Stochastic Volatility Models with Transaction Time Risk

Renault, E.; Werker, B.J.M.

Publication date:
2004

Citation for published version (APA):
Renault, E., & Werker, B. J. M. (2004). *Stochastic Volatility Models with Transaction Time Risk*. (CentER Discussion Paper; Vol. 2004-24). Finance.

CentER

Discussion Paper

No. 2004–24

STOCHASTIC VOLATILITY MODELS WITH TRANSACTION TIME RISK

By E. Renault, B.J.M. Werker

February 2004

ISSN 0924-7815

TILBURG UNIVERSITY

Stochastic Volatility Models with Transaction Time Risk

Eric Renault* and Bas J.M. Werker‡
Université de Montréal and *Tilburg University*

February 2004

Abstract

We provide a structural approach to disentangle Granger versus instantaneous causality effects from transaction durations to price volatility. So far, in the literature, instantaneous causality effects have either been excluded or cannot be identified separately from Granger type causality effects. By giving explicit moment conditions for observed returns over (random) transaction duration intervals, we are able to identify the instantaneous causality effect where news events drive simultaneously surprises in durations and surprises in volatilities. Based on ten large stocks traded at the NYSE, we conclude that instantaneous variance forecasts must be decreased by as much as one-third when not having seen the next transaction before its conditional median time. Also, taking into account the causality effects that we document, instantaneous variances are found to be much higher than indicated by standard volatility assessment procedures.

Keywords: Causality, Continuous time models, Transaction prices, Transaction times, Ultra-high frequency data.

1 Introduction

Engle (2000) defines "ultra-high frequency" data as those provided by the measurement of economic (financial) variables when all transactions or quotes are recorded. He argues that there is no higher frequency data available to econometricians. In this framework, transaction (or quote) data are described by two random variables: the first is the time at which the transaction occurs or the quote is given and the second is a vector (marks) observed at that time. To present the essence of our ideas, we consider the times at which a transaction takes place and a corresponding price, e.g., the transaction price or the mid-price of the best prevailing bid and ask at that time. Following Engle (2000), let t_i be the time at which the i-th trade occurs and let $\Delta t_{i+1} = t_{i+1} - t_i$ be the duration between the $i+1$-th and i-th trade. The so-called marks describe the actual event (trade) that occurs at time t_i and consist of a k-vector y_i at this time. Engle (2000) states that "the relevant economic questions can all be determined" from the densities:

$$p\left(y_{i+1}, \Delta t_{i+1} | \mathcal{G}_{t_i}\right) = p\left(y_{i+1} | \Delta t_{i+1}, \mathcal{G}_{t_i}\right) p\left(\Delta t_{i+1} | \mathcal{G}_{t_i}\right), \tag{1.1}$$

which decomposes the joint conditional density of $(y_{i+1}, \Delta t_{i+1})$ given the natural past in discrete time, i.e., given $\mathcal{G}_{t_i} = \sigma(y_j, \Delta t_j : j \leq i)$.

*Université de Montréal, Montréal, CIRANO, and CIREQ
†Finance Group and Econometrics Group, CentER, Tilburg University, Tilburg, The Netherlands.
‡The authors would like to thank Jeff Russell for very useful remarks as a discussant at the 2003 winter meetings in Washington D.C., Laura Spierdijk for kindly providing (well-organized) data, and Nour Meddahi for several suggestions. The second author gratefully acknowledges support from the EU within the MICFINMA network.

The focus of interest in the present paper is the economic interpretation of the occurrence of the current duration Δt_{i+1} in the function $p(y_{i+1}|\Delta t_{i+1}, \mathcal{G}_{t_i})$. We consider exclusively the effect of durations on prices, i.e., $y_i = S_{t_i}$ is the price prevailing at time t_i. In particular, we focus on the effect of durations on volatilities. However, as will become clear, the results can easily be extended to other marks, e.g., volume traded at time t_i. As Engle (2000), we want in particular to "admit the possibility that variations in Δt and variations in (the volatility) σ could be related to the same news events". Basically, we want to further address the issue of interest in Dufour and Engle (2000), which "lies in providing empirical evidence of the relevance of time in the process of price adjustment to information". Our contribution is to stress that the influence of durations on prices, i.e., the occurrence of Δt_{i+1} in $p(S_{t_{i+1}}|\Delta t_{i+1}, \mathcal{G}_{t_i})$, is twofold and should be split, in an identifiable way, into a temporal aggregation effect and an informational effect. Since both effects have different repercussions for risk measurement and management, this separate identification has important consequences.

We have shown in a previous paper (Meddahi, Renault, and Werker, 2003), that, even if the time sequence Δt_i , $i = 1, \ldots, n$, were purely deterministic or strongly exogenous, the current duration Δt_i would explicitly appear in the model $p(S_{t_{i+1}}|\Delta t_{i+1}, \mathcal{G}_{t_i})$ of the price dynamics, simply through a "time-to-build" effect in volatility fluctuations. This dependence is caused by two effects. On the one hand, the application of a standard discrete time volatility model in itself must consider the "volatility per unit of time", as in Engle (2000) in the context of GARCH modelling. On the other hand, the volatility clustering effect is likely to be erased by longer durations and, therefore, the model of volatility persistence must be conformable to the temporal aggregation formulas (see, e.g., Drost and Werker, 1996, Ghysels and Jasiak, 1998, or Grammig and Wellner, 2002, for proposals to apply the Drost and Nijman, 1993, formulas of temporal aggregation of weak GARCH processes). The exact formulas taking both into account are rigorously derived in Meddahi, Renault, and Werker (2003) using the Meddahi and Renault (2003) formulas for temporal aggregation of continuous time linear autoregressive volatility dynamics. Without the continuous time paradigm, the application of temporal aggregation formulas with random times has to be justified by resorting to something like a latent "normal duration GARCH process" (Grammig and Wellner, 2002) whose structural foundations are not clear.

But the most interesting economic issue, as put forward by Dufour and Engle (2000), has nothing to do with the aforementioned deterministic effects of irregular time sampling. In fact, the issue is to see the time between trades as a measure of trading activity which could affect price behavior. This is the reason why the economic interpretation of the information content of time durations, in models of price and trade dynamics, is better founded by identifying a structural continuous time model. Actually, only such a continuous time model will be able to disentangle what we have called the time-to-build effect from the genuine information effect. Typically, this structural model specifies the joint probability distribution of the price process S_t over some reference period $[0, T]$ as well as a sequence of stopping times t_i, $i = 1, \ldots, n$, over the same period. The marginal probability distribution of the price process provides, for any (fixed and deterministic) time interval h, the density function $p_h(S_{t_i+h}|\mathcal{G}_{t_i})$ of the conditional distribution of S_{t_i+h} given the natural past $\mathcal{G}_{t_i}$. Then, the economic issue of interest is the validity of the condition:

$$p_{\Delta_{t_{i+1}}}\left(S_{t_{i+1}}|\mathcal{G}_{t_i}\right) = p\left(S_{t_{i+1}}|\Delta t_{i+1}, \mathcal{G}_{t_i}\right). \tag{1.2}$$

When this equality is fulfilled, and under the additional assumption that the marginal process describing transaction times does not contain information about the structural parameters in the price dynamics, transaction times contain no genuine information regarding these asset price dynamics and there is no cost when these transaction times are considered to be deterministic, still taking into account that they are irregularly spaced. Aït-Sahalia and Mykland (2003) study the full information maximum likelihood under the maintained assumption (1.2). They also document the fact that there is, of course, an efficiency loss when one decides to integrate out the likelihood with respect to the random durations and, even worse, a misspecification bias if one incorrectly supposes that durations are fixed (i.e., $\Delta t_{i+1} = \bar{\Delta}$ for all i).

But if, on the contrary, some instantaneous causality relationship between durations and asset prices leads to a violation of equality (1.2), the incremental information content of Δt_{i+1} about $S_{t_{i+1}}$ given the past $\mathcal{G}_{t_i}$ is crucial in several respects. First, one cannot perform meaningful statistical inference about the probability distribution of the price process without taking into account the probability distribution of durations. Typically, when plugging into a likelihood function based on the densities $p_{\Delta t_{i+1}}\left(S_{t_{i+1}} \middle| \mathcal{G}_{t_i}\right)$ the observed values S_{t_i} as if the times t_i at which the trades occur were deterministic, one would introduce some kind of selection bias which may be significant. Besides statistical inference issues, the randomness of durations between trades is also of foremost importance for risk management. When equation (1.2) is violated, one cannot compute the volatility at time t_i of the asset price $S_{t_{i+1}}$ as if the duration Δt_i were deterministic or even conditionally independent (given the past $\mathcal{G}_{t_i}$) of the asset price. When the equality (1.2) is fulfilled, the total volatility of the asset return between the present trade (at time t_i) and the next (at random time t_{i+1}), may simply be computed by, in a first step, considering that the horizon $t_{i+1} = t_i + \Delta t_{i+1}$ is known and deterministic and, as a second step, integrating out the duration Δt_{i+1}, with respect to with respect to its conditional probability distribution (given $\mathcal{G}_{t_i}$). There is no selection bias induced by conditioning in the first stage on the random value of the next duration due to (1.2).

Our main contribution is to characterize the additional risk that may be introduced by random times in situation where (1.2) does not hold. We focus here on the instantaneous causality relationship between transaction durations and price volatilities, which may, for instance, imply that in the situation of exceptionally long times between two transactions, one is lead to adjust volatility forecasts downwards. This is illustrated in Figure 1.1, which is based on the empirical analysis as it is performed in Section 5. This figure shows that a present time prediction made for the instantaneous variance 60 seconds from now *conditional on not having seen a transaction by that time*, is 50% less than the unconditional prediction. Similarly, Figure 1.2 gives the update in present time instantaneous variance predictions conditional on having seen a transaction. From this figure, we see that conditionally on having seen a transaction within the next 60 seconds, the instantaneous variance prediction has to be increased by 22%.

To be more precise, we prove a decomposition of the total volatility of asset returns over (random) durations in a standard component (where the randomness in the duration is just integrated out) and an additional component which has a direct interpretation as transaction time risk. The interest of this decomposition is to provide a framework for the joint modelling of volatility and inter-transaction duration processes. As stressed by Dufour and Engle (2000), this may give useful insights in the dynamic behavior of market liquidity and thus could be used to design optimal trading and timing strategies. The focus of interest in the present paper is more to state a set of moment conditions that allows one to assess the statistical and economic significance of the aforementioned instantaneous causality relationship. For the purpose of statistical inference about the continuous time price processes, this gives an important semiparametric specification test to decide whether the non-causality assumption (1.2) is satisfied. For the purpose of risk management, this gives insights in the measurement and hedging of liquidity risk.

