Financial returns are frequently nonstationary due to the nonstationary distribution of zeros. In daily stock returns, for example, the nonstationarity can be due to an upwards trend in liquidity over time, which may lead to a downwards trend in the zero-probability. In intraday returns, the zero-probability may be periodic: It is lower in periods where the opening hours of the main financial centers overlap, and higher otherwise. A nonstationary zero-process invalidates standard estimators of volatility models, since they rely on the assumption that returns are strictly stationary. We propose a GARCH model that accommodates a nonstationary zero-process, derive a zero-adjusted QMLE for the parameters of the model, and prove its consistency and asymptotic normality under mild assumptions. The volatility specification in our model can contain higher order ARCH and GARCH terms, and past zero-indicators as covariates. Simulations verify the asymptotic properties in finite samples, and show that the standard estimator is biased. An empirical study of daily and intraday returns illustrate our results. They show how a nonstationary zero-process induces time-varying parameters in the conditional variance representation, and that the distribution of zero returns can have a strong impact on volatility predictions.

1. Introduction

Financial returns are frequently zero. This can be due to liquidity issues (e.g., low trading volume), price discreteness or rounding error, data issues (e.g., imputation due to missing values), market closures, and other market-specific characteristics and developments.

A number of approaches accommodate the occurrence of zeros. In continuous time approaches, for example, zeros occur when the assumed underlying price process is not observed. Lesmond, Ogden, and Trzcinka (1999) used this idea to construct a popular measure of liquidity based on observed zeros. More recently, the role of zeros has been repositioned in a continuous time framework by, amongst others, Bandi, Pirino, and Reno (2017), and Bandi et al. (2020). Building on these developments, Buccheri et al. (2020) derived a bias-correction for realized volatility (RV), and Buccheri, Pirino, and Trapi (2020) proposed a way to improve portfolio management in the presence of zeros. Appendix D (supplementary material) outlines the connection between continuous time approaches to zeros and the model proposed here. In a second body of literature, zeros naturally occur due to the discreteness of price changes. Hausman, Lo, and MacKinlay (1992) proposed an ordered probit model for discrete price changes. Russell and Engle (2005) proposed an Autoregressive Conditional Multinomial (ACM) model in combination with their Autoregressive Conditional Duration (ACD) model from Engle and Russell (1998). Liesenfeld, Nolte, and Pohlmeier (2006) criticized this approach, and proposed instead a dynamic integer count model.

This was extended to the multivariate case in Bien, Nolte, and Pohlmeier (2011). Rydberg and Shephard (2003) proposed a model where the price increment is decomposed multiplicatively into three components: Activity, direction and integer magnitude. Catania, Di Maria, and Santucci de Magistris (2020) proposed a discrete mixture approach to discrete price changes. In a third body of literature, price changes are continuous except at zero. Hautsch, Malec, and Schienle (2013) proposed a zero-inflated model for volume. Kümm and Küsters (2015) proposed a zero-inflated model, where zeros occur either because there is no information available or because of rounding. In Harvey and Ito (2020), zeros occur due to censoring of an underlying continuous variable. Finally, the generalized autoregressive conditional heteroscedasticity (GARCH) class of models provides a fourth body of literature, since it accommodates zero-returns as long as the innovation can be zero, see the discussion in Sucarrat and Gronneberg (2020). In particular, if the standardized innovation is stationary, the parameters of a GARCH specification can be consistently estimated by the standard quasi-maximum likelihood estimator (QMLE) even when the conditional zero-probability is time-varying, see, for example, Escanciano (2009).

While the aforementioned contributions accommodate zeros in one or another way, very few of them pay attention to the fact that the zero-process can be nonstationary. This is striking, since the zero-process is frequently nonstationary. In daily stock returns, for example, a downwards (upwards) trend in the zero-probability can be due to an upwards (downwards) trend in liquidity over time, or an upwards (downwards) trend in the price level of the stock. Sucarrat and Gronneberg (2020) found
widespread evidence of a trend in the zero-probability of daily stock returns at the New York Stock Exchange (NYSE). (We revisit a selection of their stocks in Section 5.1.) In intraday returns, the zero-probability is often nonstationary periodic: It is lower in periods with low liquidity (e.g., when the opening hours of the main financial centers do not overlap), and higher in periods with high liquidity (e.g., in hours where the main financial centers are open at the same time). An example is Kolokolov, Livieri, and Pirino (2020), who find clear evidence of a periodic zero-probability in intraday stock returns.

Here, in this article, we propose volatility models that accommodate nonstationary zeros, where the zero-probability can be trend-like or periodic in nature, or both. To this end, volatility is specified as a generic scale (i.e., the conditional variance is a special case). We derive a modified QMLE, which we label the zero-adjusted QMLE, and prove its consistency and asymptotic normality. We start with the standard GARCH(1,1) model for which the regularity conditions are more explicit, then we extend the results to more general models which allow for higher order lags, asymmetries and also indicators of lagged zero returns. In the stationary case, our regularity conditions coincide with the sharpest assumptions given in the literature for CAN of the QMLE. Our asymptotic results mainly rely on the ergodic theorem for nonstationary processes introduced in Francq and Gauthier (2004). Variations of it have also been used in Azrak and Mélard (2006), Phillips and Xu (2006), and Regnard and Zakoian (2010). Section 2 is devoted to the simple zero-inflated GARCH(1,1) model of the form

\[
\epsilon_t = \sigma_t \eta_t I_t^0 + \sigma_t \eta_t I_t^1 + \beta_0 \sigma_{t-1}^2,
\]

with a sequence \((\eta_t)\) of nondegenerated real random variables, and nonnegative parameters \(\omega_0, \alpha_0, \beta_0\). Note that, for the moment, we do not make any precise assumption on the model. In particular, the sequence \((I_t)\) can be the realization of a nonstationary sequence. Therefore, the model (2.1) can be considered as being semiparametric. Moreover, if a solution of Equation (2.1) exists, in general it is nonstationary. The following proposition gives a condition for the existence of such a solution.

**Proposition 2.1.** Given sequences \((\eta_t)\) and \((I_t)\), and parameters \(\omega_0 > 0, \alpha_0 \geq 0\) and \(\beta_0 \geq 0\), there exists a (unique) (nonanticipative) finite solution to Equation (2.1) if

\[
\gamma_t := \limsup_{k \to \infty} \frac{1}{k} \sum_{i=1}^k \log(\alpha_0 \eta_{t-i}^2 I_{t-i} + \beta_0) < 0 \quad \text{a.s.} \quad \forall t.
\]

This condition is satisfied if for all \(t\) there exists \(s > 0\) such that \(\limsup_{k \to \infty} \frac{1}{k} \sum_{i=1}^k |\alpha_0 \eta_{t-i}^2 I_{t-i} + \beta_0| < 1\) a.s. There exists no finite solution if \(\gamma_t > 0\) for some \(t\).

In the previous proposition, a nonanticipative solution means that \(\sigma_t\) is measurable with respect to the sigma-field \(\mathcal{F}_{t-1}\) generated by \([\eta_{t-u}, I_{t-u} \leq t < s]\), and a finite solution means that \(\sigma_t < \infty\) a.s. for all \(t\), and \((\sigma_t)\) is bounded in probability, in the sense that \(\gamma < 0\), \(\exists M > 0\) and \(n > 0\) such that \(P(\sigma_t > M) < \varepsilon\) \(\forall t > n\).

Recall that the necessary and sufficient strict stationarity condition of the standard GARCH(1,1) model is

- A1. \(\gamma := E \log(\alpha_0 \eta_t^2 + \beta_0) < 0\).

