Completion Time Structures of Stock Price Movements*

ASGER LUNDE

Department of Marketing, Informatics and Statistics,
The Aarhus School of Business, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark

Email: alunde@asb.dk

ALLAN TIMMERMANN

Department of Economics, University of California, San Diego,
9500 Gilman Drive La Jolla, CA 92093-0508, USA

Email: atimmerm@ucsd.edu

June 2004

ABSTRACT

This paper proposes to model movements in more than a century of daily US stock prices as the outcome of a multi-state marked point process and studies the time it takes for stock prices to complete an up or a down move of a certain size. We present a new econometric specification for a class of dynamic models that account for autoregressive conditional duration effects. We also present a method to account for the effect of time-varying state variables that may change within a duration. We find strong evidence of dynamic dependencies in the direction and speed of stock price movements. Past interest rates are also found to affect the speed and direction of completion times. Out-of-sample prediction results show that forecasts of the direction of moves in stock prices can be greatly improved by including covariates such as interest rates and allowing for dynamics in the econometric specification.

Journal of Economic Literature Classification Numbers: C41, G1.

*We thank an anonymous referee, an associate editor, Rob Engle, Mark Machina and Ruth Williams for helpful conversations. We are grateful to INQUIRE UK for financial support for this research.

Completion Time Structures of Stock Price Movements[a]

November 2004

Abstract

This paper proposes to model movements in more than a century of daily US stock prices as the outcome of a multi-state marked point process and studies the time it takes for stock prices to complete an up or a down move of a certain size. We present a new econometric specification for a class of dynamic models that accounts for autoregressive conditional duration effects and for the effect of time-varying state variables that may change within a duration. We find strong evidence of dynamic dependencies in the direction and speed of stock price movements. Past interest rates are also found to affect the speed and direction of completion times. Out-of-sample prediction results show that forecasts of the direction of moves in stock prices can be greatly improved by including covariates such as interest rates and allowing for dynamics in the econometric specification.

Journal of Economic Literature Classification Numbers: C41, G1.

[a] We thank an anonymous referee, an associate editor, Rob Engle, Mark Machina and Ruth Williams for helpful comments. We are grateful to INQUIRE UK for financial support for this research.

1. INTRODUCTION

The extent to which stock market returns can be predicted has long been a key question in finance. Predictable patterns in prices map into changes in investors' optimal portfolio holdings, so different models for the evolution in stock prices translate naturally into different asset allocations. Most notably, in the absence of predictability of stock returns, investors face constant investment opportunities and their optimal stock holdings become independent of the investment horizon, c.f. Merton (1969) and Samuelson (1969).

A vast body of empirical work on asset prices has found evidence that stock returns are partially predictable either through past movements in stock prices—due to the presence of a mean-reverting component, c.f. Fama & french (1988) and Poterba & Summers (1988)—or by means of time-varying predictor variables such as interest rate spreads, default premia, the dividend yield, the earnings-price ratio or similar variables, c.f. Campbell (1987), Campbell & Shiller (1988), Fama & french (1989), Ferson & Harvey (1991) and Keim & Stambaugh (1986).

Another strand of this literature has documented predictability in higher order moments of stock returns such as volatility (Glosten, Jagannathan & Runkle (1993)), skewness and kurtosis (Guidolin & Timmermann (2004)). In the presence of such time-variations in the conditional return distribution, investors' optimal asset allocation will generally depend on state variables that track current and future investment opportunities. Investors' time horizon will also matter for their asset allocation, as will the frequency of any rebalancing opportunities. This holds both if investors' objectives are to maximize expected utility or if they are simply interested in controlling risk.

Common to the vast majority of work on predictability of stock returns is that it studies price moves over a fixed holding period such as a day, a week or a month. Similarly, measures of risk are commonly based on quantiles or moments of the return distribution defined over a fixed holding period. Assets with high mean return and low volatility or low probability of incurring a large loss over a pre-specified investment horizon are viewed as more desirable by risk averse investors.

Rather than fixing the holding period and studying how far prices moved over a certain time interval, asset price movements can be analyzed by fixing the size of the movement required by the asset price and treat the resulting completion time as an unknown, random variable. This yields a complementary perspective to the analysis of fixed-horizon movements in asset prices. For example, how desirable an asset is to a risk averse investor can equally well be characterized in terms of how long the investor expects to have to wait before the asset pays off a pre-determined return such as 5% and how uncertain this waiting time is. The higher the variance of this completion time, or duration, the riskier the asset can be considered to be. Likewise, if the mean duration is long, the expected return per unit of time will tend to be low.

Duration-based risk measures are useful in many finance applications. For example, consider the asset allocation decision of a pension fund manager in charge of an underfunded pension scheme with known future liabilities. Regulatory rules often require that an existing funding gap (i.e. the difference between the projected value of the liabilities and the market value of the assets) is eliminated within a specified period of time such as a year. The pension fund manager is thus faced with the risk of not covering the underfunding within a specific period of time. To ensure compliance with such regulation, the fund manager would need a forecast of the expected time before the value of the assets under management move up by the funding gap, say $\delta\%$. The expected completion time should be well under the maximum allowed time, or else the portfolio holdings need to be changed. Our paper presents new methods that allow the fund manager to compute such probability estimates not just initially, but in real time as a function of a set of underlying state variables that predict future price movements and that are allowed to change during the duration spell. As a second example, traders in financial markets often have contracts that pay a pre-specified bonus provided that they reach a performance mark within a certain period of time (e.g., a calendar year) and duration-based probability measures are ideally suited to characterize the associated probabilities that such a threshold will be reached.

In many cases clearly a duality exists between asset return characteristics measured over a fixed horizon vis-a-vis the distribution of durations or completion times associated with a fixed return. Most notably, when asset prices are generated by a geometric random walk process with constant drift and volatility−as assumed in much of modern finance−there is a one-to-one relationship between standard measures of risk and returns that condition on a fixed holding period and duration measures based on completion times that condition on a fixed return.

Outside this framework, however, no general results exist on the relationship between risk measures based on fixed-horizon returns and measures based on fixed-return durations so one would expect that they can uncover very different features of the underlying data generating process. This becomes particularly important in the presence of nonlinearities in returns. Recent empirical work suggests that nonlinearities in asset returns are important. For example, Christoffersen & Diebold (2003) and Guidolin & Timmermann (2004) suggest complicated dynamic dependencies in the term structure of risk and in the probabilities of up and down moves, Value at Risk (VaR) or expected shortfall. Detemple, Garcia & Rindisbacher (2003) also document strong evidence of nonlinear dynamics in asset prices.

Theoretical models in the asset pricing literature also suggest complicated nonlinear dynamics in stock prices arising as a result of speculative bubbles (McQueen & Thorley (1994)), over-reaction to recent news

(De Bondt & Thaler (1985)), learning dynamics (Timmermann (1996, 2001), Veronesi (1999)), habit persistence (Campbell & Cochrane (1999)) or non-linear dynamics in the dividend process (Donaldson & Kamstra (1996)). The prevalence of such theories means that it is informative to use alternative measures of risk and different modes of analysis to identify possibly complicated dynamics in asset prices.

As an example of a nonlinearity, suppose that log-stock prices fluctuate within a narrow band for a long period of time reflecting, perhaps, support and resistance levels as hypothesized by many technical trading models, c.f. Brock, Lakonishok & Lebaron (1992) and Sullivan, Timmermann & White (1999). This type of sample path could result from a model with time-varying serial correlations that vary as a function of the distance to the bands. The resulting low volatility would make the asset seem attractive even though, over a longer horizon, the asset return can be expected to be low since investors have to wait longer than expected before earning a positive return sufficiently high to require prices to move beyond the band.[1] Such nonlinear dependencies would be reflected in the distribution of durations of up and down movements in stock prices provided that a range of thresholds $\{\pm\delta_1, \pm\delta_2, ... \pm \delta_j\}$ for the change in the log asset prices is studied. In analogy with the fixed income literature, we refer to the joint distribution of expected completion times for different values of δ as the completion time structure. Completion time measures can also capture other features commonly associated with stock returns. For example, outliers in the return distribution will lead to more short durations as will volatility clustering. Similarly, the completion time of an up move will be shorter (longer) if more large and positive (negative) than large and negative (positive) shocks occur, so duration models can also identify skews in asset price dynamics.

The contributions of our paper are three-fold. First, we present new applications to stock price movements of models for completion time dynamics using more than a century of daily stock market prices. While duration features of stock price movements have been explored in the study of market microstructure effects (c.f. the literature survey by Engle & Russell (2002)), little work has been undertaken on the duration of movements in stock prices of the size or frequency considered here.[2] Secondly, we make a methodological contribution by showing how to account for the effect of time-varying state variables that may change within a duration. This is particularly important in finance applications where predictor variables such as

[1] Suppose for example that this band is given by $\log(p_0) \pm \delta/2$, where p_0 is the initial asset price and δ is the width of the band. Then the expected rate of return within the band would be bounded by δ/T, where T is the investment period.

[2] One exception is the paper by Lunde & Timmermann (2004) which studies bull and bear states. However, that paper adopts a very different methodology that does not fix the size of the required stock price movement, but instead defines underlying (latent) states according to the size of log-price changes since the previous local peak or trough. This is an important distinction since our framework here has the random walk model as a natural benchmark whereas this is not the case in the setup with (endogenous) bull and bear states.

the interest rate are likely to be subject to important time-variations. Such inter-duration variations are likely to affect the time of completion. Hence, by only conditioning on information available at the time of the beginning of a new duration spell, an important part of the behavioral relation between stock prices and the relevant state variables may be missed or at worst reversed. Our approach allows investors to update, in real time, the expected value of the remaining time until the completion of a duration spell as the values of the relevant state variables change. Third, we present new empirical evidence of predictability in stock price movements using both in-sample and out-of-sample experiments, with the latter set up to replicate the real-time forecasts of an investor.

The plan of the paper is as follows. In the context of the geometric random walk model Section 2 shows the duality between conventional measures of risk and returns over a fixed time horizon and completion time measures and also presents some initial empirical results. Section 3 presents the econometric methods used to characterize the distribution of completion times. Section 4 reports estimation results for a range of empirical specifications, while Section 5 evaluates the duration models in an out-of-sample forecasting experiment and Section 6 concludes. Technical details are provided in two appendices at the end of the paper.

2. Completion Times under the Geometric Random Walk

Suppose we are interested in studying the time it takes for the logarithm of an asset price, P, to complete a move of a certain size. To establish a benchmark for the distribution of such durations, it is useful to first consider this question for the standard geometric random walk model. In this model the logarithm of stock prices, $p_t = \log(P_t)$, evolves according to the following stochastic differential equation:

$$dp_t = \mu dt + \sigma dW_t, \tag{1}$$

where W is a standard one-dimensional Wiener process. To measure durations, we have to define a random variable that tracks the completion times. To this end we study how long it takes for the logarithm of the stock price index to move up by a fixed amount, $\delta > 0$.[3] This can be achieved by defining a chain of first passage times measuring the time it takes for log-stock prices to cross a single barrier δ units away from the

[3] Lunde & Timmermann (2004) define bull and bear markets as states that partition the data according to sequences of local peaks and troughs. This definition reflects the most common use of these terms but does not lead to a tractable benchmark duration model since the classification of the current state depends on both past and future movements in stock prices.

initial log-price $p_{t_{0,i}}$:

$$\tau^*_{\text{up}_i}(\delta, p_{t_{0,i}}) = \inf\{t \geq t_{0,i} : p_t - p_{t_{0,i}} \geq \delta\}, \tag{2}$$

where the notation $t_{0,i}$ reflects that $p_{t_{0,i}}$ is the log-price at the origin of the ith first passage time.

The first passage time, t, for the event that some barrier δ return units away from the starting point, $p_{t_{0,i}}$, is crossed has the following Inverse Gaussian density, $f(t)$, and survival function, $S(t)$:[4]

$$f(t) = \frac{\delta}{\sigma\sqrt{2\pi t^3}} \exp\left(\frac{-(\delta - \mu t)^2}{2\sigma^2 t}\right), \tag{3}$$

$$S(t) = \Phi\left(\frac{\delta - \mu t}{\sigma\sqrt{t}}\right) - \exp\left(\frac{2\mu\delta}{\sigma^2}\right)\Phi\left(\frac{-\delta - \mu t}{\sigma\sqrt{t}}\right), \tag{4}$$

where Φ denotes the standard Gaussian cumulative density function. Thus, under the geometric random walk model (1), for a given barrier, δ, the duration distribution has a convenient closed form that is completely characterized by the drift (μ) and volatility (σ) of asset prices. If $\mu \leq 0$ no positive moments of T exist, while for $\mu > 0$ it holds that:

$$\begin{aligned} E[T] &= \frac{\delta}{\mu}, \\ Var(T) &= \frac{\delta\sigma^2}{\mu^3}. \end{aligned} \tag{5}$$

As expected, these moments depend on the drift and volatility parameters $\theta = (\mu\ \sigma)'$ as well as the distance to the barrier, δ. These parameters can either be retrieved from the duration distribution through maximum likelihood estimation based on (3) and (4) or, as is traditionally done, from the moments of the returns data sampled at a given calendar frequency.

