CEIS TOR VERGATA Research Papers

Università di Roma Tor Vergata

ISSN 2610-931X

CEIS Tor Vergata

RESEARCH PAPER SERIES

Vol. 19, Issue 2, No. 506 – March 2021

A Structural Model of Market Friction with Time-Varying Volatility

Giuseppe Buccheri, Stefano Grassi and Giorgio Vocalelli

A Structural Model of Market Friction with Time-Varying Volatility*

Giuseppe Buccheri† Stefano Grassi‡ Giorgio Vocalelli§

March 23, 2021

Abstract

We propose a model of price formation in which the trading price varies only if the value of the information signal is large enough to guarantee a profit in excess of transaction costs. Using transaction data only, we extract: (i) the conditional volatility of the underlying security, which is thus cleaned out by market frictions, (ii) an estimate of transaction costs. Our analysis reveals that, after correcting for frictions, the risk of illiquid securities is substantially different from what predicted by traditional volatility models. Furthermore, in periods of high volatility, our estimate of transaction costs remains highly correlated with bid-ask spreads, whereas alternative illiquidity proxies, such as the fraction of zero returns, loose their explanatory power.

Keywords: Illiquidity; Market Microstructure; Volatility; Risk assessment.
JEL Classification: B26; C22; C58

*We thank Fulvio Corsi, Andrea De Polis, Andrew Harvey, Roberto Renò and Luca Trapin for comments and suggestions on this paper. We also thank seminar participants at University of Rome 'Tor Vergata' and participants of the CFE 2020 conference in London, the ICEEE 2021 conference in Cagliari and the XXII Workshop on Quantitative Finance in Verona. Stefano Grassi gratefully acknowledges financial support from the University of Rome 'Tor Vergata' under Grant "Beyond Borders" (CUP: E84I20000900005). All errors are our own.

†University of Rome 'Tor Vergata' - Department of Economics and Finance. Via Columbia, 2, 00133 Rome, Italy. Email: giuseppe.buccheri@uniroma2.it

‡University of Rome 'Tor Vergata' - Department of Economics and Finance. Via Columbia, 2, 00133 Rome, Italy. Email: stefano.grassi@uniroma2.it

§University of Rome 'Tor Vergata' - Department of Economics and Finance. Via Columbia, 2, 00133 Rome, Italy. Email giorgio.vocalelli@uniroma2.it

1 Introduction

Real financial markets are not perfectly liquid. Several frictions prevent investors from instantaneously buying or selling a security at a given point in time. Bid-ask spreads, commission fees, price impact are few examples of relevant costs that must be paid to trade common stock. Illiquidity, intended as the cost needed to buy or sell a security, has a key role in modern finance research, having implications on asset pricing, market efficiency and corporate finance; see Amihud and Mendelson (1987), Longstaff (2009), Gryglewicz (2011) among others.

In this paper, we deal with the problem of extracting the volatility of a financial security in a market with frictions. From an econometric viewpoint, this problem is relevant because illiquidity affects the time-series behaviour of financial prices. When transactions are costly, investors may be unable to afford the trading costs, deciding not to trade. Lack of trading is a well-known phenomenon in finance entailing periods of market inactivity where prices are not frequently updated over time. One of the main consequences of lack of trading is the existence of a significant number of zero returns in data, whose effect is to "smooth out" transaction prices. Because of such smoothing, real markets behave differently from what expected based on traditional modelling assumptions. The hypothesis of a zero likelihood of observing a zero return, which is implicit in stochastic volatility and GARCH-type models, is violated in reality. We thus expect that neglecting illiquidity may have serious implications when extracting volatility from financial time-series.

The problem is relevant even from a financial economics viewpoint. The model proposed in this paper delivers an estimate of transaction costs given a time-series of trading prices. The novelty of this transaction cost measure is that it gauges not only the number of zero returns, but even the volatility of the information signal. There is today a solid literature linking illiquidity to zero returns; see e.g. Lesmond et al. (1999), Lesmond (2005), Bekaert et al. (2007), Goyenko et al. (2009), Naes et al. (2011), Bandi et al. (2017), Bandi et al. (2020). Our analysis reveals that using zero returns alone, as a proxy for market illiquidity, can lead to misleading results in periods of market distress characterized by high volatility. Trading activity typically increases in such periods as a result of market sentiment, for instance, due to massive fire sales of financial securities. Liquidity may thus appear to be spuriously large if measured only in terms of zero returns. As will be shown, our model does not need to estimate artificially large liquidity, because the high trading activity can be explained by a more volatile information signal.

The approach adopted in this paper is rooted in the adverse selection theory of Glosten and Milgrom (1985) and Kyle (1985). This microstructure framework establishes a connection among the absence of trading, transaction costs and the co-existence of investors with different layers of information on the asset fundamental value. We postulate the existence of an efficient martingale process that instantaneously incorporates newly available information. The observed price does not necessarily coincide with the value of the information signal. Newly available information is incorporated into the observed price only if the efficient price change is larger, in module, than a given threshold. In that case, net of transaction costs, the observed return coincides with the efficient price change. If instead the threshold is not hit, the observed price return is equal to zero. The threshold parameter acts as a friction which prevents the signal from being incorporated into the trading price if it is not sufficiently informative.

This price generation mechanism is similar to the limited dependent variable model of Lesmond et al. (1999). There is, however, a fundamental difference between the two modelling approaches. In Lesmond et al. (1999), efficient price changes are assumed to follow the Capital Asset Pricing Model (CAPM), which implies a decomposition of the *unconditional* volatility in terms of systematic and idiosyncratic components. In our model, efficient price changes are unobserved martingale difference sequences with time-varying *conditional* volatility. This assumption allows us to extract volatility from past observed returns and disentangle it from market frictions. We are thus able to track the evolution of volatility over time, a result that would not be achievable under the CAPM assumption. Similarly to Lesmond et al. (1999), our model delivers an estimate of transaction costs that is computable using solely transaction data. As argued by Bekaert et al. (2007), this constitutes an important advantage in cases where information other than transaction data is unavoidable, for instance in emerging markets.

The proposed model has a microstructure interpretation resting upon well-established financial theories of price formation. However, to use it in practice, it is necessary to set up a statistical methodology allowing to filter the volatility of efficient price changes and to estimate the threshold parameter. Score-driven models (Creal et al. 2013, Harvey 2013) provide a natural framework for this purpose. They allow building a filter for the time-varying parameters accounting for the full shape of the conditional likelihood. Our model is formulated in terms of a discrete-continuous mixture with a probability mass in zero. The update of volatility, being based on the score of such distribution, is robust to zero returns. In the climate literature, Harvey and Ito (2020) propose a similar score-driven specification for positive observations, such as daily rainfall levels; see also Harvey and Liao (2019), who propose a class of related censored Tobit models driven by the score of the conditional likelihood.

The objective of our empirical analysis is twofold. On the one hand, we aim to understand whether the threshold parameter in our model can be regarded as an estimate of transaction costs, as its microstructure interpretation suggests. On the other hand, we aim to assess the effect of illiquidity on the estimation and forecast of volatility. The dataset is composed of trades and quotes of 190 stocks traded in the NYSE from the beginning of 2006 to the end of 2014. The analysis, performed on the entire cross-section of stocks, shows that the threshold parameter is significantly correlated with the main determinants of (il)liquidity, namely trading volume, number of transactions, bid-ask spreads. An interesting result is that these correlations remain significant in periods of market distress, like the 2008 global financial crisis, whereas they tend to vanish when replacing the threshold parameter with the number of zero returns. Indeed, the high trading activity observed in these periods is not imputable to a decrease in trading costs, but rather to the widespread uncertainty prevailing in the market. In our model, larger volatility determines a larger likelihood of making a profit in excess of transaction costs, and thus a larger probability of trading. The high trading activity can thus be explained by an increase in the volatility of the information signal. This leads to a better assessment of trading costs when measuring the latter in terms of cross-sectional correlations between the threshold parameter and bid-ask spreads. The second goal of our empirical analysis is to understand whether illiquidity affects the estimated risk of observed returns. We compare the Value-at-Risk back-testing results of different volatility models. They show that neglecting illiquidity has a significant impact on traditional risk management applications involving illiquid securities. These results may have concrete policy implications in light of the role played by risk measures in the current banking regulatory framework and the whole finance industry.

In the financial econometric literature, the recent paper of Sucarrat and Grønneberg (2020) represents an attempt to estimate volatility in the presence of infrequent trading. They propose a zero-augmented model in the spirit of Hautsch et al. (2013) to assign a non-zero probability to the occurrence of zero observations. The approach proposed here differs in two main aspects. First, our model has a precise microstructure interpretation. The probability of observing a zero return depends on trading opportunities, which are in turn related to volatility and transaction costs. The threshold parameter in our model is an estimate of round-trip costs, as in Lesmond et al. (1999). From this perspective, our paper also relates to Bandi et al. (2017), who estimates a microstructure model with transaction costs and asymmetric information using moment conditions involving the daily fraction of zero returns. Second, in our approach zero returns affect not only the conditional distribution of returns but even the dynamics of volatility. As underlined by Harvey and Ito (2020), the score allows to properly choose the weight assigned to zero returns in the update of volatility. Assuming a weight equal to zero or treating zero returns as missing values is not consistent when they depend on non-zero returns. Other work which relates to our paper includes Rydberg and Shephard (2003) and Catania et al. (2020), who propose a joint distribution for modeling tick-by-tick data leading to stale prices. Even in this case, the main difference with the approach proposed here is that our model has solid microstructure foundations, allowing us to extract the volatility of efficient price changes and an estimate of round-trip costs.