Note that, when \((\eta_t)\) is supposed to be a stationary and ergodic sequence, A1 implies (2.2) (simply because \(\log(\alpha_0 \eta_t^2 I_t + \beta_0) \leq \log(\alpha_0 \eta_t^2 + \beta_0)\)). The condition is not necessary, however, because when \(\beta_0 = 0\) and \((I_t)\) is well fed in zeros, it is easy to see that (2.2) is satisfied without any restriction on \(\alpha_0\), in particular even when \(\gamma = E \log(\alpha_0 \eta_t^2) > 0\). Note also that we cannot conclude when \(\gamma_t > 0\) because we do not make assumptions of the distributions of the zeros in \((I_t)\).

For stationary GARCH models with iid innovations, it is known that the strict stationarity condition \(\gamma < 0\) entails the existence of a marginal moment (see Berkes, Horváth, and Kokoszka 2003b, lem. 2.3). The following proposition is a direct extension of that result.

**Proposition 2.2.** If \((\eta_t)\) is iid, \(E|\eta_t|^r < \infty\) for \(r > 0\) and A1 holds, then the finite solution to (2.1) is such that \(sup_t E \sigma_{t}^{2s} < \infty\) and \(sup_t E|\epsilon_t|^{2s} < \infty\) for some \(s > 0\).

Assume that

- A2. \(P(\eta_t = 0) = 0\).

Under this assumption, \(I_t = 0\) if and only if \(\epsilon_t = 0\), and the sequence \((I_t)\) is then observable whenever \((\epsilon_t)\) is observed. Given observations \(\epsilon_1, \ldots, \epsilon_n\), it is then possible to estimate the parameter \(\theta_0 = (\omega_0, \alpha_0, \beta_0) \in \Theta \subset (0, \infty)^2 \times [0, \infty)\) by

\[
\hat{\theta}_n = \arg \min_{\theta \in \Theta} \hat{L}_n(\theta), \quad \hat{L}_n(\theta) = \frac{1}{n} \sum_{t=r_0+1}^n \tilde{\ell}_t(\theta),
\]

\[
\tilde{\ell}_t(\theta) = \begin{cases}
\epsilon_t^2 / \hat{\sigma}_t^2(\theta) + \log \hat{\sigma}_t^2(\theta), & \text{if } \epsilon_t \neq 0, \\
0, & \text{if } \epsilon_t = 0.
\end{cases}
\]

**2. Structure and Estimation of the GARCH(1,1) Specification**

Let \((I_t)\) a bitstream sequence, that is, a sequence valued in \([0, 1]\). This bitstream sequence is said to be well fed in zeros (resp. ones) if, for all \(t\), there exists \(u \leq t\) such that \(I_u = 0\) (resp. \(I_u = 1\)). The value \(I_t = 0\) indicates a zero return and \(I_t = 1\) indicates a nonzero return at time \(t\). Conditionally on \((I_t)\), we will consider time series \((\epsilon_t)\) such that \(\epsilon_t = 0\) if \(I_t = 0\) and \(\epsilon_t\) follows a non degenerated GARCH-type model when \(I_t = 1\). First consider a simple zero-inflated GARCH(1,1) model of the form

\[
\epsilon_t = \sigma_t \eta_t I_t^0 + \sigma_t \eta_t I_t^1 + \beta_0 \sigma_{t-1}^2,
\]

with a sequence \((\eta_t)\) of nondegenerated real random variables, and nonnegative parameters \(\omega_0, \alpha_0, \beta_0\). Note that, for the
where \( r_0 \geq 1 \) is a fixed integer and \( \tilde{\sigma}^2(\theta) = \omega + \alpha \varepsilon_{t-1}^2 + \beta \tilde{\sigma}^2(\theta) \), with a fixed initial value \( \tilde{\sigma}^2(\theta) \), to show the consistency of this modified version of the QMLE, we need additional assumptions. We would like to deal with situations where the occurrence of the zeros may be random or/and periodic of period \( T \in \mathbb{N}^* \) (\( T = 1 \) meaning no periodicity). To this aim, we assume that \( I_t \) is determined by a realization of a \( T \)-dimensional stationary process, at least for large \( t \). Each date \( t = (N - 1)T + v \) corresponds to a cycle \( N = N_t \in \mathbb{Z} \) and a season \( v = v_t \in \{1, \ldots, T\} \). More precisely, we have \( N = \lfloor t/T \rfloor \), where \( \lfloor \cdot \rfloor \) denotes the ceiling function.

A3. Let \((\eta_t)_{t\in\mathbb{Z}}\) and \((S_N)_{N\in\mathbb{Z}}\) be two independent stationary and ergodic processes defined on some probability space \((\Omega, \mathcal{A}, \mathbb{P})\), respectively, valued in \(\mathbb{R}\) and \(S := \{0, 1\}^T\). Let \(S_N = (S_{(N-1)T+1}, \ldots, S_{(N-1)T+T})'\). Assume that there exists an almost surely finite random time \(t_0\) such that, with probability one, \(I_t = S_t\) for all \(t \geq t_0\).

For daily returns, there is usually no seasonality in the zero-process, and Figure 1 shows that the frequency of zeros stabilizes after a certain point for the stocks studied in Section 5.1. For these series, it therefore seems reasonable to assume A3 with \( T = 1 \) and \(t_0(\omega)\) corresponding to a certain date.

It is important to emphasize that in model (2.1), the sequence \(I_t\) is given. Therefore, even when \(I_t\) is the realization of a stationary process, that is, in A3 \(T = 1\) and \(I_t = S_t(\omega)\) for all \(t\), conditionally on \((I_t)\), the sequence \((\varepsilon_t)\) is not stationary. Indeed, it is clear that \(\varepsilon_t\) and \(\varepsilon_{t+1}\) can not have the same distribution when \(I_t \neq I_{t+1}\). We will work with random variables of the form \(f(I_t, I_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots)\) which, conditionally on \((I_t)\), are not stationary. The following lemma shows that a kind of law of large numbers can however be applied to such nonstationary sequences under A3. Similar results appear in Azrak and Mélard (2006), Francq and Gauthier (2004), Phillips and Xu (2006), and Regnard and Zakoian (2010).

**Lemma 2.1.** Let \( f(\cdot) : \{0, 1\}^N \times \mathbb{R}^N \to \mathbb{R} \) be a measurable function. Assume that for \( t = 1, \ldots, T \) we have \( Ef^+ (S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) < \infty \). Then, given any sequence \((I_t)\) satisfying A3, we have

\[
\frac{1}{n} \sum_{t=1}^{n} f(I_t, I_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) \\
\to \frac{1}{T} \sum_{t=1}^{T} Ef(S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) \in [-\infty, \infty)
\]

almost surely. If the condition \( Ef^+ (S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) < \infty \) is replaced by \( Ef^- (S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) < \infty \), then the limit belongs to \((-\infty, \infty]\).

**Example 2.1 (Trivial application of Lemma 2.1).** Let \((\eta_t)\) be an independent sequence of \(\mathcal{N}(0, 1)\)-distributed random variables, \(0 < \sigma(0) < \sigma(1), T = 2, \) and \(\pi_1 = P(S_1 = 1) = 1 - P(S_1 = 0) \) and \(\pi_2 = P(S_2 = 1) = 1 - P(S_2 = 0)\). Given a sequence \((I_t)\) satisfying A3, defined the process \(X_t = \sigma(I_t)\eta_t\). Reasoning conditionally to \((I_t)\), the sequence \((X_t)\) is not stationary because the distribution of \(X_t\) is either \(\mathcal{N}(0, \sigma^2(0))\) or \(\mathcal{N}(0, \sigma^2(1))\). We have however the almost sure convergence

\[
\frac{1}{n} \sum_{t=1}^{n} X_t^2 \to \frac{1}{2} \left\{ \pi_1 \sigma^2(1) + (1 - \pi_1) \sigma^2(0) \right\} + \frac{1}{2} \left\{ \pi_2 \sigma^2(1) + (1 - \pi_2) \sigma^2(0) \right\} \text{ as } n \to \infty.
\]

For proper definition of \(\hat{\theta}_0\) and identifiability of the GARCH parameters, assume the following assumptions, which are also

![Figure 1](source-url) Movement Average (MA) estimates of the daily zero-probability \(\pi_{0t}\) for a subset of NYSE stocks (see Section 5). The moving average is computed as \(\hat{\pi}_{0t} = 500^{-1} \sum_{i=t-250}^{t-2} (1 - \hat{\delta}_i)\), \(t = 250, \ldots, n - 250\). Data source: Bloomberg.
required for consistency of the QMLE of stationary GARCH models.