The single barrier definition (2) only indirectly reflects down moves in stock prices in as far as these show up as longer durations before stock prices move up by a fraction δ. An alternative is to introduce duration measures that more directly reflect down moves by tracking the time it takes for log-prices to either decline or rise by δ:

$$\tau^*_{\text{down}_i}(\delta, p_{t_{0,i}}) = \inf\{t \geq t_{0,i} : p_t - p_{t_{0,i}} \leq -\delta\}, \tag{6}$$

$$\tau^*_{\text{up}_i}(\delta, p_{t_{0,i}}) = \inf\{t \geq t_{0,i} : p_t - p_{t_{0,i}} \geq \delta\}. \tag{7}$$

Only the introduction of (6) is changed here since (7) is identical to (2). However, the definitions (6)-(7) now give rise to a two-barrier problem since we measure the time it takes before one of two barriers is crossed by the stock price index. We capture this information through a sequence of marked durations

$$\tau^*_i = \{\min(\tau^*_{\text{down}_i}, \tau^*_{\text{up}_i}), d^*_i\}, \tag{8}$$

[4] See, e.g., Lancaster (1990, page 119).

where d_i^* is an indicator variable that equals one if the upper barrier is crossed first, and otherwise is zero.

Many of the methods used in this paper are relatively new to the finance literature−with the exception of studies of market microstructure effects – so we first look at the data before considering the more detailed econometric analysis in the subsequent sections.

2.1. Data

To investigate the properties of up and down movements in stock prices along the definition proposed in Section 2, we construct a data set of daily US stock prices from 2/17/1885 to 12/31/1997. From 2/17/1885 to 2/7/1962 the nominal stock price index is based on the daily returns provided by Schwert (1990). These returns include dividends. From 3/7/1962 to 12/31/1997 the stock price index is constructed from daily returns on the Standard & Poors 500 price index, again including dividends and obtained from the CRSP tapes. Combining these series generates a time series of 31,412 daily nominal stock prices.

Inflation has varied considerably over the sample period and the drift in nominal prices does not have the same interpretation during low and high inflation periods. To deal with this issue, we construct a daily inflation index as follows. We use monthly data on the consumer price index taken from Shiller (2000) and convert it into daily inflation rates by solving for the daily inflation rate such that the daily price index grows smoothly - and at the same rate - between subsequent values of the monthly consumer price index.[5] Finally we divide the nominal stock price index by the consumer price index to get a daily index for real stock prices.

2.2. Durations

An informal comparison of the observed durations with those implied by equations (3) and (4) provides a first way of detecting deviations from the Geometric random walk model (1) that assumes independently distributed, Gaussian price increments.

Figure 1 accomplishes this by plotting the distribution of the residuals from the inverse Gaussian model fitted to US stock price durations. Another natural benchmark is a GARCH specification, so we also show the distribution of duration residuals for a price index constructed so that the price increments are normalized by the conditional volatility. This latter price index thus adjusts for persistence in the conditional volatility.[6]

[5] Since the volatility of daily inflation rates is likely to be only a fraction of that of daily stock returns, normalizing by the inflation rate has the effect of a time-varying drift adjustment. Lack of access to daily inflation data is unlikely to affect our results in any important way.

[6] The volatility adjustment was accomplished by means of the following components GARCH model proposed

In order to make the plots more comprehensible we plot residuals with a uniform distribution that are defined as follows. If T has survivor function $S(t)$, then $S(T)$ is uniformly distributed. Both for short and long durations the residuals of both series are systematically underrepresented. Irrespectively of whether stock prices are adjusted for time-variation in volatility, the figure shows that the Inverse Gaussian model does not fit the data particularly well.

To demonstrate how time-series dependence in stock prices shows up in the durations, Figure 2 plots the sequence of marked durations based on a barrier or 'filter', $\delta = 0.10$. The length of each bar measures the time it took (in trading days) before stock prices either moved up by 10% or down by 10%, whichever happened first. Durations with $\tau^*_{down} < \tau^*_{up}$ are plotted below zero while those with $\tau^*_{up} < \tau^*_{down}$ are shown above zero. If duration spells were independently distributed over time, there should be no information in the sign or length of a given duration for the evolution in future duration spells. Clearly, however, this is not the pattern observed in Figure 2.[7] Short (long) durations tend to follow short (long) durations. This is related to the well-documented phenomenon of volatility clustering which gives rise to sequences of very short durations when volatility is high and the price index moves fast. Furthermore, there appears to be dependence in the direction of the market or, equivalently, in the sign of the change in the stock price: Both up and down moves have a tendency to cluster in time suggesting that if the previous move was up (down), then there is a higher probability that the following move will also be up (down). This indicates that there is predictability not simply in the volatility but also in the direction of the market, consistent with recent findings by Christoffersen & Diebold (2003).

Using equations (6-8) we computed durations, measured in calender days, for δ ranging from 1% to 15%. Varying the size of the barrier, δ, is equivalent to varying the return horizon, as is frequently done in the literature on variance ratios. It allows us to map out different degrees of dynamic dependencies for different filter sizes. Summary statistics for the up and down durations are presented in Table 1.

Two caveats should be mentioned before interpreting the results. First, we only measure stock prices at the end of each day so our observations are reported in integer days. For small filter sizes such as $\delta = 1\%$,

---

by Engle & Lee (1999) and extended to include an ARCH-in-mean effect: $r_t = \mu + \beta r_{t-1} + \gamma\sigma_t + \varepsilon_t,\ \sigma_t^2 = q_t + \alpha(\varepsilon^2_{t-1} - q_{t-1}) + \beta(\sigma^2_{t-1} - q_{t-1}),\ q_t = \omega + \rho(q_{t-1} - \omega) + \phi(\varepsilon^2_{t-1} - \sigma^2_{t-1})$, with $\varepsilon_t \sim N(0, \sigma_t^2)$. Using this specification we compute ARCH-adjusted, normalized returns $r_i^* = \hat{\varepsilon}_i / \hat{\sigma}_i$ where $\hat{\varepsilon}_i = r_i - \hat{\mu} - \hat{\beta} r_{i-1} - \hat{\gamma}\hat{\sigma}_i$ is the residual from the ARCH model and $\hat{\sigma}_i$ its estimated volatility. We then construct a new price index that is adjusted for first-order autocorrelation (reflecting the effects of asynchronous trading) and ARCH effects in the volatility and drift. Such adjustments ensure that the new price index has the same average drift and volatility as the original one.

[7] This is also supported by the Ljung-Box statistic for serial correlation. Using 15 lags this test statistic equals 110.6 for the T series and 477.2 for the $\ln(T)$ series. Under the null of no serial correlation in the durations the 5% critical value is 25.

it is likely that the data is measured with some error since the price index may well have moved by more than δ percent during the day and could even have generated two or more durations within the same day. Such measurement errors are difficult to deal with given our data limitations, but they are unlikely to be important for slightly larger filters.[8] It should also not be forgotten that the standard analysis based on returns measured, say, on a monthly basis is faced with the same problem. Calendar months do not have the same number of trading days: February typically has fewer trading days than most other months and the markets are closed some days during December. Second, the sequence of durations depends on the starting point of our analysis, which is February 17, 1885. We experimented with different starting days and found that the results are very robust with respect to the starting point, so this does not appear to be a problem.[9]

Returning to the results, as the required size of a price move increases, the number of completed durations declines. Thus, for a filter size, δ, of 1%, there are 4,713 up-moves and 3,853 down-moves. These numbers decline to 119 up-moves and 54 down-moves for the largest filter of 15%. Hence the length of the historical sample clearly constrains the statistical analysis for large filter sizes since the power of any statistical tests declines as fewer duration spells are completed. Another feature of the data is that the proportion of up-moves is higher, the larger the filter size.[10] This is to be expected since the positive drift in the stock price index is more important relative to the volatility of price movements for the durations based on the larger filter sizes. Equally unsurprising is the finding that the standard deviation of the durations increases as the filter size goes up as we would expect from equation (5).

3. Econometric Framework

The empirical results reported so far demonstrate the limitations to the standard geometric random walk model for stock prices in its most basic form or extended to allow for ARCH effects. In this section we therefore propose a framework for analyzing and modeling the distribution of durations of stock price movements that is sufficiently general to account for dynamic dependencies in completion times and the effects of time-varying predictor variables (covariates).

To analyze the dynamics in completion times for stock prices, we propose a framework that extends the

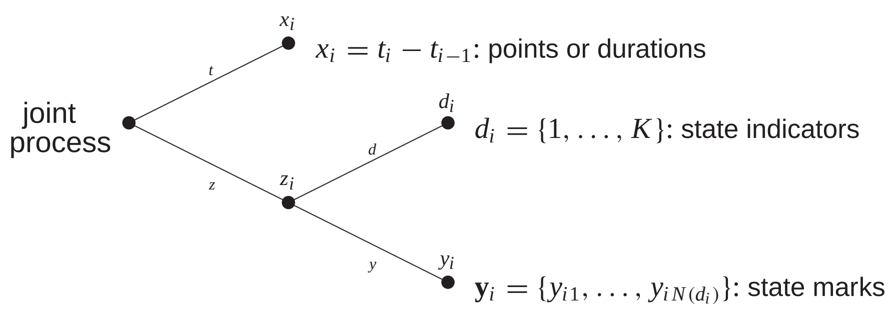
[8] To investigate the effect on maximum likelihood estimates of the parameters μ, σ implied by the observed durations, we conducted a Monte Carlos study for the model (1). We found that the effect on the parameter estimates was relatively small and limited to the smallest filters.

[9] Again this has an analogy when modeling returns over a fixed horizon such as one month. One could start the monthly horizon on any one day during a month and could also treat week-end and holiday effects in different ways.

[10] For the filter size of 1%, the proportion of down moves (as a percentage of the total number of up and down moves) is 45%, while it is 31% for the 15% filter.

models proposed by Lancaster (1990) and Engle (2000). Related but alternative models have also recently been suggested by Russell (2001), Rydberg & Shephard (1999), Kamionka (2000) and Bowsher (2002). These models build on the multivariate point process framework treated in e.g. Cox & Lewis (1972). In this setting up and down movements are treated as the marginal event processes. Kamionka (2000) and Bowsher (2002) suggest looking at the so-called pooled process of all durations. However, this framework is not appropriate for our purpose, where a down excursion often is censored by an up move. Hence our approach is much closer related to competing risk models.

Both notations and econometric methods are specific to the duration literature, so we first briefly introduce the underlying stochastic process, transition probability models and likelihood-based estimation methods. The timing of events such as the completion of a first passage time gives rise to a simple point process, $T = \{t_0, t_1, \ldots, t_n, \ldots\}$. The sequence of points in time, $t_0 < t_1 < \ldots < t_n < \ldots$ are the arrival times from which durations, $x_i = t_i - t_{i-1}$ can be defined. Since up and down moves in asset prices typically arrive at irregularly spaced points in time, the point process is an ideal modelling device for our purpose. Associated with every point, t_i, are M marks $\mathbf{z}_i = \{z_{i1}, \ldots, z_{iM}\}$ which often will be the variables of primary interest as they identify and measure the event that occurred.[11] The mark process can be decomposed into a univariate discrete process identifying the event that occurred (state indicators, d_i) and a vector process containing the marks that measure the variable defined at the associated event (state marks, $\mathbf{y}_i$):[12]



DECOMPOSITION OF A MULTI-STATE MARKED POINT PROCESS

Using this decomposition, movements in asset prices can be characterized through the stochastic process $\{(\mathbf{Y}_i, D_i, X_i);\ i = 1, \ldots, N\}$, where N is the total number of completed duration spells. We call this a

11 An example of this is Engle (2000) who analyzes data comprising two types of random variables such as the time of a financial transaction and the collection of variables observed as the trade takes place.

12 In our analysis the logarithm of the price (index) is a state mark, $\ln(P_{t_i}) = p_{t_i}$, and by construction $\left|p_{t_i} - p_{t_{i-1}}\right| \geq \delta$.

Multi-State Marked Point process, or in short an MSMP process. The joint distribution of the i'th observation conditional on the past filtration of $(\mathbf{Y}_i, D_i, X_i)$, $\mathcal{F}_{i-1}$, is given by

$$\mathcal{P}\left\{\mathbf{Y}_i \leq \mathbf{y}_i, D_i = d_i, X_i \leq x_i \mid \mathcal{F}_{i-1}; \boldsymbol{\omega}_i\right\}, \tag{9}$$

where $\boldsymbol{\omega}_i$ are the parameters determining the distribution of the i'th duration. Modelling this distribution directly can be very complicated and fortunately a simpler approach is available. Without loss of generality, the joint density can be factored into the product of the marginal density of the durations (parameterized by $\boldsymbol{\omega}_{1i}$) and the state indicators times the conditional distribution of the state marks (parameterized by $\boldsymbol{\omega}_{2i}$), all conditional on the past filtration:

$$\begin{aligned} & \mathcal{P}\left\{\mathbf{Y}_i \leq \mathbf{y}_i, D_i = d_i, X_i \leq x_i \mid \mathcal{F}_{i-1}; \boldsymbol{\omega}_i\right\} \\ = \quad & \mathcal{P}\left\{\mathbf{Y}_i \leq \mathbf{y}_i \mid D_i = d_i, X_i = x_i, \mathcal{F}_{i-1}; \boldsymbol{\omega}_{1i}\right\} \mathcal{P}\left\{D_i = d_i, X_i \leq x_i \mid \mathcal{F}_{i-1}; \boldsymbol{\omega}_{2i}\right\} \end{aligned} \tag{10}$$

It is obviously of separate interest to investigate how certain marks are related to D and X. On a macro scale interesting marks to consider would be GDP, inflation rates, interest rates etc. On a micro scale several interesting market microstructure questions can be addressed such as how volume, depth, spreads and liquidity relate to the distribution of D and X. We defer such issues to future research.

