a non-zero waiting-time (column "In").

Co	Out	In	All Trades	τ_o^e
GE	227,431	84,404	620,370	3.96
IBM	103,391	75,121	528,195	4.27
GM	115,967	61,966	419,264	5.27
MO	63,480	34,527	364,331	5.98
PG	60,038	29,458	365,800	5.54
AMD	89,449	30,209	333,248	6.59
SLB	48,283	30,200	356,341	5.41
KO	53,113	23,066	342,880	5.61
BA	52,328	26,201	323,436	6.12
AA	47,733	19,267	298,566	6.43
FNM	39,579	22,055	296,854	6.13
FDX	3,0545	21,407	260,044	7.31
CL	23,235	9,948	201,127	8.93
FPL	16,015	10,344	188,586	9.27
CAL	22,243	5,895	164,403	10.94
CAG	14,707	7,674	167,293	10.71
T	13,892	5,249	156,005	11.58
PCO	10,159	7,640	155,465	11.56
VC	18,366	6,756	130,115	14.45
HNZ	10,552	3,242	132,931	13.19
NI	8,294	3,144	105,780	16.42
POM	2,407	2,132	69,986	24.51
GTI	3745	979	62,016	27.51

Table 1

Empirical waiting-time data. The second column, under the heading "Out", indicates the number of data points, for each stock, that were discarded because a zero wait was also accompanied by a zero price change. The third column, under the heading "In", shows the number of data points which were kept because although there was a zero wait, price changes were not zero. The fourth column indicates therefore the number of data points used as duration between trades in our study. Finally the fifth column is the average waiting time (in seconds) for the data set.

3.1. Shifted-Mittag-Leffler survival function

The most conspicuous message from Figure 1 is the presence of relatively 'long' durations. These long durations are impossible to capture with an exponential waiting-time distribution, and, as we shall see below, the presence of these long waits between trades is not unique to GM. The appendix shows 22 other companies that exhibit broadly the same shaped survival function as GM. Hence, we will justify a choice of waiting-time distribution by specifying a model that can capture the right tail of the survival function, i.e. long waits.

The first step is to observe that the shape of the right tail of the survival function, in log-log space, in Figure 1 closely resembles that of a straight line with a negative slope. It is straightforward to see that this linear behavior in a log-log plot is equivalent to observing the behavior of data that is changing with a power law. In other words the (ln-)tail of the survival function shows the behavior

$$\ln \Upsilon(t) \sim -\beta \ln t + \ln a + \cdots, \qquad \text{as} \quad t \to \infty, \tag{18}$$

where $\beta > 0$ and a are constants.[9] Since from (1) we obtain the pdf of the waiting times by differentiating the survival function

$$\upsilon(t) = -\frac{d}{dt}\Upsilon(t),$$

we can use (18) to find the tail behavior of the pdf of the waiting-time distribution:

$$\ln \upsilon(t) \sim -(\beta + 1)\ln t + \ln(a\beta) + \cdots, \qquad \text{as} \quad t \to \infty. \tag{19}$$

Now that we are able to capture the crucial behavior of long waits via (19), or equivalently via (18), we take the second step and justify the choice of a waiting-time distribution. We recall that we want to be able to use our model for stock dynamics in order to price European-style options. In addition, we would like to specify a waiting-time distribution so that expression (16) in Proposition 1 can be performed analytically.

Instead of working with the tail expression of $v(t)$ given by (19), we look at its Laplace transform. Hence, we can write the tail of the waiting-time distribution in Laplace space as[10]

$$\tilde{\upsilon}(s) \sim 1 - (\tau_o s)^\beta + o(s^\beta), \qquad \text{for } 0 < \beta \le 1, \tag{20}$$

where $\tau_o > 0$ is a constant.

However, we are still left with the question of finding a suitable waiting-time distribution since we have only specified the functional form of the tail to capture the long waits. Note that there are many waiting time distributions that could exhibit a slow decay of the right tail, as shown in (20). However not all of them will deliver mathematically tractable expressions capable of being employed by standard pricing tools, and more importantly, will not facilitate the calibration of risk-neutral parameters to observed vanilla option prices (see for example Carr and Madan (1999)). Hence, below we specify $v(t)$ for all $t \ge 0$ by choosing a distribution function that allows us to calculate the characteristic function (16).

We proceed by noting that one possible choice of $\tilde{\upsilon}(s)$, consistent with (20), is given by

$$\tilde{\upsilon}(s) = \frac{1}{1 + (\tau_o s)^\beta}, \qquad \text{for } 0 < \beta \le 1. \tag{21}$$

Moreover, the Laplace transform of the survival function is given by

$$\tilde{\Upsilon}_{ML}(s) = \frac{1 - \tilde{\upsilon}(s)}{s} = \tau_o \frac{(\tau_o s)^{\beta - 1}}{1 + (\tau_o s)^\beta}, \qquad \text{for } 0 < \beta \le 1, \tag{22}$$

and by taking the inverse Laplace transform of (22), see equation (A7) in the appendix, the survival function becomes

$$\Upsilon_{ML}(t) = \sum_{j=0}^{\infty} (-1)^j \frac{(t/\tau_o)^{\beta j}}{\Gamma(\beta j + 1)}, \qquad \text{for } 0 < \beta \le 1, \tag{23}$$

which is known in the literature as the Mittag-Leffler (ML), or as a generalized, exponential function. Furthermore, we make the important observation that when $\beta = 1$ the waiting-time distribution becomes the exponential with expected value $\mathbb{E}[\tau] = \tau_o$. Hence, we can view the ML survival function as a generalization of the exponential survival function that accommodates long waits between trades when $\beta < 1$; something an exponential waiting-time distribution is unable to capture.

We employ a slight modification of (23), by including a shift parameter τ_s in the time-domain of the survival function. The intuition behind this trivial modification is to recognize that the time-stamps in our data are rounded to the nearest second. Consequently the data set are left-truncated, which therefore makes it reasonable to include a shift in the domain of the survival function to improve the statistical fitting of the ML survival model. Figure 2 shows empirical and fitted survival functions. We show (shifted) ML and exponential functions. As expected, the exponential function is not capable of capturing the long waits. Moreover, Table 2 shows the results of fitting the shifted ML parameters to all the stocks studied here and the appendix depicts the fitted distributions.

Another route to study empirical waiting times has been to restrict the data set to trading hours between 9.30am and 4.00pm and focus only on trades via NYSE. For example, in this restricted case, the IBM data set would consist of 331,057 trades as opposed to the 528,195 when all exchanges are taken into account and trading before 9.30am and after 4.00pm is also considered. Moreover, previous studies focusing on this restricted data set have found that the Weibull distribution is a good model, however it is not capable of capturing long waits. Moreover, we point out that our main objective is to explicitly model durations and to study their impact on option prices. Therefore, the choice of risk-neutral survival function is what matters when measuring the impact durations have on derivatives pricing.
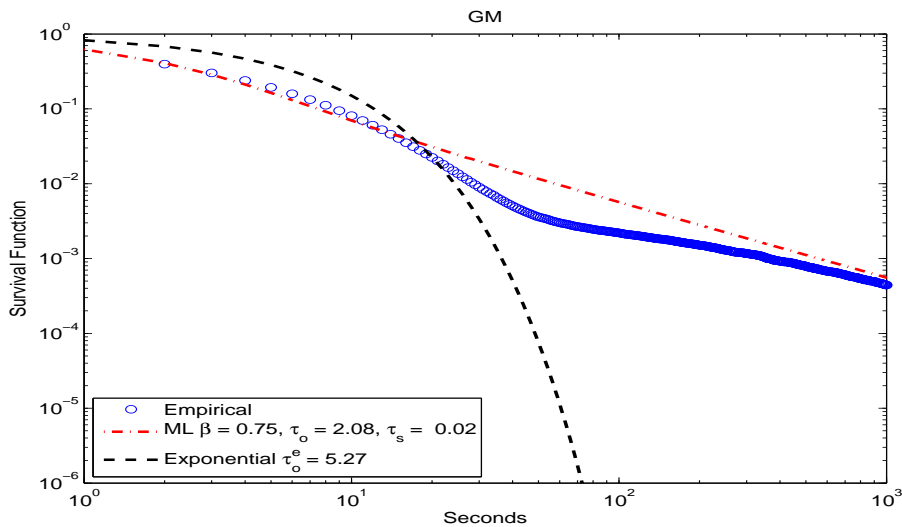

Figure 2. Fitted survival functions for GM

20

Co	τ_o	95% CI	τ_s	95% CI	β	95% CI
GE	1.5060	(1.4717, 1.5403)	0.1103	(0.0998, 0.1208)	0.8423	(0.8293, 0.8554)
IBM	1.7391	(1.6934, 1.7849)	0.0805	(0.0662, 0.0949)	0.8185	(0.8032, 0.8338)
GM	2.0860	(2.0178, 2.1543)	0.0257	(0.0043, 0.0470)	0.7584	(0.7385, 0.7784)
MO	2.6705	(2.6056, 2.7354)	0.0430	(0.0204, 0.0657)	0.8143	(0.8003, 0.8282)
PG	2.8441	(2.7791, 2.9092)	0.0523	(0.0280, 0.0765)	0.8579	(0.8453, 0.8705)
AMD	3.2402	(3.1589, 3.3214)	0.0519	(0.0227, 0.0812)	0.8135	(0.7990, 0.8281)
SLB	2.5343	(2.4663, 2.6022)	0.0833	(0.0608, 0.1059)	0.8000	(0.7841, 0.8159)
KO	2.8949	(2.8278, 2.9621)	0.0747	(0.0505, 0.0990)	0.8398	(0.8266, 0.8530)
BA	2.6259	(2.5444, 2.7074)	0.0773	(0.0516, 0.1030)	0.7556	(0.7363, 0.7750)
AA	3.2311	(3.2065, 3.2556)	0.2206	(0.2179, 0.2233)	0.6452	(0.6408, 0.6497)
FNM	2.8925	(2.7935, 2.9915)	0.0647	(0.0325, 0.0969)	0.7583	(0.7371, 0.7795)
FDX	2.9691	(2.8446, 3.0937)	0.0431	(0.0056, 0.0806)	0.6847	(0.6565, 0.7128)
CL	4.5001	(4.3526, 4.6477)	0.2319	(0.2319, 0.2319)	0.7585	(0.7351, 0.7819)
FPL	4.6416	(4.4736, 4.8096)	0.2349	(0.2349, 0.2349)	0.7351	(0.7086, 0.7616)
CAL	5.2955	(5.1344, 5.4566)	0.2268	(0.2268, 0.2268)	0.7389	(0.7167, 0.7611)
CAG	5.5407	(5.3650, 5.7165)	0.2340	(0.2340, 0.2340)	0.7610	(0.7382, 0.7837)
T	6.1676	(6.0003, 6.3349)	0.2368	(0.2368, 0.2368)	0.7786	(0.7595, 0.7978)
PCO	4.5137	(4.3330 4.6944)	0.2258	(0.2258, 0.2258)	0.6039	(0.5707, 0.6372)
VC	5.8712	(5.6332, 6.1093)	0.2076	(0.2076, 0.2076)	0.6260	(0.5929, 0.6591)
HNZ	7.2854	(7.0743, 7.4964)	0.2345	(0.2345, 0.2345)	0.7791	(0.7585, 0.7997)
NI	9.0244	(8.7679, 9.2809)	0.2409	(0.2409, 0.2409)	0.7573	(0.7366, 0.7780)
POM	14.2969	(13.8032, 14.7907)	0.2439	(0.2439, 0.2439)	0.7518	(0.7262, 0.7775)
GTI	14.7941	(14.3078, 15.2803)	0.2403	(0.2403, 0.2403)	0.7200	(0.6950, 0.7451)

Table 2
Shifted ML parameter estimates for τ_o, τ_s (in seconds) and β using ultra-high-frequency data for the trading period April 1st through June 30th 2005.