A byproduct of our framework is the possibility to fruitfully revisit the conclusions of some models previously proposed in discrete time for irregularly spaced financial data. Starting from the seminal Engle and Russell (1998) autoregressive conditional duration (ACD) model, Ghysels and Jasiak (1998) have proposed the ACD-GARCH model to jointly model the volatility and inter-transaction duration processes. This joint modelling issue has since been studied in more detail by several authors, including Engle (2000) and Grammig and Wellner (2002). A crucial issue for all these papers (see also Dufour and Engle, 2000) is the treatment of causality relationships between asset price volatility and durations between trades. Both Engle (2000) and Dufour and Engle (2000) maintain as "a simplifying operative assumption" that durations are not Granger caused by prices. This allows them to estimate a simple ACD model where durations are forecasted only from their

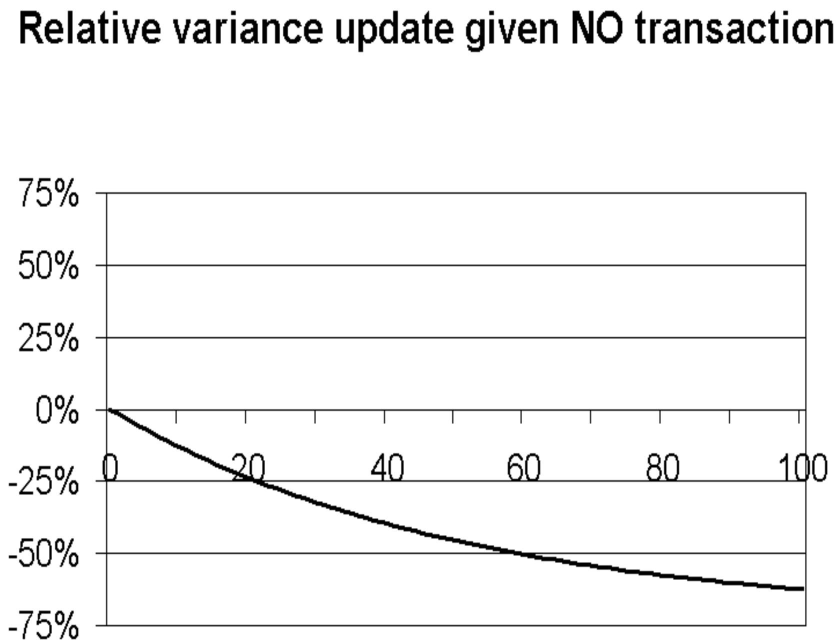

Figure 1.1: Relative update in instantaneous variance prediction due to *not* having seen a transaction by the time (in seconds) indicated on the horizontal axis. Mean transaction duration is 50 seconds and graph is based on an exponential duration distribution and the average estimates from Section 5.

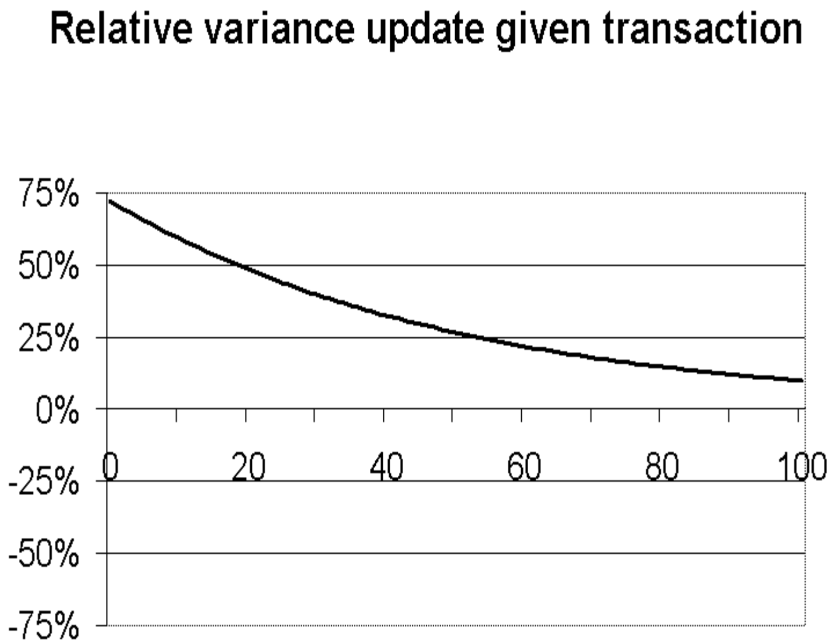

Figure 1.2: Relative update in instantaneous variance prediction when having seen a transaction by the time (in seconds) indicated on the horizontal axis. Mean transaction duration is 50 seconds and graph is based on an exponential duration distribution and the average estimates from Section 5.

own past (Engle, 2000) and to compute univariate impulse response functions for durations (Dufour and Engle, 2000). However, Dufour and Engle (2000) provide some convincing empirical evidence to show that "a buy transaction arriving after a long time interval has a lower price impact than a buy transaction arriving right after a previous trade". This gives support to a significant causality relationship from durations towards prices (through their realized volatility) consistent with the Easley and O'Hara (1992) microstructure model: long durations are likely to indicate an absence of news received by the traders, and thus to be associated with low volatility of returns. However, the direction of causality appears to be quite controversial. While Dufour and Engle (2000) focus on the Granger causality relation from times to prices, Aït-Sahalia and Mykland (2003) put forward on the contrary a negative causality effect from the volatility of returns (through their absolute value) towards the duration. They find the Granger causality effect so striking that they go even further by considering it more important than the clustering effect of durations (long waits tend to be followed by long wait times). This ambiguity about the direction of causality may be a signal that some bi-directional instantaneous causality has been overlooked. Although our focus in the present paper is on instantaneous causality relations between durations and transaction prices, we do not exclude, nor impose, Granger type causality relations in either direction.

The incremental information of the current duration Δt_{i+1} in the function $p\left(S_{t_{i+1}}|\Delta t_{i+1}, \mathcal{G}_{t_i}\right)$, in excess of the deterministic time-to-build effect, is typically neglected in the current literature. The ACD-GARCH model as proposed by Ghysels and Jasiak (1998) or Grammig and Wellner (2002) use the temporal aggregation formulas for weak GARCH processes as derived by Drost and Nijman (1993) with time-varying aggregation period (expected duration). This setup does not allow for a parameter taking into account instantaneous causality between durations and transaction prices. For example, the volatility equation of Grammig and Wellner (2002), which just takes into account the temporal aggregation effect in a "normal duration GARCH process", implicitly assumes that this "normal" regime is not influenced by unexpected durations. In spite of its name ("interdependent duration-volatility model") the model of Grammig and Wellner (2002) cannot capture any instantaneous causality relationship between volatility and duration since both the volatility equation and the duration equation are only about conditionally expected squared returns and expected durations given the past. This is the reason why the only discrete time model which can be compared with the discrete time moment restrictions that we derive from our continuous time structural model is the one of Engle (2000). In this model, according to (1.1), the conditional expectation of squared returns is computed given not only the past but also given the current duration. While volatility depends on past durations through the reciprocal of the past conditional expectation of the current duration, the dependence on the current duration goes not only through the reciprocal of the current duration but also through "surprises in durations", as measured by the relative difference between the current duration and its past conditional expectation. While the first of the three duration/volatility causality effects is typically a Granger one, the two others, and especially the last one, are more focused on instantaneous causality relationships. The general conclusion is that longer (shorter) durations lead to lower (higher) volatility. However, it is important to note that the instantaneous causality and the Granger causality relationships may play in opposite directions. We find that our continuous time structural model is useful for disentangling precisely the two causality effects, making tests of various microstructure models possible (e.g., those of Easley and O'Hara, 1992, or that of Admati and Pfeiderer, 1988). Actually, it allows to test without ambiguity the significance and the sign of an instantaneous causality relationship between duration and volatility, in the presence of, but separate from, possible Granger causality.

The paper is organized as follows. In Section 2, we present our general continuous time framework for joint modelling of transaction times and prices. We allow in our semiparametric framework for causality from transaction times towards asset volatilities that is in line with parametric models such as considered by Duffie and Glynn (2001). In Section 3, we discuss the informational content of transaction durations by making explicitly the difference between volatility conditional only on the past and volatility expected given the past and the current duration. As mentioned before, we derive a decomposition in two terms for the total volatility of asset returns over transaction intervals. Section 4 gives explicit moment conditions for observed returns that can be derived under some more specific assumptions. These are the moment conditions that we use in Section 5 where we present an empirical study to illustrate the economic and statistical significance of the instantaneous causality from trade durations to volatilities. Section 6 discusses four possible extensions to relax some assumptions of previous subsections. Without going in all details, we show that the same kind of empirical study as we carried out in Section 5 could be performed in more general settings. Section 7 concludes.

2 A general framework for modelling transaction times and volatilities

We introduce our framework for the analysis of continuous time price processes observed at random transaction times. This framework allows us to identify separately the marginal price volatility process, the marginal process for the transaction times, and the interaction between both.

An often-used approach, see, e.g., Engle (2000), is to model the marginal distribution of transaction times and the conditional distribution of transaction prices given the transaction times. This, clearly, requires a priori information on the form of the conditional distribution of returns given (future) transaction times. We feel that it is more natural to model the marginal process for transaction prices, as the majority of the empirical finance literature so far deals with this marginal price processes. We show that, given the (marginal) distributions of transaction times and prices, we can model possible causality relations between both using a simple (conditional) regression coefficient. This regression coefficient is sufficient to derive observable moment conditions. In Section 3 we use these results to identify the noncausality assumptions made in previous papers. We want to stress that not all previous papers assume noncausality of transaction times to transaction prices (e.g., Engle, 2000, and Duffie and Glynn, 2001). However, we think that the present paper is the first to explicitly address the question of (non)causality in a structural way and does not rely on ad hoc reduced form specifications.

The basis of our model is the filtration that generates the information accumulation in the market. Following the majority of the literature, we suppose that this information structure is exogenously given and that it satisfies the so-called 'usual conditions' with respect to the underlying probability space (see, e.g., Protter, 1995, p. 3).

Assumption A The information flow in the market is described by the filtration $(\mathcal{F}_t)_{t\geq 0}$ that is supposed to satisfy the usual conditions.

All stochastic processes that appear in the sequel of this paper are assumed to be adapted to the filtration $(\mathcal{F}_t)$, unless explicitly stated otherwise. Note that the filtration $(\mathcal{F}_t)$ is generally not completely observed by the econometrician. The econometrician's information, as described in the introduction, is denoted by $(\mathcal{G}_{t_i})$, with i referring to the i-th transaction and where t_i denotes the transaction time to be introduced in Assumption C. We assume that $\mathcal{G}_{t_i} \subset \mathcal{F}_{t_i}$.

Consider a financial asset with price at time t given by S_t. The evolution of the price S_t is supposed to be given by $S_0 = 1$ and

$$\mathrm{d}\log S_t = \sigma_{t-}\mathrm{d}L_t,\ t \geq 0. \tag{2.1}$$

Following Engle (2000), we ignore a possible drift which is also usual practise in short horizon risk management. Section 6.1 discusses the consequences of including a non-zero drift in the model (2.1). In our specification, (σ_t) is an arbitrary predictable processes and (L_t) is a Lévy process. In particular, we do not assume that the volatility process (σ_t) is continuous or Markovian. Clearly, in order to derive moment conditions, we need some assumptions on the existence of moments. We assume the following.

Assumption B The innovation process (L_t) is assumed to be a locally square-integrable local martingale, with respect to $(\mathcal{F}_t)$, whose compensated quadratic variation is time, i.e. $\mathrm{d}\langle L, L\rangle_t = \mathrm{d}t$. The volatility process (σ_t) is assumed to be predictable with respect to the filtration $(\mathcal{F}_t)$ and square-integrable. We assume that second-order moments of the process $\log S_t$ exist. For any stopping time T, with respect to the filtration $(\mathcal{F}_t)$, we write E_T for the conditional expectation operator given the σ-field $\mathcal{F}_T$ (Protter, 1995, p. 5). Moreover, we define

$$\xi_T(u) = \mathrm{E}_T\left\{\sigma^2_{T+u}\right\}. \tag{2.2}$$

We denote by Ξ_T the primitive of ξ_T, with the normalization that $\Xi_T(0) = 0$.

Note that Assumption B implies that (S_t) is a semimartingale adapted to the filtration $(\mathcal{F}_t)$. In fact, this provides a desirable price model since it is well-known that ruling out arbitrage possibilities (in the appropriate way) in continuous time, implies that the price processes are semimartingales (Delbaen and Schachermayer, 1997). The assumption $\mathrm{d}\langle L, L\rangle_t = \mathrm{d}t$ is a normalization that identifies σ_t as the volatility process. Assuming that L is continuous would, by Lévy's characterization theorem (Protter, 1995, p. 79), imply that L is a Brownian motion. A Brownian motion for L is the only way to exclude jumps in S. The function Ξ_T of the conditional variance predictor ξ_T will appear in the moment condition that we derive below for returns observed over (random) durations between transactions.