A4. \( \theta_0 = (\omega_0, \alpha_0, \beta_0) \in \Theta \subset (0, \infty)^2 \times \{0, 1\} \) and \( \Theta \) is compact.

A5. \( \beta < 1 \) for all \( \theta \in \Theta \).

A6. Conditionally on \((I_t)\), the sequence \((\eta_t)\) is iid, \(E\eta_t^2 = 1\) and \(P(\eta_t^2 \neq 1) \neq 1\).

**Remark 2.1 (Interpretation of \( \sigma_t \)).** To facilitate interpretation, suppose that \( E\eta_t = 0 \), as is generally the case for GARCH processes. Under A6 and (2.2), we have \( \sigma^2_t = \text{Var}(\epsilon_t | F_{t-1}, I_t = 1) \). Thus, \( \sigma_t \) corresponds to the volatility of \( \epsilon_t \) when this return is nonzero. When \( I_t = 0 \), the variable \( \sigma_t \) does not have such an interpretation. Given the observations \( \epsilon_1, \ldots, \epsilon_n \), one can thus interpret \( \sigma_{n+1} \) as the volatility of the future return \( \epsilon_{n+1} \) under the scenario that the latter is nonzero. Since \( I_t \) is taken as exogenous variable, our model is not sufficient to predict \( \epsilon_t^2 \). Indeed, we have \( E(\epsilon_t^2 | F_{t-1}) = \sigma^2_t E(I_t = 1 | F_{t-1}) \), see Section 5.

Obviously, to be able to estimate the parameter of the volatility process \( \sigma_t \), it is also necessary to assume that \((I_t)\) is well fed in 1s. We even have to assume that if \( I_t = 1 \), then \( I_{t-1} \) is not always equal to zero, otherwise, in the ARCH(1) case, \( l_1\tilde{\sigma}^2(\theta) = I(\omega + \sigma^2_{t-1}\eta_{t-1}^2 - 1) \) would not depend on \( \alpha \). In A3, we thus assume that

A7. For some \( j_0 \in \{1, \ldots, T\}, P(S_{j_0} = 0) \neq 1 \) and \( P(S_{j_0-1} = 0) \neq 1 \).

In general, \( S_t \) is not independent of \( F_{t-1}^s \), where \( F_{t-1}^s \) denotes the sigma-field generated by \( \{S_u, u; \in \leq t\} \). We however assume that the conditional distribution of \( S_t \) given \( F_{t-1}^s \) is not degenerated in the following sense.

A8. for \( j_0 \) defined in A7 and a constant \( \tau > 0 \), we have \( E(S_{j_0} | F_{j_0-1}^s) \geq \tau \) a.s.

**Theorem 2.1.** Let \((I_t)\) be a given bitstream sequence, and \((\epsilon_t)\) a zero-inflated GARCH(1,1) model satisfying (2.1). Under A1–A8, the estimator defined by Equation (2.3) satisfies \( \hat{\sigma}_n \rightarrow \sigma_0 \) almost surely.

For the asymptotic normality of the QMLE, it is necessary to assume the following.

A9. \( \theta_0 \in \Theta \), where \( \Theta \) denotes the interior of \( \Theta \).

A10. \( \kappa := E\eta_t^2 < \infty \).

Assumptions A9 and A10 are also required to show the asymptotic normality of the QMLE of standard stationary GARCH models (see, e.g., Francq and Zakoian 2019, theor. 7.2). Under A5, let \( \tilde{\sigma}^2_t(\theta) = \sum_{i=0}^{\infty} \beta^i \omega + \alpha \epsilon_{t-i}^2 \). Let \( \ell_t(\theta) \) and \( I_n(\theta) \) be defined by substituting \( \sigma_t(\theta) \) for \( \tilde{\sigma}_t(\theta) \) in \( \ell_t(\theta) \) and \( I_n(\theta) \). Note that for some measurable functions \( \ell : \Theta \times \{0, 1\}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^3 \) and \( \ell : \Theta \times \{0, 1\}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^3 \times \mathbb{R}^3 \).

\[
I = \frac{1}{T} \sum_{t=1}^{T} \ell(\theta_0; S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) \times \ell'(\theta_0; S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots),
\]

\[
J = \frac{1}{T} \sum_{t=1}^{T} \ell'(\theta_0; S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots).
\]

**Theorem 2.2.** Let \((I_t)\) be a given bitstream sequence, and \((\epsilon_t)\) a zero-inflated GARCH(1,1) model satisfying Equation (2.1). Under A1–A10, the estimator defined by Equation (2.3) satisfies

\[
\sqrt{n}(\hat{\theta}_n - \theta_0) = J^{-1} \frac{1}{n} \sum_{t=1}^{n} I_t (\eta_t^2 - 1) \left( \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \right) + o_p(1),
\]

\[
\frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \sim N(0, \Sigma), \quad \Sigma = (\kappa - 1)^{-1}.
\]

The matrix \( \Sigma \) can be consistently estimated by \( \hat{\Sigma} = (\hat{\kappa} - 1)^{-1} \), where

\[
\hat{\kappa} = \frac{\sum_{t=\tau+1}^{T} \eta_t^2}{\sum_{t=\tau+1}^{T} 1}, \quad \hat{J} = \frac{1}{n} \sum_{t=\tau+1}^{T} I_t \left( \frac{1}{\sigma_t^2(\theta_0)} \frac{\partial \sigma_t^2(\theta_0)}{\partial \theta} \right)
\]

The derivatives involved in \( \hat{J} \) can be computed recursively by

\[
\frac{\partial \sigma_t^2(\hat{\theta}_n)}{\partial \theta} = \left( \frac{1}{\sigma_t^2(\hat{\theta}_n)} \right) + \beta \frac{\partial \sigma_{t-1}^2(\hat{\theta}_n)}{\partial \theta} \text{ for } t = 2, \ldots, n,
\]

with

\[
\frac{\partial \sigma_1^2(\hat{\theta}_n)}{\partial \theta} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
\]
in the case \( \sigma_1^2(\theta) = \omega \).

### 3. Extension to General Volatility Models and Model Checking Tests

We now extend Model (2.1) by considering the general zero-inflated volatility model

\[
\epsilon_t = \sigma_t \eta_t I_t, \quad \sigma_t = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots; \theta_0), \quad (3.1)
\]

where \( \theta_0 \in \Theta \subset \mathbb{R}^d \) and \( \sigma : \mathbb{R}^\infty \times \Theta \rightarrow (0, \infty) \). Note that this general formulation includes all GARCH(\( p, q \)) models, as well as numerous volatility models with asymmetries, such as the Asymmetric Power ARCH model (APARCH) of Ding, Granger, and Engle (1993). Given observations \( \epsilon_1, \ldots, \epsilon_n \) and arbitrary initial values \( \tilde{\epsilon}_t \) for \( t \leq 0 \), for \( \theta \in \Theta \) let

\[
\tilde{\sigma}_t(\theta) = \sigma(\epsilon_{t-1}, \epsilon_{t-2}, \ldots, \epsilon_0, \tilde{\epsilon}_{-1}, \ldots; \theta).
\]

Define \( \hat{\theta}_n \) by Equation (2.3) and assume the following.