A further feature of our model is that it is not necessarily the case that the parameter β goes to one and that asset prices become Markovian when trading occurs more frequently. One way to see this is to look at the expected number of trades between time s and t, $s < t$. Assume for simplicity that the shift parameter is $\tau_s = 0$ and that at time s a trade has just occurred, i.e. that the time elapsed since the last trade is zero. Hence, we have that[11]

$$\mathbb{E}\left[N_T \middle| \mathcal{F}_s\right] = \frac{((T-s)/\tau_o)^\beta}{\Gamma(\beta+1)}, \tag{24}$$

and it is possible to observe, for fixed $\beta < 1$, that the expected number of trades can become arbitrarily large (by assuming an arbitrarily small τ_o) while the non-Markovianity property of the model is preserved because $\beta < 1$.

3.1.1. European-style options with ML survival function

If we assume that, under the risk-neutral measure, the survival function has the form (21) then the problem of pricing European-style options (see Proposition 1) reduces to deriving (16). Furthermore, in this particular case, calculations get simplified if we assume that a trade just happened, i.e. $H(0) = 0$, and for simplicity we also assume that $\tau_s = 0$. Given the high frequency of trade arrivals, assuming $H(0) = 0$ is reasonable. The following Theorem shows how European-style options are priced when the survival function of the waiting times is ML.

Theorem 3 *Assume that the prerequisites from Proposition 1 hold. Additionally, assume that the survival function is ML, with $\tau_s = 0$, and that $H(0) = 0$. Then the value of the European-style option is given by*

$$V(0) = \frac{e^{-rT}}{2\pi} \int_{-\infty+i\hat{\xi}}^{\infty+i\hat{\xi}} e^{-i\xi \ln S(0)} e^{-i\xi(r-D)T} E_{\beta,1} \left[-\left(1 - e^{\Psi(\xi)}\right) (T/\tau_o)^\beta \right] \hat{G}(\xi) d\xi. \tag{25}$$

For a proof see appendix A.

Regarding the choice of $\hat{\xi}$ in the integration limits in Theorem 3, we require $E_{\beta,1}\left[-\left(1 - e^{\Psi(\xi)}\right) (T/\tau_o)^\beta \right]$ to be analytic in a strip that intersects the strip where the (complex) Fourier transform of the $G(\cdot)$ exists. The ML function (A6) is an entire function; therefore it is analytic where $e^{\Psi(-\xi)}$ is analytic. Thus, the restrictions on $\hat{\xi}$ are the same as those required in the particular case when $\beta = 1$, i.e. when pricing with Lévy processes.[12] For example, if we let $\beta = 1$, we can verify that the price of a European call option with strike K and maturity T, using (25), is given by

$$V(0;K,T) = -\frac{e^{-rT}K}{2\pi} \int_{-\infty+i\hat{\xi}}^{\infty+i\hat{\xi}} e^{-i\xi \ln S(0) + T\left[-i\xi(r-D) + (\Psi(-\xi)-1)\tau_o^{-1}\right]} \frac{K^{i\xi}}{\xi^2 - i\xi} d\xi,$$

for $\hat{\xi} > 1$.[13]

4. Estimation of risk-neutral parameters

In this section we present results obtained from calibrating risk-neutral parameters to IBM option prices. We obtained data on IBM American options from OptionMetrics (Ivy DB) via WRDS. The data include the spot price, strike, maturity, dividend yield, interest rate and the Black-Scholes implied volatility. The algorithm employed by OptionMetrics to calculate the risk-neutral parameters and the sensitivities for American options is based on the Cox, Ross, and Rubinstein (1979) binomial tree model. The interest rate used in the calculations is based on BBA Libor rates and settlement prices of CME Eurodollar futures. For example, for a given option, the appropriate interest rate input corresponds to the zero-coupon rate that has maturity equal to the options expiration, and is obtained by linearly interpolating between the two closest zero-coupon rates of the curve.[14]

We used the parameters from the American options to devise a new data set of European options. We then used the algorithm employed in Carr and Wu (2003) to estimate the risk-neutral parameters of our model by considering two cases. In the first case, we assume that price revisions possess a Gaussian distribution and that the waiting-time survival function is the ML function. In the second case, we still assume that the waiting-time survival function is the ML function but now suppose that price revisions possess an FMLS distribution (Carr and Wu (2003)). Moreover, in order to compare the performance of our model to a well known benchmark, we also calibrate the standard Black-Scholes model.

The tables in Appendix C show the results for every trading day from April 1 through May 6 2005. In any given day we have IBM options for different strikes and for different maturities. We show the results of the calibration for the lot of IBM options with shortest maturity (including all strikes), then we add to these results the next lot, which includes those options with second shortest maturity, and so on.[15] For example, the first row in Table 4 shows risk-neutral parameters obtained from 6 options (trading in April 1 2005) that expired in 10 working days (i.e. the first lot). For this lot, the resulting implied volatility in the classical Black-Scholes model is $\sigma_{bs} = 0.1446$, the volatility of Gaussian price revisions and the beta of the model are $\sigma = 0.0951$ and $\beta = 0.718300$ respectively, and for FMLS price revisions $\alpha = 1.99827$, $\sigma = 0.06713$ and $\beta = 0.720381$.[16] In the second row, we show the results of the calibration procedure when we take into account the options that expire between 10 and 35 working days. For ease of presentation of the results for the volatility parameter σ we show it in an 'annualized'

	$T \le 20$	$T \le 40$	$T \le 60$	$T \le 80$	$T \le 100$	$T \le 120$	$T \le 140$	$T \le 160$	All
Black-Scholes	0.35530	0.26448	0.25026	0.25533	0.25533	0.25325	0.25083	0.25429	0.26222
Gaussian	0.29878	0.23766	0.22370	0.21240	0.21240	0.21105	0.20228	0.19489	0.22696
FMLS	0.16171	0.13827	0.12882	0.12830	0.12830	0.12560	0.11656	0.11239	0.13406

Table 3
RMSE for lots of options with up to different expiries. For example, column 2 shows the square root of the weighted average of the MSE shown in the table in Appendix C, for the three models, where the weights are the number of options with expiry date less than or equal to 20 days for every day over the period 1 April to 6 May 2005. The RMSEs in the other columns are calculated in a similar way.

form. For example, if for a given lot of options, with different expiries and strikes, the calibrated parameter is $\hat{\sigma}$, the table shows

$$\sigma = \hat{\sigma}\sqrt{\frac{(T/\tau_o)^\beta}{\Gamma(\beta+1)}}$$

with $T = 1$ year and $\tau_o = 2,112,780$.[17]

Table 3 shows the performance, by looking at the RMSEs, of the models with Gaussian or FMLS shocks where durations are captured by the ML waiting-time survival function. The table also shows the RMSEs for the classical Black-Scholes model. Each column in the table shows (for options with expiries up to $T = \{20, 40, 60, 80, 100, 120, 140, 160, \text{all}\}$) the square root of the weighted average of the MSE where the weights are the number of options that in every day of the sample expire before the chosen threshold. An interesting observation to note is that the performance of the 'Gaussian with ML waiting-times' model is always better than that of the Black-Scholes. For instance: for options with expiries $T \le 20$, the RMSE is 0.35530 for the Black-Scholes model and is 0.29878 for the 'Gaussian with ML waiting-times' model; and for the whole sample (i.e. all expiries) the RMSE is 0.26222 and 0.22696 for the Black-Scholes and the 'Gaussian with ML waiting-times' models respectively. Furthermore, the table also shows that for the different sets of expiry dates the RMSE for the FMLS is always lower than that of the Gaussian and Black-Scholes models.

One of the messages implied by the results is that the effect of long durations (captured by the parameter β) on option prices prevails across all maturities. It is interesting to note that this is true for both the Gaussian and FMLS cases and although the βs are not the same for both models, they do not appear to be too dissimilar for each particular day and surface we calibrate to. We interpret this as a

good sign since, especially in the Gaussian example we study, the parameter β could be accommodating for kurtosis of the risk-neutral distribution which, is produced by the spacial shocks, and is 'picked up' by the parameter β. In the next section we see how the presence of long durations ($\beta < 1$) increases the kurtosis of the risk-neutral distribution of spot prices.

5. Numerical examples: the impact of waiting times on option prices

In the previous section, we looked at the calibration of risk-neutral parameters for models that explicitly include waiting times between trades. Here, to gain more insight into the consequences of including durations, we present two examples of how waiting times affect option prices. These are calculated by choosing plausible risk-neutral parameters, so that we can focus on the effects of assuming the ML survival function. The first example assumes that the spacial shocks are Gaussian and the second example assumes that spacial shocks possess a CGMY distribution (see Carr, Geman, Madan, and Yor (2002)). In all examples we assumed that $\tau_s = 0$ and that $\tau_o = 1/1,200,000$, (i.e. that there are, on average, $100,000^\beta/\Gamma(\beta+1)$ trades per month, see (24)).

5.1. Gaussian price revisions and ML waiting-times

Figure 3 shows implied volatility (IV) when it is assumed that spacial shocks are Gaussian with mean zero and volatility $\sigma = 0.3\sqrt{\tau_o}$. With this choice of volatility, and letting $\beta = 1$, the model is asymptotically equivalent to assuming a Black-Scholes model with volatility $\sigma_{bs} = 0.30$. The Figure shows IV for different waiting times by choosing $\beta = \{0.98, 0.96, 0.94, 0.92\}$ whilst all other parameters remain unchanged. It is possible to see that the steeper IV becomes for out-of-the-money and in-the-money values the further away the parameter β is from the exponential case $\beta = 1$. This is interesting since it shows that the inclusion of waiting times that are not exponential, gives rise to the commonly observed convexity of the IV in the Black-Scholes framework despite the fact that spacial shocks are Gaussian.[18] Note that the waiting time affects the convexity of the IV in a symmetric way and does not reproduce smirks or skewed IVs. In our framework, market participants include a premium, over and above the classical Black-Scholes price for out-of-the-money values, to price in the duration times between trades.