The process (S_t) is not observed in continuous time by the econometrician. If it would be, the inference problems that follow become extremely different and in some sense degenerated. We assume that S_t is only observed at some particular (random) times $t_1, t_2, \ldots$.

Assumption C The times $t_1, t_2, \ldots$ form an increasing sequence of bounded stopping times with respect to the filtration $(\mathcal{F}_t)$. We denote durations by $\Delta t_{i+1} = t_{i+1} - t_i$. Finally, F_{t_i} denotes the distribution function of the conditional distribution of Δt_{i+1} given $\mathcal{F}_{t_i}$, i.e.,

$$F_{t_i}(u) = \mathbb{P}\left\{\Delta t_{i+1} \le u \middle| \mathcal{F}_{t_i}\right\}. \tag{2.3}$$

In this paper, t_i will refer to transaction times. The stopping time assumption merely states that, at time t, all transactions up to time t have been observed. For notational convenience we define $t_0 = 0$. Under (2.1), returns on the asset S, as they are observed over the interval $(t_i, t_{i+1}]$, are given by

$$R_{t_i:t_{i+1}} = \log\frac{S_{t_{i+1}}}{S_{t_i}} = \int_0^{\Delta t_{i+1}} \sigma_{t_i+u-}\mathrm{d}L_{t_i+u},\ i = 0, 1, 2, \ldots. \tag{2.4}$$

Note that, under the assumptions stated, $R_{t_i:t_{i+1}}$ is a martingale stopped at time Δt_{i+1}, so that Doob's optional sampling theorem (Protter, 1995, p. 10) implies

$$\mathrm{E}_{t_i}\left\{R_{t_i:t_{i+1}}\right\} = 0,\ i = 0, 1, 2, \ldots. \tag{2.5}$$

The following proposition relates the conditional variance of observed returns $R_{t_i:t_{i+1}}$ to the variance predictor Ξ_{t_i}, to the distribution function of the transaction durations F_{t_i}, and to some regression coefficient that we denote $\beta_{t_i}(\cdot)$ and formally define below.

Proposition 2.1 *Under Assumptions A-C we have the following observable moment condition:*

$$\begin{aligned}\mathrm{Var}_{t_i}\{R_{t_i:t_{i+1}}\} &= \mathrm{E}_{t_i}\{R^2_{t_i:t_{i+1}}\} \\ &= \int_0^\infty \Xi_{t_i}(u)dF_{t_i}(u) + \int_0^\infty \beta_{t_i}(u)F_{t_i}(u)(1-F_{t_i}(u))du,\end{aligned} \tag{2.6}$$

where $\beta_{t_i}(\cdot)$ is the (conditional) regression coefficient (given $\mathcal{F}_{t_i}$)

$$\beta_{t_i}(u) = \frac{\mathrm{Cov}_{t_i}\{\sigma^2_{t_i+u}, I_{(0,\Delta t_{i+1}]}(u)\}}{\mathrm{Var}_{t_i}\{I_{(0,\Delta t_{i+1}]}(u)\}}, \tag{2.7}$$

where $I_{(0,\Delta t_{i+1}]}$ denotes the indicator function of the (random) interval $(0, \Delta t_{i+1}]$.

Note that, since the conditional expectation given an indicator coincides with affine regression, we can write

$$\mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\middle| I_{(0,\Delta t_{i+1}]}(u)\right\} - \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\} = \beta_{t_i}(u)\left(I_{(0,\Delta t_{i+1}]}(u) - [1 - F_{t_i}(u)]\right). \tag{2.8}$$

From (2.8), we see that the β function characterizes by how much an instantaneous variance assessment is influenced by the information that no transaction occurred for some time. It is then not surprising that this information matters as well for measuring the volatility of returns between two consecutive transaction times as in (2.6). Observe, moreover, that the times t_i are not restricted to be transaction times, but can also refer to, e.g., times at which either the bid or ask quote is revised or any other sequence of stopping times. Generally speaking, when returns are considered over random time intervals $(t_i, t_{i+1}]$, the duration Δt_{i+1} between two consecutive stopping times may convey (through a non-zero coefficient β) some relevant information about the risk borne at time t_i over the horizon Δt_{i+1}. A general discussion of this informational content of these stopping times is provided in the next section.

3 Informational content of transaction durations

The volatility decomposition (2.6) allows us to characterize the triple role of the current value Δt_{i+1} of the duration between subsequent trades in the measurement of $\mathrm{Var}_{t_i}\{R_{t_i:t_{i+1}}\}$. Roughly speaking these three roles are:

1. What we have called in the introduction the time-to-build effect which is nothing but the deterministic effect of irregular sampling. When the duration Δt_{i+1} is random, one has to integrate out this random variable in order to define an average risk, but this has nothing to do with causality effects.

2. The filtering effect due to stochastic volatility. Our model is a stochastic volatility one. The information $\mathcal{F}_{t_i}$ that defines the conditioning in the risk measurement $\mathrm{Var}_{t_i}\{R_{t_i:t_{i+1}}\}$ does contain the current latent value σ_{t_i} of the spot volatility process. Then, if one wants to specify a GARCH type model that characterizes the dynamics of the conditional variance given the smaller information set defined only from the past observations of the asset price ($\mathcal{G}_{t_i}$), one has to reproject the above conditional variance on this smaller information set. If the current value Δt_{i+1} of the transaction duration is added, as, e.g., in Engle (2000), to this smaller information set, it may have an informational content, just as way to better filter the past values of the volatility process. This informational content may occur even when the regression coefficient β is zero. This would be akin to some indirect Granger causality effect from durations to prices through volatility (see, e.g., Renault, Sekkat, and Szafarz, 1998) and does not correspond to the instantaneous causality relationship between duration and volatility that is the focus of the present paper.

3. The instantaneous causality effect between the duration and the volatility is encapsulated in the second part of the right-hand side of (2.6) when the regression coefficient β is non-zero. It is typically this effect that may capture "the possibility that variations in durations and variations in the volatility could be related to the same news events". Besides its relevance for microstructure theory, this effect is also important for risk measurement. Typically, neglecting it would amount to overlooking a liquidity component of the risk borne by an investor who wonders at time t_i how risky the investment in this asset is over the next period.

The main advantage of the continuous time framework used in this paper is to allow one to clearly disentangle the afore described three roles of durations in volatility measurement. Let us now discuss more explicitly each of them.

Effect 1: The time-to-build effect

This effect is encapsulated in the first term of the right-hand side of the decomposition (2.6). This term can be seen as an expected integrated volatility imposing non-causality between transaction times and prices. To be more precise, note that

$$\int_0^{\infty} \Xi_{t_i}(\Delta) \mathrm{d}F_{t_i}(\Delta) = E_{t_i}^{\otimes}\left\{\int_0^{\Delta t_{i+1}} \sigma_{t_i+u}^2 \mathrm{d}u\right\}. \tag{3.1}$$

Here, $\otimes$ indicates that the expectation is taken with respect to the product measure of the marginal (yet conditional on $\mathcal{F}_{t_i}$) distributions of Δt_{i+1} and $(\sigma_{t_i+u} : \ u \geq 0)$, i.e., the measure ignoring possible instantaneous causality relations. By application of Fubini's theorem, it can then be seen as the expectation with respect to the marginal distribution F_{t_i} of Δt_{i+1} of the expected integrated volatility as computed for a deterministic duration Δ:

$$\Xi_{t_i}(\Delta) = \int_0^{\Delta} \mathrm{E}_{t_i}\{\sigma_{t_i+u}^2\} \mathrm{d}u, \tag{3.2}$$

therefore

$$\int_0^{\infty} \Xi_{t_i}(\Delta) \mathrm{d}F_{t_i}(\Delta) = E_{t_i} \Xi_{t_i}(\Delta t_{i+1}). \tag{3.3}$$

The conditional expectation of integrated volatility for deterministic duration Δ as in (3.2) has been studied in detail in Bollerslev and Zhou (2002).

Effect 2: The filtering effect

Following Engle (2000) and the decomposition (1.1), an alternative route amounts to renounce to integrate out the stochastic duration and to consider directly for the return $R_{t_i:t_{i+1}}$ over the interval $(t_i, t_{i+1}]$ a volatility measurement given the augmented σ-field $\mathcal{F}_{t_i}^* = \mathcal{F}_{t_i} \vee \sigma(\Delta t_{i+1})$, that is not only given past observations on prices, volatilities, and transaction times, but also the time needed for the next transaction to come. Actually, the Engle (2000) volatility models (39) and (40), p. 18, can be interpreted as reprojections of this volatility measure on the smaller σ-field $\mathcal{R}_{t_i}^* = \mathcal{R}_{t_i} \vee \sigma(\Delta t_{i+1})$ where $\mathcal{R}_{t_i}$ is the sub-σ-field of $\mathcal{F}_{t_i}$ defined by the econometrician's information about past returns and durations only. Then, even if the regression coefficient β is zero, the volatility measure must depend on the duration (Δt_{i+1}) as conditional expectation of $\Xi_{t_i}(\Delta t_{i+1})$ given $\mathcal{R}_{t_i}^*$. This dependence may go through not only the aforementioned time-to-build effect but also through $\mathrm{E}\{\sigma_{t_i}^2 | \mathcal{R}_{t_i}^*\}$.

This is the reason why we argue that, in this framework, even an additional significant role (besides the time-to-build effect) of the duration in the volatility measurement does not really prove that "variations in durations and variations in the volatility are related to the same news events". Of course, the empirical evidence documented by Engle (2000) is fairly convincing. The functional forms (39) and (40) in that paper are sufficiently specific to make it difficult to imagine that the significant role of the duration (Δt_{i+1}) is just a filtering effect. However, we do consider that, to fully disentangle the filtering effect from the instantaneous causality effect of interest, the stochastic volatility framework in continuous time is better suited. The filtering effect is not an issue when considering the moment condition (2.6), since this is not conditional on the current duration Δt_{t+1}.

Effect 3: The instantaneous causality effect

Since we focus on the volatility process, we consider that this effect is significant when β is non-zero, that is when the fact to know that $\Delta t_{i+1} > u$ modifies our optimal forecast at time t_i of the future spot volatility $\sigma^2_{t_i+u}$. Note that this effect is well in line with the doubly-stochastic counting processes as a model for transaction arrivals as proposed in Duffie and Glynn (2001). If the events arriving at time t are viewed as a Poisson process with time varying intensity $\lambda(S_t, \sigma_t)$, it is not surprising that the information that no event occurs between dates t_i and $t_i + u$ is relevant to modify the forecast of the state vector $(S_{t_i+u}, \sigma_{t_i+u})$. Note however that, in contrast with Duffie and Glynn (2001), we are not interested here in the inference issue about a parametric model from observations of this Markov process. Actually, we do not consider that the econometrician observes the volatility process but we perform a semiparametric inference based on the conditional variance of observed returns.