3.1. State Transitions

We are concerned with modelling the duration distribution of up and down moves in stock prices. Our focus is therefore on modelling the marginal distribution of D_i and X_i conditional on the past of the joint process:

$$\mathcal{P}\left\{D_i = d_i, X_i \leq x_i \mid \mathcal{F}_{i-1}; \boldsymbol{\omega}_{2i}\right\}. \tag{11}$$

An estimate of the density associated with this distribution can be obtained through maximum likelihood estimation based on $\{(D_i, X_i \mid \mathcal{F}_{i-1})\}_{i=1}^{N}$. We begin by considering the distribution of the completion time for the point process, T, i.e. $F_i(t) = \mathcal{P}\left\{X_i \leq t - t_{i-1} \mid \mathcal{F}_{i-1}; \boldsymbol{\omega}_{2i}\right\}$ with no concern for the alternating states. Following standard definitions the survivor function, S_i, and the hazard function, λ_i, are given by:[13]

$$S_i(t) = 1 - \mathcal{P}\left\{X_i \leq t - t_{i-1} \mid \mathcal{F}_{i-1}; \boldsymbol{\omega}_{2i}\right\}, \quad \text{for} \quad t_{i-1} \leq t < t_i,$$

and

$$\begin{aligned} \lambda_i(t) &= \lim_{h \to 0} \frac{\mathcal{P}\left\{t - t_{i-1} \leq X_i < t - t_{i-1} + h \mid X_i \geq t - t_{i-1}, \mathcal{F}_{i-1}; \boldsymbol{\omega}_{2i}\right\}}{h} \\ &= \frac{f_i(t)}{S_i(t)} = -\frac{d}{dt}\log(S_i(t)), \quad \text{for} \quad t_{i-1} \leq t < t_i. \end{aligned} \tag{12}$$

[13] As is common, we assume that $\mathcal{P}$ is absolutely continuous with respect to Lebesgue measure ς and that the density of X_i, $f = d\mathcal{P}/d\varsigma$, is the Radon-Nikodyn derivative of $\mathcal{P}$ with respect to ς.

The hazard rate, $\lambda_i(t)$, is the probability that the stock price passes through a barrier in any instant of time. For any $t > t_{i-1}$, the probability of this event is conditioned on both the fact that there has not been an event since t_{i-1} and on all past events.

Ignoring for the moment the conditioning on the past information, $\mathcal{F}_{i-1}$, and the parameters, $\boldsymbol{\omega}_{2i}$, and assuming that there are K possible states, we can use the fundamental law of probability to write

$$\mathcal{P}\left\{X_i \leq t - t_{i-1}\right\} = \sum_{k=1}^{K} \mathcal{P}\left\{X_i \leq t - t_{i-1} \mid D_i = k\right\} \mathcal{P}\left\{D_i = k\right\}.$$

Letting $\pi_{ik} = \mathcal{P}\{D_i = k\}$ and $F_{ik}(t) = \mathcal{P}\{X_i \leq t - t_{i-1} \mid D_i = k\}$, this simplifies to an expression that is easier to interpret:

$$F_i(t) = \sum_{k=1}^{K} \pi_{ik} F_{ik}(t).$$

Here the π_{ik}'s are the transition probabilities, i.e. the probabilities that the i'th event will be a move to state k. These sum to one across the K states. $F_i(t)$ is therefore a finite mixture distribution, $\pi_{i1}, \ldots, \pi_{iK}$ are the mixing weights and $F_{i1}(t), \ldots, F_{iK}(t)$ are the component distributions of the mixture. The specification is quite flexible as the transition probabilities are allowed to change from event to event as might be the case due to institutional shifts or changing market behavior. Using this notation, the survivor function can be written as follows:

$$S_i(t) = \sum_{k=1}^{K} \pi_{ik} S_{ik}(t). \tag{13}$$

Note that $\pi_{ik} S_{ik}(t)$ is the joint probability that no event occurs between t_{i-1} and t *and* that the eventual departure is to state k. Using (12) and (13), the transition intensities for the K states are given by

$$\begin{aligned} \lambda_i(t) &= -\frac{d}{dt} \log(S_i(t)) = \sum_{k=1}^{K} \frac{\pi_{ik} f_{ik}(t)}{S_i(t)} \\ &= \sum_{k=1}^{K} \lim_{h \to 0} \frac{\mathcal{P}\left\{t - t_{i-1} \leq X_i < t - t_{i-1} + h, D_i = k \mid X_i \geq t - t_{i-1}\right\}}{h} = \sum_{k=1}^{K} \lambda_{ik}(t), \end{aligned} \tag{14}$$

for $t_{i-1} \leq t < t_i$. Hence $\lambda_{ik}(t)dt$ is the probability of departure to state k in the short interval $(t, t + dt)$ given that there has been no event from t_{i-1} to t.[14]

[14] As pointed out by Lancaster (1990), $\lambda_{ik}(t)$ is not equal to $f_{ik}(t)/S_{ik}(t)$ as it would be if $\lambda_{ik}(t)$ were a hazard function conditional on departure to state k. The conditioning event for λ_{ik} is that no event occurs from t_{i-1} to t, not that no event occurs between t_{i-1} and t and that departure is to state k.

These events and their probabilities can be related to the observed data which is characterized by the probability $\mathcal{P}\{\text{departure to } k \text{ in } (t, t+dt)\}$. This can be factored out as follows

$$\begin{aligned} S_i(t)\lambda_{ik}(t)dt &= \mathcal{P}\{\text{departure to } k \text{ in } (t, t+dt) \mid \text{no events in } (t_{i-1}, t)\}\,\mathcal{P}\{\text{no events in } (t_{i-1}, t)\} \\ &= \pi_{ik} f_{ik}(t)dt, \end{aligned} \tag{15}$$

where $f_{ik}(t)dt$ is the probability of departure at time t given that the departure will be to state k. This equation shows the close connection between mixing weights and transition probabilities. Integrating (15) over t yields the following expression for the transition probabilities

$$\pi_{ik} = \int_0^\infty S_i(s)\lambda_{ik}(s)ds. \tag{16}$$

Once the transition intensities, λ_{ik}, have been estimated, the transition probabilities follow from (16).

Using these definitions, Appendix A sets up the likelihood function as a function of the transition intensities, while Appendix B provides extensions to include time-varying covariates.

4. Empirical Results

In this section we report empirical results for particular specifications of the transition intensity, moving from an autoregressive conditional duration specification to more flexible models that incorporate the effect of time-varying state variables (covariates).

4.1. Autoregressive Conditional Duration Models

The autoregressive conditional duration (ACD) model suggested by Engle & Russell (1998) and the log-ACD suggested by Bauwens & Giot (2000) provides a flexible dynamic specification. In our setup it takes the following form:

$$\lambda_{ik} = \exp\left[\lambda_{k1} + \lambda_{k2}\ln(\lambda_{i-1,k}) + \lambda_{k3}x_{i-1}(\lambda_{i-1,1} + \lambda_{i-1,2})\right] \quad \text{for } k = 1, 2. \tag{17}$$

We refer to this model as MSMP-dy. It allows the past transition intensity, $\lambda_{i-1,k}$, to influence the subsequent intensity in an autoregressive manner. In addition, we include an innovation term from the part of the past realized duration that was unpredictable ex-ante, $x_{i-1}(\lambda_{i-1,1} + \lambda_{i-1,2})$.

Estimation results are reported in Table 2. λ_{12} and λ_{22} identify the persistence in the transition intensities linked to the autoregressive term. Estimates of these parameters—shown in Panel A—are uniformly very high across all filter sizes and for both up- and down-moves. Furthermore, they only decline slowly from 0.98 to around 0.95 as the filter size increases from 1% to 15%. Clearly the expected duration is highly correlated over time. The higher the previous transition intensity, the more likely it is that the current intensity will also be high. Moreover, the significant−if somewhat smaller−coefficient estimates, $\hat{\lambda}_{13}$ and $\hat{\lambda}_{23}$, suggest that the previous innovation to the transition intensity also affects the subsequent transition intensity. These estimates are significant across all filter sizes.

The effect of past duration innovations on current transition intensities is negative across filter sizes both in up- and down-states. Transition intensities are therefore lower in both up- and down-states if the past duration was unexpectedly long. Interestingly, the largest effect of an innovation to the previous duration is found for the mid-sized filters (4-7%) in up-states but for the largest filters (14, 15%) in the down-states. Furthermore, an interesting asymmetry can be detected by comparing the estimates of λ_{13} and λ_{23}. If the last duration was surprisingly long, then the subsequent transition intensity goes down by more in up-markets than in down-markets for filters up to 7%, while the converse result holds for larger filter sizes. This suggests that it takes disproportionately longer for the market to move down by a large amount (corresponding to a large filter size, δ) following a long duration than it takes for the market to move up by the same amount. This effect is similar, but not identical, to mean reversion in asset prices in the following sense: a long previous duration (for a fixed filter size) translates into a smaller rate of return and this seems to be followed by a relative slow-down if the stock price moves down compared to if it moves up.[15]

Panel B reports diagnostic tests for this model. The estimated model is a mixture of exponential distributions, so the residuals can be obtained from the realized durations scaled by the transition intensities:

$$\xi_i = x_i(\lambda_{i1} + \lambda_{i2}) \quad \text{for} \quad i = 1, \ldots, N. \tag{18}$$

If the model fits the data, the residuals $\{\xi_i : i = 1, \ldots, N\}$ should be identically and independently exponentially distributed with a mean and a standard deviation of one. None of the filters produce any evidence of serial correlation in the residuals from this model. The only evidence against this model is the over-dispersion in the residuals based on filters, δ, under 5%.

[15] As the filter size, δ, increases, the standard errors of the parameter estimates increase and the absolute values of the associated t-statistics (for similar values of the estimated regression coefficients) decrease systematically. This a result of the fewer completed duration spells for the larger values of δ, c.f. Table 1, and the associated decline in power. The estimates in Tables 2-4 suggest that, on grounds of the power of the statistical tests, it is not advisable to consider filter sizes greater than 10-15%, depending on how many parameters the model requires estimating.

4.2. Interest Rate Effects

4.2.1. *Data*

A popular state variable frequently used in the analysis of stock market returns is the interest rate which is known to closely track cyclical movements in the economy and forecast returns. To make our results comparable to the existing literature, we therefore consider the effect of time-varying interest rates on transition intensities. Since there is no continuous data series on daily interest rates from 1885 to 1997, we construct our data from four separate sources. From 1885 to 1889 the source is again Shiller (2000). From 1890 to 1925, we use the interest rate on 90-day stock exchange time loans as reported in Banking and Monetary Statistics, Board of Governors of the Federal Reserve System (1943). These rates are provided on a monthly basis and we convert them into a daily series by simply applying the interest rate reported for a given month to each day of that month. From 1926 to 1954 we use the one-month T-bill rates from the risk-free rates file on the CRSP tapes, again reported on a monthly basis and converted into a daily series. Finally, from July 1954 to 1997 we use the daily Federal Funds rate. These three sets of interest rates are concatenated to form a single time series covering the full sample.

4.2.2. *Interest Rate Effects in the MSMP-dy Model*

Before turning to the model with time-varying covariates, first consider a dynamic model with constant covariates. This model extends the MSMP-dy specification (17) by including as covariates the state of the previous duration (d_{i-1}) which equals 1 if the $i-1$'th duration was an up-move and is 2 otherwise, the interest rate level at the beginning of the duration ($int(t_{i-1})$), and the change in the interest rate over the past duration ($\Delta int(t_{i-1})$). Hence, the model is specified as follow

$$\lambda_{ik}(t) = \exp\big[\lambda_{k1} + \lambda_{k2}\ln(\lambda_{i-1,k}) + \lambda_{k3}x_{i-1}(\lambda_{i-1,1} + \lambda_{i-1,2}) + \beta_{k1}(d_{i-1}-1) + \beta_{k2}int(t_{i-1}) + \beta_{k3}\Delta int(t_{i-1})\big] \quad \text{for } k = 1, 2 \tag{19}$$

We refer to this model as MSMP-dyc.