Another important feature of Figure 3 is the fact that the IV decreases as β decreases. For example, when $\beta = 0.98$ and expiry is $T = 20$ days, IV is roughly within $[0.265, 0.27]$ whereas when $\beta = 0.94$ and $T = 20$, IV is in the range $[0.21, 0.22]$. This result is not surprising, and is in line with the findings of Engle (2000) and Dufour and Engle (2000). Indeed in our model, the market will exhibit less activity (understood here as number of trades over a time period) and lower IV the lower β is. This is also clear in Figure 4 where, still with Gaussian spacial shocks, we fix expiry dates and vary $\beta = \{1, 0.98, 0.96, 0.94, 0.92\}$ where the exponential case is included.
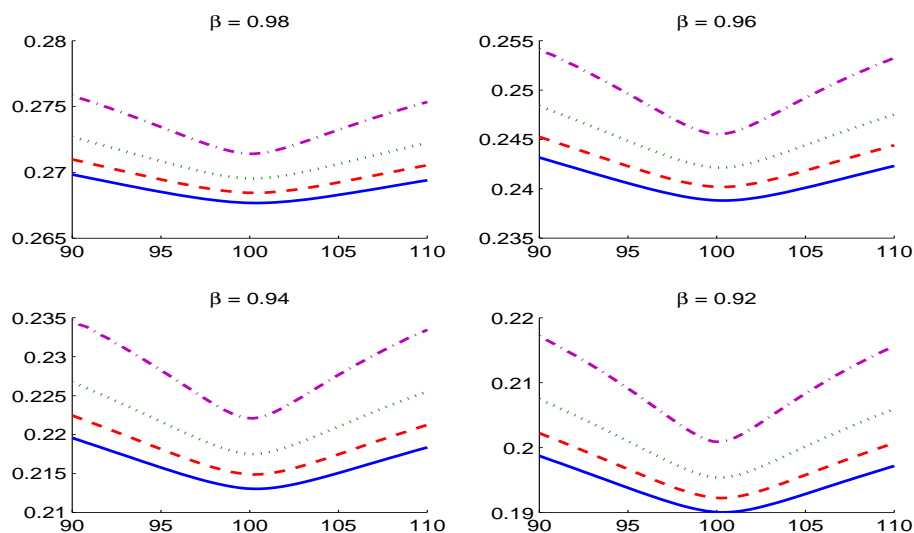

Figure 3. IV across strikes for conditionally Gaussian model with waiting times for $\beta = \{0.98, 0.96, 0.94, 0.92\}$. The volatility of the zero-mean Gaussian price revisions is $\sigma = 0.3\sqrt{\tau_o}$, and the parameters for option pricing are $r = 5\%$, $D = 0$ and $S_0 = 100$. The dash-dotted line corresponds to $T = 5$ days, the dotted line $T = 10$ days, the dashed line $T = 15$ days, and the solid line $T = 20$ days.

5.2. CGMY price revisions and ML waiting-times

In this subsection we produce the same results as above, but we allow the distribution of price revisions to exhibit fatter tails than the Gaussian distribution by choosing price revisions with a CGMY distribution, see Carr, Geman, Madan, and Yor (2002). In our examples below, we assumed that $C = 1.8750 \times 10^{-7}$, $Y = 1.5$, $G = 10$, $M = 20$, this implies that the distribution of the spacial shocks has negative asymmetry because $G < M$, and both the left and right tails of the distribution of spacial shocks are heavier than those of a Normal distribution.
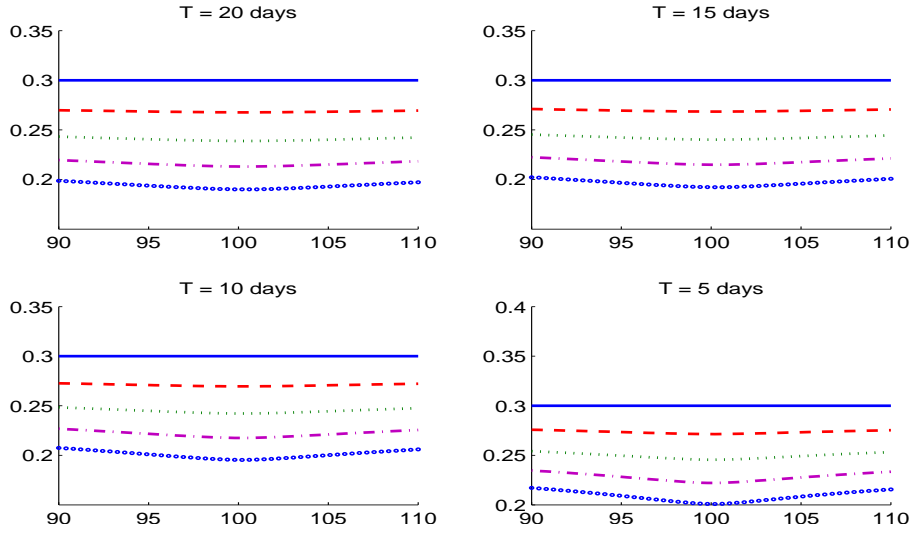
Figure 4. IV across strikes for conditionally Gaussian model with waiting times for different days to maturity $T = \{20, 15, 10, 5\}$ and varying $\beta = \{1, 0.98, 0.96, 0.94, 0.92\}$. The volatility of the zero-mean Gaussian price revisions is $\sigma = 0.3\sqrt{\tau_o}$, and the parameters for option pricing are $r = 5\%$, $D = 0$ and $S_0 = 100$. Each panel shows how IV varies when expiry remains fixed and β varies. The solid line represents $\beta = 1$, the dashed line corresponds to $\beta = 0.98$, the dotted line corresponds to $\beta = 0.96$, the dash-dotted line corresponds to $\beta = 0.94$, and circles corresponds to $\beta = 0.92$.

In Figure 5 we show IVs across strikes and maturities for different waiting time distributions by choosing $\beta = \{0.98, 0.96, 0.94, 0.92\}$. In this case we observe that the IVs are more pronounced, as well as skewed, than those observed in the Gaussian case; a finding more in line with what we observe in the financial markets. We also observe that for out-of-the-money options the shorter the maturity is, the steeper the IVs are. Moreover, in Figure 6 we fix the maturity of the option and show that when we move away from the Markovian case $\beta = 1$ the convexity of the volatility smile becomes more prominent and considerably more pronounced than those obtained with Gaussian price revisions depicted in, for example, Figure 4.
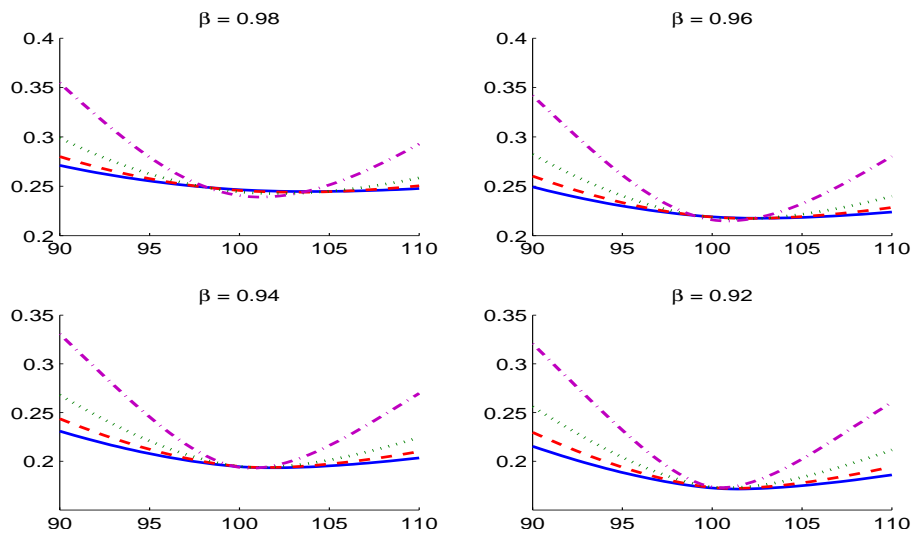
Figure 5. IV across strikes for conditionally CGMY model with waiting times for $\beta = \{0.98, 0.96, 0.94, 0.92\}$. The dash-dotted line corresponds to $T = 5$ days, the dotted line $T = 10$ days, the dashed line $T = 15$ days, and the solid line $T = 20$ days.
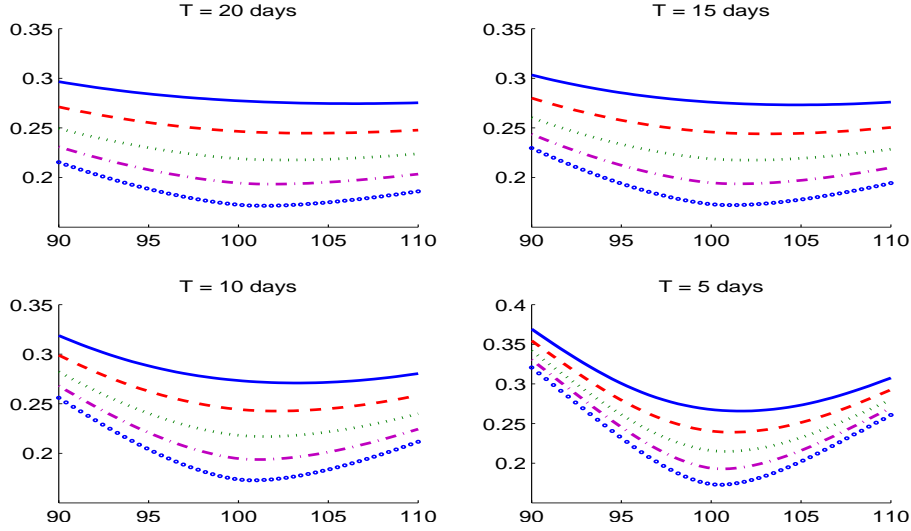
Figure 6. IV across strikes for CGMY with waiting times for different days to maturity $T = \{20, 15, 10, 5\}$ and varying $\beta = \{1, 0.98, 0.96, 0.94, 0.92\}$. Each panel shows how IV varies when expiry remains fixed and β varies. The solid line represents $\beta = 1$, the dashed line corresponds to $\beta = 0.98$, the dotted line corresponds to $\beta = 0.96$, the dash-dotted line corresponds to $\beta = 0.94$, and circles corresponds to $\beta = 0.92$.

6. Conclusions

Until now, the financial literature has only considered the question of how waiting-times or duration between trades affect the dynamics of stock prices. The question of how this random duration affects derivative prices, has not received much attention. In this article we propose a model that explicitly incorporates these waiting-times by directly exploiting the arrival of trades in the model. Besides capturing duration between trades, our model also captures key behavioral characteristics recorded in the empirical literature such as the non-Markovianity of stock prices, Easley and O'Hara (1992). Furthermore, by directly employing information given by time-stamps of trades, our approach provides a direct link between the literature on stochastic time changes and business time (see Clark (1973)) and, at the same time, highlights the link between number and time of arrival of transactions with IV and stochastic volatility models.