B1: There exists a finite solution to Model (3.1), which is of the form \( \epsilon_t = \epsilon(\theta_0; I_t, I_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) \) for some function \( \epsilon : \mathbb{R} \rightarrow \mathbb{R} \). Moreover sup \( E\tilde{\epsilon}_t^2 < \infty \) and sup \( E\tilde{\epsilon}_t^2(\theta) = \infty \) for some \( s > 0 \) and all \( \theta \in \Theta \), where \( \tilde{\epsilon}_t(\theta) = \sigma(\epsilon_t, \epsilon_{t-2}, \ldots; \theta) \) with \( \epsilon_t = \epsilon(\theta_0; S_t, S_{t-1}, \ldots; \eta_t, \eta_{t-1}, \ldots) \).
B2: For any real sequence \( (x_t) \), the function \( \theta \mapsto \sigma(x_1, x_2, \ldots, \theta) \) is continuous on \( \Theta \) and belongs to \( (\omega, \infty) \) for all \( \theta \in \Theta \) and for some \( \omega > 0 \). Moreover, \( S_1(\xi_i^2(\theta) - \xi_i^2(\theta_0)) = 0 \) for \( t = 1, \ldots, T \) if \( \theta = \theta_0 \).

B3: There exist a random variable \( K \) measurable with respect to \( \{\epsilon_m, u \leq 0\} \) and a constant \( \rho \) in \( (0, 1) \) such that
\[
\sup_{\theta \in \Theta} |\sigma_1(\theta) - \sigma_1(\theta)| \leq K \rho^t
\]
where \( K \) and \( \rho \) are as in B3 and \( V(\theta_0) \) is some neighborhood of \( \theta_0 \).

B4: There exist no nonzero \( \lambda \in \mathbb{R}^d \) such that \( S_1 \lambda^T \hat{\xi}(\theta_0)^T \) is not equal to zero a.s. for \( t = 1, \ldots, T \).

B5: The function \( \theta \mapsto \sigma(x_1, x_2, \ldots, \theta) \) has continuous second-order derivatives, and
\[
\sup_{\theta \in \Theta} \left\| \frac{\partial \sigma_1(\theta)}{\partial \theta} - \frac{\partial \sigma_1(\theta)}{\partial \theta} \right\| \leq K \rho^t
\]
where \( K \) and \( \rho \) are as in B3 and \( V(\theta_0) \) is some neighborhood of \( \theta_0 \).

B6: There exists a neighborhood \( V(\theta_0) \) of \( \theta_0 \) such that, for \( t = 1, \ldots, T \), the following variables have finite expectation:
\[
\sup_{\theta \in \Theta} \left\| \frac{1}{\xi^2(\theta)} \frac{\partial \xi(\theta)}{\partial \theta} \right\| \leq K \rho^t,
\]
\[
\sup_{\theta \in \Theta} \left\| \frac{1}{\xi^2(\theta)} \frac{\partial^2 \xi(\theta)}{\partial \theta^2} \right\| \leq K \rho^t,
\]
\[
\sup_{\theta \in \Theta} \left\| \xi(\theta) \right\| \leq K \rho^t.
\]
For Model (2.1), we have seen that B1 is satisfied under A1, and the first part of B2 is satisfied under A4 and A6. The identifiability condition in B2 and B4 are entailed by A7 and A8. Assumption A5 entails B3 and B5, as well as the existence of \( \sigma_1(\theta) \) and its derivatives for all \( \theta \in \Theta \). Relations (A.8) and (A.10) of the proof of Theorem 2.2 show that B6 also holds true under the assumptions of Theorem 2.2.

Theorem 3.1. Let \( (I_t) \) be a given bitstream sequence, and \((\epsilon_t)\) a zero-inflated volatility model satisfying Equation (3.1). Under B1–B3, A2, A3 and A6, we have \( \hat{\theta}_n \to \theta_0 \) a.s. Assume in addition B4–B6 and A9 and A10, then the convergence in distribution (2.4) holds true.

A model of the form Equation (3.1) of particular interest is
\[
\epsilon_t = \sigma_t \eta_t I_t, \quad \sigma_t^2 = \omega_0 + \sum_{i=1}^{r} \tau_0 \xi_{t-i} + \alpha_0 \epsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2,
\]
with \( \sum_{i=1}^{r} \tau_0 > -\omega_0 \) and the same constraints and notations as for (2.1). If \( \tau_0 > 0 \), then zero returns tend to increase the volatility, as could be expected when zero returns reflect liquidity issues, but we do not impose this sign constraint a priori. It is clear that, for identifiability of the \( \tau_0 \) coefficients, it is necessary to assume that \( (I_t) \) is well fed in zeros and ones. There exist less trivial reasons for nonidentifiability of the parameters. For example, if \( P(\epsilon_t = 1|\epsilon_{t-1} = 0) = 1 \) and \( P(\epsilon_t = 0|\epsilon_{t-1} = 1) = 1 \) then all the pairs \( (\tau_0_1, \tau_0_2) \) such that \( \tau_0_1 + \tau_0_2 \) is fixed are equivalent. We thus reinforce A7 and A8 by assuming that
\[
A^* \text{ for } j_0 \in \{1, \ldots, T\} \text{ and } \tau > 0, \text{ we have } E(S_{j_0-1|\xi_{j_0-1}^2}^2) \in [\tau, 1-\tau] \text{ a.s. for } i = 0, 1, \ldots, r \vee 1.
\]

Note that, by convention, Model (3.2) with \( r = 0 \) corresponds to (2.1). In this case, A8* reduces to the conditions \( E(S_{j_0-1|F_{j_0-1}^\theta}^2) \in (0, 1) \) and \( E(S_{j_0-1|F_{j_0-2}^\theta}^2) \in (0, 1) \) a.s., which is an alternative to A7 and A8.

Corollary 3.1. Let \( \hat{\theta}_n \) be the zero-adjusted QMLE of the parameter \( \theta_0 = (\omega_0, \alpha_0, \beta_0, \tau_1, \ldots, \tau_r)^T \) of Model (3.2). Under A1–A6, A8*, with obvious changes in A4, in particular assuming
\[
\Theta \supseteq \{ (\omega, \alpha, \beta, \tau_1, \ldots, \tau_r) \in \mathbb{R}^{3+r} : \omega > 0, \alpha > 0, \beta \in (0, 1), \sum_{i=1}^{r} \tau_i + \omega > 0 \},
\]
we have \( \hat{\theta}_n \to \theta_0 \) a.s. Assume in addition A9 and A10, then the convergence in distribution (2.4) holds true.

It is common to assess the adequacy of a time series model by testing the whiteness of the residuals, plotting their empirical (partial) autocorrelations of using formal portmanteau tests, see the monograph by Li (2004). To test the goodness of fit of volatility models, Li and Mak (1994) proposed portmanteau tests based on the autocovariances of the squares of the residuals. The asymptotic distribution of these tests has been studied in particular by Berkes, Horváth, and Kokoszka (2003a) for the standard GARCH models, Carbon and Francq (2011) for APARCH models, Francq, Wintenberger, and Zakoïan (2018) for Log-GARCH and EGARCH models.