The rest of the paper is organized as follows. Section (2) introduces the model and discusses its main properties. Section (3) shows the results of Monte Carlo experiments and discusses their implications on volatility estimation and forecasting. Section (4) illustrates the empirical results. Section (5) concludes. We report in the Appendix the computations related to the derivation of the score formula.

2 Methodology

2.1 The microstructure model

In a frictionless and informationally efficient market, the price of a financial security can be described as a martingale process that instantaneously incorporates all available information. If transactions are costly, and the market is informationally efficient, the martingale price still reflects all information, but it does not necessarily coincide with the price observed in the market. The asset fundamental value can differ from the observed price for several reasons. For example, an investor might not react to new information if transaction costs are judged to be too large; similarly, an informed trader might decide not to trade if the value of the information signal is not large enough to cover the trading costs. These examples are direct implications of the adverse selection theory of Kyle (1985) and Glosten and Milgrom (1985), which relates lack of trading to transaction costs and asymmetric information.

Trading opportunities are not only determined by transaction costs. The volatility of the information signal also plays a key role in the price formation mechanism. If volatility is low, the efficient price varies only by a small amount. Consequently, it is difficult for the informed investor to set up trading strategies guaranteeing a profit in excess of transaction costs. If in contrast volatility is large, an informed investor can exploit the big fluctuations of the asset fundamental value to realize net profits. The first example describes a situation where prices are unlikely to be updated over time, whereas in the second example we expect that the observed price will change to reflect newly available information. We also know that volatility can vary significantly over time. The time-varying nature of volatility can thus be responsible for dynamic changes in the trading activity. We aim to formulate a microstructure model which explains the interrelation among lack of trading, transaction costs, and volatility.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let x_t, y_t be two random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. To fix ideas, x_t can be thought of as representing the return of the efficient price at time t, whereas y_t can be regarded as the return observed in the market at time t. We assume that, conditionally on a given filtration, which will be specified below, the process x_t is a martingale difference sequence with conditional volatility denoted by σ_t. We also assume that the observed returns y_t are generated as follows:

$$y_t = (x_t - c)I_{x_t > c} + (x_t + c)I_{x_t < -c}, \tag{1}$$

where $c \in [0, +\infty[$ is constant and $I_{\{\cdot\}}$ denotes the indicator function. The microstructure interpretation of the model in Equation (1) is immediate. If x_t lies inside the interval $[-c, c]$, then the observed return is equal to zero. In this situation, an informed trader decides not to react to the information signal, as the profit coming from buying or selling the underlying security would not be enough to offset the round-trip cost. That includes bid-ask spreads, commission fees, price impact costs, and is given by $2c$ in our model. If in contrast, the efficient price change is outside the interval $[-c, c]$, the observed price is updated to reflect new information. In this case, the observed return matches the efficient price change in excess of round-trip costs.

The model in Equation (1) implicitly defines a certain likelihood of observing a zero return that depends on the threshold c and the volatility parameter σ_t. If transaction costs are large, then trading is more likely to be inconvenient. Similarly, if volatility is small, it is less likely that a strategy exploiting efficient price movements will generate profits in excess of transaction costs. Note that large trading costs do not necessarily imply absence trading. If the volatility of the information signal is large as well, then trading may be convenient. Note also that the zero returns generated by the model are not temporally independent. If the volatility process σ_t is persistent, the model can generate periods of high volatility characterized by large trading activity, and periods of low volatility characterized by reduced trading activity.

It is worth highlighting that the proposed model does not imply any structural relationship between transaction costs and volatility, which are treated as two independent variables. Several theoretical and empirical studies have demonstrated the explanatory power of volatility in determining bid-ask spreads; see e.g. Stoll (1978b), Stoll (1978a), Amihud and Mendelson (1987), Amihud and Mendelson (1989), Stoll (2000). In our framework, such dependencies can only be detected *a posteriori*, i.e. after estimating the model on real data; see Section (4.3).

Our price generation mechanism shares the same logic of the microstructure model of Lesmond et al. (1999). There is, however, a fundamental difference in the approach adopted to model efficient price changes. In Lesmond et al. (1999), they are assumed to follow the CAPM. Their methodology provides an estimate of transaction costs, together with market betas and idiosyncratic volatilities; see Bekaert et al. (2007), Mei et al. (2009), Goyenko et al. (2009), Griffin et al. (2010) for applications of the illiquidity measure of Lesmond et al. (1999). As detailed below, we assume that x_t is a martingale difference sequence with time-varying volatility. Our methodology provides a filter for σ_t, which represents the conditional volatility of the underlying security, corrected to account for transaction costs. Extracting σ_t from data would not be possible within the framework of Lesmond et al. (1999), as it only implies a decomposition of the unconditional volatility in terms of systematic and idiosyncratic components.

2.2 Estimation and filtering

To use in empirical applications the model in Equation (1), it is necessary to specify the dynamic of the volatility process σ_t and to devise a filtering procedure to extract it from data. The score methodology of Creal et al. (2013) and Harvey (2013) is particularly useful for the problem at hand. In score-driven models, the time-varying parameters are fully determined by past observations. Their law of motion is of autoregressive type, with the score of the conditional density acting as a forcing variable. Let $\mathcal{F}_{t-1} = \sigma\{y_j\}_{j=-\infty}^{t-1} \subseteq \mathcal{F}$ denote the σ-algebra generated by past observations of the process y_t. We assume that, conditionally on $\mathcal{F}_{t-1}$, the law of x_t is absolute continuous with respect to the Lebesgue measure. Let us denote its Radon-Nikodym derivative by $p_X(x_t; \mathcal{F}_{t-1}, \sigma_t)$ and let σ_t be measurable with respect to $\mathcal{F}_{t-1}$. By writing $\sigma_t = e^{\lambda_t}$, the score based update of λ_{t+1} given observations up to time t is:

$$\lambda_{t+1} = \omega + \phi\lambda_t + \kappa u_t, \tag{2}$$

where

$$u_t = \frac{\partial \log p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t)}{\partial \lambda_t}, \tag{3}$$

with $p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t)$ denoting the conditional density of the observed returns.

There are several reasons which motivate the use of the score as a general methodology to update the time-varying parameters. It is possible to show that the score update is optimal based on information theoretic criteria; see Blasques et al. (2015). Furthermore, as λ_{t+1} is $\mathcal{F}_t$-measurable, the conditional likelihood can be written down in closed form, and thus parameters are easily estimated by standard maximum likelihood. Popular non-linear time-series models, like the GARCH model of Bollerslev (1986) and the ACD model of Engle and Russell (1998), are particular instances of the class of score-driven models; see Creal et al. (2013) and Harvey (2013).

The model in Equation (1) implies that, conditionally on $\mathcal{F}_{t-1}$, the density of the observed returns can be written as:

$$p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t) = \begin{cases} p_X(y_t + c; \mathcal{F}_{t-1}, \lambda_t) & y_t > 0 \\ F_X(c; \mathcal{F}_{t-1}, \lambda_t) - F_X(-c; \mathcal{F}_{t-1}, \lambda_t) & y_t = 0 \\ p_X(y_t - c; \mathcal{F}_{t-1}, \lambda_t) & y_t < 0, \end{cases} \tag{4}$$

where $F_X(\cdot, \mathcal{F}_{t-1}, \lambda_t)$ denotes the law of x_t, conditionally on $\mathcal{F}_{t-1}$. Thus, the observed returns are described by a discrete-continuous mixture obtained by shifting to the left (right) the graph of $p_X(x_t; \mathcal{F}_{t-1}, \lambda_t)$ by an amount c for $x_t > 0$ ($x_t < 0$) and then censoring the negative (positive) values. The probability mass in zero is set equal to the mass of the censored returns.

An example is illustrated in Figure (1), which shows $p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t)$ and $p_X(x_t; \mathcal{F}_{t-1}, \lambda_t)$ in the case where $p_X(\cdot)$ is a Student-t density. Harvey and Ito (2020) use a similar discrete-continuous mixture with score-driven scale parameter to model daily rainfall levels. The variable x_t in their model is described by a probability density with positive support, which is thus shifted to the left.

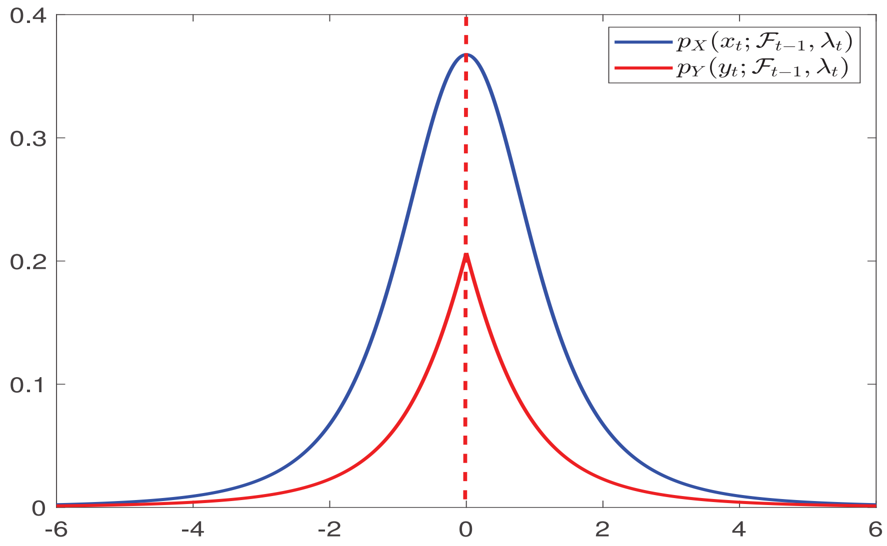

Figure 1: Conditional density of efficient price changes (blue line) and discrete-continuous mixture of observed returns (red line). The latter is obtained from the former by shifting to the left (right) of an amount $c = 1$ the branch of the graph corresponding to positive (negative) returns, and censoring the negative (positive) part. The probability mass in zero, represented by the dashed red line, is equal to the mass of the censored returns. The conditional density of efficient price changes is chosen to be a standardized Student-t with $\nu = 3$ degrees of freedom.