In our model we make the working assumption that waiting-times and spacial shocks are independent. Although this assumption is not endorsed by empirical data, it allows us great flexibility in the modeling of spacial shocks; for example it allows us to assume that price revisions have an infinitely divisible distribution. For this general case, we are able to price European-style options by solving an

integro-pde where the standard Lévy-based models (assuming exponentially distributed duration) are a particular case.

We propose the use of the ML survival function as a candidate to model waiting times. One of the main advantages is that with the ML it is straightforward to use the usual transform methods employed in the Lévy process literature relating to finance to price options. As an example, we calibrated risk-neutral parameters, using IBM options data, to a model with ML waits and Gaussian price revisions and to a model with ML waits and FMLS price revisions. In both cases the effects of durations were captured by risk-neutral βs, which were in the vast majority of cases less than one.

As another illustration of our model, we chose to isolate the effect of the waits by calculating options prices with ML waits and Gaussian revision and with ML waits and CGMY price revisions. We saw that for different maturities the inclusion of waiting-times that are not exponentially distributed contribute to the IV observed in financial markets. In particular, when we assume that price revisions are Gaussian, as described by the classical BS framework, the inclusion of waiting-times ($\beta < 1$) is solely responsible for the emergence of the convexity in the volatility 'smile'. Moreover, we see that the level of activity (as the number of trades, over a given time period) is higher the larger β is. We also observe that, ceteris paribus, IV decreases in β a finding in line with those of Engle (2000) and Dufour and Engle (2000) which links the relationship of levels of activity and volatility for stock prices.

References

Aït-Sahalia, Yacine, and Per A. Mykland, 2003, The Effects of random and discrete sampling when estimating continuous-time diffusions, *Econometrica* 71, 483–549.

Andersen, Torben G., 1996, Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility, *The Journal of Finance* 51, 169–204.

Barndorff-Nielsen, Ole E., and Neil Shephard, 2001, Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics, *Journal of the Royal Statistical Society* Series B, 167–241.

Bossaerts, Peter, Eric Ghysels, and Christian Gouriéroux, 1997, Arbitrage-Based Pricing When Volatility is Stochastic, Social Science Working Paper 977, Caltech.

Cameron, A. Colin, and Pravin K. Trivedi, 1996, Count data models for financial data, in G.S. Maddala, and Calyampudi Rao, eds.: *Handbook of Statistics* (Elsevier, North-Holland, Amsterdam,).

Carr, Peter, Hélyette Geman, Dilip Madan, and Marc Yor, 2002, The fine Structure of asset returns: an empirical investigation, *Journal of Business* 75, 305–332.

Carr, Peter, Hélyette Geman, Dilip B. Madan, and Marc Yor, 2003, Stochastic Volatility for Lévy Processes, *Mathematical Finance* 13, 345–382.

Carr, Peter, and Dilip Madan, 1999, Option Valuation Using the Fast Fourier Transform, *Journal of Computational Finance* 2, 61–73.

Carr, Peter, and Liuren Wu, 2003, The Finite Moment Logstable Process and Option Pricing, *The Journal of Finance* LVIII, 753–777.

Carr, Peter, and Liuren Wu, 2004, Time-Changes Lévy Processes and Option Pricing, *Journal of Financial Economics* 71, 113–141.

Clark, Peter K., 1973, A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, *Econometrica* 41, 135–155.

Cox, John C., Stephen A. Ross, and Mark Rubinstein, 1979, Option pricing: A simplified approach, *Journal of Financial Economics* 7, 229–263.

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

Dufour, Alfonso, and Robert F. Engle, 2000, Time and the Price Impact of a Trade, *The Journal of Finance* LV, 2467–2498.

Easley, David, and Maureen O'Hara, 1992, Time and the Process of Security Price Adjustment, *The Journal of Finance* XLVII, 577–605.

Engle, Robert F., 2000, The Econometrics of Ultra-High-Frequency Data, *Econometrica* 68, 1–22.

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

Epps, Thomas W., and Mary Lee Epps, 1976, The Stochastic Dependence of Security Price Changes and Transaction Volumes: Implications for the Mixture-of-Distributions Hypothesis, *Econometrica* 44, 305–21.

Gallant, A. Ronald, Peter E. Rossi, and George Tauchen, 1992, Stock Prices and Volume, *Review of Financial Studies* 5, 199–242.

Jacobsen, Martin, 2006, *Point Process Theory and Applications*. (Birkhäuser).

Jacod, Jean, and Albert N. Shiryaev, 2002, *Limit Theorems for Stochastic Processes* vol. 228 of *A Series of Comprehensive Studies in Mathematics*. (Springer).

Jones, Charles M., Gautam Kaul, and Marc L. Lipson, 1994, Transactions, volume and volatility, *The Review of Financial Studies* 7, 631–651.

Karpoff, Jonathan M., 1987, The Relation between price changes and trading volume: a survey, *Journal of Financial and Quantitative Analysis* 22, 109–126.

Mainardi, Francesco, Marco Raberto, Rudolf Gorenflo, and Enrico Scalas, 2000, Fractional Calculus and Continuous-time Finance II: the waiting-time distribution, *Physica A* 287, 469–481.

Osborne, M.F.M., 1959, Brownian Motion in the stock market, *Operations Research* 7, 145–173.

Podlubny, Igor, 1999, *Fractional Differential Equations* vol. 198 of *Mathematics in Science and Engineering*. (Academic Press San Diego, California) 1st edn.

Ross, Stephen A., 1989, Information and Volatility: The No-Arbitrage Martingale Approach to Timing and Resolution Irrelevancy, *Journal of Finance* 44, 1–17.

Tauchen, George E., and Mark Pitts, 1983, The Price Variability-Volume Relationship on Speculative Markets, *Econometrica* 51, 485–505.

Appendix A. Proofs of propositions and the ML function

Proof Theorem 1.

We show that $(X(t), H(t))$ is described by a stochastic differential equation (SDE), whose coefficients only depend on the process itself. Then it is well known that $(X(t), H(t))$ is a time homogenous Markov process.

In between trades, the backward recurrence time $H(t)$ defined in (6) evolves linearly in t and reverts to zero each time there is a jump in $X(t)$. Therefore $H(t)$ follows the dynamics given by the SDE

$$dH(t) = dt - H(t^-)\,dN_t = dt - \int_{\mathbb{R}_0} H(t^-)z N^1(dt, dz).$$

where $N^1(\omega, dt, dz) = N^1(dt, dz)$ denotes the integer valued random measure that represents the jump measure of the counting process N_t. The intensity of the counting process N_t is given by $u(H(t))$ (see e.g. Jacobsen (2006)) where the hazard function $u(t)$ is given by (2). We can write the predictable compensating measure of $N^1(dt, dz)$ as

$$\nu^1(\omega, dt, dz) = u(H(t))\,dt\delta_1(dz), \tag{A1}$$

where $\delta_1(dz)$ is the Dirac measure centered at 1.

Then it follows that the multivariate dynamics of the two-dimensional process $(X(t), H(t))$ is described by

$$\begin{aligned}
\begin{pmatrix} dX(t) \\ dH(t) \end{pmatrix} &= \begin{pmatrix} r-D \\ 1 \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & -H(t^-) \end{pmatrix} \begin{pmatrix} d\sum_{i=1}^{N_t} Y_i \\ dN(t) \end{pmatrix} \\
&= \begin{pmatrix} r-D \\ 1 \end{pmatrix} dt + \int_{\mathbb{R}_0^2} \begin{pmatrix} z_1 \\ -H(t^-)z_2 \end{pmatrix} N^2(dt, dz_1, dz_2)
\end{aligned} \tag{A2}$$

where $N^2(\omega, dt, dz_1, dz_2) = N^2(dt, dz_1, dz_2)$ denotes the jump measure of the two-dimensional process $(X(t), N(t))$ on $\mathbb{R}_+ \times \mathbb{R}^2 \setminus \{0\}$. Since the two processes $X(t)$ and $N(t)$ jump at exactly the same times, but with independently distributed jump sizes, the predictable compensator of $N^2(dt, dz_1, dz_2)$ is given by

$$\nu^2(\omega, dt, dz_1, dz_2) = u(H(t))g(z_1)\,dt dz_1\delta_1(dz_2). \tag{A3}$$

Thus the two-dimensional process $(X(t), H(t))$ is described by SDE (A2) with Lipschitz continuous coefficients and predictable compensator that only depend on the process $(X(t), H(t))$ itself (more precisely, on the second component $H(t)$). Then it is well known that $(X(t), H(t))$ is a time-homogenous Markov process.

∎

Proof Proposition 1.

We will denote the Fourier transform of a function $g(x)$ by

$$\mathcal{F}[g(x)] = \hat{g}(\xi) = \int_{-\infty}^{\infty} e^{ix\xi} g(x)dx,$$

where $\xi \in \mathbb{C}$. Hence, assuming the pay-off $G(\cdot)$ is such that we can invert its Fourier transform,

$$
\begin{aligned}
V(t) &= e^{-r(T-t)}\mathbb{E}^Q[G(X(T))\,|\,\mathcal{F}_t] \\
&= e^{-r(T-t)}\mathbb{E}^Q\left[\frac{1}{2\pi}\int_{-\infty+i\xi_i}^{\infty+i\xi_i} e^{-i\xi X_T}\hat{G}(\xi)d\xi\,|\,\mathcal{F}_t\right] \\
&= \frac{e^{-r(T-t)}}{2\pi}\int_{-\infty+i\xi_i}^{\infty+i\xi_i} e^{-i\xi \ln S(t)}e^{-i\xi(r-D)(T-t)}\mathbb{E}^Q\left[e^{i\xi \sum_{i=N_t+1}^{N_T} Y_i}\,|\,\mathcal{F}_t\right]\hat{G}(\xi)d\xi,
\end{aligned}
\tag{A4}
$$

where $\mathbb{E}^Q$ denotes the risk-neutral expectation operator.

$\blacksquare$

Proof Theorem 3.