Our focus of interest is the causality property which makes β non-zero, that is

$$\mathrm{E}_{t_i}\{\sigma^2_{t_i+u} | \Delta t_{i+1} > u\} \neq \mathrm{E}_{t_i}\{\sigma^2_{t_i+u}\}. \tag{3.4}$$

The equality

$$\mathrm{E}_{t_i}\{\sigma^2_{t_i+u} | \Delta t_{i+1} > u\} = \mathrm{E}_{t_i}\{\sigma^2_{t_i+u}\}, \tag{3.5}$$

is actually a testable implication of the non-causality property:

$$\mathrm{E}_{t_i}\{\sigma^2_{t_i+u} | \Delta t_{i+1}\} = \mathrm{E}_{t_i}\{\sigma^2_{t_i+u}\}. \tag{3.6}$$

Following the Florens and Fougère (1996) terminology (more precisely, their Definition 2.1, p. 1197), (3.6) means that the filtration $\mathcal{F}^*_t = \mathcal{F}_t \vee \sigma(\Delta t_{n_t+1})$ does not weakly globally cause the volatility process, given $\mathcal{F}_t$, where $n_t = \max(i :\ t_i \leq t)$ denotes the number of transactions up to time t. In more intuitive terms, the next transaction time to come does not weakly (i.e., in expectation) cause the spot volatility process. Note that, given the absence of a drift function, (3.6) would imply also that $(\mathcal{F}^*_t)$ does not weakly instantaneously cause the price process given $(\mathcal{F}_t)$ in the Granger sense (Florens and Fougère, 1996, Definition 3.1., p. 1202), insofar as it does not cause the innovation process L in (2.1). Then, the price process remains a martingale with respect to the augmented filtration $(\mathcal{F}^*_t)$. If we knew more generally that the Doob-Meyer decomposition would not change for any $(\mathcal{R}_t)$-adapted special semimartingale, we would say (Florens and Fougère, 1996, Definition 3.2., p. 1203) that $(\mathcal{F}^*_t)$ does not strongly instantaneously cause the price process given $(\mathcal{F}_t)$ in the Granger sense. In this case, for any function of the price process, the Doob-Meyer decomposition is not modified by the knowledge of the next transaction time. This strong instantaneous non-causality property in the Granger sense is obviously implied by the strong global non-causality property (Florens and Fougère, 1996, Definition 2.2., p. 1197):

$$\mathcal{F}^*_t \text{ and } \mathcal{R}_{t+h} \text{ are conditionally independent given } \mathcal{F}_t, \text{ for all } h > 0. \tag{3.7}$$

The converse is less clear. Theorem 3.1, p. 1203, in Florens and Fougère (1996) states that "strong global non-causality" and "strong instantaneous non-causality in the Granger sense" are equivalent when $\mathcal{F}_t = \mathcal{R}_t$, that is typically not our case since a stochastic volatility process has been added to the filtration $(\mathcal{R}_t)$ of past returns to define the filtration $(\mathcal{F}_t)$. The additional instantaneous causality effects in continuous time to consider to get strong global non-causality in the context of stochastic volatility are sketched in Comte and Renault (1996). The reason why strong global non-causality of transaction times towards the price process is not guaranteed, even when strong instantaneous non-causality is, is that the Doob-Meyer decomposition of the volatility process itself might also be modified by the knowledge of transaction times. Testing for this later causality effect is beyond the scope of the present paper. We define below a simple test for the hypothesis $\beta = 0$, which is an implication of the weak instantaneous non-causality of transaction times towards volatility.

4 Explicit moment conditions

In this section we show how the parameters of interest in the present paper can be identified from moment restrictions concerning conditional variances of observed returns. Clearly, we want to avoid as far as possible any additional assumption about the joint probability distribution of future volatility and durations. We know from Proposition 2.1 that two terms are involved. First, a time-to-build effect

$$TB_{t_i} = \int_0^\infty \Xi_{t_i}(u)\mathrm{d}F_{t_i}(u), \tag{4.1}$$

and, secondly, an instantaneous causality effect

$$IC_{t_i} = \int_0^\infty \beta_{t_i}(u)F_{t_i}(u)\left(1 - F_{t_i}(u)\right)\mathrm{d}u. \tag{4.2}$$

We first notice from (4.1) that any forecasting formula for the volatility process which implies a nonlinear primitive function $\Xi_{t_i}(v) = \int_0^v \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\}\mathrm{d}u$ will involve higher order conditional moments of future durations Δt_{i+1} to compute the time-to-build effect TB_{t_i}. In order to avoid any additional assumption about the duration process, we assume that $\mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\}$ is constant with respect to the horizon u, that is that the (squared) volatility process is a martingale:

$$\mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\} = \sigma^2_{t_i}, \text{ for all } u \geq 0. \tag{4.3}$$

Note that such an integrated volatility process is often found to be empirically reasonable for high-frequency financial data (see Hansen, 1995, and the references therein). We will see in Section 6.2 that linear mean reversion in the volatility process can be accommodated at the cost of additional assumptions on the conditional distribution of duration. Moreover, one may expect that a small level of mean reversion would not significantly modify the inference conclusions about causality that we derive under the simplifying assumption (4.3). In this context, we have $\Xi_{t_i}(u) = \sigma^2_{t_i}u$ and the time-to-build effect is nothing but

$$TB_{t_i} = \sigma^2_{t_i}\psi_{t_i}, \tag{4.4}$$

where $\psi_{t_i} = \mathrm{E}_{t_i}\left\{\Delta t_{i+1}\right\}$ denotes the expected value of the duration Δt_{i+1}, given all the information available at time t_i.

In order to characterize the instantaneous causality effect (4.2) by some simple parameters, it is convenient to assume that the normalized durations $\Delta t_{i+1}/\psi_{t_i}$ are independent of past information $\mathcal{F}_{t_i}$, i.e.,

$$F_{t_i}(u) = F\left(\frac{u}{\psi_{t_i}}\right), \tag{4.5}$$

for some fixed distribution function F. Such an assumption is commonly made in the ACD literature. Note that this does not preclude instantaneous causality relationships between durations and volatility, that is durations and future volatility can be, conditionally on $\mathcal{F}_{t_i}$, dependent. Moreover, thanks to the martingale assumption (4.3) on the volatility process, we are able to handle the time-to-build effect without a parametric specification of the conditional probability distribution of durations and, hence, we remain fully non-parametric with respect to distribution F of rescaled durations. In that respect, we adopt a semiparametric ACD approach, much along the lines of Drost and Werker (2003).

However, to write explicit moment conditions for the instantaneous causality effect

$$IC_{t_i} = \int_0^\infty \beta_{t_i}(u)F(u/\psi_{t_i})\left(1 - F(u/\psi_{t_i})\right)\mathrm{d}u,$$

we also need to extend the ACD kind of assumption to the specification of the regression coefficient function $\beta_{t_i}(\cdot)$. Our maintained assumption will be

$$\beta_{t_i}(u) = \beta\left(\frac{u}{\psi_{t_i}}\right) \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\}, \tag{4.6}$$

for some given function $\beta(\cdot)$. Combining with the martingale assumption (4.3), this leads to

$$\beta_{t_i}(u) = \beta\left(\frac{u}{\psi_{t_i}}\right) \sigma^2_{t_i},$$

and thus

$$IC_{t_i} = \beta^* \sigma^2_{t_i} \psi_{t_i}, \tag{4.7}$$

with

$$\beta^* = \int_0^\infty \beta(v) F(v)(1 - F(v)) \mathrm{d}v \tag{4.8}$$

When testing for instantaneous causality between volatility and durations, β^* will actually be our key parameter of interest. Clearly, rejection of the null hypothesis $H_0 : \ \beta^* = 0$, implies that there exist causality effects for some rescaled horizon u/ψ_{t_i}.

It is worth noticing that our maintained assumption (4.6) is quite natural in an ACD context. More precisely, we have the following lemma, whose proof is omitted.

Lemma 4.1 *Under the maintained assumption (4.5) the two following properties are equivalent:*

1. $\beta_{t_i}(u) = \beta(u/\psi_{t_i}) \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\}$ *for all* $u \geq 0$, *for some given function* $\beta(\cdot)$;
2. $\mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u} I_{(0,\Delta t_{i+1})}(u)\right\} = G(u/\psi_{t_i}) \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\}$ *for all* $u \geq 0$, *for some given function* G.

In this case, the function $\beta(\cdot)$ *is identically equal to zero if and only if* $G = 1 - F$.

To get more feeling for our chosen specification of the causality term, which actually measures the relation between surprises in volatility and surprises in durations, let us rewrite (2.8) as

$$\frac{\mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u} \middle| \Delta t_{i+1} \geq u\right\}}{\mathrm{E}_{t_i} \sigma^2_{t_i+u}} = 1 + \beta\left(\frac{u}{\psi_{t_i}}\right) F\left(\frac{u}{\psi_{t_i}}\right). \tag{4.9}$$

Using the conditional quantile transform $u = \psi_{t_i} F^{-1}(\alpha)$, for some quantile $\alpha \in (0, 1)$, (4.9) states that the revision in the instantaneous variance estimate, assuming that the next transaction will take more time than the α-th quantile of the conditional probability of the duration, is given by

$$\frac{\mathrm{E}_{t_i}\left\{\sigma^2_{t_i+\psi_{t_i} F^{-1}(\alpha)} \middle| \Delta t_{i+1}/\psi_{t_i} \geq F^{-1}(\alpha)\right\}}{\mathrm{E}_{t_i} \sigma^2_{t_i+\psi_{t_i} F^{-1}(\alpha)}} = 1 + \alpha\beta(F^{-1}(\alpha)).$$

Such an assumption essentially means that surprises in durations have the same relative effect on volatility forecasts, independent of the actual state $\mathcal{F}_{t_i}$ of the market.

Summarizing, we will maintain the following additional assumptions in our empirical analysis in Section 5 below.

Assumption D We assume the following in our empirical analysis.

1. Rescaled durations are independent of the past, i.e.,

$$F_{t_i}(u) = F\left(\frac{u}{\psi_{t_i}}\right);$$

2. Instantaneous volatility-duration causality depends, in relative terms, on the rescaled waiting time only, i.e.,
$$\beta_{t_i}(u) = \beta\left(\frac{u}{\psi_{t_i}}\right) \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\};$$
3. Volatility is integrated, i.e.,
$$\mathrm{E}_{t_i}\sigma^2_{t_i+u} = \sigma^2_{t_i}.$$

Assumption D is maintained to get simple explicit moment conditions, allowing us to identify the instantaneous volatility-duration causality relationship, in a semiparametric ACD kind of framework. More precisely, we deduce from Proposition 2.1, jointly with (4.4), and (4.7)

$$\mathrm{Var}_{t_i}\left\{R_{t_i:t_{i+1}}\right\} = (1+\beta^*)\sigma^2_{t_i}\psi_{t_i}. \tag{4.10}$$

In other words, the causality effect manifests itself through a multiplicative factor applied to the time-to-build effect $\sigma^2_{t_i}\psi_{t_i}$. This multiplicative factor is smaller than one if and only if the information that the next transaction will take more time than some α-th quantile will lead to a downward revision in the instantaneous variance prediction. Condition (4.7) paves the way for feasible GMM inference insofar as $\sigma^2_{t_i}\psi_{t_i}$ can be related to a conditional expectation of a known function of observables, that is returns and durations. Various tricks may be imagined to meet this requirement. While we propose one approach in Section 6.3 that is tightly related to the standard way to write duration-volatility models in discrete time, we choose here to focus on the causality issue, through an assumed linear relationship
$$\sigma^2_{t_i} = \alpha_0 + \alpha_1\psi_{t_i}. \tag{4.11}$$
We remark already that such a specification will allow us to avoid parametric relations on the way expected durations ψ_{t_i} depend on $\mathcal{F}_{t_i}$, as would be generally required in duration analysis.

In this specification, $\alpha_0 + \alpha_1\mathrm{E}\psi_{t_i}$ measures the unconditional instantaneous variance while α_1 measures the sensitivity of instantaneous volatility with respect to expected duration. Given that larger volatility usually goes together with expecting more trades, that is smaller expected durations, we expect α_1 to be negative. However, note that this volatility-expected duration relationship has nothing to do with the instantaneous causality effect between volatility and durations as measured by β^*. While the former will generate a kind of Granger causality effect from past durations to current volatility (see, e.g., Dufour and Engle, 2000), the latter relates instantaneously surprises in durations to surprises in volatility. From Relation (4.11) we deduce that

$$\sigma^2_{t_i}\psi_{t_i} = \alpha_0\psi_{t_i} + \alpha_1\psi^2_{t_i} = \mathrm{E}_{t_i}\left\{\alpha_0\Delta t_{i+1} + \alpha_1\varphi\left(\Delta t_{i+1}\right)^2\right\},$$

where $\varphi \in (0,1]$ is a parameter measuring the duration dispersion

$$\mathrm{Var}_{t_i}\left\{\frac{\Delta t_{i+1}}{\psi_{t_i}}\right\} = \frac{1}{\varphi} - 1,$$

in other words
$$\varphi = \left[\int_0^\infty v^2 \mathrm{d}F(v)\right]^{-1}. \tag{4.12}$$
Some well-documented empirical evidence of overdispersion of durations with respect to the benchmark exponential distribution would lead to expect that $\varphi < \frac{1}{2}$. We will see in the empirical section that incorrectly excluding instantaneous causality when in fact $\beta^* < 0$, leads to upward biased estimates for φ.