In order to capture the fat-tail behaviour of real financial returns, we set $p_X(x_t; \mathcal{F}_{t-1}, \lambda_t)$ as a Student-t density[1]:

$$p_X\left(x_t; \mathcal{F}_{t-1}, \lambda_t\right) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left(\pi\nu e^{2\lambda_t}\right)^{1/2}} \left(1 + \frac{x_t^2}{\nu e^{2\lambda_t}}\right)^{-\frac{\nu+1}{2}}, \tag{5}$$

where $\Gamma(\cdot)$ denotes the Gamma function. We assume that $\nu > 2$, so that the conditional variance of x_t exists and is given by $\mathrm{Var}[x_t|\mathcal{F}_{t-1}, \lambda_t] = \frac{\nu}{\nu-2}e^{2\lambda_t}$. For $x_t > 0$, the law $F_X(\cdot, \mathcal{F}_{t-1}, \lambda_t)$ of x_t, conditionally on $\mathcal{F}_{t-1}$, is given by:

$$F_X(x_t, \mathcal{F}_{t-1}, \lambda_t) = \frac{1}{2} + \frac{1}{2}\beta(b_{x_t}; 1/2, \nu/2) \tag{6}$$

[1]Following Harvey and Lange (2017), it is also possible to set $p_X\left(x_t; \mathcal{F}_{t-1}, \lambda_t\right)$ equal to a generalized t distribution, which would allow the distribution of x_t to be skewed and asymmetric. For simplicity, we limit ourselves to consider a Student-t density here, however the formulation of the model can easily be generalized using a more flexible distribution.

where $b_{x_t} = \frac{x_t^2/(\nu e^{2\lambda_t})}{x_t^2/(\nu e^{2\lambda_t})+\nu}$ and $\beta(\cdot;\cdot,\cdot)$ is the regularized incomplete beta function; see e.g. Abramowitz and Stegun (1972). For $x_t < 0$, the expression of $F_X(x_t, \mathcal{F}_{t-1}, \lambda_t)$ can be recovered by symmetry. The probability of observing a zero return, given by the mass of the censored values of x_t, can easily be computed as:

$$F_X(c; \mathcal{F}_{t-1}, \lambda_t) - F_X(-c; \mathcal{F}_{t-1}, \lambda_t) = \beta(b_{c,t}; 1/2, \nu/2) \tag{7}$$

where $b_{c,t} = \frac{c^2/(\nu e^{2\lambda_t})}{c^2/(\nu e^{2\lambda_t})+1}$. The conditional variance of the observed return, $\mathrm{Var}[y_t|\mathcal{F}_{t-1}, \lambda_t]$, is characterized by the following proposition, proved in Appendix A:

Proposition 2.1. *Let $p_X(x_t; \mathcal{F}_{t-1}, \lambda_t)$ be the Student-t density in Equation (5) and assume $\nu > 2$. Then, the conditional variance of the observed return exists and is given by:*

$$\begin{aligned} Var[y_t|\mathcal{F}_{t-1}, \lambda_t] = &\frac{\nu}{(\nu-2)} e^{2\lambda_t} - \frac{\nu e^{2\lambda_t} B(b_{c,t}; 1+1/2, \nu/2-1)}{B(1/2, \nu/2)} + c^2[1-\beta(b_{c,t}; 1/2, \nu/2)] \\ &- 2\frac{c\sqrt{\nu}\sigma}{B(1/2,\nu/2)} \frac{\left(1+\frac{c^2}{\nu e^{2\lambda_t}}\right)^{-\frac{(\nu-1)}{2}}}{\frac{(\nu-1)}{2}}, \end{aligned} \tag{8}$$

where $B(\cdot;\cdot,\cdot)$ is the incomplete beta function and $B(\cdot,\cdot)$ is the beta function. Moreover, $\lim_{c\to+\infty} Var[y_t|\mathcal{F}_{t-1}, \sigma_t] = 0$.

Figure (2) shows $\mathrm{Var}[y_t|\mathcal{F}_{t-1}, \lambda_t]$ as a function of the threshold parameter c. In this example, the number of degrees of freedom of the Student-t density is set as $\nu = 3$, whereas the value of λ_t is chosen in such a way that the efficient price change has a variance equal to one, i.e. $\mathrm{Var}[x_t|\mathcal{F}_{t-1}, \lambda_t] = \frac{\nu}{\nu-2} e^{2\lambda_t} = 1$. As we can see, in absence of transaction costs, the conditional variance of the observed return coincides with $\mathrm{Var}[x_t|\mathcal{F}_{t-1}, \lambda_t]$. If instead c is nonzero, the observed return is subject to a smoothing effect generated by market frictions, with a conditional variance becoming substantially smaller than $\mathrm{Var}[x_t|\mathcal{F}_{t-1}, \lambda_t]$ for large values of c.

For $y_t \neq 0$, the α-th conditional quantile of y_t, denoted by $q_{t,\alpha}$, is obtained from the α-th conditional quantile of x_t by shifting the latter of an amount c, that is:

$$q_{t,\alpha} = \begin{cases} F_X^{-1}(\alpha; \mathcal{F}_{t-1}, \lambda_t) - c, & \text{for } y_t > 0 \\ F_X^{-1}(\alpha; \mathcal{F}_{t-1}, \lambda_t) + c, & \text{for } y_t < 0, \end{cases} \tag{9}$$

where $F_X^{-1}(\alpha; \mathcal{F}_{t-1}, \lambda_t)$ can be computed using the inverse of the regularized incomplete beta function. We use this expression to compute the Value-at-Risk in Sections (3) and (4).

To implement the law of motion in Equation (2), we need to compute the score of $p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t)$ with respect to λ_t. In Appendix B, we prove the following result:

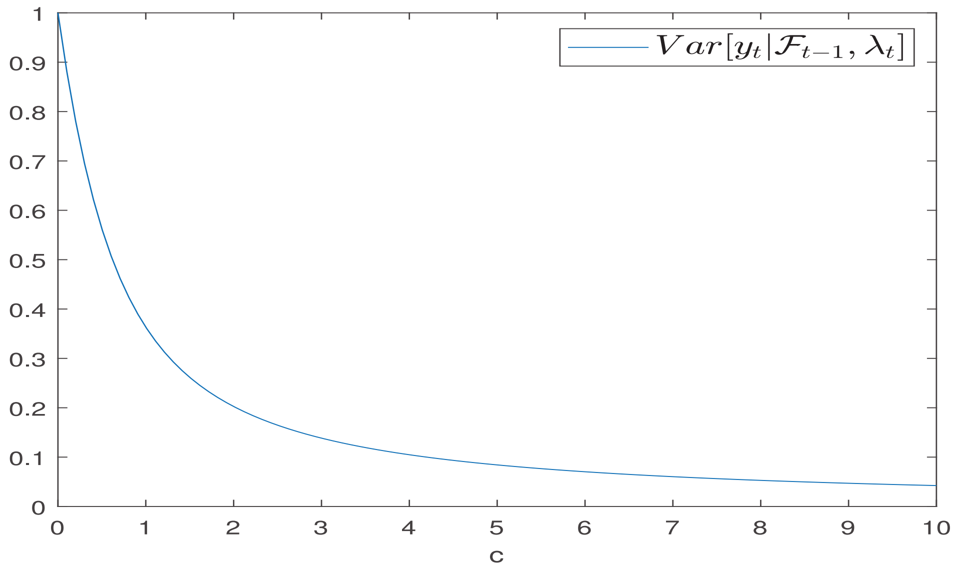

Figure 2: Conditional variance of the observed return y_t as a function of the threshold parameter. The curve is drawn using the closed-form expression of $Var[y_t|\mathcal{F}_{t-1}, \lambda_t]$ reported in Proposition 2.1. The number of degrees of freedom of the Student-t density is set as $\nu = 3$. The scale parameter λ_t is instead chosen in such a way that the efficient price change x_t has conditional variance equal to one.

Proposition 2.2. *Let $p_X(x_t; \mathcal{F}_{t-1}, \lambda_t)$ be the Student-t density in Eq.*(5). *Then, the score defined in Equation* (3) *is given by:*

$$u_t = (1 - I_{y_t=0})[(\nu + 1)b_t - 1] - I_{y_t=0}\left[\frac{2b_{c,t}^{1/2}(1 - b_{c,t})^{\nu/2}}{\beta(b_{c,t}; 1/2, \nu/2)B(1/2, \nu/2)}\right], \tag{10}$$

where $b_t = \frac{z_t}{z_t+1}$, with $z_t = x_t^2/(\nu e^{2\lambda_t})$.

The proof follows Harvey and Ito (2020), who perform a similar computation in the case where $p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t)$ is obtained by shifting to the left a generalized beta distribution of second kind (GB2) and censoring negative values. If $c = 0$, $p_Y(\cdot; \mathcal{F}_{t-1}, \lambda_t)$ reduces to $p_X(\cdot; \mathcal{F}_{t-1}, \lambda_t)$ and the score becomes:

$$u_t = (\nu + 1)b_t - 1. \tag{11}$$

In absence of trading costs, the proposed specification reduces to a dynamic scale model driven by the score of the Student-t density, known as t-GAS or Beta-t-EGARCH[2]; see Creal et al. (2013) and Harvey (2013). The non-linear update in Equation (11) is robust to fat-tails since it downweights volatility estimates when extreme price movements occur. For this reason, it has been extensively applied in the finance and economic literature; see also Creal et al. (2011),

[2] Hereafter, we will use the nomenclature Beta-t-EGARCH to denote the Student-t model based on the update in Equation (11).