We will denote the Laplace transform of a function $f(t)$ by

$$\mathcal{L}[f(t)] = \tilde{f}(s) = \int_0^{\infty} e^{st}f(t)dt.$$

Further, we assume $H(0) = 0$, i.e. a trade just happened. It will be useful to have an expression for the probability density function $P(n,t)$ of observing n trades during the time interval $[0,t]$. Using the survival function (1) the probability that a trade does not take place before time t is given by

$$P(n=1,t) = \int_0^t \upsilon(s)\Upsilon(t-s)ds = (\upsilon \star \Upsilon)(t),$$

where $\star$ denotes convolution. Then the probability of observing n trades over the interval $[0,t]$ is given by $(\upsilon^n \star \Upsilon)(t)$ and taking its Laplace transform yields

$$\tilde{P}(n,s) = \tilde{\upsilon}(s)^n\tilde{\Upsilon}(s) = \tilde{\upsilon}(s)^n\frac{1-\tilde{\upsilon}(s)}{s}.\tag{A5}$$

Therefore, from Proposition 1, we need to calculate

$$
\begin{aligned}
\hat{q}(\xi, 0, T) &= \mathbb{E}^{Q}\left[e^{i\xi \Sigma_{i=1}^{N_T} Y_i}\right] \\
&= \mathbb{E}^{Q}\left[e^{(N_T)\Psi(\xi)}\right] \\
\mathcal{L}\{\hat{q}(\xi, 0, T)\} &= \mathcal{L}\left\{\mathbb{E}^{Q}\left[e^{(N_T)\Psi(\xi)}\right]\right\} \\
&= \mathcal{L}\left\{\sum_{0}^{\infty} P(n, T) e^{n\Psi(\xi)}\right\} \\
&= \sum_{0}^{\infty} \mathcal{L}\{P(n, T)\} e^{n\Psi(\xi)} \\
&= \sum_{0}^{\infty} \tilde{P}(n, s) e^{n\Psi(\xi)} \\
&= \sum_{0}^{\infty} \tilde{\upsilon}(s)^n \frac{1 - \tilde{\upsilon}(s)}{s} e^{n\Psi(\xi)} \\
&= \frac{1 - \tilde{\upsilon}(s)}{s} \sum_{0}^{\infty} \tilde{\upsilon}(s)^n e^{n\Psi(\xi)} \\
&= \frac{1 - \tilde{\upsilon}(s)}{s} \frac{1}{1 - e^{\Psi(\xi)}\tilde{\upsilon}(s)},
\end{aligned}
$$

where $\tilde{\upsilon}$ is given by (21). Then

$$
\begin{aligned}
\hat{q}(-\xi, 0, T) &= \mathcal{L}^{-1}\left\{\frac{1 - \tilde{\upsilon}(s)}{s} \frac{1}{1 - e^{\Psi(-\xi)}\tilde{\upsilon}(s)}\right\} \\
&= E_{\beta,1}\left[-\left(1 - e^{\Psi(-\xi)}\right)(T/\tau_o)^\beta\right], \qquad \text{using (A7) below.}
\end{aligned}
$$

∎

The ML function

In its most general form, the two-parameter Mittag-Leffler function is given by

$$
E_{\beta,\gamma}(z) = \sum_{j=0}^{\infty} \frac{z^j}{\Gamma(\beta j + \gamma)}, \qquad \beta > 0, \quad \gamma > 0, \tag{A6}
$$

and its Laplace transform, see Podlubny (1999), by

$$
\mathcal{L}\left\{t^{\beta n + \gamma - 1} E_{\beta,\gamma}^{(n)}(\pm a t^\beta)\right\} = \frac{n! s^{\beta - \gamma}}{(s^\beta \mp a)^{n+1}}, \qquad \text{Re}(s) > |a|^{1/\gamma}, \tag{A7}
$$

where $E_{\beta,\gamma}^{(n)}(y) = \frac{d^n}{dy^n} E_{\beta,\gamma}(y)$. This distribution has previously been proposed in the context of financial data in Mainardi, Raberto, Gorenflo, and Scalas (2000).

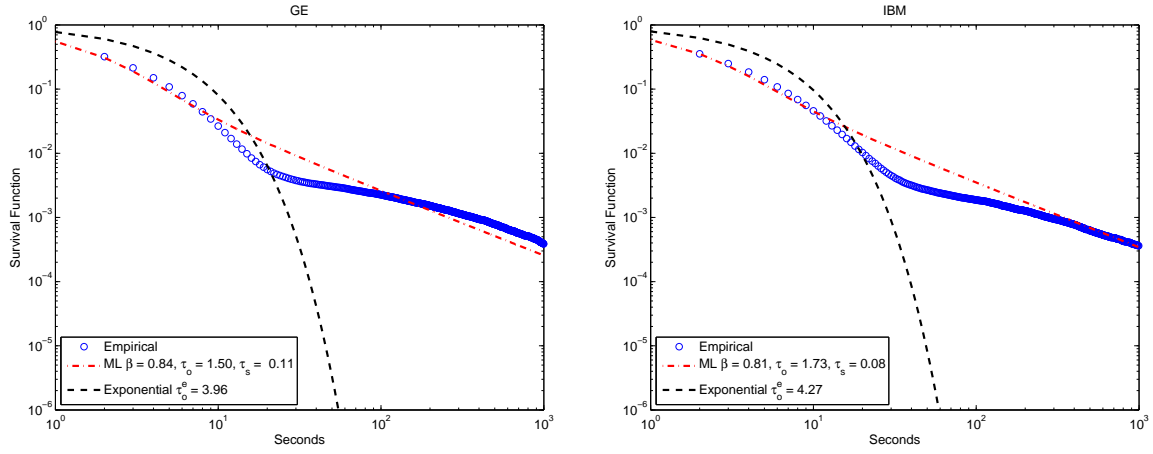# Appendix B. Empirical and fitted Shifted-Mittag-Leffler survival function
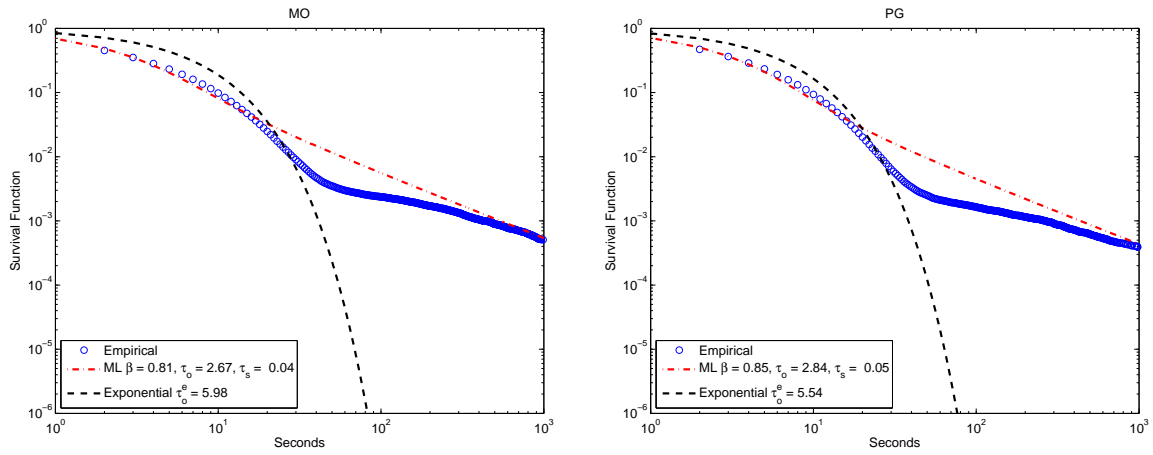

Figure 7. GE and IBM

Figure 8. MO and PG

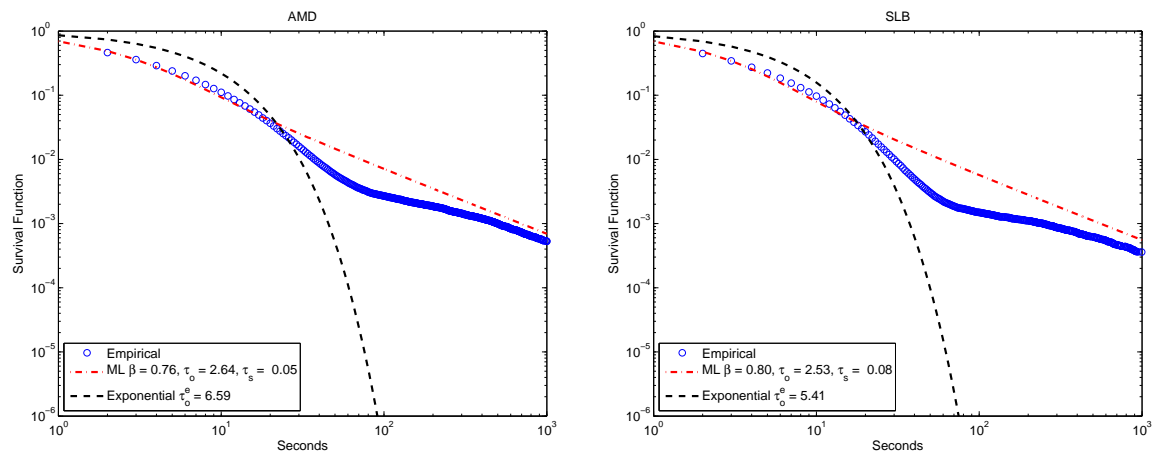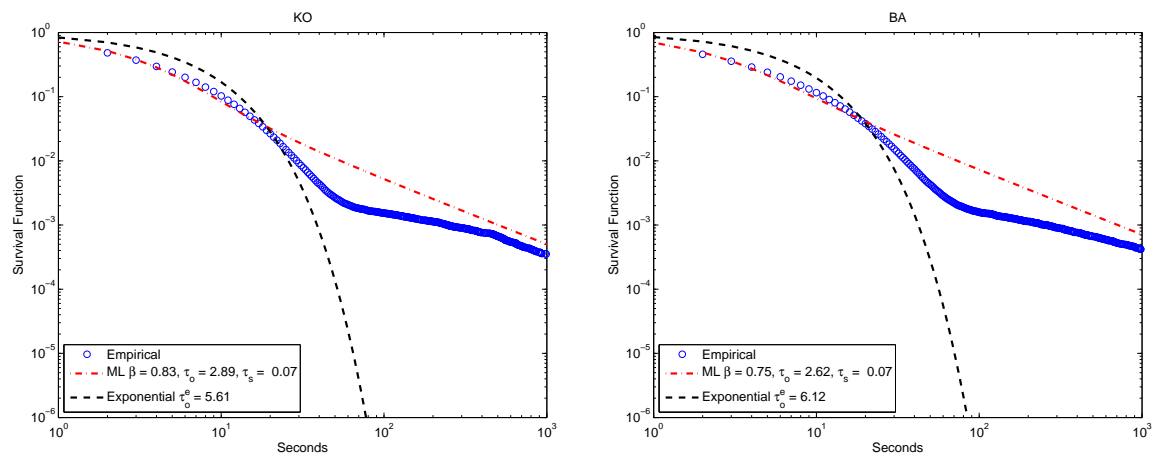**Figure 9.** AMD and SLB