Estimation results for this model are reported in panel A of Table 3. Since this and the following model have more parameters to estimate, we only report results for filter sizes, δ, up to 10%. This ensures that we have sufficiently many durations available for estimation. Ignoring the other covariate effects, variations in the conditional transition intensities, $\lambda_{ikj} = \lambda_{i|d_i=k,d_{i-1}=j}$, due to the dummy indicators can be mapped out as follows:

$$\begin{aligned} \lambda_{i11} &= \exp\left[\lambda_{11}\right] && : \text{Intensity of } \textit{ups} \text{ coming from the } \textit{up} \text{ state.} \\ \lambda_{i12} &= \exp\left[\lambda_{11} + \beta_{11}\right] && : \text{Intensity of } \textit{ups} \text{ coming from the } \textit{down} \text{ state.} \\ \lambda_{i21} &= \exp\left[\lambda_{21}\right] && : \text{Intensity of } \textit{downs} \text{ coming from the } \textit{up} \text{ state.} \\ \lambda_{i22} &= \exp\left[\lambda_{21} + \beta_{21}\right] && : \text{Intensity of } \textit{downs} \text{ coming from the } \textit{down} \text{ state.} \end{aligned} \tag{20}$$

The positive and significant estimates of β_{11} and β_{21} in Table 3 show that the transition intensities of both up- and down-moves are higher coming from a down state. This is likely to reflect the well-known leverage effect: negative shocks have a disproportionately large effect on future volatility compared with positive shocks of the same size. Negative shocks thus have the effect of increasing the 'speed' of the stock price index, leading to higher intensities. Furthermore, the estimates of β_{21} are generally higher than those of β_{11} suggesting that the effect is highest for future down-markets.

The autoregressive parameter estimates λ_{12} and λ_{22} continue to be very high (0.97) for the smallest filters but are now much smaller for the largest filter sizes. Past duration shocks continue to have a negative effect on the transition intensities as both $\hat{\lambda}_{13}$ and $\hat{\lambda}_{23}$ are negative, suggesting that an unexpectedly long previous duration gives rise to a lower transition intensity and hence a longer expected future duration.

For both up and down moves the estimated effects of the level of the interest rate on the transition intensities ($\hat{\beta}_{12}$ and $\hat{\beta}_{22}$) are generally negative and significant for about half of the filters. Higher interest rates are thus associated with lower transition intensities for both up- and down-moves and will increase the expected completion time. Turning to the effects of changes in interest rates we find some quite intuitive results. Rising interest rates are associated with slower up moves (as $\hat{\beta}_{13}$ is always negative) and faster down moves (as $\hat{\beta}_{23}$ is positive). The effect of interest rate changes on transition intensities is also much higher than the effect of the interest rate level as can be seen in the size of the coefficient estimates.[16]

Interest rates therefore influence both *how long* it takes before the market moves by a certain percentage and also *whether* the market moves up or down. Furthermore, the coefficient on the interest rate covariate goes up as the filter size increases, suggesting different long- and short-run effects of interest rate movements on stock prices.

Finally the diagnostic tests reported in Panel B of Table 3 suggest that there is no serial correlation in the residuals although some evidence of excess dispersion remains for filter sizes up to 4%.

[16] The estimated effects are not directly comparable as they appear in the table, but we also estimated the model for all filter sizes with the interest rate variables standardized to have unit variance. This resulted in even bigger differences.

4.2.3. *Time-varying Interest Rate Effects in the* MSMP-dy *Model*

The model with covariates (19) only accounts for interest rate effects in as far as these have evolved prior to the beginning of a new duration. In contrast, the following time-varying covariate model allows interest rate effects that can change *during* duration spells:

$$\lambda_{ik}(t) = \exp[\lambda_{k1} + \lambda_{k2}\ln(\lambda_{i-1,k}(x_{i-1})) + \lambda_{k3}x_{i-1}(\lambda_{i-1,1}(x_{i-1}) + \lambda_{i-1,2}(x_{i-1}))$$
$$+ \beta_{k1}(d_{i-1} - 1) + \beta_{k2}int(t) + \beta_{k3}\Delta int(t)] \quad \text{for} \;\; k = 1, 2. \tag{21}$$

Here $\Delta int(t)$ is the most recent change in the interest rate occurring just before time t. Estimation of this model proceeds by means of the techniques developed in Appendix B.

This dynamic model with time-varying covariate effects is referred to as MSMP-dytvc. This is quite a complicated model that allows the transition intensities to depend on their lagged values, past innovations to the transition intensities as well as movements in the interest rate. Results from estimation of this model are provided in Table 4. The estimates of the persistence and innovation effects do not change much as compared to the earlier model (19) that did not allow interest rates to change during the duration. The coefficients measuring interest rate level effects are also largely unchanged. In contrast the effect of changes in interest rates during a duration spell is less pronounced than before. A possible interpretation of this finding is that changes in interest rates do not have an immediate impact on stock prices and take more time to show up systematically in future stock price movements.

4.3. Comparison with Fixed-horizon Results

It is fair to ask what can be learned from these results that could not easily have been gleaned from the standard fixed-horizon analysis of predictability in stock returns. Indeed, when stock prices follow a simple linear process driven by Gaussian innovations, the same information can be obtained from either approach. Which method dominates in a particular application will then only depend on which estimation approach is most efficient. However, outside this very special case (which is strongly rejected empirically), the two methods can yield very different insights as they are affected differently by nonlinearities and dynamic dependencies in stock prices. For example, ARCH effects in stock prices are generally viewed as mainly affecting the second moment of returns with a smaller effect on expected returns. In contrast, as we saw in our analysis, the mean duration time will also be strongly affected by ARCH effects.

The tendency of future down moves to follow current or past down moves is also an interesting new finding since the equivalent measure based on a fixed-horizon analysis—namely serial correlation in returns—tends to be very weak for a broadly defined stock price index such as that considered here. Again this demonstrates that the power of a duration-based approach to identify regularities in stock returns can differ significantly from that of the traditional fixed-horizon approach.

5. Out-of-Sample Forecast Evaluation

So far we have measured model fit by means of values of the likelihood function and by inspecting diagnostics for serial correlation and over-dispersion in the residuals. For an investor, however, the value of the various completion time models depends on how well they predict movements in stock prices out-of-sample since this will ultimately measure any potential improvements over a passive investment strategy such as buy-and-hold. To provide information on this issue we compute one-step-ahead forecasts of the time-varying transition probabilities and compare these to the subsequent direction of movements in stock prices. These forecasts simulate an investor's forecasts in real time since estimates of the model used to forecast at some point in time, t, only use information up to that point in time.

To get a sense of the movements in the transition probabilities, Figure 3 plots these over the full sample period for the MSMP-dyc model. Transition probabilities follow step-like functions and can vary significantly over short periods of time such as during the period with highly volatile interest rates in the late 1970s and early 1980s.

We next undertook a formal statistical test of the ability of the various model specifications to predict the direction of future moves in the stock market. Results from an out-of-sample forecast experiment are provided in Table 5. For filter sizes between 1% and 8% we use half the sample of durations for estimating the first forecast, and progressed to the end of the sample from there, re-estimating the parameters after each completed duration. For filter sizes of 9% and 10% we use three quarters of the durations for the initial estimation. We did not perform the experiment for the larger filter sizes between 11% and 15%, as these generate too few durations to enable a meaningful analysis.

To evaluate out-of-sample performance, the table reports 2×2 contingency tables cross-listing the sign of the predicted and realized move in stock prices. We also provide values of the Pearson χ^2 statistic for independence between the predicted and realized sign of the subsequent move and the PT test of sign predictability proposed by Pesaran & Timmermann (1992).

As a benchmark we report the results from a randomized strategy that predicts an up-move with probability π_{i1} and a down move with probability $1 - \pi_{i1}$, with π_{i1} chosen to match the duration data.[17] The

[17] The randomization is necessary since the statistical tests for directional forecasting require that the model does

performance of such a strategy is averaged across 5,000 simulations.

As expected, the benchmark model does not show any ability to predict the direction of the market and generates ‘hit rates’, defined as the sum of the probabilities on the main diagonal, i.e. the proportion of correctly predicted directional moves, between 50% and 53%. This fraction increases to between 52% and 62% under the MSMP-dy model and to between 55% and 67% under the MSMP-dyc model. Notice that the percentage of correct predictions generally increases for all three models as the filter size gets larger and the average duration increases—albeit not uniformly due to random sampling variation. This is consistent with the evidence of stronger predictability in long-run durations and is particularly noticeable under the MSMP-dyc model where two-thirds of the signs of the duration spells are correctly predicted for the largest filter sizes. Turning to the formal test statistics, the MSMP-dy model shows some ability to forecast up and down moves, but only for filter sizes smaller than or equal to 3%. In contrast, the MSMP-dyc model generates highly significant out-of-sample performance statistics for all filter sizes. The decline in the test statistic for the largest filter sizes is due to the decrease in the effective sample size as the filter size increases and fewer duration spells are terminated during our historical sample. This suggests that although large up- and down-moves in stock prices appear to be easier to predict than smaller moves, it also becomes more difficult to document predictability with much statistical precision, the larger the size of the move.

These results show that out-of-sample predictions of the direction of fixed-size moves in stock prices can be improved by accounting for information on the direction of previous moves and by including covariates and dynamics in the specification of the transition intensities.

Our results are related in interesting ways to those reported by Christoffersen & Diebold (2003) for fixed-horizon asset returns. Consistent with our results, Christoffersen and Diebold find that the sign of price changes is predictable. They also find that sign predictability is horizon-specific—which is consistent with our finding of considerable variation in ‘hit rates’ across filter sizes—and tends to be strongest at intermediate horizons.

6. Conclusion

This paper proposed a new approach to modelling dynamic dependencies in stock prices by studying the completion times of up and down moves that exceed a given threshold. Our econometric approach is very general and accounts for autoregressive dynamics in transition intensities, time-varying covariates and multiple states.

not always forecast the same direction.

Our results suggest that there is strong evidence of deviations from standard models of stock price movements, some of which are related to well-known patterns while others appear to be novel. Unsurprisingly, the expected duration of a move of a given size is strongly autocorrelated. If the past duration was short, the subsequent duration is also expected to be short and vice versa. This is related to the well-known volatility clustering effects. While this type of persistence is strong for both up and down moves, there are also important asymmetries that cannot be explained by standard ARCH effects. Arriving from a down state the transition probability to another down-state is higher, while the transition probability of an up-state declines, thus increasing the likelihood of a future down move. The level of interest rates both at the outset of an event and during the event also affects the durations of up- and down-moves asymmetrically. Higher interest rates lead to slower up-moves but tend to increase the speed of down moves.

There are many interesting portfolio implications of our findings. Predictability of fixed-horizon returns has been demonstrated to translate into significant changes in investors' optimal asset allocation, c.f. Campbell, Chan & Viceira (2002) and Kandel & Stambaugh (1996). Similarly, the dynamic dependencies in the durations of up and down moves in stock prices uncovered here could be exploited in an investment rule that lowers the weight in stocks if the last duration spell was long and represented a down-move so that the probability of a subsequent short-lived up-move is lower. A more formal approach could be based on using Monte Carlo simulation methods to sample repeatedly from a particular duration model, e.g., MSMP-dyc, and then use dynamic programming methods under a specified objective function such as power utility to optimize the allocation to stocks as a function of the current state variables. This is in line with methods proposed by Detemple et al. (2003) and can readily be used even for nonlinear dynamic models for stock prices such as the ones entertained here.

For risk measurement or risk management purposes, there is important information in the duration distribution or survivor function for up and down moves. In fact, the methods developed in Appendix B would allow an investor to compute these functions in real time conditional upon the current state variables. This provides an alternative measure of the risk associated with holding stocks that naturally complements measures based on a fixed investment horizon such as Value at Risk over a one day or a one week period. When stock prices do not follow a simple stochastic process such as geometric Brownian motion, none of the standard measures of risk will be a sufficient statistic so combinations of different measures of risk will generally provide a more complete picture of the risks associated with holding stocks.

APPENDIX A: LIKELIHOOD FUNCTION

Equations (13), (14), (15) and (16) in the main text specify the key equations for the probabilities in the marginal model. Here we derive the density of the marginal process for the state indicators and durations, (11), conditional on past information, $\mathcal{F}_{i-1}$.

If we only observe that an event occurred after time t then the probability of that event is just $f_i(t)dt$. Moreover, if it is known that the event at time t was a move to state k, then the probability of this event is $f_{ik}(t)\pi_{ik}$ which again equals $\lambda_{ik}(t)\exp\left\{-\int_0^t \sum_{k=1}^K \lambda_{ik}(s)ds\right\}dt$, where K is the number of possible states. The joint p.d.f. of the indicator variables $D_{i1}, \ldots, D_{iK}$ and the duration X_i is therefore

$$f(d_{i1}, \ldots, d_{iK}, x_i) = \exp\left\{-\int_0^{x_i} \sum_{k=1}^K \lambda_{ik}(s)ds\right\} \prod_{k=1}^K \lambda_{ik}(x_i)^{d_{ik}}.$$

Hence the log-likelihood function for the i'th departure is given by

$$l(\mathbf{d}_i, x_i; \boldsymbol{\omega}_{2i} \mid \mathcal{F}_{i-1}) = \sum_{k=1}^K \left[d_{ik}\log(\lambda_{ik}(x_i)) - \int_0^{x_i} \lambda_{ik}(s)ds \right].$$

Summing over all events, N, the log-likelihood function is

$$l(\mathbf{d}, \mathbf{x}; \boldsymbol{\omega}_2 \mid \mathcal{F}_N) = \sum_{i=1}^N l(\mathbf{d}_i, x_i, \boldsymbol{\omega}_{2i} \mid \mathcal{F}_{i-1}). \tag{A.1}$$

Notice the simplicity of this likelihood function when expressed in terms of the transition intensities. Specifying the transition intensities − preferably based on economic theory − leads directly to the likelihood function. All dependence of the likelihood on parameters and past information comes through the transition intensities. We make this dependence explicit in the following notation that specifies the transition intensities conditional on the past filtration:

$$\lambda_{ik}(t) \equiv \lambda_{ik}(t \mid \mathcal{F}_{i-1}, \boldsymbol{\omega}_{2ik}) \equiv \lambda_{ik}(t, \mathbf{Y}^{i-1}, D^{i-1}, X^{i-1}; \boldsymbol{\omega}_{2ik}).$$

We continue to adopt a notation that expresses the transition intensities as a function of $\mathbf{Y}^{i-1}$, D^{i-1} and X^{i-1} because much of the empirical analysis aims to find the best functional form for the transition intensities as a function of these components.