First note that \( \hat{\theta} = 0 \) when \( I_t = 0 \), so that \( \hat{\theta} \) should only be a good proxy of \( \eta_t \) when \( I_t = 1 \). Let \( n_t = \sum_{t=r+1}^{T} I_t \) and \( t_1, \ldots, t_n \) the increasing sequence of the times \( t \in \{t_1 + 1, \ldots, n\} \) such that \( I_t = 1 \). For fixed integers \( h < n_t \) and \( m < n_t \), let
\[
\hat{\eta}_h = \frac{1}{n_t} \sum_{h=1}^{t_n} \hat{\eta}_{t-n_t}, \quad \hat{\eta}_{t-n_t} = \hat{\eta}_{t-n_t}^2 - 1, \quad \hat{\eta}_m = (\hat{\eta}_1, \ldots, \hat{\eta}_m)^T.
\]
We will determine the asymptotic distribution of the vector \( \hat{\eta}_m \) of autocovariances of the squares residuals under the null hypothesis
\[
H_0 : \text{the process } (\epsilon_t) \text{ satisfies (3.1)}.
\]

Define the \( m \times d_0 \) matrix whose \( h \)-th row is
\[
\tilde{\eta}_m(h, \cdot) = \frac{1}{n_t} \sum_{h=1}^{t_n} \hat{\eta}_{t-n_t} \frac{\partial \log \hat{\eta}_{t-n_t}}{\partial \theta^{h'}}.
\]
A random variable of the form \( \lambda \eta_{t-n_t}^2 + \mu^2 \log \hat{\eta}_{t-n_t}^2 / \partial \theta^{h'} \) is \( F_{t-n_t} \)-measurable, but in general it is not \( F_{t-n_t-1} \)-measurable. In particular, it is shown in appendix that the following assumption is satisfied under the assumptions of Corollary 3.1.

B7: If \( \lambda \eta_{t-n_t}^2 + \mu^2 / \partial \theta^{h'} \) is \( F_{t-n_t-1} \)-measurable, then \( \lambda = 0 \).

Let \( \tilde{\eta}_m \) the identity matrix of size \( m \) and \( p_1 = T^{-1} \sum_{t=1}^{T} P(S_t = 1) \) the asymptotic proportion of 1’s in the bitstream sequence, which can be estimated by \( \tilde{p}_1 = n_t / n \).

Theorem 3.2. Under \( H_0 \), the assumptions of Theorem 3.1 and B7, we have
\[
T_n := n_1 \hat{\eta}_m \hat{\eta}_m^T \overset{d}{\to} \chi_m^2,
\]
where \( \hat{\eta}_m \) is the identity matrix of size \( m \) and \( p_1 = T^{-1} \sum_{t=1}^{T} P(S_t = 1) \) the asymptotic proportion of 1’s in the bitstream sequence, which can be estimated by \( \tilde{p}_1 = n_t / n \).
It can be seen that an alternative consistent estimator of $D$ is the
empirical variance of $Y_{t_1}, \ldots, Y_{t_n}$, where

$$Y_{t_i} = \hat{s}_{t_i} \left( \hat{s}_{t_i-1} - \frac{n}{n} \tilde{\kappa}_{m} \tilde{\kappa}_{m}^{-1} \frac{\partial \log \tilde{\gamma}_{m}^{2}(\hat{\theta}_m)}{\partial \theta'} \right),$$

$$\hat{s}_{t_i-1} = (\hat{\gamma}_{t_i-1}, \ldots, \hat{\gamma}_{t_i-m})'.$$

The portmanteau test of Li and Mak (1994) consisted in
rejecting $H_0$ at the asymptotic level $\alpha \in (0, 1)$ if $T_n > \chi_m^2(1 - \alpha)$, where $\chi_m^2(\alpha)$ is the $\alpha$-quantile of the $\chi_m^2$ distribution.

4. Simulations

To study the finite sample properties of the zero-adjusted
QMLE, we undertake a set of Monte Carlo simulations. In the simulations the GARCH specifications are nested in

$$\varepsilon_t = \sigma_t I_t, \quad \eta_t \sim iid(0, 1), \quad t = 1, 2, \ldots, n,$$ (4.1)

$$\sigma_t^2 = \omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 + \tau_0 I_{t \in \{\varepsilon_{t-1} = 0\}}.$$ (4.2)

$$\left(\omega_0, \alpha_0, \beta_0, \tau_0\right) = (0.2, 0.1, 0.8, 1.0),$$ (4.3)

where Equation (4.2) is a particular case of model (3.2) for
which the zero-adjusted QMLE is studied in Corollary 3.1. The parameter values correspond (approximately) to the median
values of the estimates in Table 3. The zero-probability $\pi_{0t} = Pr(I_t = 0)$ is governed by one of the following DGPs:

DGP 1: $\pi_{0t} = 0$ for all $t$.

DGP 2: $\pi_{0t} = \begin{cases} 0.5 - (t - 0.49)/(n - 0.7) & \text{if } t \leq n - 0.7 \\ 0.05 & \text{if } t > n - 0.7 \end{cases}$

DGP 3: $\pi_{0t} = 0.1$ if $t$ is odd and $\pi_{0t} = 0.4$ if $t$ is pair.

This means $\{I_t\}$ is stationary in DGP 1, but not in DGPs 2 and 3. In DGP 2 the zero-probability $\pi_{0t}$ is downwards trending in a way that is characteristic among the daily returns of Section 5.1, see Figure 1. For $t = 1$ the probability is $\pi_{0t} = 0.5$, and then it declines until $t = n \cdot 0.7$, that is, at 70% of the sample, where $\pi_{0t} = 0.05$. Thereafter, $\pi_{0t}$ remains constant and equal to 0.05. This is in line with A3. In DGP 3 the zero-probability is periodical—as is common in intraday financial data, and varies between $\pi_{0t} = 0.1$ and $\pi_{0t} = 0.4$ as in our illustration in Section 5.2.

The results for the GARCH(1,1) model are contained in the
upper part of Table 1. For comparison, we include the results of the Standard QMLE in addition to the zero-adjusted
QMLE. Note that in DGP 1 the two QMLEs—and therefore also their results—are identical. When $n = 10,000$, the average finite sample error is 0.004 or less in absolute value for the zero-adjusted QMLE. For the Standard QMLE, by contrast, the finite sample error ranges from 0.02 to 0.13 (in absolute value) in DGP 2, and from 0.01 to 0.056 (in absolute value) in DGP 3. This can be substantial in empirical applications. The asymptotic standard errors of the zero-adjusted QMLE are contained in the columns labeled $ase(\cdot)$, see the supplemental appendix for their computation. The values correspond well to their empirical counterparts—contained in the columns labeled $se(\cdot)$, since they differ a maximum of 0.001 (in absolute value) across the DGPs. When $n = 3000$, the zero-adjusted QMLE also produces substantially less biased estimates than the ordinary QMLE, and the empirical standard errors correspond reasonably well to their asymptotic counterparts. The only exception is $\beta$ in DGP 3, where the Standard QMLE is slightly less biased.