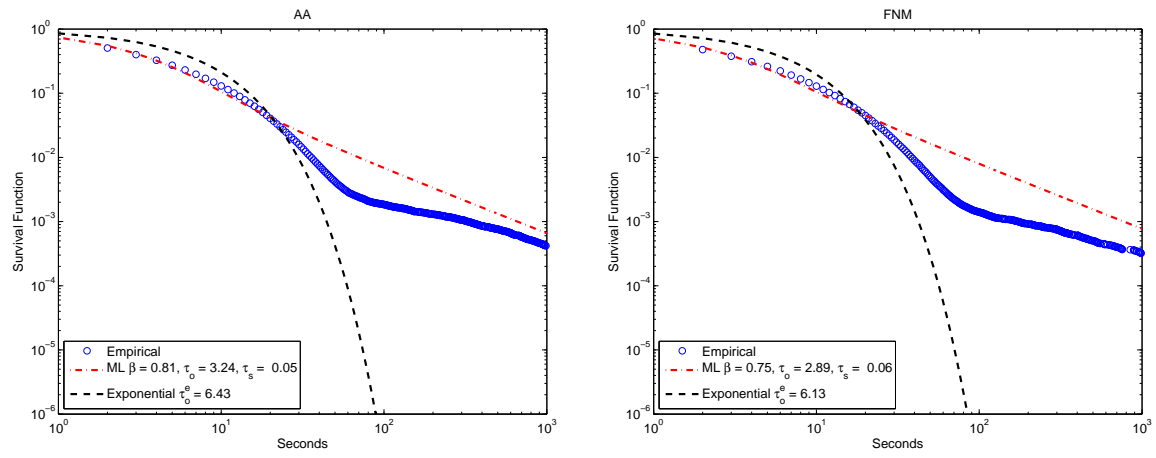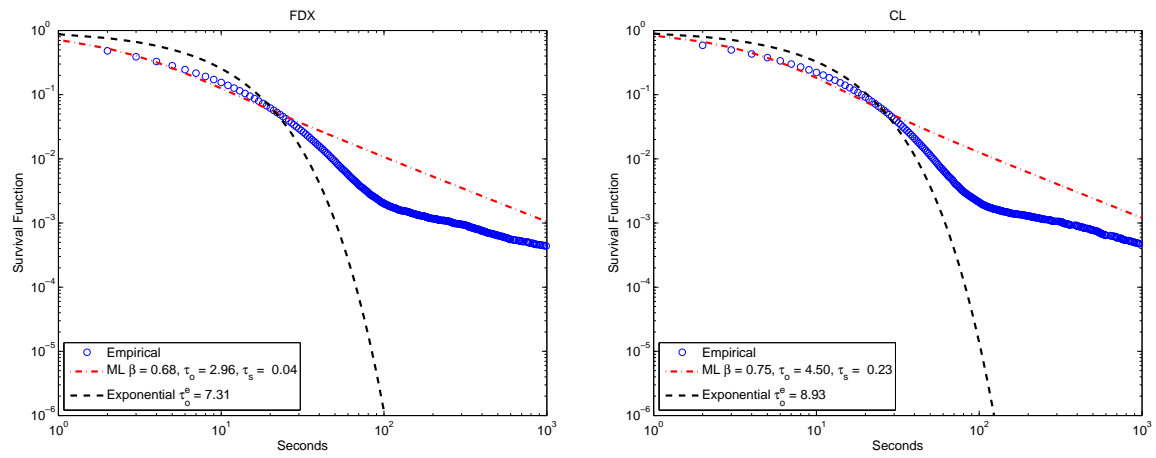**Figure 10.** KO and BA

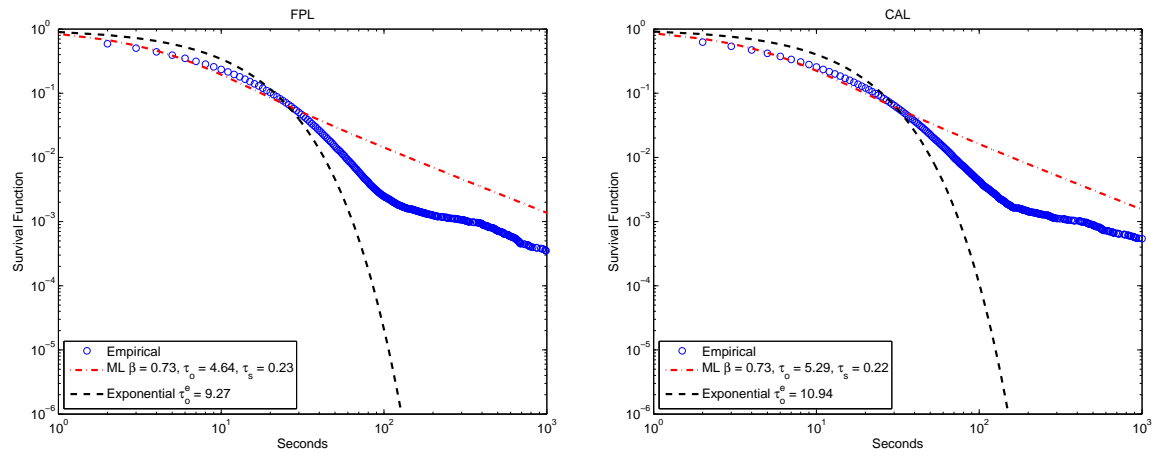**Figure 11.** AA and FNM
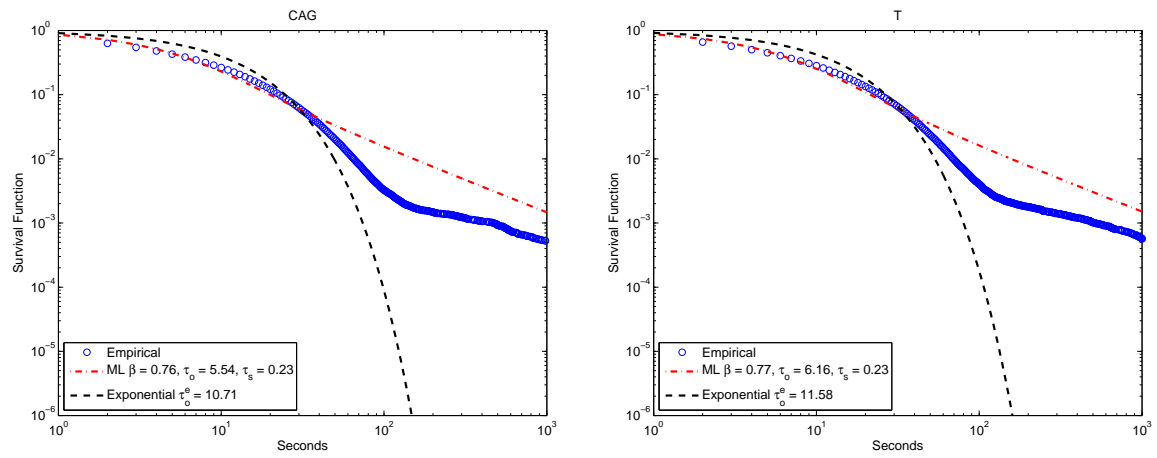