If only two events are possible−namely an up-move or a down-move−the log-likelihood function (A.1) takes the form:

$$l(\mathbf{d}, \mathbf{x}, \boldsymbol{\omega}_2 \mid \mathcal{F}_N) = \sum_{i=1}^N \{d_{i1}\log(\lambda_{i1}(x_i)) + d_{i2}\log(\lambda_{i2}(x_i)) - \int_0^{x_i} (\lambda_{i1}(s) + \lambda_{i2}(s))ds \}. \tag{A.2}$$

Estimation of this equation requires specifying a model for the transition intensities $\lambda_{i1}(x_i)$ and $\lambda_{i2}(x_i)$. Since the hazard function, $\lambda_i(t)$, and the transition intensities are instantaneous probabilities, they must be positive. To avoid introducing complicated restrictions on the parameters resulting from this constraint, we model the logarithm of $\lambda_{ik}(s),\ k = 1, 2$. The simplest specification of the transition intensities ignores any dependence on past marks and therefore takes the form

$$\lambda_{ik}(t) = \phi\left(\exp\left[\psi_i(D^{i-1}, X^{i-1}, \Lambda^{i-1}(t))\right], t\right),\ k = 1, 2, \tag{A.3}$$

for $\Lambda^{i-1}(t) = \{\lambda_{1,i-1}(t), \lambda_{2,i-1}(t), \lambda_{1,i-2}(t), \lambda_{2,i-2}(t), \ldots\}$.

Duration-invariant specifications assume that the transition intensities are independent of t. Hence they do not depend on the time since the last event and, using that $\phi(s, t) = s$, simplify to[18]

$$\lambda_{ik} = \exp\left[\psi_i(D^{i-1}, X^{i-1}, \Lambda^{i-1})\right],\ k = 1, 2. \tag{A.4}$$

This reduces the log-likelihood function to

$$l(\mathbf{d}, \mathbf{x}, \boldsymbol{\omega}_2 \mid \mathcal{F}_N) = \sum_{i=1}^{N} \{d_{i1}\log(\lambda_{i1}) + d_{i2}\log(\lambda_{i2}) - x_i(\lambda_{i1} + \lambda_{i2})\}.$$

The hazard function is simply $\lambda_i = \lambda_{i1} + \lambda_{i2}$ while, using (16), the two transition probabilities are

$$\pi_{ik} = \frac{\lambda_{ik}}{\lambda_i},\ \ k = 1, 2. \tag{A.5}$$

In this model the probability that the next event will be an up or a down move of size δ therefore does not depend on when the event occurs and only depends on the past through π_{ik}.

Appendix B: Extensions to Time-varying Covariates[19]

A potential drawback of the models formulated so far is that they only condition on variables that are determined before the beginning of a particular duration and do not allow explanatory variables to change *within* a duration. While such models are of separate interest, it is clearly interesting to develop a framework that allows us to address the effect of a change in variables such as the interest rate on the expected duration of the event as this would allow investors to update their estimates in real time. This effect is summarized by

[18] We take duration-dependence to mean that the hazard rates for a given event change with the length of the associated duration and thus depend on t. Time-dependence means that the hazard rates depend on i, the event number. Using these definitions the intensities defined by (A.4) are duration-independent, but time-dependent.

[19] This material was previously contained in a paper entitled A Generalized Gamma Autoregressive Conditional Duration Model.

the change in the mean residual lifetime at the point in time where the covariate changes. The mean residual lifetime at x measures the expected remaining lifetime of the most recent duration:

$$mrl(x) = E(X - x|X > x) = \frac{\int_x^\infty (t - x) f(t)dt}{S(x)} = \frac{\int_x^\infty S(t)dt}{S(x)} \tag{A.1}$$

Clearly this expectation changes with the covariates through the hazard function. In the exponential case the mean residual lifetime is simply given by

$$mrl_i(x) = \frac{\int_x^\infty \exp(-\theta_i(x)u)\, du}{e^{-\theta_i(x)x}} = \theta_i(x)^{-1}.$$

To deal with this complication we extend the MSMP model class to incorporate time-varying covariates building on earlier work by Petersen (1986) and Hamilton & Jorda (2002).[20] To keep the notation as simple as possible, we present the likelihood for a non-negative continuous random variable, T, that measures the inter-event arrival time, and its state indicator, d. The time spent waiting for the next event depends on a set of exogenous covariates, $\mathbf{V}(t)$, that follow step-functions or deterministic continuous functions of time. Hence the observed duration, t, may be partitioned into intervals $0 = a_0 < \ldots < a_k = t$ where $\mathbf{V}$ stays constant at $\mathbf{V}(a_{j-1})$ in $[a_{j-1}; a_j)$ and jumps to $\mathbf{V}(a_j)$ at a_j, for $j = 1, \ldots k$. In this setting the hazard rate conditional on the path of the covariates up to time t is defined as

$$\begin{aligned}\lambda(t|\mathbf{V}(t)) &= \lambda_1(t|\mathbf{V}(t)) + \lambda_2(t|\mathbf{V}(t)) \\ &= \lim_{h\downarrow 0} \frac{\mathcal{P}(t \le T < t+h, D=1|T \ge t, \mathbf{V}(t))}{h} + \lim_{h\downarrow 0} \frac{\mathcal{P}(t \le T < t+h, D=0|T \ge t, \mathbf{V}(t))}{h}.\end{aligned}$$

Based on the partitioning sequence, $\{a_j\}_0^k$, the cumulative hazard is

$$\Lambda(t) = \int_0^t \lambda(u|\mathbf{V}(u))du = \sum_{j=1}^{k} \int_{a_{j-1}}^{a_j} \lambda(u|\mathbf{V}(a_{j-1}))du. \tag{A.2}$$

The assumption of step-function covariates makes integration easy because the total duration can always be partitioned into sub-periods during which all the covariates stay constant. As usual the survivor function is given by $S(t) = \exp(-\Lambda(t))$, while the likelihood of the i'th duration is given by

$$L_i = \lambda_{i1}(t_i|\mathbf{V}(t_i^-))^{d_{i1}} \lambda_{i2}(t_i|\mathbf{V}(t_i^-))^{d_{i2}} S(t_i), \tag{A.3}$$

[20] One way of incorporating time-varying covariate effects is to utilize the Cox regression approach suggested by Cox (1972) and detailed in Cox (1975). The Cox regression is a semiparametric method which in a first step estimates the effect of the covariates parametrically and then, in the second step, estimates a nonparametric baseline hazard. This method assumes that the observations are independent, and it is not clear how one would extend it to the case with sequences of dependent durations.

and the log-likelihood function of the full set of events for the ith duration is

$$\begin{aligned} \ln L_i &= d_{i1} \ln \lambda_{i1}(t_i|\mathbf{V}(t_i^-)) + d_{i2} \ln \lambda_{i2}(t_i|\mathbf{V}(t_i^-)) + \ln S(t_i) \\ &= d_{i1} \ln \lambda_{i1}(t_i|\mathbf{V}(t_i^-)) + d_{i2} \ln \lambda_{i2}(t_i|\mathbf{V}(t_i^-)) - \sum_{j=1}^{k} \int_{a_{j-1}}^{a_j} \lambda(u|\mathbf{V}(a_{j-1}))du. \end{aligned} \quad \text{(A.4)}$$

In practice we have a sample of events at times $t_0, \ldots, t_i, t_{i+1}, \ldots, t_N$ and wish to consider the duration from the t_{i-1}'th to the t_i'th event, $x_i = t_i - t_{i-1}$. The covariates, x_i, change at times $a_{i0} < \ldots < a_{ik} \in [t_{i-1}; t_i)$ with $a_{i0} = t_{i-1}$ and $a_{ik} = t_i$, where i refers to the duration number and k refers to the partition. The log-likelihood for the full sample is thus given by[21]

$$\ln L = \sum_{i=1}^{N} \left(d_{i1} \ln \lambda_{i1}(t_i|\mathbf{V}(t_i^-)) + d_{i2} \ln \lambda_{i2}(t_i|\mathbf{V}(t_i^-)) - \sum_{j=1}^{k} \int_{a_{ij-1}}^{a_{ij}} \lambda(u|\mathbf{V}(a_{ij-1}))du \right) \quad \text{(A.5)}$$

and the model is fully specified through the functional form of the hazard function.

Transition intensities that include dynamic autoregressive effects and time-varying covariates now take the form

$$\ln(\lambda_{ik}(t)) = \lambda_{k1} + \lambda_{k2} \ln(\lambda_{i-1,k}(x_{i-1})) + \lambda_{k3} x_{i-1} \left(\lambda_{i-1,1}(x_{i-1}) + \lambda_{i-1,2}(x_{i-1})\right) + \boldsymbol{\beta}_k \mathbf{V}(t) \quad \text{(A.6)}$$
$$\text{for } k = 1, 2,$$

where $\lambda_{i-1,k}(x_{i-1})$ is the transition intensity at the end of the previous duration. Notice that the hazard rate only depends on t through time-variations in the covariates:

$$\lambda(t - t_{i-1}|\mathbf{V}(t)) = \lambda_{i1}(t) + \lambda_{i2}(t).$$

Utilizing the step-function assumption, from (A.5) we get the following log-likelihood:

$$\ln L = \sum_{i=1}^{N} \left(d_{i1} \ln \lambda_{i1}(a_{ik-1}) + d_{i2} \ln \lambda_{i2}(a_{ik-1}) - \sum_{j=1}^{k} (a_{ij} - a_{ij-1}) \left[\lambda_{i1}(a_{ij-1}) + \lambda_{i2}(a_{ij-1})\right] \right). \quad \text{(A.7)}$$

We use this specification in our empirical analysis in Section 4. It provides a general framework that encompasses as special cases MSMP models with constant transition probabilities ($\lambda_{k2} = \lambda_{k3} = 0$ and $\boldsymbol{\beta}_k = \mathbf{0}$) and pure autoregressive conditional duration models ($\boldsymbol{\beta}_k = \mathbf{0}$).

[21] Here we assume that there are no censored observations, i.e. that all events have finished by the end of the sample. This assumption is of course harmless since it at most affects one duration spell (the one at the end) in our sample.

References

Bauwens, L. & Giot, P. (2000), 'The logartihmic ACD model: an application to the bid-ask quote process of three NYSE stocks', *Annales d'Economie et de Statistique* **60**(1), 117–149.

Bowsher, C. G. (2002), Modelling security market events in continuous time: Intensity based, multivariate point process models. unpublished manuscript, Nuffield College, Oxford.

Brock, W., Lakonishok, J. & Lebaron, B. (1992), 'Simple technical trading rules and the stochastic properties of stock returns.', *Journal of Finance* **47**, 1731–1764.

Campbell, J. (1987), 'Stock returns and the term structure', *Journal of Financial Economics* **18**, 373–399.

Campbell, J., Chan, Y. & Viceira, L. (2002), 'A multivariate model of strategic asset allocation', *Journal of Financial Economics* **67**, 41–80.

Campbell, J. & Cochrane, J. (1999), 'By force of habit: A consumption-based explanation of aggregate stock market behavior', *Journal of Political Economy* **107**, 205–251.

Campbell, J. & Shiller, R. (1988), 'The dividend price ratio and expectations of future dividends and discount factors', *Review of Financial Studies* **1**, 195–228.

Christoffersen, P. F. & Diebold, F. X. (2003), 'Financial asset returns, direction-of-change forecasting, and volatility dynamics', *NBER working paper 10009* .

Cox, D. R. (1972), 'Regression models and life-tables (with discussion)', *Journal of the Royal Statistical Society* **B 34**, 187–220.

Cox, D. R. (1975), 'Partial likelihood', *Biometrika* **62**(2), 269–276.

Cox, D. R. & Lewis, P. A. W. (1972), Multivariate point processes, *in* L. M. Le Cam, J. Neyman & E. L. Scott, eds, 'Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability', Vol. III, University of California Press, Berkeley and Los Angeles, pp. 401–448. volume III.

De Bondt, W. F. M. & Thaler, R. (1985), 'Does the stock market overreact?', *Journal of Finance* **40**, 793–805.