To summarize, we have the following result.

Proposition 4.2 *Under Assumptions A-D and the a constant sensitivity of volatility with respect to expected durations, $\sigma^2_{t_i} = \alpha_0 + \alpha_1\psi_{t_i}$, returns and transaction times are related by the conditional moment restriction*

$$\mathrm{E}_{t_i}\left\{R^2_{t_i:t_{i+1}} - \alpha_0(1+\beta^*)\Delta t_{i+1} - \alpha_1(1+\beta^*)\varphi\left(\Delta t_{i+1}\right)^2\right\} = 0 \tag{4.13}$$

Not surprisingly, the two causality parameters α_1 and β^* cannot be separately identified from transaction data alone. The two sensitivity factors, α_1 (for Granger causality) and $1+\beta^*$ (for instantaneous causality) play multiplicativ roles in the moment condition (4.13). But this problem can be easily resolved by adding extra identifying moment restrictions based on *deterministic* duration intervals. By definition, the instantaneous causality effect is no longer at stake when observing returns over fixed time intervals of length h, say h corresponding to five minutes intervals. Then, Proposition 2.1 applied to deterministic durations of length h leads straightforwardly to the moment conditions

$$\mathrm{E}_{t_i}\left\{R^2_{t_i:t_i+h} - \alpha_0 h - \alpha_1 h\Delta t_{i+1}\right\} = 0. \tag{4.14}$$

Now, the full set $(\alpha_0, \alpha_1, \varphi, \beta^*)$ of unknown parameters can be identified from the conditional moment restrictions (4.13) and (4.14). This is the basis of our empirical study in the following section.

5 Empirical illustration

In order to assess the economic and statistical relevance of our discussion on instantaneous causality between transaction durations and volatility, we estimate the causality parameter β^* as introduced in Section 4, for ten liquid stocks traded at the NYSE. Although the scope of this empirical exercise is limited, it does show that possible instantaneous causality effects need to be taken into account in ultra-high frequency inference and risk-management. We first discuss in Section 5.1 the ten stocks that we analyse and, subsequently, in Section 5.2 we show that, at least for these stocks and the time period we study, instantaneous causality effects from durations to volatilities are statistically and economically significant.

5.1 Data description

We consider ten stocks traded at NYSE. We use price/duration data from the TAQ dataset for 64 days from August 2, 1999, until October 29, 1999. No trade took place on September 6 due to Labor day. Zero durations are removed from the data set. The only other pre-analysis data cleaning we performed was to replace returns above 100 basis points (in absolute value) by the average return. For the present data set, 100 basis points corresponds to at least five, but usually more than ten, standard deviations and at most 69 observations are affected. Durations are measured below in seconds (*sec*) and returns in basis points (*bp*). Returns at transaction times are calculated using the best prevailing bid and ask price at the time of each transaction. Note that no seasonal adjustment is made to either prices or durations. We do not a priori exclude seasonal effects, but relation (4.11) does impose that seasonal effects in volatilities and durations are synchronized through time.

The ten stocks we use, with ticker symbol in parentheses, are Dillard's (DDS), Federated (FD), IBM (IBM), JCPenney (JCP), Mattel (MAT), May (MAY), McDonald's (MCD), Saks (SKS), Schlumberger (SLB), and Walmart (WMT). Summary statistics are in Table 1. The first row in Table 1 gives, for each of the ten stocks, the number of observations that are available in the estimation. For a fairly illiquid stock like Dillard's, we still have more than 14000 observations available. For the most liquid stock (IBM), we have almost ten times as many. The difference in liquidity also follows from the second row, that gives the average duration (in seconds) between subsequent transactions for each stock, ranging from 11 seconds again for IBM to more than one

	DDS	FD	IBM	JCP	MAT	MAY	MCD	SKS	SLB	WMT
Observations	14389	24344	137796	26444	40907	23079	56618	14259	62799	88432
Average dur.	102	60	11	56	36	64	30	101	23	17
Stand.dev. dur.	134	78	10	72	48	81	31	130	28	19
Robustified ret.	51	19	24	9	25	18	15	69	21	26
Average ret.	0.0	-0.1	0.0	-0.2	-0.2	0.0	0.0	-0.1	0.0	0.0
Stand.dev. ret.	18.7	9.6	3.6	10.5	10.5	9.3	5.6	19.1	5.1	5.2

Table 1: Summary statistics for durations and returns for ten stocks from TAQ database August 2, 1999, until October 29, 1999. The rows of the table present from top to bottom the number of observations, the average duration between transactions, the standard deviation of durations, the number of returns that have been replaced by the average for robustness reasons, the average return between transactions, the standard deviation of returns between transactions. All durations are in seconds, returns in basis points.

and a half minute for Dillard's. The standard deviation of durations is usually slightly above the average, indicating unconditional excess dispersion with respect to the Exponential distribution. As mentioned before, transaction returns over 100 basis points have been replaced by average returns for reasons of robustness. The fourth row in Table 1 shows, for each of the ten stocks, how many returns have been affected. For all stocks this number is below 0.5% of the total number of observations available, and for most stocks below 0.1%. Finally, we present the average transaction returns and standard deviations. Note that there is a clear positive relationship between average durations and standard deviations of transaction returns, due to, in particular, the time-to-build effect.

5.2 Empirical results

We present estimation results on the causality effect from duration to volatility based on the moment conditions detailed in Section 4 and the ten stocks described above. Following standard GMM practice, the conditional moment conditions (4.13) and (4.14) are transformed into unconditional ones using various instruments. It is well-known that both durations and squared returns are autocorrelated. Therefore, the obvious candidates for our instruments, besides the constant, are Δt_i and $R^2_{t_{i-1}:t_i}$. We use the standard optimal weighting matrix for weighting the unconditional moment conditions. The use of both returns over transaction intervals and deterministic intervals of length h induces a overlapping samples problem, since clearly $R_{t_i:t_{i+1}} = R_{t_i:t_i+\Delta t_{i+1}}$ and $R_{t_i:t_i+h}$ are correlated. To resolve this problem, we estimate the variance of the unconditional moment conditions using a Newey-West estimator with a fixed number of lags. The number of lags is fixed at 30. Given the average durations in Table 1, this number of lags ensures that we always cover at least $h = 300$ seconds of overlap, which is the length of the deterministic intervals that we use in our moment conditions.

The estimation results are in Table 2. The parameter α_0 determines the level of the instantaneous variance. Given the average durations in Table 1 and the estimated values for α_1, we can easily derive the average level of the instantaneous variance for each of the ten stocks. These values are presented in the rows with label "Average variance". Clearly, there are some variations in riskiness, ranging from 1.6 to 4.8 bp^2/sec^2. A typical value of 2.0 is equivalent to $\sqrt{2.0 \times 3600 \times 8 \times 252}/100 = 38$ percent volatility on an annual basis, using the standard rule of thumb for temporally aggregating volatilities (which is strictly speaking, of course, not applicable in this case). The parameter α_1 is estimated to be negative in all cases. Recall that α_1 measures the relation between instantaneous volatility and expected durations. Consequently, a higher instantaneous volatility indeed goes together with smaller expected durations.

	DDS		FD		IBM		JCP		MAT	
Parameter	est.	t-val	est.	t-val	est.	t-val	est.	t-val	est.	t-val
α_0	7.25	8.41	4.87	8.16	2.05	6.80	5.98	9.58	5.38	7.22
$\alpha_1(\%)$	-2.46	-3.47	-4.04	-4.75	-4.52	-1.80	-5.56	-6.23	-6.26	-4.32
β^*	-0.07	-0.59	-0.32	-3.46	-0.06	-0.42	-0.22	-2.50	0.37	1.99
φ	0.55	6.08	0.41	9.46	0.91	2.46	0.43	15.03	0.52	10.68
Average variance	4.8		2.5		1.6		2.9		3.1	
Variance update	-7%		-32%		-6%		-22%		37%	
Underestimation	8%		47%		6%		28%		-27%	
	MAY		MCD		SKS		SLB		WMT	
α_0	6.05	9.17	3.24	13.20	7.21	3.87	4.53	16.54	2.53	4.07
$\alpha_1(\%)$	-5.65	-5.95	-6.46	-7.99	-2.62	-1.68	-10.56	-10.92	-4.20	-1.56
β^*	-0.49	-6.82	-0.13	-1.78	0.14	0.44	-0.48	-14.57	0.53	1.68
φ	0.37	10.18	0.46	17.38	0.63	3.10	0.41	24.13	0.95	2.26
Average variance	2.5		1.6		4.6		2.1		1.8	
Variance update	-49%		-13%		14%		-48%		53%	
Underestimation	95%		15%		-13%		93%		-35%	

Table 2: Point estimates and asymptotic t-values for the relation between instantaneous volatility and expected durations (α_0 and α_1), the instantaneous causality parameter (β^*), and the duration dispersion parameter (φ). The last three lines in each panel refer to the average instantaneous variance of the stock prices (in bp^2/sec^2), the implied variance update due to not having seen a transaction by the predicted median duration, and the variance underestimation due to not taking into account the instantaneous causality. See main text for details.

However, the key interest in the present paper is instantaneous causality between future volatilities and surprises in durations as measured by β^*. This parameter is estimated negative for seven out of ten stocks, while for four stocks the estimate is significantly negative at the 5% level. This gives ample support for our claim that the instantaneous causality effects are statistically significant, at least for several stocks. If we combine all the estimates for β^*, we find a (precision weighted) average value of $\beta^* = -36\%$, with a t-value of -14.9 (which is only indicative since it is based on the assumption that the estimates for the different stocks are independent). Figures 1.1 and 1.2 in the introduction are based on this estimate of β^*, considering, for illustrative purposes only, the typical case of exponentially distributed durations, martingale volatility, and a constant function $\beta(\cdot)$. Using (4.8) and since for the exponential distribution $\int F(v)[1 - F(v)]\mathrm{d}v = 1/2$, we find that the function β in (4.6) is equal in size to $-0.36 \times 2 = -0.72$. Now consider the event that, after waiting the (conditional) median duration, we have not seen the next transaction yet. Then, according to (4.9), we should update our current instantaneous variance prediction with $\beta(u/\psi_{t_i})F(u/\psi_{t_i})$ for $u = F_{t_i}^{-1}(1/2) = \psi_{t_i}F^{-1}(1/2)$. We find that we update the current instantaneous variance prediction with $\beta(F^{-1}(1/2))/2 = \beta^* = -0.36$, i.e., a 36% decrease.

An alternative interpretation of β^* is based on the formula for the variance of observed returns (4.10). Incorrectly putting $\beta^* = 0$, would lead to an estimate for the instantaneous variance per unit of time of $\mathrm{Var}_{t_i}\{R_{t_i:t_{i+1}}\}/\psi_{t_i}$. Such a formula is often used implicitly in microstructure research. However, Equation (4.10) shows that, in case of negative β^*, this would underestimate instantaneous volatility significantly. For the average β^* of -36% the actual instantaneous volatility would be $-\beta^*/(1+\beta^*) = 56\%$ higher than the estimated one. Clearly, this may have important repercussion for risk management.