Figure 12. FDX and CL

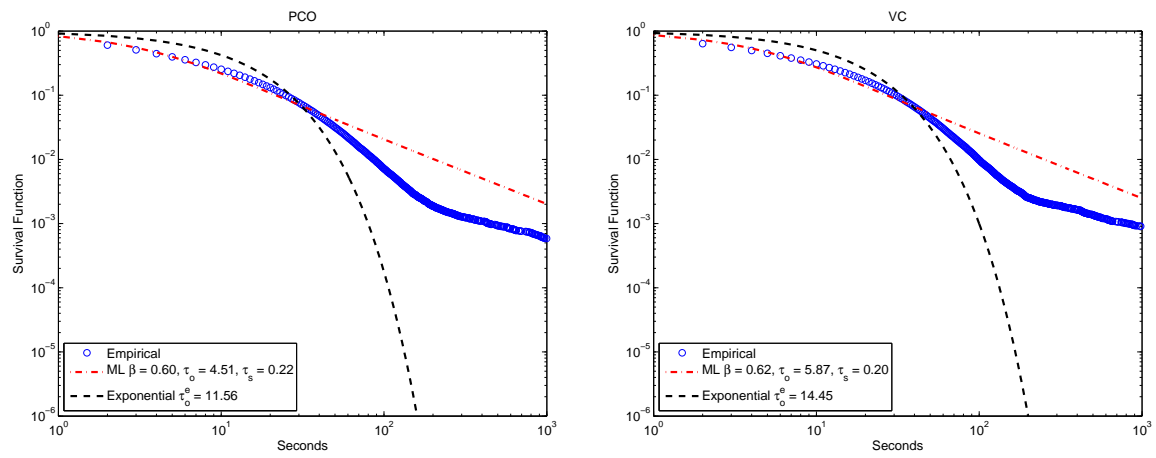**Figure 13.** FPL and CAL

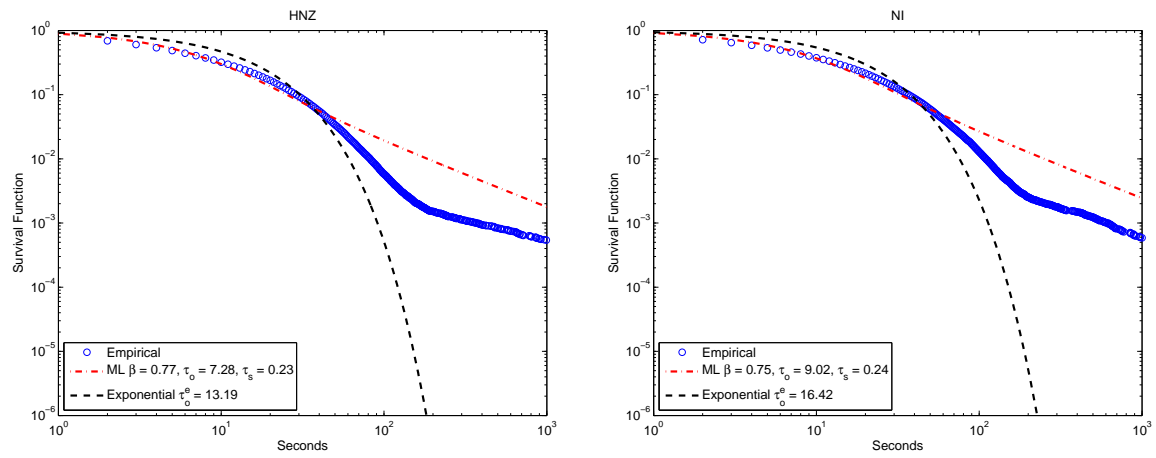**Figure 14.** CAG and T

Figure 15. PCO and VC
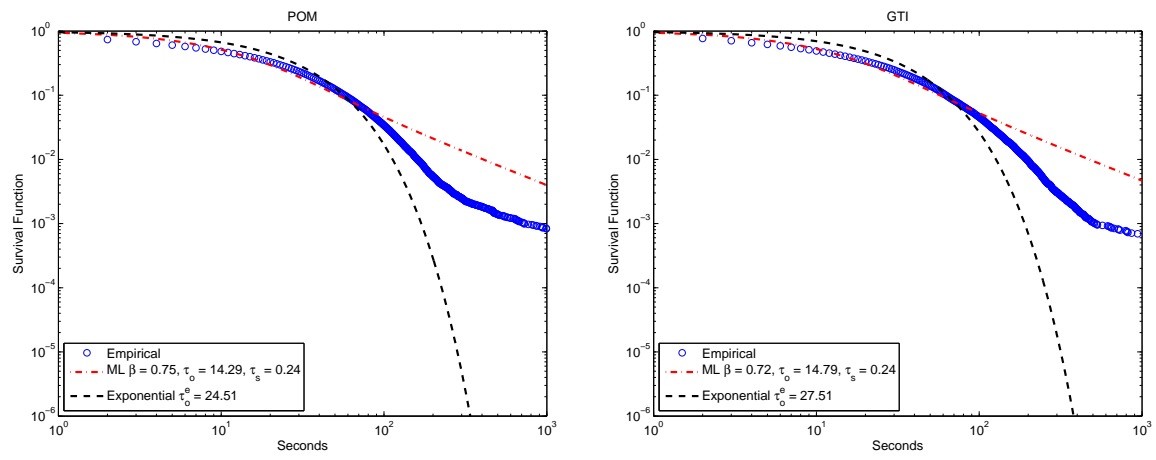

Figure 16. HNZ and NI

Figure 17. POM and GTI

Appendix C. Calibration of risk-neutral parameters: IBM

Date	Days to Expiry	N	Black-Scholes		Gaussian Price Revisions			FMLS Price Revisions			
			σ_{bs}	RMSE	σ	β	RMSE	α	σ	β	RMSE
01-Apr	10	6	0.1446	0.12140	0.0951	0.718300	0.10034	1.99827	0.06713	0.720381	0.10032
	10, 35	17	0.1491	0.14639	0.1497	0.999999	0.14639	1.93180	0.08140	0.999999	0.11759
	10, 35, 73	30	0.1477	0.15214	0.1448	0.976027	0.15079	1.91621	0.07435	0.963982	0.10618
	10, 35, 73, 142	45	0.1498	0.17743	0.1497	1.000000	0.17742	1.89157	0.06914	0.992953	0.10723
	10, 35, 73, 142, 185	74	0.1567	0.24712	0.1570	0.999996	0.24712	1.86177	0.06250	1.000000	0.16001
04-Apr	9, 34	16	0.1563	0.18016	0.1570	0.999999	0.18016	1.90989	0.07796	0.984135	0.14663
	9, 34, 72	29	0.1529	0.16933	0.1462	0.947338	0.16224	1.91061	0.07354	0.920267	0.12182
	9, 34, 72, 141	44	0.1537	0.18544	0.1514	0.978771	0.18397	1.88749	0.06877	0.948755	0.11744
	9, 34, 72, 141, 184	73	0.1597	0.24606	0.1454	1.000000	0.24606	1.86465	0.06541	1.000000	0.15745
05-Apr	8, 33	15	0.1572	0.20434	0.1570	1.000000	0.20434	1.88950	0.07413	1.000000	0.12611
	8, 33, 71	29	0.1537	0.19048	0.1364	0.961793	0.18714	1.89544	0.07053	0.932977	0.12318
	8, 33, 71, 140	44	0.1544	0.19609	0.1536	0.984390	0.19533	1.88100	0.06690	0.951237	0.11161
	8, 33, 71, 140, 183	74	0.1608	0.25307	0.1613	1.000000	0.25307	1.86080	0.05814	1.000000	0.15452
06-Apr	7	6	0.1459	0.27556	0.0326	0.117378	0.24506	1.88500	0.06800	0.974588	0.22364
	7, 32	18	0.1585	0.22855	0.1454	1.000000	0.22854	1.88467	0.07268	1.000000	0.16242
	7, 32, 70	32	0.1538	0.20704	0.1378	0.972736	0.20546	1.90052	0.07208	0.937845	0.15908
	7, 32, 70, 139	49	0.1546	0.20260	0.1541	0.990995	0.20236	1.88617	0.06306	0.954191	0.13761
	7, 32, 70, 139, 182	78	0.1609	0.24998	0.1613	1.000000	0.24998	1.86709	0.06541	1.000000	0.16340
07-Apr	6, 31	13	0.1610	0.21408	0.1613	1.000000	0.21408	1.87824	0.07268	1.000000	0.13080
	6, 31, 69	27	0.1556	0.19003	0.1545	0.988900	0.18970	1.90214	0.07322	0.936459	0.14579
	6, 31, 69, 138	42	0.1551	0.19028	0.1544	0.985119	0.18951	1.88835	0.07008	0.937595	0.12588
	6, 31, 69, 138, 181	71	0.1611	0.23895	0.1613	1.000000	0.23895	1.87563	0.06871	0.997840	0.15981
08-Apr	5, 30	14	0.1742	0.13930	0.1744	1.000000	0.13930	1.94588	0.10175	1.000000	0.11079
	5, 30, 68	28	0.1623	0.20799	0.1483	0.900850	0.17395	1.94092	0.08495	0.892743	0.15418
	5, 30, 68, 137	43	0.1599	0.20617	0.1525	0.917617	0.17954	1.91483	0.07830	0.904480	0.13924
	5, 30, 68, 137, 180	72	0.1644	0.23653	0.1631	0.980267	0.23537	1.88554	0.07354	0.964589	0.17446
11-Apr	4, 29	14	0.2017	0.41293	0.1461	0.754089	0.18112	1.95482	0.09013	0.746634	0.16867
	4, 29, 67	29	0.1813	0.42280	0.1401	0.729480	0.16594	1.93149	0.08061	0.717946	0.13922
	4, 29, 67, 136	45	0.1731	0.39049	0.1486	0.769292	0.19610	1.90293	0.07713	0.754112	0.15083
	4, 29, 67, 136, 179	75	0.1729	0.34599	0.1627	0.840561	0.26786	1.87363	0.07401	0.823721	0.20864

Date	Days to Expiry	N	Black-Scholes		Gaussian Price Revisions			FMLS Price Revisions			
			σ_{bs}	RMSE	σ	β	RMSE	α	σ	β	RMSE
12-Apr	3, 28	13	0.2008	0.25103	0.1804	0.919815	0.21630	1.91042	0.09258	0.905976	0.17132
	3, 28, 66	29	0.1835	0.32599	0.1567	0.839862	0.21727	1.90285	0.07945	0.821585	0.17540
	3, 28, 66, 135	46	0.1835	0.32599	0.1567	0.839862	0.21727	1.88379	0.07447	0.819946	0.14963
	3, 28, 66, 135, 178	76	0.1762	0.30630	0.1690	0.897703	0.26706	1.86262	0.07191	0.872714	0.19297
13-Apr	27	10	0.2096	0.16953	0.1559	0.716529	0.13426	1.92714	0.09780	0.816297	0.07761
	27, 65	25	0.1897	0.28542	0.1415	0.627967	0.14072	1.92726	0.08311	0.631265	0.10060
	27, 65, 134	42	0.1805	0.29148	0.1565	0.738285	0.17043	1.90569	0.08299	0.730365	0.11626
	27, 65, 134, 177	72	0.1795	0.26729	0.1710	0.850863	0.22443	1.90288	0.08134	0.837758	0.16236
14-Apr	26	10	0.2250	0.24240	0.1578	0.667694	0.20749	1.90260	0.09784	0.835731	0.12398
	26, 64	25	0.2005	0.34639	0.1429	0.579937	0.16601	1.92461	0.08451	0.586031	0.12399
	26, 64, 133	42	0.1889	0.34845	0.1599	0.699355	0.19184	1.90220	0.08510	0.693662	0.13512
	26, 64, 133, 176	72	0.1858	0.30738	0.1747	0.806823	0.23449	1.88370	0.08389	0.794803	0.16881
15-Apr	25	9	0.2334	0.09495	0.2340	1.000000	0.09495	1.98573	0.15556	0.989830	0.09103
	25, 63	25	0.2198	0.18612	0.1883	0.809590	0.11980	1.96368	0.11910	0.810394	0.10748
	25, 63, 132	42	0.2093	0.23323	0.1878	0.806724	0.12579	1.94344	0.11106	0.801054	0.09952
	25, 63, 132, 175	72	0.2064	0.21517	0.1977	0.872084	0.16560	1.91808	0.10508	0.861568	0.12335
18-Apr	24	12	0.2376	0.37424	0.1102	0.309411	0.33296	1.79584	0.06571	0.691509	0.12503
	24, 43	28	0.2254	0.32644	0.1577	0.621009	0.27300	1.82344	0.07052	0.663036	0.10927
	24, 43, 62	45	0.2202	0.29667	0.1755	0.722059	0.24913	1.83441	0.07546	0.731041	0.10122
	24, 43, 62, 131	62	0.2136	0.30080	0.1896	0.806386	0.24552	1.83282	0.07643	0.788706	0.09655
	24, 43, 62, 131, 174	92	0.2101	0.28104	0.2002	0.872586	0.24550	1.83692	0.07889	0.846766	0.10924
19-Apr	23	9	0.2174	0.16136	0.1721	0.792012	0.14433	1.93218	0.10918	0.892435	0.09793
	23, 42	22	0.2122	0.18431	0.1791	0.823310	0.15987	1.90761	0.09662	0.857251	0.08700
	23, 42, 61	39	0.2114	0.17755	0.1916	0.881652	0.16245	1.89752	0.09616	0.895252	0.07924
	23, 42, 61, 130	56	0.2069	0.19356	0.1936	0.890708	0.16604	1.88927	0.09234	0.880713	0.07561
	23, 42, 61, 130, 173	86	0.2049	0.19220	0.1993	0.926979	0.17585	1.88065	0.09053	0.909808	0.07799
20-Apr	22	10	0.2936	0.29079	0.2151	0.729837	0.26879	1.87858	0.12246	0.893683	0.12163
	22, 41	24	0.2753	0.33749	0.1931	0.637909	0.23785	1.88100	0.10203	0.664748	0.12103
	22, 41, 60	41	0.2664	0.32226	0.2058	0.694919	0.22032	1.87869	0.10384	0.701882	0.10863
	22, 41, 60, 129	58	0.2518	0.38465	0.2105	0.719306	0.22056	1.86959	0.10121	0.709487	0.10886
	22, 41, 60, 129, 172	88	0.2375	0.40030	0.2138	0.740331	0.22797	1.85754	0.09785	0.724599	0.11778