Detemple, J., Garcia, R. & Rindisbacher, M. (2003), 'A monte carlo method for optimal portfolios', *Journal of Finance* **58**, 401–446.

Donaldson, G. & Kamstra, M. (1996), 'new dividend forecasting procedure that rejects bubbles in asset prices: The case of 1929s stock crash', *Review of Financial Studies* **9**, 333–383.

Engle, R. F. (2000), 'The econometrics of ultra-high frequency data', *Econometrica* **68**(1), 1–22.

Engle, R. F. & Lee, G. G. J. (1999), A permanent and transitory component model of stock return volatility, *in* R. F. Engle & H. White, eds, 'Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger', Oxford University Press, Oxford, pp. 475–497.

Engle, R. F. & Russell, J. R. (1998), 'Autoregressive conditional duration: a new model for irregularly spaced transaction data', *Econometrica* **66**(5), 1127–1163.

Engle, R. F. & Russell, J. R. (2002), Analysis of high frequency and transaction data", *in* Y. Aït-Sahalia & L. P. Hansen, eds, 'forthcoming in Handbook of Financial Econometrics', Vol. I, Elsevier-North Holland, Amsterdam.

Fama, E. F. & french, K. R. (1988), 'Permanent and temporal components of stock prices', *Journal of Political Economy* **96**(2), 246–273.

Fama, E. F. & french, K. R. (1989), 'Business conditions and expected returns on stocks and bonds', *Journal of Financial Economics* **29**, 23–49.

Ferson, W. & Harvey, C. (1991), 'The variation of economic risk premiums', *Journal of Political Economy* **99**, 385–415.

Glosten, L. R., Jagannathan, R. & Runkle, D. E. (1993), 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance* **48**, 1779–1801.

Guidolin, M. & Timmermann, A. (2004), 'Strategic asset allocation and consumption decisions under multivariate regime switching', *Manuscript, University of Virginia and UCSD* .

Hamilton, J. D. & Jorda, O. (2002), 'A model of the Federal Funds rate target', *Journal of Political Economy* **111**(5), 1135–1167.

Kamionka, T. (2000), 'La modélisation des données haute fréquence', *Journal de la Société Française de Statistique* **141**(1-2), 167–211.

Kandel, S. & Stambaugh, R. (1996), 'On the predictability of stock returns: An asset-allocation perspective', *Journal of Finance* **51**, 385–424.

Keim, D. & Stambaugh, R. (1986), 'Predicting returns in the stock and bond markets', *Journal of Financial Economics* **17**, 357–390.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Cambridge university press.

Lunde, A. & Timmermann, A. (2004), 'Duration dependence in stock prices: An analysis of Bull and Bear markets', *Journal of Business & Economic Statistics* **22**(3), 253–273.

McQueen, G. R. & Thorley, S. (1994), 'Bubbles, stock returns and duration dependence', *Journal of Financial and Quantitative Analysis* **29**, 379–401.

Merton, R. C. (1969), 'Lifetime portfolio selection under uncertainty: The continuous-time case', *Review of Economics & Statistics* **51**, 247–257.

Pesaran, M. H. & Timmermann, A. (1992), 'A simple non-parametric test of predictive performance', *Journal of Business & Economic Statistics* **10**, 461–465.

Petersen, T. (1986), 'Fitting parametric survival models with time-dependent covariates', *Applied Statistics* **35**(3), 281–288.

Poterba, J. & Summers, L. (1988), 'Mean reversion in stock returns: evidence and implications', *Journal of Financial Economics* **22**, 27–60.

Russell, J. R. (2001), Econometric modelling of multivariate irregulary-spaced high-frequency data. Unpublished manuscript, University of Chicago, Graduate School of Business.

Rydberg, T. H. & Shephard, N. (1999), Modelling trade-by-trade price movements of multiple assets using multivariate compound Poisson processes. Unpublished Manuscript, Nuffield College, Oxford.

Samuelson, P. (1969), 'Lifetime portfolio selection by dynamic stochastic programming', *Review of Economics & Statistics* **51**, 239–246.

Schwert, G. W. (1990), 'Indexes of united states stock prices from 1802 to 1987', *Journal of Business* **63**, 399–426.

Shiller, R. J. (2000), *Irrational Exuberance*, Princeton University Press.

Sullivan, R., Timmermann, A. & White, H. (1999), 'Data-snooping, technical trading rule performance, and the bootstrap', *Journal of Finance* **54**, 1647–1691.

Timmermann, A. (1996), 'Excess volatility and predictability of stock returns in autoregressive dividend models with learning', *Review of Economic Studies* **63**(4), 523–557.

Timmermann, A. (2001), 'Structural breaks, incomplete information and stock prices', *Journal of Business & Economic Statistics* **19**, 299–315.

Veronesi, P. (1999), 'Stock market overreaction to bad news in good times: A rational expectations equilibrium model', *Review of Financial Studies* **5**, 975–1007.

Appendix: Tables and Figures

Table 1: Summary Statistics for the duration of up and down movements.

Series	Filter	no. obs	Mean	Median	Standard dev.	Min.	Max.
	1	4713	4.07	3	3.58	1	35
	2	2145	9.59	7	9.28	1	86
	3	1261	16.83	12	16.10	1	126
	4	831	25.91	18	25.55	1	221
	5	596	36.77	27	37.31	1	373
	6	458	47.52	34	44.96	1	280
	7	361	62.33	48	56.23	1	377
Up excursions (days)	8	287	80.72	58	78.53	3	510
	9	240	98.46	70	89.69	2	519
	10	220	109.27	74	108.71	2	704
	11	183	127.87	88	110.97	3	645
	12	167	147.41	108	145.54	3	1059
	13	145	165.12	112	159.80	2	862
	14	130	179.59	134	152.06	3	702
	15	119	211.04	151	177.58	4	999
	1	3853	3.17	2	2.96	1	28
	2	1681	6.45	4	7.21	1	88
	3	961	10.57	6	12.53	1	119
	4	607	16.23	9	18.68	1	117
	5	411	23.10	12	27.70	1	169
	6	301	31.92	18	38.71	1	296
	7	227	39.08	22	46.52	1	329
Down excursions (days)	8	175	46.44	21	57.81	1	317
	9	137	55.88	27	66.04	1	339
	10	128	56.67	28	69.24	1	323
	11	96	82.34	35	137.27	2	850
	12	86	77.74	40	114.86	1	866
	13	71	103.63	67	150.59	2	904
	14	62	127.81	77	167.07	1	884
	15	54	115.48	76	161.84	5	861

This table reports summary statistics for up and down excursions. The durations are computed using the definition of barriers given by equations (6-8) and are measured in trading days.

Table 2: ESTIMATING MSMP-dy

Panel A: Parameter Estimates

Filter	Parameters					
	λ_{11} ($t_{\lambda_{11}=0}$)	λ_{12} (*SE*)	λ_{13} ($t_{\lambda_{12}=0}$)	λ_{21} ($t_{\lambda_{21}=0}$)	λ_{22} (*SE*)	λ_{23} ($t_{\lambda_{22}=0}$)
1	0.076 (11.49)	0.984 (0.002)	-0.085 (-11.62)	0.064 (8.72)	0.989 (0.002)	-0.072 (-8.74)
2	0.124 (9.61)	0.976 (0.005)	-0.137 (-9.40)	0.118 (7.97)	0.979 (0.005)	-0.131 (-7.87)
3	0.142 (7.57)	0.970 (0.009)	-0.156 (-7.21)	0.157 (7.86)	0.974 (0.007)	-0.174 (-7.83)
4	0.212 (8.78)	0.940 (0.013)	-0.239 (-8.78)	0.199 (6.54)	0.962 (0.008)	-0.223 (-6.70)
5	0.221 (7.65)	0.918 (0.021)	-0.254 (-7.62)	0.173 (6.10)	0.958 (0.009)	-0.202 (-6.58)
6	0.185 (5.22)	0.928 (0.032)	-0.213 (-4.92)	0.131 (4.80)	0.967 (0.009)	-0.155 (-5.38)
7	0.172 (2.96)	0.927 (0.061)	-0.201 (-2.49)	0.139 (4.81)	0.966 (0.013)	-0.165 (-5.60)
8	0.082 (3.53)	0.985 (0.014)	-0.088 (-3.45)	0.123 (4.56)	0.983 (0.014)	-0.138 (-4.37)
9	0.079 (2.81)	0.981 (0.025)	-0.086 (-2.84)	0.141 (4.14)	0.978 (0.025)	-0.161 (-3.83)
10	0.142 (2.79)	0.975 (0.032)	-0.150 (-2.62)	0.190 (4.44)	0.976 (0.026)	-0.210 (-3.94)
11	0.105 (3.11)	0.975 (0.026)	-0.113 (-3.10)	0.164 (4.06)	0.969 (0.027)	-0.195 (-3.96)
12	0.090 (3.08)	0.976 (0.016)	-0.097 (-3.43)	0.155 (3.78)	0.965 (0.023)	-0.189 (-3.81)
13	0.096 (3.08)	0.977 (0.016)	-0.101 (-3.64)	0.171 (4.71)	0.973 (0.021)	-0.198 (-4.91)
14	0.093 (2.56)	0.970 (0.016)	-0.102 (-2.91)	0.217 (4.27)	0.951 (0.025)	-0.271 (-4.50)
15	0.081 (1.68)	0.964 (0.026)	-0.091 (-2.14)	0.173 (2.94)	0.938 (0.035)	-0.244 (-4.34)

Panel B: Diagnostics

Filter	$LB(x, 15)$	$LB(\xi, 15)$	$s_\xi - 1$	E-R EDT	Max Like	LR-bench.
1	3965.40	11.55	0.212	-12.40	-13929.00	1094.90
2	2402.30	18.86	0.123	-5.05	-6014.30	909.28
3	1687.90	13.55	0.116	-3.63	-3420.30	665.55
4	865.86	13.19	0.116	-2.94	-2187.10	479.48
5	488.43	8.30	0.041	-0.894	-1530.40	327.38
6	335.61	14.03	0.014	0.277	-1165.50	212.25
7	268.24	17.11	0.032	-0.540	-907.78	153.04
8	104.51	15.18	0.052	0.811	-721.96	94.69
9	125.75	7.72	0.014	0.196	-581.12	92.07
10	116.52	14.38	0.022	0.290	-522.72	109.80
11	89.02	14.85	0.037	0.440	-422.04	71.34
12	49.08	16.97	0.097	1.14	-389.97	51.32
13	66.06	9.47	0.091	0.984	-322.24	62.60
14	79.89	15.97	0.089	-0.833	-293.54	38.97
15	54.61	11.53	0.005	0.048	-266.67	26.80

Table 2 presents the results of estimating the dynamic MSMP model given in equation (17). Panel A shows the parameter estimates. The first column gives the filter size; columns 2-4 show the parameter estimates for the up intensities and columns 5-7 show the parameter estimates for the down intensities. T-statistics or standard errors based on a Quasi-Likelihood are shown in parentheses.

Panel B reports diagnostics for the fitted models. LB(x,15) and LB(ξ,15) denotes the Ljung Box statistic with 15 lags, computed on raw durations and the residuals defined in (18) ($\chi^2_{0.95}(15) = 25$). Column 4 shows the amount of over-dispersion in the residuals. In column 5, E-R EDT is a test for over-dispersion with respect to the exponential distribution suggested by Engle & Russell (1998). The last column shows a likelihood ratio test of the null $\lambda_{12} = \lambda_{22} = 0$.