The results for the GARCH(1,1) model with the lagged zero-indicator as covariate are contained in the lower part of Table 1. Note that simulations under DGP 1 is not possible due to exact colinearity. Qualitatively, the simulation results are similar to those of the plain GARCH(1,1). When $n = 10,000$, the average finite sample bias is low in absolute value for the zero-adjusted QMLE (0.006 or less), whereas it is high for the Standard QMLE (0.010 to about 0.504 in absolute value). The largest bias is for
5. Empirical Illustrations

Standard estimators of volatility, for example, the Standard QMLE, provide estimates of the conditional variance. The volatility $\sigma_t^2$ in our model, by contrast, is not at the same scale-level. To facilitate comparison, the conditional variance representation of our model is therefore obtained as follows:

$$E(\epsilon_t^2 | F_{t-1}) = \sigma_t^2 = \sigma_{1t}^2 \pi_{1t}, \quad (5.1)$$

where $\pi_{1t} = \Pr(I_t = 1 | F_{t-1})$, recall Remark 2.1. In other words, the conditional variance representation can be written as a GARCH with time-varying parameters. In particular, in the case of a GARCH(1,1) with the lagged 0-indicator as covariate, the conditional variance representation is

$$\sigma_{0,0adj}^2 = \omega_0 + \alpha_0 \epsilon_{t-1}^2 + \beta_0 (\sigma_{1,0adj}^2 - \pi_{1t}), \quad (5.2)$$

where

$$\sigma_{1,0adj}^2 = \pi_{1t} \sigma_{0,0adj}^2, \quad \omega_0 = \pi_{1t} \omega_0, \quad \alpha_0 = \pi_{1t} \alpha_0, \quad \beta_0 = \frac{\pi_{1t} \beta_0}{\pi_{1t-1}}, \quad \tau_{0t} = \pi_{1t} \tau_0. \quad (5.3)$$

A higher value on the zero-probability $\pi_{0t} = 1 - \pi_{1t}$ thus implies a lower "volatility-level" $\omega_0$, a lower "sensitivity" $\alpha_0$ to nonzero price increments in the previous period, a lower impact $\beta_0$ from the conditional variance (i.e., $\sigma_{1,0adj}^2$) in the previous period, and a lower impact $\tau_{0t}$ from a zero-return in the previous period. Note also that, when the change in $\pi_{1t}$ from $t-1$ to $t$ is sufficiently small, then $\beta_{0t} \approx \beta_0$.

5.1. Daily Returns at the NYSE

We revisit a subset of the NYSE stocks studied in Sucarrat and Grønneberg (2020). The subset of stocks, 24 in total, together with descriptive statistics of their daily returns, are contained in Table 2. The daily returns are computed as $\epsilon_t = 100 \cdot (\ln S_t - \ln S_{t-1})$, where $S_t$ is the closing price of the stock in question at day $t$. The datasource is Bloomberg. To be included in the subset, the NYSE stock must satisfy four criteria. First, at least $n = 1000$ daily price observations must be available over the period 3 January 2007–4 February 2019. Second, the proportion of zero returns must be greater than 10% over the available sample. Third, a moving average ($n = 500$) estimate of the zero-probability should clearly indicate that the zero-process is nonstationary. Graphs of the moving averages are contained in Figure 1. One of our anonymous reviewers suggested that a trend-like evolution in the zero-probability may be due to a corresponding trend-like evolution in the price level: The higher (lower) the nominal price, the lower (higher) the zero-probability due to discrete price changes. Plots of the prices (see the supplemental appendix) suggest such an effect may indeed be present in several of the stocks. Finally, the fourth criterion is that the graphs suggest assumption A3 holds.

GARCH estimates of the daily returns obtained with the zero-adjusted QMLE are contained in Table 3. As noted above, the estimates are not directly comparable to standard GARCH estimates—recall (5.1) and (5.2), and must therefore be adjusted...
Table 3. GARCH estimates (zero-adjusted QMLE) of the daily NYSE returns (Section 5.1).

| Ticker | \( \omega \) (s.e.) | \( \alpha_1 \) (s.e.) | \( \beta_1 \) (s.e.) | \( \tau \) (s.e.) | 95% CI for \( \tau \) | \( \chi^2(2) \) (p-val) |
|--------|-----------------|-----------------|-----------------|-------------|-----------------|-----------------|
| ARR    | 0.099 (0.0274)  | 0.777 (0.0355)  | -0.099 (0.0359) | 0.094 (0.0366) | 0.094 - 0.945  |
| BKT    | 0.009 (0.0105)  | 0.881 (0.0385)  | 0.016 (0.0359)  | 0.009 (0.0366) | 0.009 - 0.817  |
| CO     | 3.127 (1.2138)  | 0.743 (0.1648)  | -0.885 (1.2726) | 3.127 (1.2945) | 0.000 - 6.784  |
| CPS    | 1.871 (0.375)   | 0.217 (0.1648)  | 1.702 (1.2726)  | 1.034 (1.3964) | 0.000 - 3.648  |
| CUBI   | 0.278 (0.0388)  | 0.872 (0.0385)  | 2.813 (0.1648)  | 0.278 (0.1648) | 0.000 - 6.948  |
| DOOR   | 0.218 (0.0399)  | 0.092 (0.0385)  | 2.746 (0.1648)  | 0.036 (0.1648) | 0.000 - 6.948  |
| EROS   | 0.308 (0.0899)  | 0.890 (0.0385)  | 9.306 (0.1648)  | 4.057 (0.1648) | 0.000 - 15.556 |
| ESTE   | 0.250 (0.0428)  | 0.928 (0.0399)  | 0.449 (0.0385)  | 0.250 (0.0399) | 0.000 - 4.388  |
| EVF    | 0.076 (0.0141)  | 0.843 (0.0309)  | -0.103 (0.0309) | 0.022 (0.0309) | 0.000 - 0.844  |
| FF     | 0.106 (0.0267)  | 0.673 (0.0309)  | 2.954 (0.0309)  | 0.009 (0.0309) | 0.000 - 0.450  |
| FSB    | 0.166 (0.0509)  | 0.903 (0.0499)  | 1.390 (0.0499)  | 0.166 (0.0499) | 0.000 - 0.002  |
| FTS    | 0.015 (0.0267)  | 0.663 (0.0309)  | 0.298 (0.0309)  | 1.027 (0.0309) | 0.000 - 0.555  |
| GNK    | 0.047 (0.0136)  | 0.160 (0.0114)  | 0.190 (0.0114)  | 0.047 (0.0114) | 0.000 - 0.000  |
| GPRK   | 0.699 (0.0615)  | 0.799 (0.0615)  | 10.000 (0.0615)| 1.119 (0.0615)| 0.000 - 2.980  |
| GSP    | 0.617 (0.0499)  | 0.940 (0.0499)  | 3.648 (0.0499)  | 1.125 (0.0499)| 0.000 - 2.254  |
| ICL    | 0.000 (0.0100)  | 0.957 (0.0100)  | 0.453 (0.0100)  | 0.091 (0.0100) | 0.000 - 0.975  |
| NOMD   | 0.126 (0.0364)  | 0.897 (0.0399)  | 1.971 (0.0399)  | 1.126 (0.0399) | 0.000 - 6.256  |
| NVGS   | 0.696 (0.0849)  | 0.849 (0.0849)  | 10.000 (0.0849) | 0.696 (0.0849) | 0.000 - 1.830  |
| OSB    | 0.019 (0.0114)  | 0.945 (0.0114)  | 0.445 (0.0114)  | 0.067 (0.0114) | 0.000 - 1.791  |
| PARR   | 0.068 (0.0273)  | 0.968 (0.0273)  | 0.085 (0.0273)  | 0.068 (0.0273) | 0.000 - 0.060  |
| TARO   | 1.123 (0.360)   | 0.400 (0.1939)  | 5.048 (0.1939)  | 1.778 (0.1939) | 0.000 - 3.817  |
| TIER   | 0.014 (0.0136)  | 0.947 (0.0136)  | 1.515 (0.0136)  | 0.035 (0.0136) | 0.000 - 1.354  |
| TU     | 0.012 (0.0073)  | 0.962 (0.0073)  | 0.450 (0.0073)  | 0.037 (0.0073) | 0.000 - 0.256  |
| WCN    | 0.012 (0.0132)  | 0.963 (0.0132)  | 0.295 (0.0132)  | 0.012 (0.0132) | 0.000 - 0.217  |

NOTE: zero-adjusted QMLEs of \( \sigma^2 = \omega + \alpha_1 \sigma^2_t + \beta_1 \sigma^2_{t-1} + \tau \), s.e. standard error of estimate. Upper bound of 95% CI for \( \tau \) computed as \( \tau + \text{s.e.}(\tau) \cdot 1.96 \), where \( \text{s.e.}(\tau) \) is the standard error of \( \tau \). Lower bound computed as \( \text{max}(\omega, 0) \), where \( \omega \) is the parameter of interest. To avoid explosive volatility-paths, the upper bound of \( \tau \) is imposed during estimation, \( \chi^2(2) \), the results from the portmanteau test of Section 3 of autocorrelation up to and including order 2 of \( \eta^2 \) (p-value in parentheses).

