Journal of Econometrics 180 (2014) 233–250

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

The dynamic mixed hitting-time model for multiple transaction prices and times

Eric Renault[a], Thijs van der Heijden[b,*], Bas J.M. Werker[c]

[a] Department of Economics, Brown University, 64 Waterman Street, Providence RI 02912, USA
[b] Department of Finance, The University of Melbourne, 198 Berkeley Street, Carlton VIC 3010, Australia
[c] Econometrics and Finance Group, Netspar, Tilburg University, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands

ARTICLE INFO

Article history:
Received 14 September 2012
Received in revised form
11 July 2013
Accepted 7 January 2014
Available online 12 March 2014

JEL classification:
C41
G12

Keywords:
Duration modeling
Hitting time
Trading intensity
Market microstructure

ABSTRACT

We propose a structural model for durations between events and (a vector of) associated marks, using a multivariate Brownian motion. Successive passage times of one latent Brownian component relative to random boundaries define durations. The other, correlated, Brownian components generate the marks. Our model embeds the class of stochastic conditional (SCD) and autoregressive conditional (ACD) duration models, which impose testable restrictions on the relation between the conditional expectation and conditional volatility of durations. We strongly reject the SCD and ACD specifications for both a very liquid and less liquid NYSE-traded stock, and characterize causality relations between volatilities and durations.

1. Introduction

This paper provides a general econometric framework to perform likelihood inference about asset returns when they are observed at endogenous random times. Observation times are endogenous precisely because they are closely related to the price and volatility process. The endogeneity of time complicates likelihood analysis in two respects.

First, the sequence of observed random times conveys relevant information about the price process and, thus, a (parametric) model is needed to describe the dynamics of consecutive durations (the amount of time elapsed between two observation times). Second, to write down the likelihood for observed asset returns at endogenous random times, one cannot simply import standard stochastic volatility or GARCH-type models and generalize them for observations that occur at unequally spaced time intervals (see Ghysels and Jasiak (1998), Meddahi et al. (2006) and references therein for such a generalization). The correct likelihood specification tightly depends upon a specification of the causality relations between prices and durations.

Following Engle's (2000) strategy, we first focus on a dynamic duration model while the distribution of asset returns given contemporaneous durations will be specified in a second stage. Engle and Russell (1998) have proposed a model for the arrival times of trades, termed the autoregressive conditional duration (ACD) model. Let t_i be the ith transaction time where $0 = t_0 < t_1 < \cdots < t_n$, and $\Delta t_{i+1} = t_{i+1} - t_i$ be the duration between trades. The basic assumption of the ACD model is

$$\frac{\Delta t_{i+1}}{\psi_{t_i}} \sim \text{i.i.d. } \mathcal{L}(\theta), \quad \text{where } \psi_{t_i} = E\left[\Delta t_{i+1} \mid \Delta t_i, \Delta t_{i-1}, \ldots\right] \quad (1)$$

is the conditional mean duration between the ith and the $(i+1)$th trade (generally parametrically specified) and $\mathcal{L}(\theta)$ is a probability distribution on $[0, +\infty)$ with unit mean and parameter vector θ.

The ACD model assumes the error term to be multiplicative and all past information to enter the current duration through the conditional mean duration ψ_{t_i} only. The flexibility of the ACD model lies in the rich family of candidates for the distribution $\mathcal{L}(\theta)$ as well as the specification of the dynamic structure of the conditional mean duration ψ_{t_i}. In their analysis of

* Corresponding author. Tel.: +61 390353442.
E-mail addresses: eric_renault@brown.edu (E. Renault), thijsv@unimelb.edu.au (T. van der Heijden), Werker@TilburgUniversity.nl (B.J.M. Werker).

http://dx.doi.org/10.1016/j.jeconom.2014.01.009
0304-4076/

an IBM duration series, Engle and Russell (1998) use for $\mathcal{L}(\theta)$ both Exponential and Weibull distributions. Regarding the specification of the conditional mean duration, the initial ACD(1,1) model is conformable to linear autoregressive dynamics, similar to a GARCH(1,1) model. With their so-called log-ACD specification, Bauwens and Giot (2000) prefer to use a log-transformation on the durations, in order to alleviate the parameter restrictions imposed by the linear ACD model.

According to the general definition (1) above, the log-ACD model is still an ACD model. One of the most flexible parametric ACD specifications has been proposed by Zhang et al. (2001). Their threshold ACD model is a powerful generalization of the ACD model that allows subregimes to have different persistence, conditional means, and error distributions. These differentials can potentially generate a rich collection of nonlinear dynamics.

By allowing the conditioning information in the definition of the conditional mean duration ψ_{t_i} to depend on some latent information, Bauwens and Veredas (2004) go one step further. While the ACD model is the analog for positive time series of the GARCH model, their so-called Stochastic Conditional Duration model (SCD) is the analog of stochastic volatility models (see also Ghysels et al. (2004)).