Linton and Wu (2020). The score of a zero return in the Beta-t-EGARCH reduces to $u_t = -1$. As shown in the numerical examples in Section (3), the trivial update $u_t = -1$ significantly underrates volatility estimates in the presence of zero returns.

If instead c is non-zero, the score of zero returns depends on the value of λ_t. Figure (3) shows on the left panel the second term in Equation (10) as a function of the scale parameter $\sigma_t = e^{\lambda_t}$. We note that the score of zero returns ranges between -1 and zero, approaching -1 when volatility is large and zero when volatility is small. As can be seen from the right panel of Figure (3), which shows the probability of zero returns as a function of the scale parameter, large volatility leads to a very small probability of zero returns. In this circumstance, the discrete-continuous mixture behaves similarly to the Beta-t-EGARCH, and thus the two models feature the same score. On the other hand, updating zero returns by setting the score equal to zero can be regarded as treating them as missing values. The "setting to zero" method is inspired by an analogy with the missing value problem in linear-Gaussian models, and it is commonly employed in score-driven models; see e.g. Lucas et al. (2016), Buccheri et al. (2020). Recently, Blasques et al. (2020) have shown that this method leads to inconsistent maximum likelihood estimates. The score of zero returns in Equation (10) is always in-between the Beta-t-EGARCH weight ($u_t = -1$) and the "setting to zero" weight ($u_t = 0$). These three update schemes are compared in the numerical examples in Section (3).

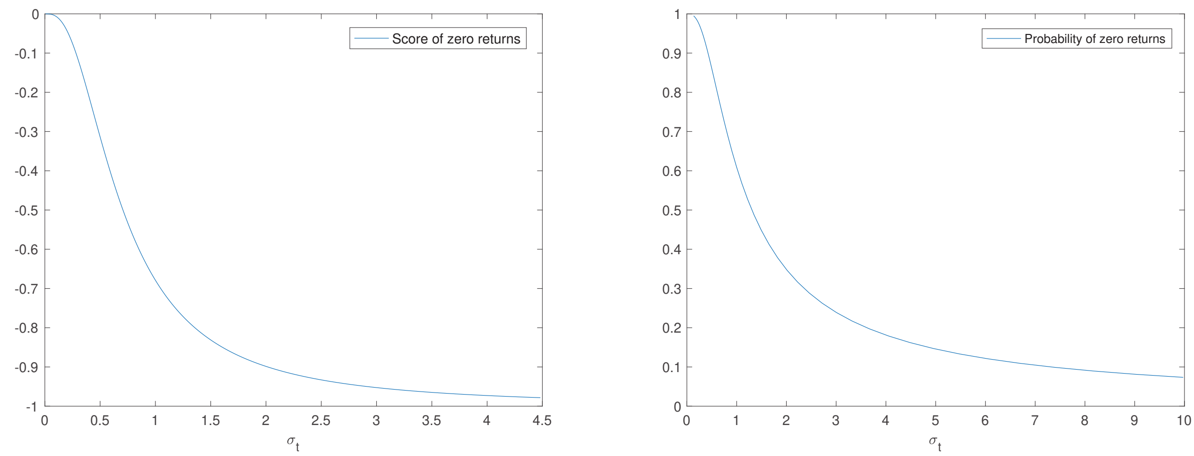

Figure 3: Left: score of zero returns, given by the second term in Equation (10)*, as a function of the scale parameter* σ_t*. Right: probability of zero returns, computed as* $F_X(c; \mathcal{F}_{t-1}, \lambda_t) - F_X(-c; \mathcal{F}_{t-1}, \lambda_t)$*, as a function of the scale parameter* σ_t*. In both cases, the conditional density is a Student-t with* $\nu = 3$ *degrees of freedom. Moreover, we set* $c = 1$*.*

In general, under regularity conditions, the score has conditional expectation equal to zero, and it is therefore a martingale difference sequence. This result motivates the use of the score as an innovation process in the update of volatility. The fact that, in our model, $p_Y(y_t; \mathcal{F}_{t-1}, \lambda_t)$

has a discontinuity in zero does not compromise this general property of the score. Indeed, in Appendix C, we prove the following:

Proposition 2.3. *Let* $p_X(x_t; \mathcal{F}_{t-1}, \lambda_t)$ *be the Student-t density in Equation* (5). *Then, the score defined in Equation* (3) *has zero conditional expectation, i.e.* $\mathrm{E}[u_t|\mathcal{F}_{t-1}, \lambda_t] = 0$.

Even in this case, the proof follows Harvey and Ito (2020), who prove a similar result for a GB2 conditional density.

3 Monte-Carlo analysis

When the probabilistic law describing financial returns admits a continuous density, the likelihood of observing a zero return is virtually zero. This is the case in GARCH-type models, where the conditional density is normal or Student-t, and even in score-driven models, which are based on Student-t and other continuous non-normal densities; see Creal et al. (2013) and Harvey (2013). As underlined above, the assumption of a continuous density conflicts with the significant amount of zero returns observed in real financial markets. In this section, we aim to study the consequences of misspecifying the price generation process when returns have a non-zero likelihood of being equal to zero. We focus here on three different aspects that have a relevant role in our understanding of the volatility process: the maximum likelihood estimates of the model parameters, the filtered volatility, and the Value-at-Risk (VaR) forecasts.

3.1 Maximum likelihood estimates

Standard volatility models are characterized by a set of static parameters which govern the dynamics of volatility. We question whether their maximum likelihood estimates are biased in the presence of market frictions. Intuitively, the answer to this question is positive, because the effect of zero returns is to smooth out the observed prices, which thus appear less volatile.

We simulate the information signal x_t as a martingale difference sequence with volatility σ_t generated by a Beta-t-EGARCH, which is obtained by setting $c = 0$ in our discrete-continuous mixture model. As such, the conditional density of efficient returns is a Student-t and the update of volatility is the same as in Equation (2), with the score u_t given by Equation (11). The static parameters are set as follows: $\omega = 0.01$, $\kappa = 0.05$, $\phi = 0.95$, $\nu = 3$. The observed price returns y_t are generated from the simulated returns x_t using the microstructure mechanism given in Equation (1), with $c = \{0.00, 0.25, 0.40\}$.

We simulate $N = 1000$ time-series of $n = 1000$ observations of the return processes y_t. The discrete-continuous mixture model is estimated on each of the simulated time-series, obtaining $N = 1000$ maximum likelihood estimates of the parameters ω, κ, ϕ, ν. To assess the effect of misspecification, the parameters are also estimated using the Beta-t-EGARCH model. Zero returns are handled in the Beta-t-EGARCH using two alternative methods. The first trivially sets $y_t = 0$, which leads to $u_t = -1$ in Equation (11). The second treats zero returns as missing observations using the "setting to zero" method of Lucas et al. (2016). The method is implemented by setting $u_t = 0$ whenever a zero return occurs. To distinguish between these two different methods, we denote by Beta-t-EGARCH(0) the implementation of the "setting to zero" method of Lucas et al. (2016).

Figure (4) shows kernel density estimates of the model parameters obtained using the discrete-continuous mixture model and the two different implementations of the Beta-t-EGARCH. In absence of transaction costs ($c = 0.00$), all the models lead to identical maximum likelihood estimates, centred around the true values. As c increases, the two Beta-t-EGARCH implementations provide significantly biased estimates. In particular, the parameters ω, ν are downward biased, κ is upward biased, whereas the estimates of ϕ are still centred around the true value. As expected, the unconditional mean of the volatility process, given by $\omega/(1-\phi)$, is underestimated due to the smoothing caused by zero returns. The fact that even the "setting to zero" method is biased agrees with the result of Blasques et al. (2020), who show that this method may lead to inconsistent maximum likelihood estimates. The strong downward bias observed on ν is a consequence of the probability mass in zero, which makes the density of returns extremely leptokurtic. A similar effect on real data is found e.g. by Hasbrouck (1999). As shown in Subsection (3.2), the bias on ν has relevant implications on the filtered volatility and VaR forecasts.

Contrary to the Beta-t-EGARCH, the discrete-continuous mixture is not affected by zero returns, remaining unbiased when trading costs come into play. This result suggests that the proposed specification is to be preferred to the Beta-t-EGARCH as a tool to extract volatility when immediate trading is not allowed due to the presence of transaction costs.