5.2. Intraday 5-min USD/EUR Returns

Intraday financial returns are frequently characterized by a periodic nonstationary zero-process; see, for example, Kolokolov, Livieri, and Pirino (2020). An example is the intraday 5-min
USD/EUR exchange rate return. Let $S_t$ denote the exchange rate at the end of a 5-min interval, and let $r_t$ denote the log-return in basis points from the end of one interval to the end of the next: $\epsilon_t = 100^2 \cdot (\ln S_t - \ln S_{t-1})$. The left graph in Figure 3 contains the returns from 2 January 2017 to 31 December 2018, a total of $n = 147,347$ returns. The source of the data is Forexite. Only
the returns are characterized by excess kurtosis relative to the Normal distribution, and first-order autocorrelation in $\epsilon_t^2$. The proportion of zero-returns over the sample is 20.3%, and the right graph of Figure 3 depicts how the zero-proportion varies intraday across the 24-hour trading day. In the beginning of the day, only the Asian markets are active, so the zero-probability is higher. As European markets open, activity increases and so does the zero-probability. Intraday activity is lower. As European markets open, activity increases and so does the zero-probability. The zero-probability remains low until the close of the European markets, and then gradually increases again as only the American markets remain active. The zero-probability reaches its peak at the close of the American markets.

The middle part of Table 5 contains the GARCH estimates. In both the standard and zero-adjusted cases, $\tau$ is estimated to be negative, and the 95% CIs for $\tau$ do not contain the value 0. In other words, the results suggest a zero-return in the previous period tends to reduce volatility in the next period at the 5-min frequency for this exchange rate during the sample period of the data at the trading platform in question. To obtain estimates of $\pi_{1t}$ and $\pi_{0t}$, we use a centered moving average of length 12— that is, one hour of trading—made up of the intradaily zero-proportions of the 5-min intervals. The zero-proportions over the trading day, together with the estimate $\hat{\pi}_{0t}$, are both depicted in the right graph of Figure 3. Note that the periodic cycle is 288. Figure 4 contains the estimates of the time-varying parameters implied by Equation (5.2) together with the estimates of the Standard QMLE. As is clear, the standard estimate of $\alpha$ is biased downwards throughout the day, and it is also outside the 95% CI throughout the day. The standard estimate of $\alpha$ is biased upwards throughout the day, and most of the time outside the 95% CI. The intraday evolution of the zero-adjusted estimate $\hat{\alpha}_t$ is similar to that of $\hat{\alpha}_t$: It is at its highest in the middle of the day when trading is at its highest, and at its lowest in the beginning and end of the day when trading is at its thinnest. The zero-adjusted estimate of $\beta_t$ oscillates about the Standard estimate of 0.857, and in only a couple of instances is the Standard estimate outside the 95% CI. The estimates of $\tau$ are both negative. The standard estimate is biased upwards, but it is always within the 95% CI of the zero-adjusted estimate. So they are not significantly different from each other at 5%.

One of our anonymous reviewers asked us to compare the estimates of the zero-adjusted GARCH, which is of observed return, with those of a GARCH model of the efficient return process as defined in Bandi et al. (2020). There, zeros occur when the efficient return process is unobserved. To this end,
Figure 3. Intraday 5-min USD/EUR log-returns in basis points (left graph) from 2 January 2017 to 31 December 2018 ($n = 147,347$), and the proportion of zero-returns in each intraday 5-min interval (right graph). The smoother is a centred moving average of length 12 (see Section 5.2). Data source: Forexite.

we derive a modified version of the moment-based estimator of Kristensen and Linton (2006), see Appendix D (supplementary material) for the details. The estimates are also contained in the middle part of Table 5. Note that an estimate of $\tau$ is not available for this estimator. Compared with the zero-adjusted estimates depicted in Figure 4, the $\omega$ and $\alpha$ estimates are lower, whereas the estimate of $\beta$ is higher. The $\alpha$ and $\beta$ estimates of 0.019 and 0.958, respectively, are particularly different, since they are always substantially outside the 95% CIs of the zero-adjusted estimates.

To investigate to what extent the Standard and zero-adjusted QMLEs produce different volatility estimates, we study the distance $x_t = \hat{\sigma}_{t,\text{adj}} - \hat{\sigma}_t$, just as in Section 5.1. The lower part of Table 5 reports the properties of $x_t$. Again the test of whether $E(x_t) = 0$ or not is implemented via the regression $x_t = \mu + u_t$ with Newey and West (1987) standard errors. The average of $x_t$ is $-0.037$, and a two-sided test with 0 as null is rejected at all the usual significance levels. Accordingly, the results suggests the Standard QMLE produces conditional volatilities that are too high, on average. Unconditionally, the value of $-0.037$ is not large. Conditionally, the range between the maximum and minimum values of $x_t$ suggests the discrepancy can be large on a day-to-day basis. Figure 5 contains the graph of $x_t$. Most of the time $x_t$ lies between 0.3 and $-1.0$. Recalling that the 5-min returns are expressed in basis points, these differences do not appear to be large in economic terms.

6. Conclusions

Financial time series are frequently nonstationary due to a nonstationary zero-process. In these situations, standard estimators are not consistent. We propose a GARCH model that accommodates a nonstationary zero-process, and derive a zero-adjusted QMLE. The nonstationary zero-process can either be trend-like in nature, as is common in daily data, or periodic, as is common in intraday data, or both. The volatility specification in our model can contain higher order ARCH and GARCH terms, asymmetry terms (“leverage”) and past zero-indicators as covariates. The latter is of special interest in the current context, since it enables us to study the effect of a zero return on volatility in the subsequent period. Consistency and asymptotic normality of the zero-adjusted QMLE is proved under mild assumptions. Moreover, under stationarity of the zero-process the estimator will still be CAN, so there is no harm in applying our estimator under stationarity. Finite sample simulations verify that the estimator has good finite sample properties, and confirm that the Standard QMLE is biased when the zero-
process is nonstationary. Two empirical studies illustrate our results. One is on 24 daily stock returns at NYSE, and one is on intraday 5-min USD/EUR exchange rate returns. In both studies we find that the time-varying zero-probability affects the dynamics in substantial ways, that the fitted volatilities can differ significantly, and that a zero-return in the previous day...
can have a substantial effect on volatility in the subsequent day. Interestingly, however, we do not always find that the effect is positive.

While a nonstationary zero-process is frequent in financial time-series, only recently have researchers directed their attention toward this characteristic. Several lines of future research suggest themselves. First, the extension to more general volatility models outlined in Section 3 accommodates models with asymmetry ("leverage"). An interesting line of further research is to study how the evolution of the zero-probability impacts on the effect of asymmetry. Second, it is well known that financial time series—both daily and intraday—can be nonstationary due to changes in the level of the unconditional volatility. How frequent are such changes due to a nonstationary zero-process? To the best of our knowledge, this has not been investigated before. Third, to obtain the conditional variance representation of our model, estimates of the time-varying probabilities of a nonstationary zero-process is required. This is challenging. More research is needed to ascertain what the most suitable approach is, and under which assumptions. Fourth, as noted by one of our anonymous reviewers, the zero-process may not be the only source of nonstationarity. In addition, the volatility intercept (ω), and the ARCH and GARCH parameters, may also be time-varying. To the best of our knowledge, nobody has developed methods for situations where both types of nonstationarities are present. Finally, knowledge about the relation between observed zeros and the underlying efficient return process is limited, so more research on this is needed.