For the individual stocks, the appropriate updates in instantaneous volatility predictions and assessments as calculated above for the average value of β^*, are given in the rows with "Variance

	DDS		FD		IBM		JCP		MAT	
Parameter	est.	t-val	est.	t-val	est.	t-val	est.	t-val	est.	t-val
α_0	7.25	10.20	4.83	9.82	2.02	14.16	5.92	11.16	7.95	17.00
$\alpha_1(\%)$	-2.47	-4.17	-4.06	-5.60	-4.53	-3.63	-5.60	-7.09	-11.35	-11.40
β^*	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-
φ	0.59	12.72	0.51	25.34	0.95	7.37	0.48	41.16	0.44	59.19
	MAY		MCD		SKS		SLB		WMT	
α_0	4.06	9.01	3.23	17.69	7.18	4.96	2.84	19.93	2.54	9.79
$\alpha_1(\%)$	-2.78	-4.37	-6.46	-10.44	-2.55	-2.08	-4.81	-9.46	-4.18	-2.95
β^*	0.00	-	0.00	-	0.00	-	0.00	-	0.00	-
φ	0.59	14.03	0.50	54.00	0.58	7.76	0.65	26.07	0.64	9.00

Table 3: Point estimates and asymptotic t-values for the relation between instantaneous volatility and expected durations (α_0 and α_1) and the duration dispersion parameter (φ) when instantaneous causality between durations and volatilities is excluded a priori.

update" and "Variance underestimation" in Table 2, respectively. Overall, we find that both from a statistical and an economic point of view, the instantaneous causality effect is significant, with some variation for the individual stocks.

Finally, let us consider the parameter φ which measures the dispersion of the rescaled (by their conditional expectation) durations. For exponentially distributed rescaled durations, we have $\varphi = 1/2$. The results for the ten stocks we study vary in this respect, leading to the conclusion that some stocks exhibit overdispersion and others exhibit underdispersion for the conditional duration distribution. For illustrative purposes, we conclude this section with Table 3 that provides the estimation results for the same parameters, based on the same moment conditions, but excluding a priori the instantaneous causality effect from durations to volatilities, i.e., fixing $\beta^* = 0$. Comparing Table 3 with Table 2, we see that for most stocks imposing $\beta^* = 0$ does not affect the point estimates for the other parameters. However, the estimated standard errors are dramatically affected. An analysis of intraday prices and durations ignoring the instantaneous causality effect between durations and volatilites documented in this paper, may thus lead to misleadingly precise results. For all stocks, the estimated values of φ become larger when β^* is estimated negative, and smaller when it is estimated positive. This shows that allowing for the causality effect is not only interesting in itself, but also affects the marginal estimation of the duration process.

To the best of our knowledge, the present paper is the first one that specifically addresses empirically the origin of observed dependencies between durations and volatility. Reduced form VAR-models do not allow for disentangling dependencies between expected durations and current instantaneous volatility on the one hand, and surprises in durations and in future instantaneous volatility on the other hand. As mentioned before, the approach of Grammig and Wellner (2002) implicitly imposes that all dependence takes place through the relation between expected durations and instantaneous volatility. We confirm this effect, but find in addition that exogenous news events apparently drive both durations and volatility.

6 Possible extensions

Although the results in Section 2 require no specific assumptions on the volatility or duration model that one wishes to adopt, our empirical illustration does rely on several more specific assumptions (in particular, Assumption D). We now discuss four possible extensions which all lead to a more difficult empirical analysis than the one presented in Section 5. First of all, we discuss the possibility of including a non-zero drift μ in the log-price process. Secondly, we consider the consequences of relaxing the martingale volatility assumption towards the more common case of a volatility model with linear mean-reversion. Subsequently, we focus on the imposed linear relationship between instantaneous variance and expected durations. Finally, we show how one could add microstructure noise in the model following Bandi and Russell (2003).

6.1 Price process with a non-zero drift

If a general semi-martingale model for the price process is considered, returns over the interval $(t_i, t_{i+1}]$ would be given by

$$\begin{aligned} R_{t_i:t_{i+1}} &= \log \frac{S_{t_{i+1}}}{S_{t_i}} \\ &= \int_0^{\Delta t_{i+1}} \mu_{t_i+u} \mathrm{d}u + \int_0^{\Delta t_{i+1}} \sigma_{t_i+u-} \mathrm{d}L_{t_i+u}, \end{aligned}$$

where μ_{t_i+u} denotes the drift of the log-price process. Of course, any source of randomness in the drift term would possibly introduce other causality relationships with transactions times. For instance, a risk premium related to $\sigma^2_{t_i}$ would introduce causality in higher order moments. Let us focus here on the simplest case of a constant drift term

$$R_{t_i:t_{i+1}} = \mu \Delta t_{i+1} + \int_0^{\Delta t_{i+1}} \sigma_{t_i+u-} \mathrm{d}L_{t_i+u}. \tag{6.1}$$

It is clear from (6.1) that, due to the introduction of the drift term, two instantaneous causality effects involving transaction times are now at stake when computing conditional variances of returns. Not only, as explained in Section 3, we need to know whether the filtration $\mathcal{F}^*_t = \mathcal{F}_t \vee \sigma(\Delta t_{n_t+1})$ does weakly globally cause the volatility process. But, we also need to know whether $\mathcal{F}^*_t$ may weakly globally cause the innovation process L. If it is the case, L is not a $\mathcal{F}^*_t$- martingale and the two terms of $R_{t_i:t_{i+1}}$ will, in general, be correlated given $\mathcal{F}_{t_i}$.

While the former causality effect is the focus of our interest in this paper, its interpretation and statistical identification are simpler if the latter is precluded. This is the reason why we will maintain the following strengthened version of Assumption B.

Assumption E In addition to Assumption B, the innovation process (L_t) is assumed to be a locally square-integrable local martingale with respect to $(\mathcal{F}^*_t)$ whose compensated quadratic variation is time.

In this framework, we deduce from (6.1),

$$\mathrm{Var}_{t_i} \left\{ R_{t_i:t_{i+1}} \right\} = \mu^2 \mathrm{Var}_{t_i} \left\{ \Delta t_{i+1} \right\} + \mathrm{Var}_{t_i} \left\{ \int_0^{\Delta t_{i+1}} \sigma_{t_i+u-} \mathrm{d}L_{t_i+u} \right\},$$

which, using (4.13), immediately leads to a moment condition. Summarizing, a straightforward extension of the arguments leading to Proposition 4.2 gives the following result.

Proposition 6.1 *Under Assumptions A–E, a constant drift μ and a constant sensitivity of volatility with respect to expected durations $\sigma^2_{t_i} = \alpha_0 + \alpha_1 \psi_{t_i}$, returns and transaction times are related by the conditional moment restrictions*

$$\begin{aligned} \mathrm{E}_{t_i} \left\{ R_{t_i:t_{i+1}} - \mu \Delta t_{i+1} \right\} &= 0, & (6.2) \\ \mathrm{E}_{t_i} \left\{ R^2_{t_i:t_{i+1}} - \alpha_0 (1+\beta^*) \Delta t_{i+1} - \left[\mu^2 + \varphi \alpha_1 (1+\beta^*) \right] (\Delta t_{i+1})^2 \right\} &= 0. & (6.3) \end{aligned}$$

As in Section 4, these moment restrictions must be completed by extra identifying moment restrictions based on deterministic duration intervals of length h. A straightforward generalization of (4.14) to the case of non-zero drift gives

$$\begin{aligned} \mathrm{E}_{t_i}\left\{R_{t_i:t_i+h} - \mu h\right\} &= 0, \\ \mathrm{E}_{t_i}\left\{R^2_{t_i:t_i+h} - \alpha_0 h - \mu^2 h^2 - \alpha_1 h \Delta t_{i+1}\right\} &= 0. \end{aligned}$$

We also performed the empirical analysis of Section 5, including a non-zero but constant drift based on the above moment conditions. The conclusion is that for almost all of the ten stocks the drift is estimated small and insignificantly different from zero and the estimates of the parameters of interest remain unchanged. This justifies a posteriori our maintained assumption $\mu = 0$ in Section 5. A zero drift assumption is also very usual in the literature about market microstructure, see, e.g., Engle (2000). From our point of view, a great advantage of this assumption is that, even when causality from durations to the innovation process L is present, this does not affect our analysis focused on possible causality relationships between duration and volatility. In particular, Proposition 2.1 remains valid.

6.2 Volatility mean-reversion

The often used volatility models that exhibit linear mean-reversion can also be accommodated in our framework. They fall in the framework where there exist some deterministic functions $a(\cdot)$ and $b(\cdot)$ such that

$$\xi_{t_i}(u) = \mathrm{E}_{t_i}\{\sigma^2_{t_i+u}\} = a(u)\sigma^2_{t_i} + b(u). \tag{6.4}$$

Note that this volatility prediction formula is found in the linear autoregressive volatility model put forward in Meddahi and Renault (2003). In that case, there is a positive coefficient κ of mean reversion such that we have

$$a(u) = \exp(-\kappa u), \tag{6.5}$$
$$b(u) = \sigma^2(1 - \exp(-\kappa u)), \tag{6.6}$$

where σ^2 denotes the unconditional variance $\mathrm{E}\sigma^2_t$, which is assumed to be time-constant in these models. Formula (6.4), together with (6.5) and (6.6), is, for instance, also implied by a square-root or Ornstein-Uhlenbeck like model of volatility (Barndorff-Nielsen and Shephard, 2002). Clearly, the martingale case that we studied in Sections 4 and 5 above correspond to $\kappa = 0$, i.e., $a(u) = 1$ and $b(u) = 0$. While (6.5) and (6.6) translate the GARCH(1,1) model to a stochastic volatility setting, the martingale volatility case extends the IGARCH(1,1) model. All these models correspond to ARMA(1,1) dynamics for squared innovations of returns (see Meddahi and Renault, 2003). The log-normal stochastic volatility model (Harvey, Ruiz, and Shephard, 1994) is also conformable to (6.4) with vanishing $b(u)$.

If we denote by A and B the primitive functions of a and b, respectively, normalized to zero for $u = 0$, we deduce from (6.4)

$$\Xi_{t_i}(\Delta t_{i+1}) = A(\Delta t_{i+1})\sigma^2_{t_i} + B(\Delta t_{i+1}), \tag{6.7}$$

and, for the specification (6.5) and (6.6), we get immediately

$$A(u) = \frac{1 - \exp(-\kappa u)}{\kappa} \text{ and } B(u) = \sigma^2(u - A(u)). \tag{6.8}$$

Formulas (6.7) and (6.8) basically correspond to the formulas used by Ghysels and Jasiak (1998) or Grammig and Wellner (2002), Formula (5), p.374, when one focuses only on the volatility persistence parameter $A(\Delta t_{i+1})$ that is the sum of the two GARCH(1,1) coefficients. As already stressed, the occurrence of the current duration (Δt_{i+1}) in these formulas has nothing to do with any causal relationship between volatility and durations that would deserve a microstructure theory based interpretation. It is just a time-to-build effect that would equally occur with deterministic transaction times (see Meddahi, Renault, and Werker, 2003). Actually, to get rid of the random feature of durations, Grammig and Wellner (2002) replace the duration (Δt_{i+1}) in (6.7) by its conditional expectation computed from the distribution F_{t_i}. The above formulas show that this is not correct, since the functions A and B are in general nonlinear, and the correct formula for the first part of our volatility decomposition is

$$\mathrm{E}_{t_i}\{\Xi_{t_i}(\Delta t_{i+1})\} = \mathrm{E}_{t_i}\{A(\Delta t_{i+1})\}\sigma^2_{t_i} + \mathrm{E}_{t_i}\{B(\Delta t_{i+1})\}. \tag{6.9}$$

Clearly, for small durations, we get $\mathrm{E}_{t_i}\{A(\Delta t_{i+1})\} \approx A(\psi_{t_i})$ and $\mathrm{E}_{t_i}\{B(\Delta t_{i+1})\} \approx B(\psi_{t_i})$, but the exact implication of this approximation for estimation and testing remains unclear.