3.2 Filtered volatility and Value-at-Risk

We now investigate the effect of infrequent trading on the filtered volatility estimates. While static parameters are comparable only across nested models, filtered volatilities can be compared across completely different models. We thus include in the set of benchmark models the GARCH model of Bollerslev (1986), which is extremely popular in financial applications. Of course, to perform a fair comparison among the discrete-continuous mixture, the Beta-t-EGARCH and the GARCH, the information signal must be generated by a different volatility model. A well consolidated literature dating back to the seminal papers of Nelson (1992), Nelson and Foster (1994) and Nelson and Foster (1995) assumes that financial returns are generated by a stochastic volatility model, and that one-step-ahead predictable models, such as GARCH-type

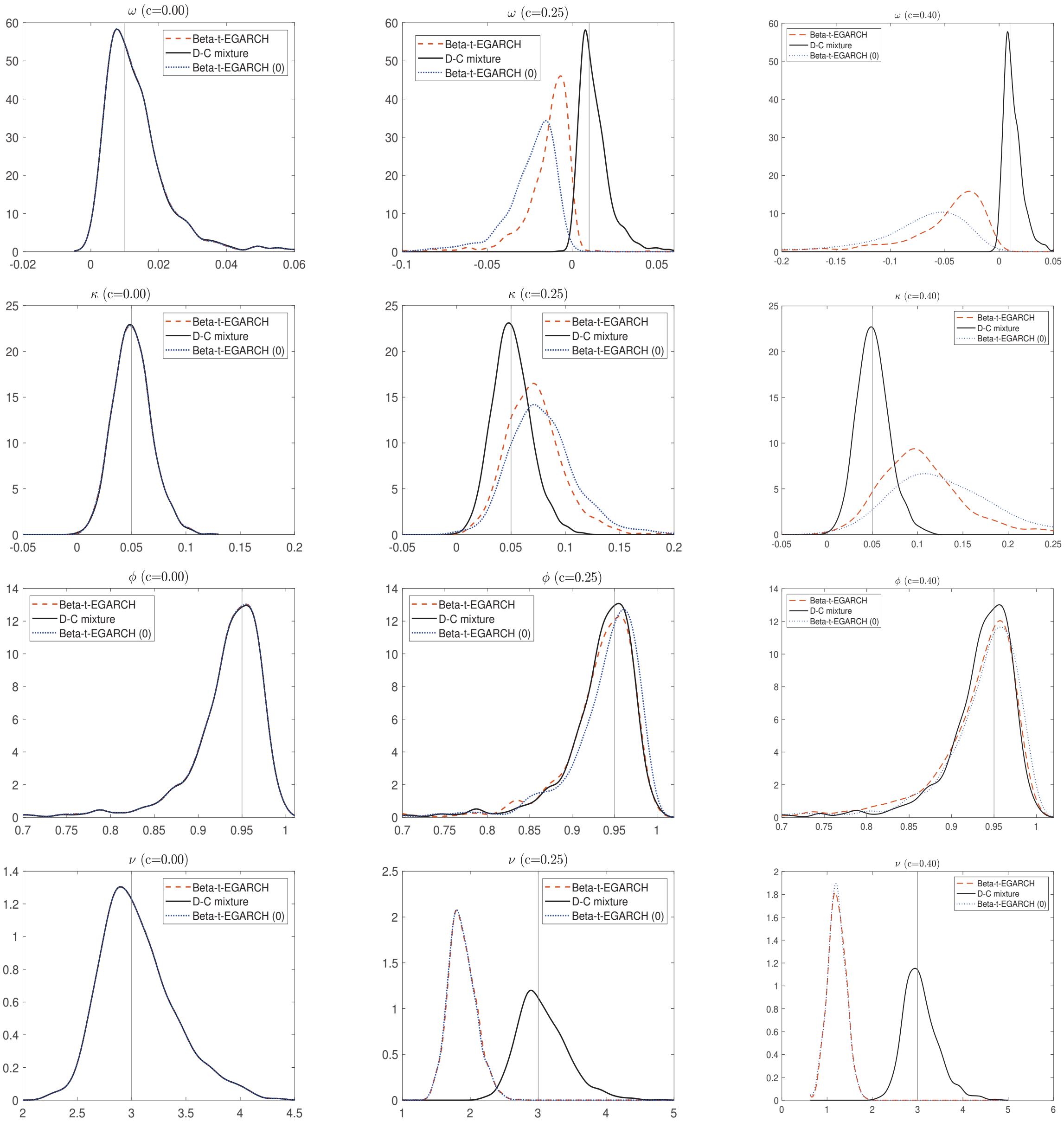

Figure 4: Kernel density estimates of parameters ω, κ, ϕ, ν based on $N = 1000$ simulated time-series of $n = 1000$ observations. The threshold parameter c is set equal to $c = 0.00$ (left column), $c = 0.25$ (middle column), $c = 0.40$ (right column). The models used to estimate the parameters are the Beta-t-EGARCH (red dashed line), the discrete-continuous mixture model described in Section (2) (black solid line) and the Beta-t-EGARCH implemented based on the "setting to zero" method of Lucas et al. (2016) (blue dotted line). Vertical lines denote the true values of the parameters.

models, are approximations to the true filtering density. We adopt this view here, and assume that the information signal is generated as follows:

$$\begin{aligned} x_t &= e^{\alpha_t}\varepsilon_t, & \varepsilon_t &\sim t_\nu, \\ \alpha_t &= \mu + \varphi(\alpha_{t-1} - \mu) + \eta_t, & \eta_t &\sim \mathrm{N}(0, \sigma_\eta^2), \end{aligned} \tag{12}$$

where ε_t, η_t are i.i.d. innovations, independent from each others. Note that we allow for a fat-tailed returns distribution by modeling the standardized residuals ε_t through a Student-t density with ν degrees of freedom. The parameters of the volatility process are close to the values that are typically estimated on financial time-series: $\mu = 0.5$, $\varphi = 0.98$, $\sigma_\eta = 0.016$; see Sandmann and Koopman (1998). As in the previous experiment, the observed returns y_t are generated from the efficient price changes using the microstructure mechanism in Equation (1), with $c = \{0.00,\ 0.50,\ 1.00\}$.

We simulate $N = 1000$ time-series of $n = 2000$ observations of the return process y_t. For each simulated time-series, the discrete-continuous mixture, the Beta-t-EGARCH and the GARCH are estimated on the first $n_{\text{in}} = 1000$ observations, whereas the out-of-sample root mean-square error (RMSE) is evaluated on the remaining sub-sample of $n_{\text{out}} = 1000$ observations. Specifically, the RMSE is computed as:

$$\mathrm{RMSE}^{(i)} = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n}\left(\hat{\sigma}_t^{(i)} - \sigma_t^{(i)}\right)^2}, \quad i = 1, \ldots, N,$$

where $\hat{\sigma}_t^{(i)}$ is the scale parameter estimated by each benchmark model, whereas the target $\sigma_t^{(i)}$ is computed as $\sigma_t^{(i)} = e^{\alpha_t^{(i)}}$, where $\alpha_t^{(i)}$ is the path of the state variable in the i-th simulation. To deal with zero returns in Beta-t-EGARCH and GARCH, we adopt the two methods described in the previous experiment. In the first method we simply set $y_t = 0$, whereas in the second we set the score equal to zero[3].

Table (1) shows the average RMSE for increasing values of c and ν. The scenario with $\nu = \infty$ is obtained by generating the residuals ε_t in Equation (12) by means of a standard normal distribution. In this scenario, the return distribution is not fat-tailed, and therefore the relative differences among models are only imputable to market frictions. We first note that the RMSE of the discrete-continuous mixture is significantly smaller compared to the Beta-t-EGARCH. The relative difference between the two models increases as c increases. Large values of c entail less frequent price updates and thus a large incidence of zero returns. In these circumstances, the Beta-t-GARCH suffers from the smoothing caused by zero returns, underrating significantly the volatility of the information signal. For $c = 1.00$, we do not report the RMSE of the Beta-t-EGARCH because its volatility estimates collapse to zero due to the strong level of misspecification.

[3]The GARCH has an ARMA representation in terms of the volatility innovations $\eta_t = y_t^2 - \sigma_t^2$. The latter coincide with the score of the normal density up to a normalization.

c	Beta-t-E.	Beta-t-E.(0)	D-C mixture	GARCH	GARCH(0)
			$\nu = 3$		
0.00	1.002	1.002	1.000	1.367	1.350
0.50	3.477	3.520	1.000	1.516	1.483
1.00	–	–	1.000	1.721	1.168
			$\nu = 9$		
0.00	1.002	1.002	1.000	1.037	1.039
0.50	3.477	3.520	1.000	1.368	1.399
1.00	–	–	1.000	1.946	1.980
			$\nu = \infty$		
0.00	1.000	1.000	1.000	1.003	0.987
0.50	3.675	3.717	1.000	1.470	1.510
1.00	–	–	1.000	2.240	2.287

Table 1: Relative RMSE of Beta-t-EGARCH, Beta-t-EGARCH(0), the discrete-continuous mixture model described in Section (2), GARCH and GARCH(0). The two models denoted by Beta-t-EGARCH(0) and GARCH(0) are estimated using the "setting to zero" method of Lucas et al. (2016). Each RMSE measure is normalized by the RMSE of the D-C mixture. The results are reported for increasing values of the degrees of freedom parameter, $\nu = \{3, 6, \infty\}$, and increasing values of the threshold parameter, $c = \{0.00, 0.50, 1.00\}$. For $c = 1.00$, the results of the Beta-t-EGARCH are not reported because its volatility estimates collapse to zero.

The GARCH is affected by two different forms of misspecification. For low ν, its volatilities are too sensitive to extreme price movements. They exhibit a sharp increase when a jump in prices occurs, followed by a slow decay toward the unconditional mean level; see e.g. Harvey (2013). On the other hand, for large c, GARCH estimates are affected by the smoothing of zero returns. It is interesting to note that, compared to the Beta-t-EGARCH, this effect is significantly less severe in the GARCH, which tends to perform better as c increases, even for low values of ν. Although this may seem counterintuitive, the relative performance of GARCH and Beta-t-EGARCH is clear in light of the results obtained in the previous experiment. Lack of trading leads to a probability mass in zero which increases the kurtosis of the returns distribution. When estimating the Beta-t-EGARCH, the higher kurtosis determines a downward bias in the degrees of freedom parameter, which in turn heavily downweights volatility estimates for large c. In other words, in the Beta-t-EGARCH, the misspecification associated with zero returns translates into a lower degrees of freedom parameter, but this deeply affects the estimation of volatility. A similar mechanism is absent in the GARCH, which thus provides less biased volatility estimates for large c. We conclude that, as a consequence of the smoothing caused by zero returns, the volatility of the information signal is downward biased if estimated using standard methodologies.