### Supplementary Materials

The supplementary file contains four appendices. Appendix A gathers all the proofs and some complementary theoretical results. Appendix B describes the computation of the asymptotic covariance of the 0-adjusted QMLE in the simulation experiments. Appendix C gives details about the NYSE stocks used in the empirical study. Appendix D discusses the possibility of estimating the volatility of the efficient returns considered in Bandi et al. (2017) and Bandi et al. (2020). The replication files are available via https://www.sucarrat.net/research/replication-files-garch-0-nonstationary.zip.

### Acknowledgments

We are grateful to the editor, the associate editor and two anonymous reviewers for their helpful comments and suggestions. Christian Francq is grateful to the Agence Nationale de la Recherche (ANR), which supported this work via the Project MLEforRisk (ANR-21-CE26-0007-01) and also thanks the labex ECODEC.

### References

Azrak, R., and Mélard, G. (2006), "Asymptotic Properties of Quasi-Maximum Likelihood Estimators for ARMA Models with Time-Dependent Coefficients," *Statistical Inference for Stochastic Processes*, 9, 279–330. [54,55]

Bandi, F. M., Kolokolov, A., Pirino, D., and Reno, R. (2020), “Zeros,” *Management Science*. DOI: 10.1287/mnsc.2019.3527. [53,62,65]

Bandi, F. M., Pirino, D., and Reno, R. (2017), “Excess Idle Time,” *Econometrica*, 85, 1793–1846. [53,65]

Berkes, I., Horváth, L., and Kokoszka, P. (2003a), “Asymptotics for GARCH Squared Residual Correlations,” *Econometric Theory*, 19, 515–540. [57]

Berkes, I., Horváth, L., and Kokoszka, P. (2003b), “GARCH Processes: Structure and Estimation,” *Bernoulli*, 9, 201–227. [54]

Bien, K., Noftle, I., and Pohlmeier, W. (2011), “An Inflated Multivariate Integer Count Hurdle Model: An Application to Bid and Ask Quote Dynamics,” *Journal of Applied Econometrics*, 26, 669–707. [53]

Buccheri, G., Livieri, G., Pirino, D., and Pollastri, A. (2020), “A Closed-Form Formula Characterization of the Epps Effect,” *Quantitative Finance*, 20, 243–254. [53]

Buccheri, G., Pirino, D., and Trapi, L. (2020), “Managing Liquidity with Portfolio Staleness,” *Decisions in Economics and Finance*. DOI: 10.1007/s10203-020-00300-z. [53]

Carbon, M., and Francq, C. (2011), “Portmanteau Goodness-of-Fit Test for Asymmetric Power GARCH Models,” *Austrian Journal of Statistics* 40, 55–64. [57]

Catania, L., R. Di Maria, and P. Santucci de Magistris (2020), “Dynamic Discrete Mixtures for High-Frequency Prices,” *Journal of Business and Economic Statistics*. DOI: 10.1080/07350015.2020.1840994. [53]

Ding, Z., Granger, C., and Engle, R. (1993), “A Long Memory Property of Stock Market Returns and a New Model,” *Journal of Empirical Finance* 1, 83–106. [56]

Engle, R. F., and Russell, J. R. (1998), “Autoregressive Conditional Duration: A New Model of Irregularly Spaced Transaction Data,” *Econometrica*, 66, 1127–1162. [53]

Escanciano, J. C. (2009), “Quasi-Maximum Likelihood Estimation of Semi-Strong GARCH Models,” *Econometric Theory*, 25, 561–570. [53]

Francq, C., and Gauthier, A. (2004), “Large Sample Properties of Parameter Least Squares Estimates for Time-Varying Arma-Models,” *Journal of Time Series Analysis* 25, 765–783. [54,55]

Francq, C., Wintenberger, O., and Zakoïan, J.-M. (2018), “Goodness-of-Fit Tests for Log and Exponential GARCH Models,” *TEST*, 27, 27–51. DOI: 10.1007/s11776-016-0506-2. [57]

Francq, C., and Zakoïan, J.-M. (2019), *GARCH Models* (2nd ed.), New York: Wiley. [56]

Harvey, A. C., and Ito, R. (2020), “Modeling Time Series When Some of the Observations are Zero,” *Journal of Econometrics*, 214, 33–45. [53]

Hausman, J., Lo, A., and MacKinlay, A. (1992), “An Ordered Probit Analysis of Transaction Stock Prices,” *Journal of Financial Economics*, 31, 319–379. [53]

Hautsch, N., Malec, P., and Schienle, M. (2013), “Capturing the Zero: A New Class of Zero-Augmented Distributions and Multiplicative Error Processes,” *Journal of Financial Econometrics*, 12, 89–121. [53]
Kolokolov, A., Livieri, G., and Pirino, D. (2020), "Statistical Inferences for Price Staleness," *Journal of Econometrics*, 218, 32–81. [54,60]

Kolokolov, A., and Reno, R. (2019), "Jumps or Staleness," working paper, DOI: 10.2139/ssrn.3208204. [54]

Kristensen, D., and Linton, O. (2006), "A Closed-Form Estimator for the GARCH(1,1) Model," *Econometric Theory*, 22, 323–337. [62,63]

Kümm, H., and Küsters, U. (2015), "Forecasting Zero-Inflated Price Changes with a Markov Switching Mixture Model for Autoregressive and Heteroscedastic Time Series," *International Journal of Forecasting*, 31, 598–608. [53]

Lesmond, D. A., Ogden, J. P., and Trzcinka, C. A. (1999), "A New Estimate of Transaction Costs," *The Review of Financial Studies*, 12, 1113–1141. [53]

Li, W. (2004). *Diagnostic Checks in Time Series*, Boca Raton, FL: Chapman and Hall. [57]

Li, W., and Mak, T. (1994), "On the Squared Residual Autocorrelations in Non-Linear Time Series With Conditional Heteroscedasticity," *Journal of Time Series Analysis*, 15, 627–636. [57,58]

Liesenfeld, R., Nolte, I., and Pohlmeier, W. (2006), "Modelling Financial Transaction Price Movements: A Dynamic Integer Count Data Model," *Empirical Economics*, 30, 795–825. [53]

Ljung, G., and Box, G. (1979), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 66, 265–270. [59,62]

Newey, W., and West, K. (1987), "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708. [60,62,63]

Phillips, P. C., and Xu, K. (2006), "Inference in Autoregression Under Heteroskedasticity," *Journal of Time Series Analysis*, 27, 289–308. [54,55]

Regnard, N., and Zakoian, J.-M. (2010), "Structure and Estimation of a Class of Nonstationary Yet Nonexplosive GARCH Models," *Journal of Time Series Analysis*, 31, 348–364. [54,55]

Russell, J. R., and Engle, R. F. (2005), "A Discrete-State Continuous-Time Model of Financial Transaction Prices and Times: The Autoregressive Conditional Multinomial-Autoregressive Conditional Duration Model," *Journal of Business and Economic Statistics*, 23, 166–180. [53,60]

Rydberg, T. H., and Shephard, N. (2003), "Dynamics of Trade-by-Trade Price Movements: Decomposition and Models," *Journal of Financial Econometrics*, 1, 2–25. [53]

Sucarrat, G., and Grønneberg, S. (2020), "Risk Estimation with a Time Varying Probability of Zero Returns," *Journal of Financial Econometrics*, Forthcoming. DOI: 10.1093/jjfinec/nbaa014. [53,59]