			Black-Scholes		Gaussian Price Revisions			FMLS Price Revisions			
Date	Days to Expiry	N	σ_{bs}	RMSE	σ	β	RMSE	α	σ	β	RMSE
21-Apr	21	8	0.2462	0.31792	0.1557	0.611305	0.29378	1.83263	0.08648	0.908230	0.05828
	21, 40	21	0.2361	0.28778	0.1839	0.747935	0.24866	1.84694	0.08301	0.779884	0.08370
	21, 40, 59	38	0.2310	0.25488	0.1942	0.797014	0.21715	1.85674	0.08691	0.802607	0.08233
	21, 40, 59, 128	55	0.2230	0.27027	0.1995	0.825142	0.20996	1.85291	0.08595	0.808750	0.07858
	21, 40, 59, 128, 171	85	0.2176	0.26146	0.2065	0.866166	0.21460	1.84636	0.08444	0.843004	0.08862
22-Apr	20	7	0.2348	0.29362	0.1377	0.557621	0.26241	1.83170	0.07641	0.803034	0.04257
	20, 39	20	0.2311	0.27635	0.1971	0.842038	0.25989	1.83001	0.07911	0.866790	0.06229
	20, 39, 58	37	0.2292	0.25720	0.2067	0.882096	0.24439	1.82901	0.07950	0.883406	0.06089
	20, 39, 58, 127	54	0.2214	0.27160	0.2017	0.856904	0.23100	1.82947	0.07832	0.834035	0.06693
	20, 39, 58, 127, 170	84	0.2177	0.25953	0.2086	0.897403	0.23183	1.82896	0.07889	0.868423	0.08358
25-Apr	19	10	0.2360	0.47356	0.0851	0.152247	0.37182	1.77059	0.05317	0.481404	0.15181
	19, 38	25	0.2312	0.34390	0.1781	0.743105	0.31217	1.82075	0.07367	0.769464	0.16106
	19, 38, 57	42	0.2295	0.30081	0.1988	0.835064	0.27982	1.83093	0.07925	0.834099	0.13867
	19, 38, 57, 126	59	0.2223	0.30816	0.2014	0.847767	0.26786	1.82662	0.07787	0.823936	0.12114
	19, 38, 57, 126, 169	89	0.2173	0.29366	0.2071	0.878604	0.25803	1.82693	0.07810	0.848881	0.11163
26-Apr	18	6	0.2191	0.13784	0.1668	0.780005	0.10583	1.99934	0.11691	0.774004	0.10588
	18, 37	18	0.2156	0.15586	0.1877	0.865566	0.13065	1.94931	0.11334	0.873660	0.10663
	18, 37, 56	34	0.2167	0.15528	0.2038	0.931347	0.14720	1.93595	0.11535	0.932495	0.11544
	18, 37, 56, 125	51	0.2108	0.19254	0.1963	0.892637	0.15659	1.91851	0.10449	0.881957	0.11172
	18, 37, 56, 125, 168	81	0.2068	0.20188	0.1996	0.908278	0.17109	1.89700	0.09728	0.891907	0.10417
27-Apr	17	6	0.2186	0.15537	0.1489	0.712010	0.13096	1.94046	0.10199	0.817184	0.08316
	17, 36	18	0.2140	0.15489	0.1846	0.860764	0.12586	1.93821	0.10781	0.867373	0.08294
	17, 36, 55	34	0.2149	0.16698	0.2004	0.922940	0.15659	1.92013	0.10756	0.922640	0.10423
	17, 36, 55, 124	51	0.2084	0.20164	0.1927	0.884829	0.15839	1.91177	0.10031	0.872865	0.10092
	17, 36, 55, 124, 167	81	0.2045	0.20725	0.1966	0.904565	0.17217	1.89718	0.09597	0.888060	0.10041
28-Apr	16	7	0.2335	0.19751	0.1392	0.603381	0.14249	1.93240	0.09535	0.721569	0.07978
	16, 35	20	0.2275	0.20405	0.1895	0.829267	0.16881	1.91423	0.10343	0.842242	0.09397
	16, 35, 54	37	0.2253	0.20123	0.1996	0.866456	0.17360	1.90432	0.10254	0.867924	0.09543
	16, 35, 54, 123	54	0.2179	0.23443	0.1980	0.858403	0.17515	1.89410	0.09736	0.845847	0.08714
	16, 35, 54, 123, 166	84	0.2119	0.23656	0.2013	0.875442	0.18031	1.88485	0.09490	0.857824	0.08400

			Black-Scholes		Gaussian Price Revisions			FMLS Price Revisions			
Date	Days to Expiry	N	σ_{bs}	RMSE	σ	β	RMSE	α	σ	β	RMSE
29-Apr	15	6	0.2215	0.12008	0.1553	0.736213	0.08523	1.99999	0.11082	0.742261	0.08525
	15, 34	17	0.2205	0.14083	0.2025	0.920912	0.12860	1.96331	0.12676	0.923427	0.11431
	15, 34, 53	33	0.2211	0.15083	0.2118	0.951754	0.14565	1.94189	0.12128	0.949849	0.11655
	15, 34, 53, 122	50	0.2151	0.18489	0.2018	0.906645	0.14983	1.93003	.111606	0.896408	0.11241
	15, 34, 53, 122, 165	80	0.2105	0.19635	0.2031	0.912868	0.16180	1.91137	0.10386	0.898212	0.10610
02-May	14	5	0.2290	0.27764	0.0823	0.261493	0.15757	1.92113	0.06444	0.445254	0.08383
	14, 33	18	0.2248	0.25097	0.1911	0.853769	0.22727	1.88115	0.09171	0.862769	0.11872
	14, 33, 52	35	0.2226	0.22346	0.2001	0.886397	0.20390	1.88495	0.09536	0.882423	0.10679
	14, 33, 52, 121	52	0.2175	0.23294	0.2022	0.896346	0.19974	1.87778	0.09249	0.878115	0.09291
	14, 33, 52, 121, 164	82	0.2128	0.23619	0.2043	0.905821	0.20286	1.86820	0.08954	0.883192	0.08246
03-May	13	5	0.2200	0.15755	0.1284	0.621416	0.10695	1.95551	0.08675	0.669594	0.07296
	13, 32	16	0.2152	0.16671	0.1856	0.871787	0.13568	1.93963	0.10793	0.872032	0.09189
	13, 32, 51	32	0.2166	0.18103	0.2041	0.935952	0.17291	1.91551	0.10623	0.931162	0.11538
	13, 32, 51, 120	49	0.2130	0.18928	0.2031	0.930869	0.17064	1.90509	0.10150	0.915566	0.10123
	13, 32, 51, 120, 163	79	0.2103	0.20033	0.2046	0.940809	0.18496	1.88715	0.09475	0.921092	0.09448
04-May	12, 31	15	0.2113	0.17191	0.1948	0.931329	0.16351	1.93085	0.10751	0.918234	0.12925
	12, 31, 50	32	0.2140	0.18295	0.2093	0.979913	0.18216	1.90714	0.10392	0.964360	0.12560
	12, 31, 50, 119	49	0.2104	0.19264	0.2030	0.947152	0.18198	1.89506	0.09677	0.925302	0.11161
	12, 31, 50, 119, 162	79	0.2080	0.20165	0.2035	0.950344	0.19090	1.88057	0.09179	0.925543	0.09917
05-May	11, 30	14	0.2137	0.20835	0.2078	0.976319	0.20751	1.91977	0.10787	0.945693	0.17370
	11, 30, 49	30	0.2180	0.21121	0.2180	1.000000	0.21121	1.90003	0.10477	0.983502	0.15584
	11, 30, 49, 118	47	0.2143	0.20566	0.2084	0.957260	0.19910	1.89242	0.09804	0.928399	0.13475
	11, 30, 49, 118, 161	77	0.2114	0.20562	0.2074	0.953734	0.19663	1.88030	0.09330	0.923950	0.11336
06-May	10	9	0.2173	0.72244	0.0575	0.090412	0.64069	1.60222	0.02533	0.331809	0.34102
	10, 29	22	0.2171	0.54819	0.1765	0.825221	0.53545	1.66691	0.03688	0.773434	0.28549
	10, 29, 48	39	0.2203	0.45614	0.2038	0.922345	0.45220	1.71998	0.04793	0.858471	0.24774
	10, 29, 48, 117	56	0.2153	0.41119	0.2028	0.914298	0.40001	1.74354	0.05310	0.848558	0.21435
	10, 29, 48, 117, 160	86	0.2119	0.36243	0.2049	0.926622	0.35100	1.77206	0.06015	0.867371	0.18471

Table 4
IBM risk-neutral parameters April-May 2006

Notes

[1]Below we discuss in detail how consolidated trades from the TAQ database were employed.

[2]In fact, a Kolmogorov-Smirnov test clearly rejects the hypothesis that the data came from an exponential survival function.

[3]The early work of Osborne (1959) already introduces the notion of number of transactions, per unit of time, as a measure of market activity.

[4]In the language of counting processes the process $\sum_{i=1}^{N_t} Y_i$ is called a (0-delayed) renewal process.

[5]In Clark (1973) price dynamics are modeled by subordinating Brownian motion to operational time which is measured by trading volume; in Carr, Geman, Madan, and Yor (2003) the persistence of volatility is formulated by evaluating Lévy processes at stochastic times. And in a broader sense one may interpret stochastic volatility models as instruments that also measure the evolution of the operational clock.

[6]This is equivalent to assuming that the distribution of durations under the physical and risk-neutral measures are different.

[7]In the rest of the paper we will not include the leverage effect.

[8]We calculate this arbitrary duration strictly greater than zero in the following way. Out of all the zero-duration trades we count how many times there were two trades within one second, three trades within one second, etc. Then we calculate a weighted average of number of trades within one second and assume that these occur within 0.5 second instead of deleting them from the sample. Furthermore, for simplicity we do not alter the duration of the trade following those zero-duration trades for which we assigned a non-negative duration.

[9]The constant a could be a function of the parameter β.

[10]To arrive at expression (20) we use the property that

$$\mathcal{L}\{v(t)\} = \mathcal{L}\left\{-\frac{d\Upsilon(t)}{dt}\right\} = -s\mathcal{L}\{at^{-\beta}\} + \Upsilon(0) = -a\gamma(1-\beta)s^\beta + 1,$$

47

and restrict $0 < \beta \leq 1$ to have a valid (monotonic) survival function.

[11]Here we use the characteristic function

$$\mathbb{E}\left[e^{i\xi \Sigma_{i=1}^{N_T} Y_i}|\mathcal{F}_s\right] \;\;=\;\; E_{\beta,1}\left[-\left(1-e^{\Psi(\xi)}\right)((T-s)/\tau_o)^{\beta}|\mathcal{F}_s\right],$$

where $e^{\Psi(\xi)}$ is the characteristic function of Y, to calculate the expectation of the number of trades.

[12]There are a number of articles in the literature that use transform techniques to price and calibrate options, see for example Carr and Wu (2003), Carr and Wu (2004).

[13]Note that we must require $e^{\Psi(-\xi)}$ to be analytic in a line that intersects $[-\infty + i\hat{\xi}, \infty + i\hat{\xi}]$ where $\hat{\xi} > 1$.

[14]For further information on methodology (dividend treatment, processing of BBA Libor rates, etc.) see http://wrds.wharton.upenn.edu and the Ivy DB File and Data Reference Manual.

[15] We do not calibrate to lots where there are less than 5 options.

[16]For simplicity we assumed that τ_o remained the same under both the physical and statistical measure (see Table 2 for IBM) and that $\tau_s = 0$.

[17]Note that $\tau_o = 4 * 528,195$ where 528,195 is in Table 1 for IBM under 'All trades' over the period April-June 2005.

[18]Figure 3 does not include the case $\beta = 1$ where IV becomes 0.30 for all expiries as expected. However, the case $\beta = 1$ can be seen in Figure 4.