We now investigate whether the risk profile of observed returns is biased as well. To study in detail this problem, we assess the out-of-sample VaR of the previously simulated returns. We compute the 90%, 95%, 99% VaR using the one-step-ahead volatility forecasts of the continuous-

c	Beta-t-E.	Beta-t-E.(0)	D-C mixture	GARCH	GARCH(0)
			$\text{VaR}_{90\%}$		
			$\nu = 3$		
0.00	0.8996	0.8996	0.8996	0.9404	0.9404
0.50	0.8549	0.8574	0.8986	0.9437	0.9437
1.00	–	–	0.8999	0.9515	0.9515
			$\nu = 9$		
0.00	0.9001	0.9001	0.9001	0.9147	0.9147
0.50	0.8517	0.8538	0.8991	0.9195	0.9195
1.00	–	–	0.8978	0.9334	0.9334
			$\nu = \infty$		
0.00	0.9005	0.9005	0.9005	0.9062	0.9062
0.50	0.8507	0.8531	0.9002	0.9117	0.9117
1.00	–	–	0.8915	0.9274	0.9274
			$\text{VaR}_{95\%}$		
			$\nu = 3$		
0.00	0.9485	0.9485	0.9485	0.9637	0.9637
0.50	0.9353	0.9391	0.9487	0.9641	0.9641
1.00	–	–	0.9490	0.9665	0.9665
			$\nu = 9$		
0.00	0.9493	0.9493	0.9493	0.9534	0.9534
0.50	0.9372	0.9418	0.9494	0.9498	0.9498
1.00	–	–	0.9490	0.9526	0.9526
			$\nu = \infty$		
0.00	0.9499	0.9498	0.9499	0.9512	0.9512
0.50	0.9346	0.9339	0.9492	0.9446	0.9446
1.00	–	–	0.9493	0.9484	0.9484
			$\text{VaR}_{99\%}$		
			$\nu = 3$		
0.00	0.9899	0.9899	0.9899	0.9833	0.9838
0.50	0.9972	0.9980	0.9895	0.9820	0.9820
1.00	–	–	0.9897	0.9816	0.9816
			$\nu = 9$		
0.00	0.9895	0.9895	0.9895	0.9873	0.9873
0.50	0.9972	0.9990	0.9893	0.9787	0.9787
1.00	–	–	0.9893	0.9755	0.9755
			$\nu = \infty$		
0.00	0.9894	0.9894	0.9894	0.9833	0.9833
0.50	0.9966	0.9990	0.9894	0.9790	0.9790
1.00	–	–	0.9894	0.9739	0.9739

Table 2: Average coverage at 90%, 95%, 99% nominal confidence levels of Beta-t-EGARCH, Beta-t-EGARCH(0), the discrete-continuous mixture model described in Section (2), GARCH and GARCH(0). The two models denoted by Beta-t-EGARCH(0) and GARCH(0) are estimated using the "setting to zero" method of Lucas et al. (2016). The results are reported for increasing values of the degrees of freedom parameter, $\nu = \{3, 9, \infty\}$, and increasing values of the threshold parameter, $c = \{0.00,\ 0.50,\ 1.00\}$. For $c = 1.00$, the results of the Beta-t-EGARCH are not reported because its volatility estimates collapse to zero.

discrete mixture, Beta-t-EGARCH and GARCH in the second sub-sample of 1000 observations. The coverage rate, defined as the fraction of times in which the observed returns do not exceed the VaR, is reported in Table (2) for each scenario. The discrete-continuous mixture provides coverage rates that perfectly match the nominal confidence levels. At 90% confidence level, the Beta-t-EGARCH heavily underestimates the risk as c increases, as can be seen from the small coverage rates provided by this model. In light of the results recovered in previous analyses, this effect is imputable to the underestimation of volatility. As the VaR confidence level increases, for nonzero c, the coverage rates of the Beta-t-EGARCH are slightly below the nominal confidence level (at 95%) and slightly above (at 99%). This particular behavior is explained by an additional effect due to the misspecification of zero returns. A downward biased degrees of freedom parameter leads to a Student-t conditional density with fatter tails, which tends to overestimate the quantiles corresponding to large confidence levels. The results of the Beta-t-EGARCH thus reflect the combination between these two effects, both imputable to market frictions.

The performance of the GARCH is additionally affected by the presence of fat-tails. To isolate the effect of market frictions, we concentrate on the scenarios with $\nu = \infty$. The nominal confidence level is overestimated by the GARCH coverage rate at 90%, whereas it is underestimated at 99%. At 95%, the coverage rate is close to the nominal confidence level. This result is a consequence of misspecifying the conditional density. As underlined above, the presence of a probability mass in zero increases the kurtosis of the return distribution, which thus appears to be extremely leptokurtic. As a consequence, a normal density will overestimate the quantiles corresponding to small confidence levels and, similarly, it will underestimate the quantiles corresponding to large confidence levels.

To summarize, zero returns deeply affect the estimation of volatility and VaR in the Beta-t-EGARCH. The GARCH behaves better in terms of VaR assessment, but it provides biased estimates of risk when fat-tails come into play. Moreover, in absence of fat-tails, GARCH VaR forecasts at 90% and 99% are affected by the presence of zero returns. The discrete-continuous mixture overcomes the difficulties of both models. It is robust to fat-tails and, at the same time, it explicitly accounts for the probability mass in zero in both the return distribution and the update of volatility.

4 Empirical analysis

4.1 Data

The dataset used in the empirical analysis is provided by Thomson Reuters and includes trade and quote data of 250 stocks listed in the New York Stock Exchange (NYSE). Data is available in the reference period from January 3, 2006, to December 31, 2014, a total of 2265 trading days. The 250 stocks have been selected from the data provider among the most frequently traded NYSE assets in the reference period. We limit our attention to those assets that have been traded each day, reducing the total number of stocks to 190. Data is timestamped at a frequency of one millisecond, however, in our analysis we sample transaction prices at a frequency of 30 seconds, using the procedures described by Barndorff-Nielsen et al. (2009). Quote, data is employed to compute bid-ask spreads, which will be used in the subsequent analyses. Associated with each transaction, the dataset provides the number of traded shares, from which we compute the daily trading volume.

We divide the 190 stocks into five groups of equal size by computing the 5-quantiles of the average trading volumes. For each group, Table (3) shows the average fraction of 30 seconds zero returns, the average bid-ask spreads, and the average number of transactions. The results in the Table show that the fraction of zero returns and bid-ask spreads increase as volume decreases, whereas the number of transactions decreases as volume decreases. These relationships among microstructure variables are well-known in the finance literature; see among others Goyenko et al. (2009).

Group	Volume ($\$10^6$)	Zero returns (%)	Bid-ask spread (%)	N. of transactions
1	219.250	29.11	0.047	8468
2	95.816	32.53	0.060	5835
3	60.415	35.64	0.063	4359
4	43.803	40.10	0.076	3672
5	25.604	48.59	0.110	2844
Ticker		Max/min		
XOM	598.640	18.47	0.024	15641
ALK	11.613	51.01	0.160	1494

Table 3: The 190 NYSE stocks are sorted based on their average daily volumes. Five groups of equal size are selected by taking the 5*-quantiles of the sorted list. For each group, we show the average daily volume, the average fraction of zero returns, the average bid-ask spread, and the average number of transactions. The average fraction of zero returns are computed from transactions sampled at 30-seconds. The bid-ask spread refers to the percentage bid-ask, computed as* $(\mathrm{ASK} - \mathrm{BID})/\mathrm{ASK}$*; see Lo et al. (2003).*

As underlined above, the stocks included in the dataset are frequently traded NYSE assets. On 30 seconds returns, the average fraction of zero returns spans from 29.11% for the most liquid assets to 48.50% for the least liquid. However, when sampling returns on a daily frequency, the number of zero returns tends to vanish for all the stocks included in the dataset. For this reason, to assess the impact of illiquidity on daily Value-at-Risk forecasts, we consider an illiquid stock not included in the dataset. We perform our analysis on Glen Burnie Bancorp (GLBZ), a small cap stock whose daily closing prices are publicly available online.

4.2 Intraday patterns and parameter estimates

Volatility has an asymmetric U-shape pattern that is high at the beginning of the trading day, low in the middle of the day and slightly higher during the last few minutes of trades; see e.g. Hasbrouck (1999). When estimating our microstructure model on intraday data, it is necessary to allow for further flexibility in the parameter space to account for the deterministic pattern of volatility. It is a common practice in the financial econometric literature to model the U-shaped volatility pattern as a sum of two exponential functions; see among others Hasbrouck (1999) and Andersen et al. (2012). We follow this strategy here and write the volatility of the information signal as the product of a deterministic and a stochastic component:

$$\sigma_t = s_t e^{\lambda_t}. \tag{13}$$

The process λ_t obeys the law of motion in Equation (2). The shape s_t is a deterministic function of time which we write as:

$$s_t = C + Ae^{-\alpha t} + Be^{-\beta(1-t)}, \tag{14}$$

where, for identification, α, β, A, B, C are normalized so that $\int_0^1 s_t^2 dt = 1$. These parameters can be estimated by maximum likelihood together with the parameters governing the dynamics of λ_t.