Irrespective of flexible parametric specifications, the basic multiplicative structure of ACD/SCD models appears to be overly restrictive both for empirical goodness of fit and for multivariate modeling. It is the purpose of this paper to relax these restrictions. We propose a more versatile structural model for durations between events and associated marks. Being a dynamic extension of the mixed hitting-time model of Abbring (2012), our model is structural in the sense that both durations and marks are generated by a multivariate underlying Brownian motion. In particular, we model the durations as the successive passage times of one specific component of this Brownian motion relative to in itself random boundaries. The randomness in the boundaries is generated by defining them as functions of past observations as well as draws from a mixing variable. The other, correlated, Brownian components generate the marks.

Interestingly enough, our duration model nests the multiplicative structure (1) and thereby all possible ACD or SCD models associated to this structure. In terms of dynamics, the hitting-time model is actually even more flexible since general forms of dependence on past information can be accommodated not only in conditionally expected durations ψ_{t_i} but also in standardized durations $\Delta t_{i+1}/\psi_{t_i}$, for example. The conditional probability distribution of the next hitting time may depend on past information both through the value of the hitting barrier as well as the Brownian motion drift term. Both are updated at each hitting time and their parametric specification paves the way for any kind of (nonlinear) autoregressive dynamics in durations.

An interesting special case occurs when conditioning past information is encapsulated in the value of past durations. In that case, the serial independence of standardized durations $\Delta t_{i+1}/\psi_{t_i}$, implied by the parametric multiplicative structure (1), corresponds to a constraint of inverse proportionality between past-dependent components of the Brownian drift and its hitting barrier. In other words, the multiplicative structure (1) becomes a testable hypothesis (within a more general parametric model) and its validity an empirical question.

Recent results in the literature on hitting time problems allow us to claim that the mixed hitting time model can capture almost any possible probability distribution for the standardized duration $\Delta t_{i+1}/\psi_{t_i}$. The problem has a long history in statistics related to the theory of Fredholm integral equations. Jaimungal et al. (submitted for publication) have shown that the distribution of the first time a Brownian motion with drift hits a randomized barrier can always be seen as an infinite mixture of gamma distributions. This allows us to approximate any target distribution arbitrarily well.

Beyond nesting the standard ACD/SCD duration models, our structural model offers the required versatility to accommodate some important modeling issues that cannot be addressed within the classical multiplicative framework.

First, we want to be able to consider several duration processes simultaneously. If each of them is driven by a sequence of events on a given market, we want to have an internally consistent structure. That is, a model specified for the richer information set of two sequences of random times considered together must keep a similar parametric form when only one sequence is taken into account. We will achieve this by resorting to a market-wide latent Brownian factor, common to the determination of all duration processes.

Second, related to the previous point, our structural approach allows us to address another drawback of ACD/SCD models. As pointed out by Hamilton and Jordá (2002), a weakness of the ACD/SCD approach is that it does not allow to incorporate new information that appeared since the previous event and could be relevant for predicting the timing of the next event. Within the Dynamic Mixed Hitting-Time model the arrival time of such new information can be considered a separate event leading to an update of the distribution of the remaining duration. Such update can then go beyond the simple survival effect that the current duration exceeds the time elapsed so far.

Third, as already stressed by several authors, trading intensity and durations between trades convey relevant information, in particular about the volatility of corresponding asset returns. Several theoretical models (e.g., Admati and Pfleiderer (1988) and Foster and Viswanathan (1990)) have been developed explaining the high (low) volatilities during exchange trading (nontrading) periods. Renault and Werker (2011) document some empirical evidence of instantaneous causality between durations and volatility through a semiparametric (GMM) approach. The present paper complements that analysis by providing a versatile parametric framework for the specification of the joint distribution of duration and return processes. Our structural approach is especially well suited since, through predictability in both drifts and volatility of the underlying correlated latent Brownian motions running between each event, any kind of Granger or instantaneous causality relation between random times t_i and returns on random intervals $[t_i, t_{i+1}]$ can be accommodated.

As an alternative to the multiplicative framework discussed above, direct modeling of hazard rates or intensities has been proposed, for example, in Russell (1999), Hamilton and Jordá (2002), and Bauwens and Hautsch (2006). The Dynamic Mixed Hitting-Time model in the current paper can be rewritten in terms of the hazard rate and hence it is a matter of mathematical convenience which form is most appropriate. In particular, the hazard rate can be computed analytically when conditioning upon the value of the mixing variable as well as past information. However, analytic expressions cease to exist when conditioning upon past information only, and therefore we prefer the framework as presented in Section 2.

Finally, we point out that the specification as currently used in the empirical analysis of Section 4 implies that the hazard rate remains constant between two consecutive observation times, which is similar to the assumptions underlying the analyses in recent work by Chen et al. (2013) and Pelletier and Zheng (2012). The latter paper is based on a joint continuous-time multivariate Ornstein–Uhlenbeck specification of the log-intensity of events and the log-volatility of prices. The duration dynamics in the Markov Switching Multi-fractal Duration (MSMD) model in Chen et al. (2013) are of the SCD type as rescaled durations are i.i.d. (see their modelization (5)). Within this assumption, intensities remain constant between events and follow a product of unobserved positive