Table 3: ESTIMATING MSMP-dyc

Panel A: Parameter Estimates

Filter	Parameters											
	$\hat{\lambda}_{11}$ (t$_0$)	$\hat{\lambda}_{12}(SE)$	$\hat{\lambda}_{13}$ (t$_0$)	$\hat{\beta}_{11}$ (t$_0$)	$\hat{\beta}_{12}$ (t$_0$)	$\hat{\beta}_{13}$ (t$_0$)	$\hat{\lambda}_{21}$ (t$_0$)	$\hat{\lambda}_{22}(SE)$	$\hat{\lambda}_{23}$ (t$_0$)	$\hat{\beta}_{21}$ (t$_0$)	$\hat{\beta}_{22}$ (t$_0$)	$\hat{\beta}_{23}$ (t$_0$)
1	.057 (8.08)	.971 (.005)	-.09 (-12)	.041 (5.82)	-.001 (-2.13)	-.033 (-1.81)	.022 (3.20)	.974 (.007)	-.08 (-7.0)	.089 (4.58)	-.001 (-1.77)	.068 (4.05)
2	.086 (6.66)	.951 (.010)	-.14 (-10)	.093 (6.41)	-.003 (-2.91)	-.012 (-0.50)	.036 (2.26)	.946 (.018)	-.15 (-6.9)	.199 (4.18)	-.003 (-2.28)	.132 (5.59)
3	.097 (4.54)	.930 (.017)	-.17 (-7.3)	.119 (5.53)	-.004 (-2.29)	-.050 (-1.71)	.039 (1.45)	.922 (.027)	-.20 (-6.7)	.283 (3.85)	-.005 (-2.11)	.142 (4.44)
4	.123 (4.00)	.890 (.019)	-.23 (-8.1)	.187 (5.74)	-.007 (-2.55)	-.064 (-1.73)	.025 (0.62)	.891 (.032)	-.24 (-6.8)	.396 (4.42)	-.010 (-2.52)	.099 (2.81)
5	.123 (3.33)	.856 (.027)	-.25 (-7.8)	.236 (5.05)	-.009 (-2.24)	-.093 (-2.19)	-.030 (-0.39)	.862 (.054)	-.23 (-7.1)	.489 (3.22)	-.012 (-2.12)	.089 (2.14)
6	.082 (1.97)	.866 (.037)	-.20 (-6.0)	.219 (3.88)	-.007 (-1.63)	-.075 (-1.63)	-.036 (-0.26)	.861 (.116)	-.21 (-3.7)	.442 (1.48)	-.012 (-1.20)	.119 (2.68)
7	.091 (1.59)	.844 (.064)	-.21 (-3.5)	.199 (3.27)	-.006 (-1.05)	-.088 (-1.47)	.019 (0.24)	.901 (.063)	-.18 (-4.3)	.282 (1.68)	-.008 (-1.17)	.138 (3.65)
8	.060 (0.72)	.648 (.091)	-.26 (-3.7)	.360 (3.44)	-.021 (-1.93)	-.075 (-1.37)	-.264 (-1.65)	.646 (.084)	-.25 (-3.8)	.854 (4.28)	-.033 (-2.79)	.151 (2.57)
9	-.035 (-0.39)	.699 (.109)	-.17 (-2.3)	.336 (3.28)	-.012 (-1.31)	-.127 (-2.05)	-.166 (-0.54)	.750 (.267)	-.24 (-2.8)	.679 (1.25)	-.022 (-0.83)	.166 (3.13)
10	.038 (0.48)	.733 (.082)	-.26 (-4.2)	.554 (3.89)	-.020 (-1.90)	-.068 (-1.01)	-.068 (-0.45)	.783 (.112)	-.28 (-4.5)	.718 (2.34)	-.029 (-1.57)	.177 (3.09)

Panel B: Diagnostics

Filter	$LB(x, 15)$	$LB(\xi, 15)$	$s_\xi - 1$	E-R EDT	Max Like	LR-bench.
1	3965.40	5.19	0.215	-12.58	-13863.00	130.79
2	2402.30	13.06	0.137	-5.59	-5937.00	154.52
3	1687.90	20.57	0.128	-3.99	-3355.60	129.44
4	865.86	13.37	0.130	-3.26	-2134.90	104.42
5	488.43	9.87	0.071	-1.54	-1485.50	89.65
6	335.61	9.91	0.027	0.526	-1145.60	39.81
7	268.24	18.29	0.031	-0.514	-889.60	36.35
8	104.51	12.11	0.009	0.136	-695.71	52.51
9	125.75	8.71	0.013	0.175	-561.07	40.10
10	116.52	15.02	0.018	-0.231	-498.23	48.97

Table 3 presents the results of estimating the dynamic MSMP model given in equation (19). Panel A shows the parameter estimates. The first column gives the filter size, columns 2-7 show the parameter estimates for the up intensity and columns 8-13 provide estimates for the down intensity. QML T-statistics/standard errors are given in parentheses. Panel B reports diagnostics for the fitted models. LB(x,15) and LB(ξ,15) denotes the Ljung Box statistic with 15 lags, computed on raw durations and the residuals defined in equation (18) ($\chi^2_{0.95}(15) = 25$). Column 4 reports the amount of over-dispersion in the residuals. In column 5, E-R EDT is a test for over-dispersion wrt the exponential distribution suggested by Engle & Russell (1998). The last column gives the likelihood ratio test of the null $\beta_{11} = \beta_{12} = \beta_{13} = \beta_{21} = \beta_{22} = \beta_{23} = 0$.

Table 4: Estimating MSMP-dytvc

Panel A: Parameter Estimates

Filter	$\hat{\lambda}_{11}$ (t_0)	$\hat{\lambda}_{12}$(*SE*)	$\hat{\lambda}_{13}$ (t_0)	$\hat{\beta}_{11}$ (t_0)	$\hat{\beta}_{12}$ (t_0)	$\hat{\beta}_{13}$ (t_0)	$\hat{\lambda}_{21}$ (t_0)	$\hat{\lambda}_{22}$(*SE*)	$\hat{\lambda}_{23}$ (t_0)	$\hat{\beta}_{21}$ (t_0)	$\hat{\beta}_{22}$ (t_0)	$\hat{\beta}_{23}$ (t_0)
1	.056 (8.50)	.970 (.004)	-.09 (-12)	.040 (6.09)	-.001 (-2.48)	-.004 (-0.32)	.020 (2.61)	.971 (.008)	-.08 (-7.8)	.099 (5.20)	-.001 (-2.10)	.022 (1.64)
2	.087 (6.78)	.944 (.008)	-.15 (-11)	.092 (6.64)	-.003 (-3.57)	-.026 (-1.02)	.023 (1.32)	.931 (.017)	-.17 (-8.8)	.240 (5.81)	-.003 (-2.94)	.048 (1.68)
3	.104 (4.39)	.915 (.019)	-.18 (-6.7)	.116 (5.28)	-.004 (-2.55)	-.061 (-1.36)	.006 (0.20)	.896 (.028)	-.21 (-7.8)	.355 (5.13)	-.006 (-2.51)	.010 (0.22)
4	.127 (3.79)	.886 (.022)	-.23 (-7.2)	.175 (5.35)	-.007 (-2.73)	-.006 (-0.15)	.002 (0.04)	.872 (.032)	-.25 (-7.6)	.449 (5.31)	-.010 (-2.95)	.027 (0.53)
5	.138 (3.41)	.849 (.028)	-.25 (-6.9)	.212 (4.54)	-.010 (-2.73)	-.027 (-0.50)	-.077 (-1.01)	.831 (.049)	-.24 (-7.8)	.574 (4.51)	-.013 (-2.71)	.048 (0.81)
6	.113 (2.46)	.882 (.043)	-.21 (-5.5)	.171 (3.11)	-.007 (-1.54)	.016 (0.32)	.004 (0.03)	.900 (.117)	-.19 (-2.8)	.326 (1.08)	-.009 (-0.91)	.116 (1.92)
7	.097 (1.37)	.872 (.082)	-.20 (-2.6)	.162 (2.74)	-.004 (-0.71)	-.020 (-0.33)	-.026 (-0.15)	.885 (.152)	-.18 (-2.7)	.357 (1.02)	-.007 (-0.68)	.080 (0.96)
8	.085 (0.90)	.490 (.202)	-.28 (-4.1)	.335 (2.79)	-.038 (-1.86)	-.232 (-1.40)	-.385 (-2.57)	.585 (.069)	-.25 (-3.6)	.958 (6.30)	-.026 (-2.22)	.269 (3.20)
9	.019 (0.19)	.688 (.206)	-.19 (-2.1)	.274 (2.58)	-.016 (-0.81)	-.151 (-0.74)	-.321 (-0.89)	.683 (.255)	-.22 (-3.0)	.853 (1.81)	-.016 (-1.09)	-.050 (-0.32)
10	.067 (0.83)	.736 (.086)	-.26 (-4.1)	.511 (3.59)	-.022 (-1.77)	-.047 (-0.36)	-.153 (-0.84)	.768 (.124)	-.26 (-4.0)	.833 (2.42)	-.023 (-1.83)	.067 (0.71)

Panel B: Diagnostics

Filter	$LB(x, 15)$	$LB(\xi, 15)$	$s_\xi - 1$	E-R EDT	Max Like	LR-bench.
1	3965.40	4.27	0.215	-12.55	-13872.00	113.35
2	2402.30	15.47	0.133	-5.43	-5946.70	135.22
3	1687.90	27.72	0.127	-3.95	-3365.80	108.98
4	865.86	15.04	0.129	-3.24	-2141.20	91.89
5	488.43	12.31	0.069	-1.50	-1490.00	80.68
6	335.61	9.28	0.016	0.305	-1149.60	31.81
7	268.24	17.21	0.028	-0.472	-896.47	22.61
8	104.51	18.07	0.004	0.066	-694.47	54.98
9	125.75	12.02	0.005	0.063	-567.04	28.16
10	116.52	15.04	0.016	-0.202	-502.48	40.46

Table 4 presents the results from estimating the dynamic MSMP-dytvc model given in equation (21). Panel A provides the parameter estimates. The first column shows the filter size, columns 2-7 show the parameter estimates for the up intensity and columns 8-13 show the estimates for the down intensity. QML T-statistics/standard errors are given in parentheses. Panel B reports diagnostics for the fitted models. LB(x,15) and LB(ξ,15) denotes the Ljung Box statistic with 15 lags, computed on raw durations and the residuals defined in equation (18) ($\chi^2_{0.95}(15) = 25$). Column 4 reports the amount of over-dispersion in the residuals. In column 5, E-R EDT is a test for over-dispersion wrt the exponential distribution suggested by Engle & Russell (1998). The last column shows the likelihood ratio test of the null $\beta_{11} = \beta_{12} = \beta_{13} = \beta_{21} = \beta_{22} = \beta_{23} = 0$.

Table 5: SUMMARY STATISTICS FOR THE OUT-OF-SAMPLE FORECAST EVALUATION

Model	Stat	Filter (Nobs) [NFC]									
		1 (8566) [4283]	2 (3826) [1913]	3 (2222) [1111]	4 (1438) [719]	5 (1006) [503]	6 (758) [379]	7 (588) [294]	8 (462) [231]	9 (372) [93]	10 (344) [86]
Benchm.	Matrix	$\begin{bmatrix} .291 & .245 \\ .252 & .212 \end{bmatrix}$	$\begin{bmatrix} .286 & .241 \\ .257 & .216 \end{bmatrix}$	$\begin{bmatrix} .298 & .231 \\ .265 & .206 \end{bmatrix}$	$\begin{bmatrix} .307 & .226 \\ .268 & .198 \end{bmatrix}$	$\begin{bmatrix} .330 & .220 \\ .270 & .180 \end{bmatrix}$	$\begin{bmatrix} .344 & .225 \\ .261 & .171 \end{bmatrix}$	$\begin{bmatrix} .349 & .218 \\ .266 & .166 \end{bmatrix}$	$\begin{bmatrix} .381 & .210 \\ .264 & .145 \end{bmatrix}$	$\begin{bmatrix} .405 & .193 \\ .272 & .130 \end{bmatrix}$	$\begin{bmatrix} .400 & .204 \\ .263 & .133 \end{bmatrix}$
	% Correct	0.503 (0.008)	0.502 (0.011)	0.504 (0.015)	0.505 (0.019)	0.510 (0.022)	0.514 (0.026)	0.516 (0.029)	0.526 (0.032)	0.535 (0.052)	0.533 (0.053)
	χ^2-stat.	1.003 (1.459)	1.011 (1.380)	1.011 (1.431)	0.993 (1.428)	0.996 (1.424)	1.057 (1.486)	0.982 (1.361)	0.999 (1.422)	1.021 (1.438)	1.026 (1.407)
	PT-test	-0.001 (1.002)	-0.014 (1.006)	0.031 (1.006)	0.020 (0.997)	-0.005 (0.999)	-0.007 (1.030)	0.006 (0.993)	-0.009 (1.002)	0.003 (1.016)	-0.008 (1.019)
MSMP-dy	Matrix	$\begin{bmatrix} .330 & .262 \\ .213 & .194 \end{bmatrix}$	$\begin{bmatrix} .320 & .247 \\ .223 & .211 \end{bmatrix}$	$\begin{bmatrix} .385 & .258 \\ .178 & .178 \end{bmatrix}$	$\begin{bmatrix} .394 & .282 \\ .182 & .142 \end{bmatrix}$	$\begin{bmatrix} .362 & .217 \\ .239 & .183 \end{bmatrix}$	$\begin{bmatrix} .354 & .208 \\ .251 & .187 \end{bmatrix}$	$\begin{bmatrix} .463 & .238 \\ .153 & .146 \end{bmatrix}$	$\begin{bmatrix} .489 & .238 \\ .156 & .117 \end{bmatrix}$	$\begin{bmatrix} .570 & .269 \\ .108 & .054 \end{bmatrix}$	$\begin{bmatrix} .558 & .314 \\ .105 & .023 \end{bmatrix}$
	% Correct	0.525	0.531	0.563	0.535	0.545	0.541	0.609	0.606	0.624	0.581
	χ^2-stat.	4.930	4.870	10.10	0.260	1.800	1.260	5.770	2.050	0.009	1.360
	PT-test	2.221	2.206	3.175	0.510	1.344	1.124	2.407	1.435	0.098	-1.174
MSMP-dyc	Matrix	$\begin{bmatrix} .301 & .205 \\ .243 & .251 \end{bmatrix}$	$\begin{bmatrix} .308 & .191 \\ .234 & .266 \end{bmatrix}$	$\begin{bmatrix} .348 & .190 \\ .215 & .247 \end{bmatrix}$	$\begin{bmatrix} .346 & .164 \\ .229 & .260 \end{bmatrix}$	$\begin{bmatrix} .380 & .141 \\ .221 & .258 \end{bmatrix}$	$\begin{bmatrix} .322 & .135 \\ .282 & .261 \end{bmatrix}$	$\begin{bmatrix} .367 & .136 \\ .248 & .248 \end{bmatrix}$	$\begin{bmatrix} .398 & .117 \\ .247 & .238 \end{bmatrix}$	$\begin{bmatrix} .495 & .151 \\ .183 & .172 \end{bmatrix}$	$\begin{bmatrix} .477 & .151 \\ .186 & .186 \end{bmatrix}$
	% Correct	0.552	0.574	0.595	0.606	0.638	0.583	0.616	0.636	0.667	0.663
	χ^2-stat.	46.00	42.80	36.90	32.40	37.70	13.60	16.40	17.60	6.160	6.040
	PT-test	6.787	6.544	6.076	5.692	6.146	3.689	4.056	4.203	2.496	2.473

Table 5 reports the results of the out-of-sample forecasting experiments. Shown below the filter size is the number of excursions in the full sample (in round brackets) and the number of observations in the out-of-sample forecasting period (in square brackets). The ”confusion matrices” sort the proportion of predicted up and down moves against the proportion of realized up and down moves. The percentage of correct forecasts is the sum of the diagonal elements from these matrices. Also shown are Fisher’s Chi-squared test and the PT-test for independence between the predicted and realized direction of stock price moves.