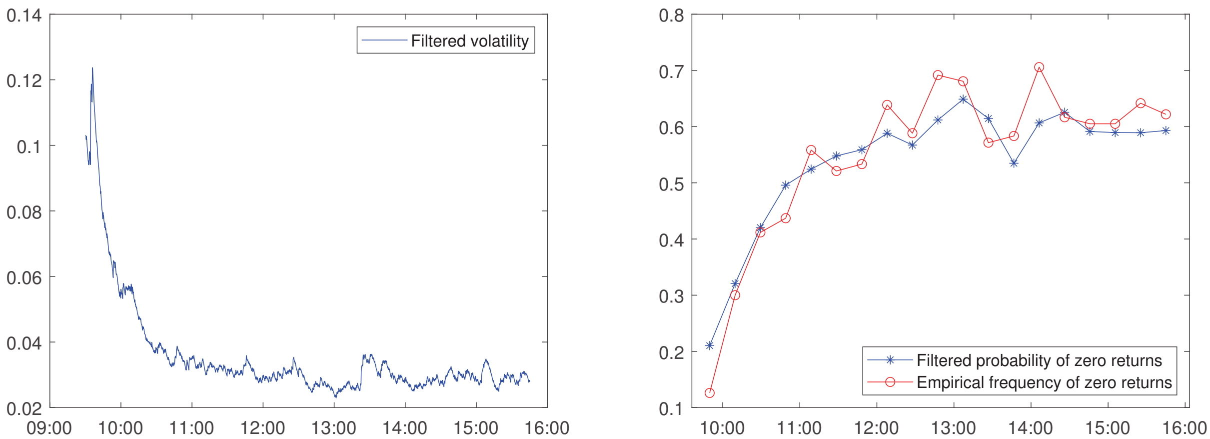

Figure 5: Left: percentage filtered volatility of JPM. Right: filtered probability of zero returns averaged over blocks of 20 minutes (blue-asterisk line) and empirical frequency of zero returns over the same blocks (red-circle line). To obtain less noisy empirical frequencies of zero returns in each block, we use data sampled at 10-seconds instead of 30-seconds.

Since in our microstructure model the probability of observing a zero return depends on volatility, the deterministic component s_t gives rise to an analogous pattern in trading activity. The larger volatility observed during the opening and closing hours determines a lower zero return probability, and thus a higher trading activity. Similarly, the smaller volatility observed in the central hours determines a higher zero return probability, and thus a lower trading activity. This pattern in trading activity is close to the one observed in real financial markets. Figure (5) shows the intraday filtered volatility of JPMorgarn Chase (JPM) on July 10, 2014. It also shows the probabilities of zero returns implied by the discrete-continuous mixture model. The latter are given by the mass of the censored observations, as indicated in Equation (4). These probabilities are averaged over blocks of 20 minutes, and are compared to the empirical frequencies of zero returns evaluated in the same blocks. The filtered probabilities closely match the empirical frequencies and both exhibit an inverse U-shaped pattern. Patterns of this kind are observed in many markets and can be viewed as a stylized property of high-frequency financial data.

For each stock, we re-estimate the discrete-continuous mixture on each day between January 3, 2006, and December 31, 2014. The intraday time-series are made up of 30 seconds percentage returns sampled from 9:30 to 15:45. Table (4) shows, in the case of four specific assets, the average of the maximum likelihood estimates over the entire sample. The four assets are JPM and ExxonMobil (XOM), which are liquid stocks belonging to Group 1 in Table (3), Alaska Air Group (ALK) and Advanced Micro Devices (AMD), which are illiquid stocks belonging to Group 5 in Table (3).

We first note that, as expected, the estimated threshold parameter c is larger in the case of the two illiquid stocks. For the two liquid stocks, the estimates of c are smaller but highly statistically significant. This implies that the time-series behaviour of these assets is significantly affected by transaction costs, even if they are extremely liquid. We also note that the estimated degrees of freedom parameter ν is large on average, though being characterized by a high variance. The fact that ν is generally large is imputable to the robustness of the discrete-continuous mixture to zero returns. As seen in the Monte-Carlo analysis of Section (3.1), in the Beta-t-EGARCH the excess kurtosis induced by zero returns leads to a downward bias on ν. In the discrete-continuous mixture, the downward bias observed in the Beta-t-EGARCH tends to vanish, and the estimates of ν are only affected by the fat-tail behavior of the return distribution. e.g. by jumps in prices. Indeed, the high variance of this parameter indicates that there are days where ν is significantly lower than average. These days are characterized by extreme price movements, such as those occurring during the 2007-2008 global financial crisis, or are associated with flash-crashes, like the one recorded on May 6, 2010. See also the application in Section (4.4), which shows that the estimate of ν is smaller when estimating the model on daily returns.

	$\hat{\omega}$	$\hat{\phi}$	$\hat{k}$	$\hat{\nu}$	$\hat{c}$	$\hat{A}$	$\hat{B}$	$\hat{C}$	$\hat{\alpha}$	$\hat{\beta}$
JPM	-0.2279	0.9209	0.0303	188.0576	0.0182	0.5576	0.4424	0.5141	11.8964	7.4099
	(0.3227)	(0.1015)	(0.0131)	(121.9671)	(0.0048)	(0.1832)	(0.1832)	(0.1947)	(20.5297)	(17.9391)
XOM	-0.2547	0.9193	0.0319	117.4667	0.0086	0.5430	0.4570	0.4295	10.4076	7.5664
	(0.4745)	(0.1500)	(0.0209)	(120.3808)	(0.0024)	(0.2068)	(0.2068)	(0.1940)	(21.1512)	(18.2337)
ALK	-0.2870	0.8621	0.0198	116.218	0.0920	0.5813	0.4187	0.6740	6.7283	8.8473
	(0.4820)	(0.2269)	(0.0169)	(123.3303)	(0.0390)	(0.0148)	(0.0148)	(0.1862)	(10.0206)	(14.2011)
AMD	-0.3390	0.7122	0.0331	279.0177	0.3390	0.6557	0.3443	0.6610	8.1491	13.6853
	(0.3887)	(0.3031)	(0.0202)	(67.8980)	(0.2736)	(0.1855)	(0.1855)	(0.2150)	(11.6654)	(22.0706)

Table 4: The Table reports for assets JPM, XOM, ALK, AMD the average of maximum likelihood estimates recovered by re-estimating the discrete-continuous mixture model on a daily basis over the period from January 3, 2006 to December 31, 2014. The values inside the parenthesis indicate the standard deviations of the estimated parameters.

4.3 The parameter $2c$ as an estimate of transaction costs

The parameter c in the model in Equation (1) represents the difference between the return of the information signal and the ex-post return of the trader. As such, $2c$ includes all the transaction costs the trader has to pay in order to buy or sell the underlying security, e.g. bid-ask spreads, commission fees, price impact, etc. The maximum likelihood estimate of c can thus be regarded as an estimate of round-trip costs and, more generally, as a proxy for asset illiquidity.

Compared to classical illiquidity measures, such as bid-ask spreads and trading volumes, this round-trip cost measure can be extracted from data using solely transaction prices. As discussed by Bekaert et al. (2007), this feature is particularly relevant in situations where quote data or other microstructure variables are not easily accessible, for instance in emerging markets. Our illiquidity measure shares such a property with the fraction of zero returns (FZR), which is widely recognized as a proxy for asset illiquidity; see e.g. Bandi et al. (2020). Despite this similarity, the two measures capture different aspects of illiquidity. FZR is uniquely determined by the number of zero returns. A small amount of zero returns leads to a low FZR, implying high liquidity. As underlined in Section (2.1), the probability of zero returns in the microstructure model in Equation (1) is determined not only by the threshold c, but even by the volatility of the information signal. A high trading activity does not necessarily entail a small round-trip cost. If c is not small, trading might be convenient provided that volatility is sufficiently large. A similar situation can occur during periods of market distress associated with broad fire sales of financial securities. These periods are characterized by high trading activity, however, the latter is not imputable to an increase in liquidity, but rather to the large uncertainty prevailing in the market.

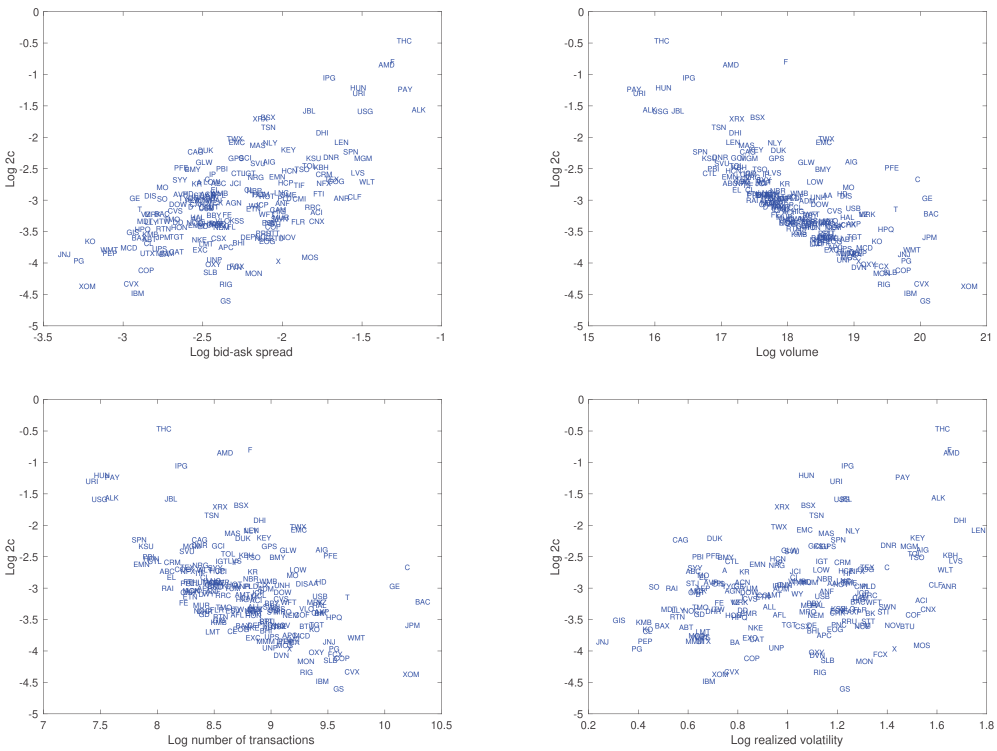

Figure 6: For each of the 190 stocks, we plot the average, over 2008, of the estimated round-trip costs against the average, over the same period, of: (a) bid-ask spreads, computed as (ASK − BID)/ASK *(top-left), (b) trading volumes (top-right), (c) number of transactions (bottom-left), (d) realized volatilities (bottom-right). All quantities are reported on a logarithmic scale.*

To highlight these differences, we perform a cross-sectional analysis in two different years, 2008 and 2014. These two years are representative of a period with high volatility (2008), and a period with low volatility (2014). In both years we compute, for each stock, the average of the daily maximum likelihood estimates of $2c$ and the average of bid-ask spreads, trading volumes, number of transactions and realized volatilities.