We focus here on the derivation of moment restrictions for the variance of returns over the period $(t_i, t_{i+1}]$ in the simplest case of the linear autoregressive volatility model defined by (6.4) with (6.5) and (6.6). In that case, we have

$$\mathrm{E}_{t_i}\left\{A\left(\Delta t_{i+1}\right)\right\} = \kappa^{-1}\left(1 - \mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\}\right), \tag{6.10}$$

$$\mathrm{E}_{t_i}\left\{B\left(\Delta t_{i+1}\right)\right\} = \sigma^2\left(\psi_{t_i} - \mathrm{E}_{t_i}\left\{A\left(\Delta t_{i+1}\right)\right\}\right). \tag{6.11}$$

For ease of exposition, we consider the time-to-build and causality effect separately.

6.2.1 Time-to-build effect for mean-reverting volatility

By comparison to the martingale volatility case, we see from (6.10) and (6.11) that linear mean reversion adds two important difficulties to the analysis, that can however be solved with standard tools.

First, the specification of $\mathrm{E}_{t_i}\left\{A\left(\Delta t_{i+1}\right)\right\}$ is tantamount to the specification of the Laplace transform $\mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\}$ of the conditional distribution of durations. Since this Laplace transform must be computed for any possible value κ of the mean reversion parameter in some parameter interval $(0, \bar{\kappa}) \subset \mathbb{R}_+$, its parametric specification is akin to a fully parametric specification of the conditional probability model of rescaled durations. This is actually the price to pay to introduce mean reversion in volatility while the martingale model allowed us to remain nonparametric with respect to the ACD specification. However, while the additional restriction of constant sensitivity of (squared) volatility with expected durations, i.e., $\sigma^2_{t_i} = \alpha_0 + \alpha_1\psi_{t_i}$ implies a martingale structure for expected durations when volatility is integrated, the autoregressive model (6.4) for volatilities applies also to expected durations, in line with the ACD model of Engle and Russell (1998).

The second difficulty that arises due to the introduction of volatility mean reversion is that, for a given specification of the Laplace transform of the conditional distribution of durations, the computation of the time-to-build effect from (6.9) will lead to products $\mathrm{E}_{t_i}\left\{A\left(\Delta t_{i+1}\right)\right\}\sigma^2_{t_i}$ that may be more difficult to translate in terms of observable moment conditions than in the martingale case, i.e., for $\kappa = 0$. However, simple Taylor expansions of the Laplace transform in the neighborhood of $\kappa = 0$ may alleviate this second complication. To see this, let us focus on the example of a Gamma conditional distribution for durations. Generally speaking, all the standard conditional duration models, as reviewed for instance in Bauwens and Giot (2001), admit simple closed-form expressions for the Laplace transform and can be used similarly. See also Darolles, Gourieroux, and Jasiak (2002).

Let us assume that the conditional probability distribution of Δt_{i+1} given $\mathcal{F}_{t_i}$ is gamma with parameters ν and $\nu\psi_{t_i}$, so that its expectation is indeed ψ_{t_i}. Then the rescaled durations $\Delta t_{i+1}/\psi_{t_i}$ are i.i.d. $\Gamma(\nu,\nu)$ and

$$\mathrm{Var}_{t_i}\left\{\frac{\Delta t_{i+1}}{\psi_{t_i}}\right\} = \frac{1}{\nu} = \frac{1}{\varphi} - 1,$$

where φ is the overdispersion parameter introduced in (4.12). Now, we find (see, e.g., Bauwens and Giot, 2001, p. 98),

$$\mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\} = \left(\frac{\nu}{\nu+\kappa\psi_{t_i}}\right)^{\nu} = 1+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{\nu^j}\frac{\Gamma(\nu+j)}{\Gamma(\nu)}\frac{\psi_{t_i}^j}{j!}, \tag{6.12}$$

where the latter series converges for $\psi_{t_i} < \nu/\kappa$. When the support of the conditional expected duration ψ_{t_i} is not included in the interval $(0, \nu/\kappa)$, the expansion in (6.12) should be truncated and understood as a Taylor expansion. For sake of notational simplicity, we always write formal series expansions. It is now easy to get observable moment restrictions since, for any positive integer j, we have

$$\mathrm{E}_{t_i}\left\{\Delta t_{i+1}\right\}^j = \frac{\psi_{t_i}^j}{\nu^j}\frac{\Gamma(\nu+j)}{\Gamma(\nu)}. \tag{6.13}$$

Together with $\sigma_{t_i}^2 = \alpha_0+\alpha_1\psi_{t_i}$, we now find the time-to-build effect in terms of observable moment conditions using

$$\begin{aligned}\mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\}\sigma_{t_i}^2 &= \alpha_0+\alpha_1\psi_{t_i}+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{\nu^j j!}\frac{\Gamma(\nu+j)}{\Gamma(\nu)}\left[\alpha_0\psi_{t_i}^j+\alpha_1\psi_{t_i}^{j+1}\right] \qquad (6.14)\\ &= \alpha_0+\alpha_i\psi_{t_i}+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{j!}\left[\alpha_0\mathrm{E}_{t_i}\left(\Delta t_{i+1}\right)^j+\alpha_1\frac{\nu}{\nu+j}\mathrm{E}_{t_i}\left(\Delta t_{i+1}\right)^{j+1}\right].\end{aligned}$$

It is worth noting that the assumption of a Gamma distribution for rescaled durations is not needed, at the cost of introducing extra parameters. A direct expansion of $\mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\}$ implies, without any specification of the probability distribution F of rescaled durations,

$$\mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\} = 1+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{j!}\mathrm{E}_{t_i}\left(\Delta t_{i+1}\right)^j = 1+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{j!}\gamma_j\psi_{t_i}^j,$$

where $\gamma_j = \mathrm{E}_{t_i}\left\{\left(\Delta t_{i+1}/\psi_{t_i}\right)^j\right\}$ denotes the j-th uncentered conditional moment of rescaled durations. Then, again imposing $\sigma_{t_i}^2 = \alpha_0+\alpha_1\psi_{t_i}$, we find

$$\begin{aligned}\mathrm{E}_{t_i}\left\{\exp\left(-\kappa\Delta t_{i+1}\right)\right\}\sigma_{t_i}^2 &= \alpha_0+\alpha_1\psi_{t_i}+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{j!}\left[\alpha_0\gamma_j\psi_{t_i}^j+\alpha_1\gamma_j\psi_{t_i}^{j+1}\right] \qquad (6.15)\\ &= \alpha_0+\alpha_1\mathrm{E}_{t_i}\left(\Delta t_{i+1}\right)\\ &\quad+\sum_{j=1}^{\infty}\frac{(-\kappa)^j}{j!}\left[\alpha_0\mathrm{E}_{t_i}\left(\Delta t_{i+1}\right)^j+\alpha_1\frac{\gamma_j}{\gamma_{j+1}}\mathrm{E}_{t_i}\left(\Delta t_{i+1}\right)^{j+1}\right].\end{aligned}$$

To summarize, the time-to-build effect as component of the conditional variance of observed returns can always be expressed in terms of observable moment conditions involving all the higher order moments of durations. However, the important advantage of a parametric specification of the probability distribution of rescaled durations is that, as exemplified by (6.14) for the Gamma case, these moment conditions do not involve additional nuisance parameters relative to the integrated volatility case. Only the overdispersion parameter φ is at play through $\nu = \varphi/(1-\varphi)$. By contrast, the general moment condition (6.15), although similar to (6.14), introduces an infinite number of nuisance parameters through the quotients γ_j/γ_{j+1} which simply collapse into $\nu/(\nu+j)$ in the Gamma case.

6.2.2 The instantaneous causality effect for mean-reverting volatility

Allowing for mean reversion in the volatility process also complicates the computation of the causality term in the conditional variance of observed returns. Using Assumption D, we get

$$IC_{t_i} = \int_0^\infty \mathrm{E}_{t_i}\left\{\sigma^2_{t_i+u}\right\} \beta\left(\frac{u}{\psi_{t_i}}\right) F\left(\frac{u}{\psi_{t_i}}\right)\left(1 - F\left(\frac{u}{\psi_{t_i}}\right)\right) \mathrm{d}u. \tag{6.16}$$

Insofar as $\mathrm{E}_{t_i}\{\sigma^2_{t_i+u}\}$ depends on the forecasting horizon u, the causality effect can no longer be characterized only by a single parameter β^* defined by (4.8). From the forecasting formula (6.4) with (6.5) and (6.6), we see that the crucial building block to characterize the instantaneous causality effect will be

$$\sigma^2_{t_i} \int_0^\infty \exp(-\kappa u)\beta\left(\frac{u}{\psi_{t_i}}\right) F\left(\frac{u}{\psi_{t_i}}\right)\left(1 - F\left(\frac{u}{\psi_{t_i}}\right)\right) \mathrm{d}u = \sigma^2_{t_i}\psi_{t_i}\tilde{\beta}_{t_i}(\kappa), \tag{6.17}$$

with:

$$\tilde{\beta}_{t_i}(\kappa) = \int_0^\infty \exp\left(-\kappa\psi_{t_i} v\right)\beta(v)F(v)\left(1 - F(v)\right) \mathrm{d}v. \tag{6.18}$$

As already announced, the particular integrated volatility case corresponds to $\kappa = 0$, i.e., $\tilde{\beta}_{t_i}(0) = \beta^*$. In order to obtain the instantaneous causality term in the decomposition of the variance of observed returns for the general case $\kappa \in \mathbb{R}_+$, we provide the following lemma.

Lemma 6.2 *Define $\tilde{\beta}_{t_i}(\kappa)$ by (6.18) and*

$$B_{t_i}(z) = \int_0^z \exp\left(-\kappa\psi_{t_i} x\right)\beta(x)\,dx.$$

Then, we have

$$\tilde{\beta}_{t_i}(\kappa) = \mathrm{E}_{t_i}\left\{B_{t_i}\left(\frac{\Delta t_{i+1}}{\psi_{t_i}}\right)\left[2F\left(\frac{\Delta t_{i+1}}{\psi_{t_i}}\right) - 1\right]\right\}.$$

Let us indicate how Lemma 6.2 may be used in the simplest case of a constant function $\beta(\cdot) \equiv \bar{\beta}$. The way the computations are performed below shows that more general functions $\beta(\cdot)$ of exponential form would also work. An application of Lemma 6.2 gives

$$\tilde{\beta}_{t_i}(\kappa) = \frac{\bar{\beta}}{\kappa\psi_{t_i}}\mathrm{E}_{t_i}\left\{\left[1 - \exp\left(-\kappa\Delta t_{i+1}\right)\right]\left[2F\left(\frac{\Delta t_{i+1}}{\psi_{t_i}}\right) - 1\right]\right\}, \tag{6.19}$$

and the building block (6.17) of IC_{t_i} becomes

$$\sigma^2_{t_i}\mathrm{E}_{t_i}\frac{\bar{\beta}}{\kappa}\left\{\left[1 - \exp\left(-\kappa\Delta t_{i+1}\right)\right]\left[2F\left(\frac{\Delta t_{i+1}}{\psi_{t_i}}\right) - 1\right]\right\}.$$

It is then manifest that, insofar as the distribution function F of rescaled durations is simply expressed from exponential functions (as it is the case for the exponential ACD model), the instantaneous causality effect will be expressed from products of $\sigma^2_{t_i}$ times the Laplace transform of rescaled durations. Therefore, the instantaneous causality effect will be expressed in terms of observable moment conditions in the same way that it has been done above for the time-to-build effect.