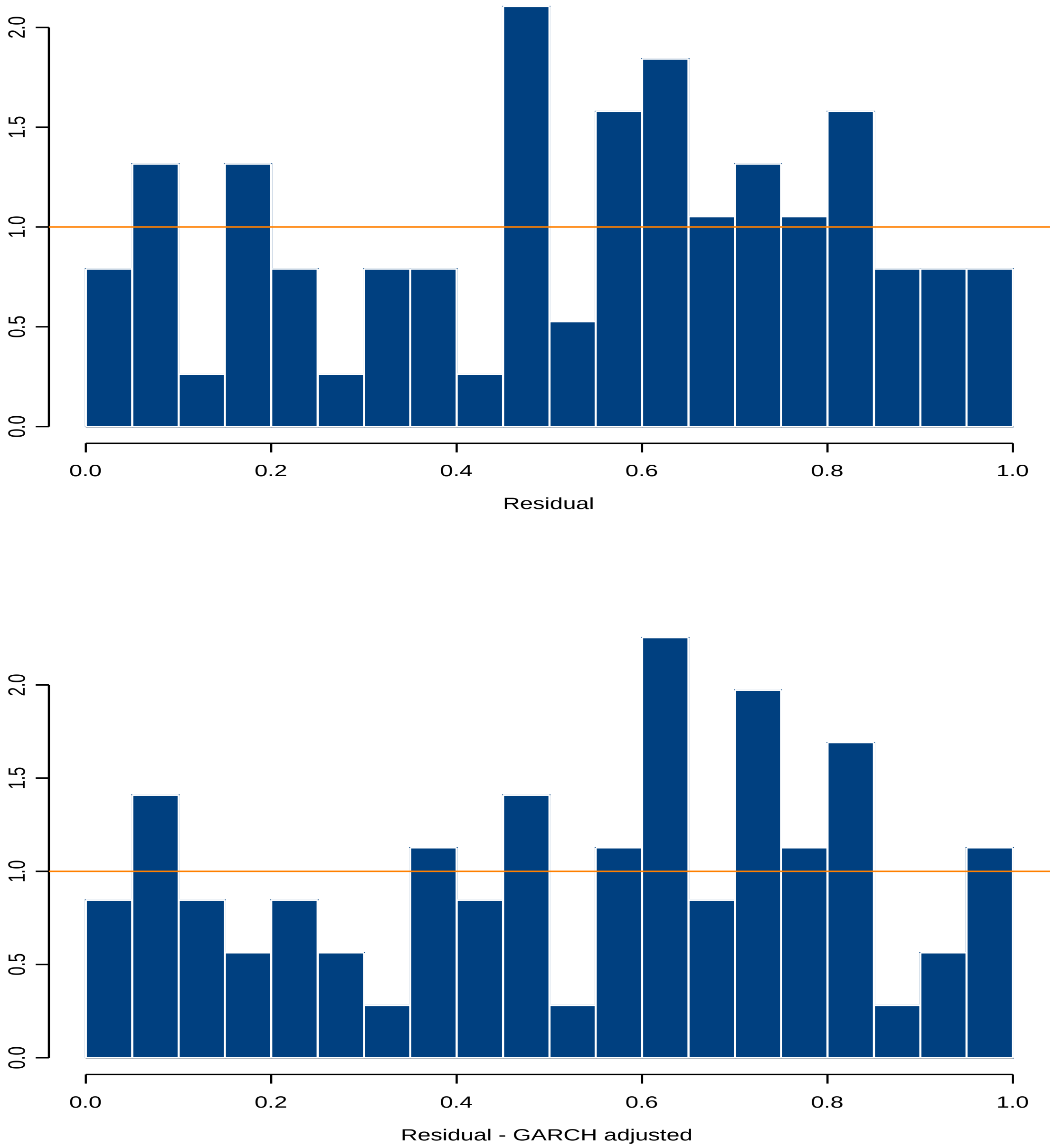


Figure 1: Residuals from the inverse Gaussian distribution fitted to the distribution of first passage times ($\delta = 0.10$).

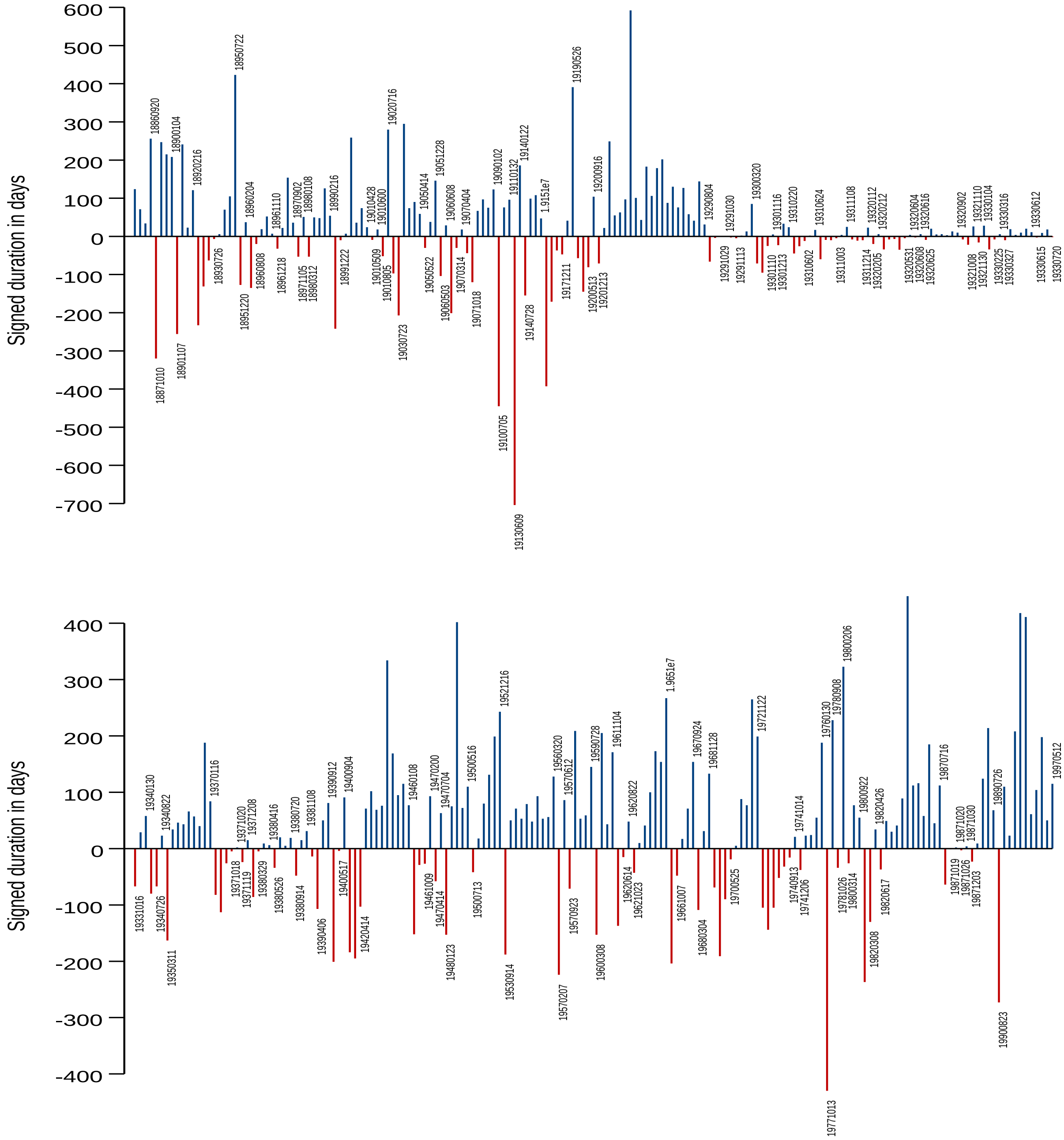


Figure 2: Sequence of first passage times before ±10 percent barrier is crossed.

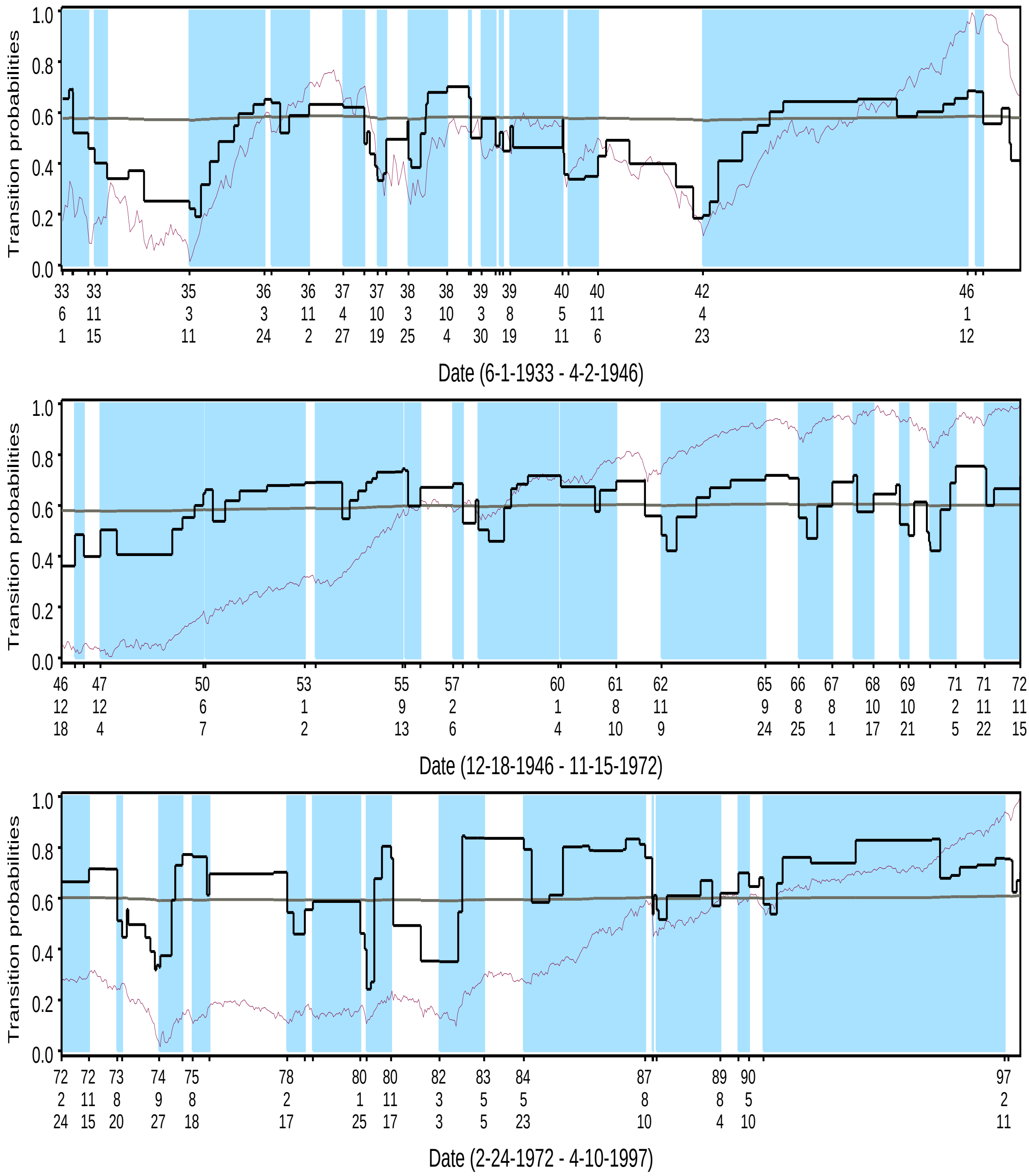


Figure 3: Transition probabilities (black line) from the MSMP-dyc model, for an 8 % filter, plotted against log-prices (thin line). The nearly horizontal gray line tracks the unconditional transition probabilities. The shaded areas are periods of up durations.