6.3 The relation between instantaneous volatility and expected durations

Our empirical analysis is simplified significantly by the assumption that instantaneous variance and expected durations are perfectly linearly related. Note, once more, that in turn we do not rely on any specification that relates expected durations to past durations, past volatilities, or past expected durations. Moreover, the assumption also implies that we do not need to specify and possible intra-day seasonality for either volatilities or durations. However, a possible alternative would be, under the maintained assumption of martingale volatility and zero expected returns, to impose the common assumption that for some horizon h, rescaled returns $R_{t_i:t_i+h}/\sqrt{h\sigma^2_{t_i}}$ and scaled durations $\Delta t_{i+1}/\psi_{t_i}$ are jointly independent of the past. Clearly, such an assumption would imply

$$\mathrm{Cov}_{t_i}\left\{\frac{R^2_{t_i:t_i+h}}{h\sigma^2_{t_i}}, \frac{\Delta t_{i+1}}{\psi_{t_i}}\right\} = \delta, \tag{6.20}$$

where δ is some constant (i.e., not depending on the information available at time t_i). Now, condition (6.20), together with $\mathrm{E}_{t_i} R^2_{t_i:t_i+h} = h\sigma^2_{t_i}$, implies

$$\begin{aligned}
\mathrm{E}_{t_i}\left\{R^2_{t_i:t_i+h}\Delta t_{i+1}\right\} &= h\sigma^2_{t_i}\psi_{t_i} + \mathrm{Cov}\left\{R^2_{t_i:t_i+h}, \Delta t_{i+1}\right\} && (6.21)\\
&= (1+\delta)h\sigma^2_{t_i}\psi_{t_i}. && (6.22)
\end{aligned}$$

This result can be used to turn the moment condition (4.10) in a condition in terms of observables only. To be more precise, we clearly have

$$\mathrm{Var}_{t_i}\left\{R_{t_i:t_{i+1}}\right\} = \frac{1+\beta^*}{(1+\delta)h}\mathrm{E}_{t_i}\left\{R^2_{t_i:t_i+h}\Delta t_{i+1}\right\}.$$

Observe that this moment condition can also be used to determine the horizon h at which the scaled returns and durations are indeed independent of the past. Moreover, the argument is easily extended to the situation of a non-zero constant drift μ, as in Section 6.1.

6.4 The role of independent microstructure noise

Bandi and Russell (2003) have recently pointed out how to take into account possible contaminations in observed prices due to market microstructure effects. Their approach consists in postulating that observed asset prices $\tilde{S}_t$ are related to "fundamental" asset prices S_t via

$$\log \tilde{S}_t = \log S_t + \eta_t, \tag{6.23}$$

where the contaminations η_t, $t \in \mathbb{R}_+$, are i.i.d., mean zero, and independent of past and future values of the fundamental price process (S_t) and the transaction times (t_i). We assume that η_t has finite variance σ^2_η. Observed returns over the transaction interval $(t_i, t_{i+1}]$ are now given by

$$\tilde{R}_{t_i:t_{i+1}} = R_{t_i:t_{i+1}} + \eta_{t_{i+1}} - \eta_{t_i}. \tag{6.24}$$

Moment conditions in terms of the observed contaminated returns $\tilde{R}_{t_i:t_{i+1}}$ are now easily derived from those obtained for $R_{t_i:t_{i+1}}$, upon noting that

$$\begin{aligned}
\mathrm{E}_{t_i}\left\{\tilde{R}_{t_i:t_{i+1}}\right\} &= \mathrm{E}_{t_i}\left\{R_{t_i:t_{i+1}}\right\}, && (6.25)\\
\mathrm{Var}_{t_i}\left\{\tilde{R}_{t_i:t_{i+1}}\right\} &= \mathrm{Var}_{t_i}\left\{R_{t_i:t_{i+1}}\right\} + 2\sigma^2_\eta. && (6.26)
\end{aligned}$$

Combining these with the moment conditions for "fundamental" returns $R_{t_i:t_{i+1}}$ as for instance derived in Proposition 4.2, we can estimate both the parameters of interest in this paper and the variance of the microstructure noise, i.e., σ^2_η. We applied these moment conditions to the same data as in Section 5. It turns out that the estimates for the parameters of interest in the present paper are hardly affected. At the same time, the variance of microstructure noise is estimated very imprecisely. The exact reasons for this have still to be established, but they have no repercussions for the present analysis.

7 Concluding remarks

The present paper considers a structural continuous time model for the analysis of instantaneous causality relations between transaction price volatility and transaction durations, in addition to possible Granger causality. We argue that these instantaneous causality effects are significant and that failure to take them into account may lead to severely biased volatility estimates and, consequently, inappropriate risk management.

We identify the instantaneous causality effects using appropriate moment conditions. These conditions (see, Proposition 2.1) are sufficiently general to be applicable for a wide range of statistical and economic model specifications. The analysis does not yet take into account relevant microstructure variables, like volume and quote revision times. Since our results for the variance of observed returns is based on a specification of volatility predictions given all current information (the function ξ_T in Assumption B), volume can easily be included. Also, while we focus on an interpretation of t_i as transaction times, this is not required in our Propositions 2.1 and 4.2. As such, interesting empirical application could include situations where quote revisions times are studied or cross-causality effects where surprises in transaction durations for one stock, may cause instantaneous volatility in an other stock.

Appendix: Proofs

Proof of Proposition 2.1: We consider the conditional expectation of squared observed returns. Note that, using the Doob-Meyer decomposition applied to $\left(\int_0^{t\wedge\Delta t_{i+1}} \sigma_{t_i+u-}\mathrm{d}L_{t_i+u}\right)^2$ (Protter, 1995, p. 94),

$$\mathrm{E}_{t_i}\left\{R^2_{t_i:t_{i+1}}\right\} = \mathrm{E}_{t_i}\int_0^{\Delta t_{i+1}} \sigma^2_{t_i+u}\mathrm{d}u.$$

Consequently,

$$\begin{aligned}
\mathrm{E}_{t_i}R^2_{t_i:t_{i+1}} &= \mathrm{E}_{t_i}\int_0^\infty I_{(0,\Delta t_{i+1}]}(u)\sigma^2_{t_i+u}\mathrm{d}u \\
&= \int_0^\infty \mathbb{P}_{t_i}\{\Delta t_{i+1}\geq u\}\xi_{t_i}(u)\mathrm{d}u + \int_0^\infty \mathrm{Cov}_{t_i}\{I_{(0,\Delta t_{i+1}]}(u),\sigma^2_{t_i+u}\}\mathrm{d}u \\
&= \int_0^\infty \Xi_{t_i}(u)\mathrm{d}F_{t_i}(u) + \int_0^\infty \beta_{t_i}(u)F_{t_i}(u)(1-F_{t_i}(u))\mathrm{d}u,
\end{aligned}$$

where the first equality follows from the optional sampling theorem (Protter, 1995, p. 10), the compensated quadratic variation of L, and the fact the Δt_{i+1} is a stopping time for the filtration $(\mathcal{F}_{t_i+u}:\ u\geq 0)$.

Proof of Lemma 6.2 The proof is based on the following textbook formula

$$\mathrm{E}h(X) = \int_0^\infty h'(x)(1-F(x))\mathrm{d}x,$$

where X is a positive random variable with distribution function F and h a differentiable function with derivative h', and $h(0)=0$. Applying this formula to $h(x)=B(x)F(x)$ and, thus, $h'(x) = B'(x)F(x)+B(x)f(x)$, where f denotes the density function of X, we deduce

$$Eh(X) = \int_0^\infty B'(x)F(x)(1-F(x))\mathrm{d}x + \mathrm{E}\left\{B(X)(1-F(X))\right\}.$$

Therefore, from the definition of h, we get

$$\int_0^\infty B'(x)F(x)(1-F(x))\mathrm{d}x = \mathrm{E}\left\{B(X)(2F(X)-1)\right\}.$$

With $B = B_{t_i}$ defined as in Lemma 6.2, we deduce with $X = \Delta t_{i+1}/\psi_{t_i}$

$$\int_0^\infty \exp\left(-\kappa\psi_{t_i}v\right)\beta(v)F(v)(1-F(v))\mathrm{d}v = \mathrm{E}_{t_i}\left\{B_{t_i}\left(\frac{\Delta t_{i+1}}{\psi_{t_i}}\right)\left(2F\left(\frac{\Delta t_{i+1}}{\psi_{t_i}}\right)-1\right)\right\}.$$

This proves the lemma.

References

ADMATI, A., and P. PFLEIDERER (1988), A theory of intraday patterns: volume and price variability, *The Review of Financial Studies*, **1**, 3–40.

AÏT-SAHALIA, Y. and P. MYKLAND (2003), The effects of random and discrete sampling when estimating continuous-time diffusions, *Econometrica*, **71**, 483–549.

BANDI, F.M. and J.R. RUSSELL (2003), Microstructure noise, realized volatility, and optimal sampling, working paper, University of Chicago.

BARNDORFF-NIELSEN, O. and N. SHEPHARD (2002), Integrated OU processes and non-Gaussian OU-based stochastic volatility models, *Scandinavian Journal of Statistics*, forthcoming.

BAUWENS, L. and P. GIOT (2001), Econometric modelling of stock market intraday activity, Kluwer Academic Publishers, Dordrecht.

BOLLERSLEV, T., and H. ZHOU (2002), Estimating stochastic volatility diffusion using conditional moments of integrated volatility, *Journal of Econometrics*, **109**, 33–65.

COMTE, F. and E. RENAULT (1996), Noncausality in continuous time models, *Econometric Theory*, **12**, 215–256.

DAROLLES, S., C. GOURIEROUX, and J. JASIAK (2002), Structural Laplace transforms and compound autoregressive models, preprint.

DELBAEN, F., and W. SCHACHERMAYER (1997), Non-Arbitrage and the Fundamental Theorem of Asset Pricing: Summary of Main Results, preprint.

DROST, F.C., and TH.E. NIJMAN (1993), Temporal aggregation of GARCH processes, *Econometrica*, **61**, 909–927.

DROST, F.C., and B.J.M. WERKER (1996), Closing the GARCH gap: continuous time GARCH modelling, *Journal of Econometrics*, **74**, 31–57.

DROST, F.C., and B.J.M. WERKER (2003), Semiparametric duration models, *Journal of Business and Economic Statistics*, forthcoming.

DUFFIE, D. and P. GLYNN (2001), Estimation of Continuous-time Markov processes sampled at random time intervals, preprint.

DUFOUR, A., and R.F. ENGLE (2000), Time and the price impact of a trade, *The Journal of Finance*, **55**, 2467–2498.

EASLEY, D., and M. O'HARA (1992), Time and the process of security price adjustment, *The Journal of Finance*, **47**, 577–605.

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

ENGLE, R.F., and J.R. RUSSELL (1998), Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica*, **66**, 1127–1162.

Florens, J.P. and D. Fougère (1996), Noncausality in continuous time, *Econometrica*, **64**, 1195–1212.

Ghysels, E., and J. Jasiak (1998), GARCH for irregularly spaced financial data: The ACD-GARCH model, *Studies in Nonlinear Dynamics and Econmetrics*, **2**, 133–149.

Grammig, J. and M. Wellner (2002), Modelling the interdependence of volatility and inter-transaction duration processes, *Journal of Econometrics*, **106**, 369–400.

Hansen, B.E. (1995), Regression with nonstationary volatility, *Econometrica*, **63**, 1113–1132.

Harvey, A., E. Ruiz, and N. Shephard (1994), Multivariate stochastic variance models, *Review of Economic Studies*, **61**, 247–264.

Meddahi, N. and E. Renault (2003), Temporal aggregation of volatility models, *Journal of Econometrics*, forthcoming.

Meddahi, N., E. Renault, and B.J.M. Werker (2003), GARCH and irregularly spaced data, working paper.

Protter, P. (1995), Stochastic Integration and Differential Equations: A New Approach, Springer-Verlag, Berlin.

Renault, E., K. Sekkat, A. Szafarz (1998), Testing for spurious causality in exchange rates, *Journal of Empirical Finance*, **5**, 47–66.