Figures (6) and (7) show the scatter plots of the estimated round-trip costs against the latter four quantities in 2008 and 2014, respectively. In both years, there exists a clear dependence structure between our estimate of transaction costs and the other microstructure variables. Round-trip costs tend to increase with increasing bid-ask spreads and volatilities, whereas they tend to decrease with increasing trading volumes and the number of transactions. The behaviour of our round-trip cost measure appears to be natural in light of the widely accepted view that large transaction costs are associated with large bid-ask spreads and small trading volumes. Also, the fact that assets with large volatility are characterized by a lower liquidity agrees with both the empirical literature and the market microstructure theory; see e.g. Stoll (1978b), Stoll (1978a), Amihud and Mendelson (1987), Amihud and Mendelson (1989) and

Stoll (2000).

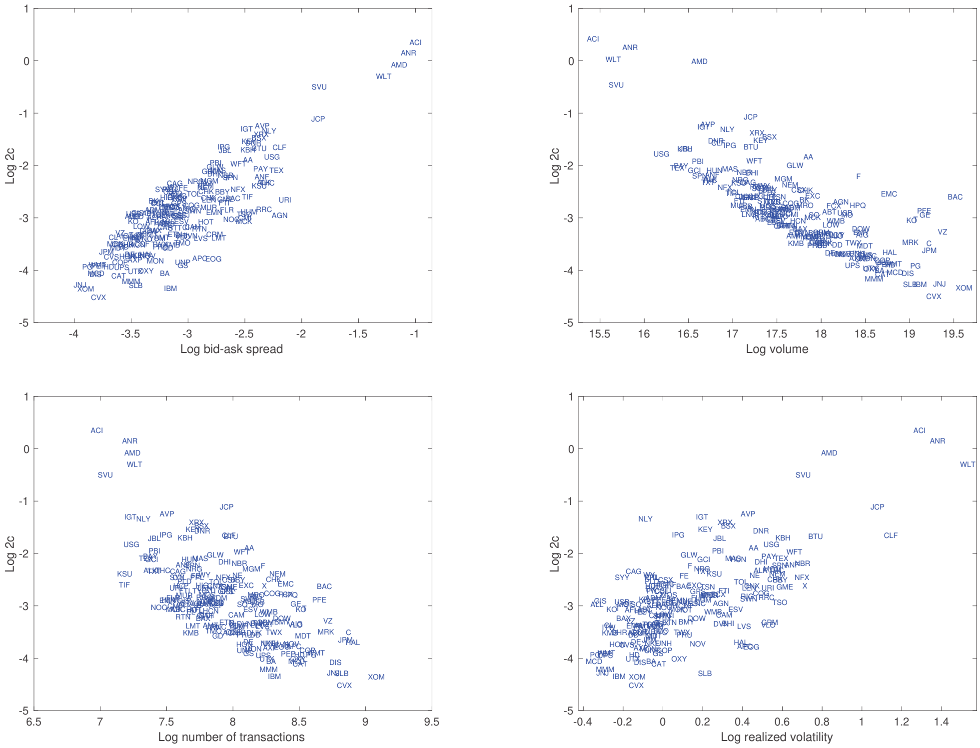

Figure 7: For each of the 190 stocks, we plot the average, over 2014, of the estimated round-trip costs against the average, over the same period, of: (a) bid-ask spreads, computed as (ASK − BID)/ASK *(top-left), (b) trading volumes (top-right), (c) number of transactions (bottom-left), (d) realized volatilities (bottom-right). All quantities are reported on a logarithmic scale.*

In the top panel of Table (5), we report the sample correlations among our estimate of round-trip costs and the other microstructure variables. In both years, correlations are large in absolute value and highly statistically significant. In the bottom panel, we report the sample correlations between the four microstructure variables and the FZR measure, computed as the average fraction of 30 seconds zero returns. While the correlations among FZR, trading volume and number of transactions are similar to the analogous correlations of $2c$, the correlations among FZR, bid-ask spreads and realized volatility are significantly smaller. In particular, in 2008, FZR and bid-ask spreads have a sample correlation of 21.83%, which is considerably smaller compared to the analogous correlation of the estimated round-trip costs. The correlation between FZR and realized volatility, in 2008, is the least significant.

To further highlight the results in Table (5), we report in Figure (8) the scatter plot of FZR against bid-ask spreads and realized volatility in 2008. Only a few very illiquid assets, such as ALK and AMD, are characterized by significantly large FZR, bid-ask spread and realized volatility. For the remaining assets, it is difficult to detect a clear dependence structure in the

	Bid-ask spread	Volume	N. of transactions	Realized volatility
	Estimated Round-trip costs ($2c$)			
2008	0.6095	-0.7638	-0.5483	0.3368
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
2014	0.8409	-0.8061	-0.6755	0.6967
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
	FZR			
2008	0.2183	-0.7008	-0.5891	-0.1628
	(0.0025)	(0.0000)	(0.0000)	(0.0248)
2014	0.6214	-0.7293	-0.7311	0.3400
	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Table 5: In the top panel, we report the cross-sectional correlation among the average round-trip cost estimates and the average of bid-ask spreads, trading volumes, number of transactions and realized volatilities in 2008 and 2014. In the bottom panel, we report the cross-sectional correlation among the fraction of zero returns (FZR) measure and the latter four quantities. We show inside the parenthesis the p-values associated to each correlation coefficient.

scatter plot.

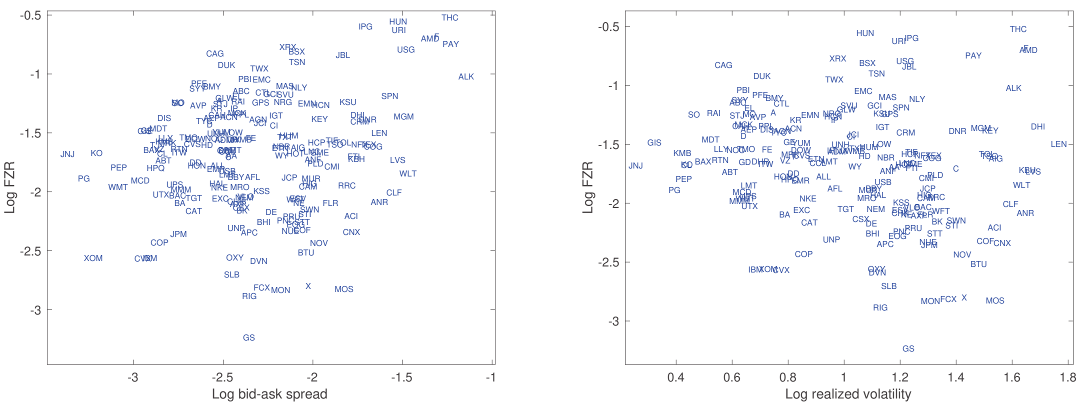

Figure 8: For each of the 190 stocks, we plot the average, over 2008, of the FZR measure against the average, over the same period, of: (a) bid-ask spreads, computed as $(\mathrm{ASK}-\mathrm{BID})/\mathrm{ASK}$ *(left) and (b) realized volatilities (right). All quantities are reported on a logarithmic scale.*

To shed light on this result, we plot in Figure (9) the daily bid-ask spread and the daily number of transactions, both obtained by averaging these quantities over the 190 stocks. The average number of transactions increases during the period of the global financial crisis. The higher trading activity is not associated with an increase in liquidity, as shown by the fact that bid-ask spreads tend to spike in 2008. It is rather imputable to the selling pressure induced by the overall market sentiment. Since the increase in trading activity naturally leads to a smaller FZR, liquidity appears to be spuriously large when assessed through FZR. As a consequence, the positive correlation between FZR and bid-ask spreads found in other periods tends to vanish here. Even the correlation between FZR and volatility disappears because volatility spikes in this sub-sample whereas FZR does not.

The round-trip cost estimate behaves differently. From the scatter plots in Figure (6) and from the sample correlations reported in Table (5), we still find evidence of a significant dependence structure between $2c$, bid-ask spreads and volatility in 2008. The discrete-continuous mixture captures the high volatility spikes observed during the global financial crisis, which in turn determine a lower probability of observing a zero return. The estimated round-trip costs need not be spuriously small to gauge the higher trading activity observed in the market. This particular feature of the model has a clear financial interpretation. In a bear market, investors are willing to sell financial securities amid widespread uncertainty. Higher volatility triggers further sales, leading to an increase in trading activity even in the presence of large transaction costs.

Using the FZR measure to track trading costs may thus overlook such an important dimension of (il)liquidity. A viable proxy of market (il)liquidity should take into account not only the number of zero returns but even the volatility of the information signal. From this point of view, our approach is close to Bandi et al. (2017), which estimate a similar microstructure model using a set of moment conditions involving the Excess Idle Time (EXIT) estimator. The latter represents a generalization of the FZR measure which formally characterizes the deviation of the observed price from the traditional assumption made in the continuous-time finance literature of a Brownian semimartingale process.

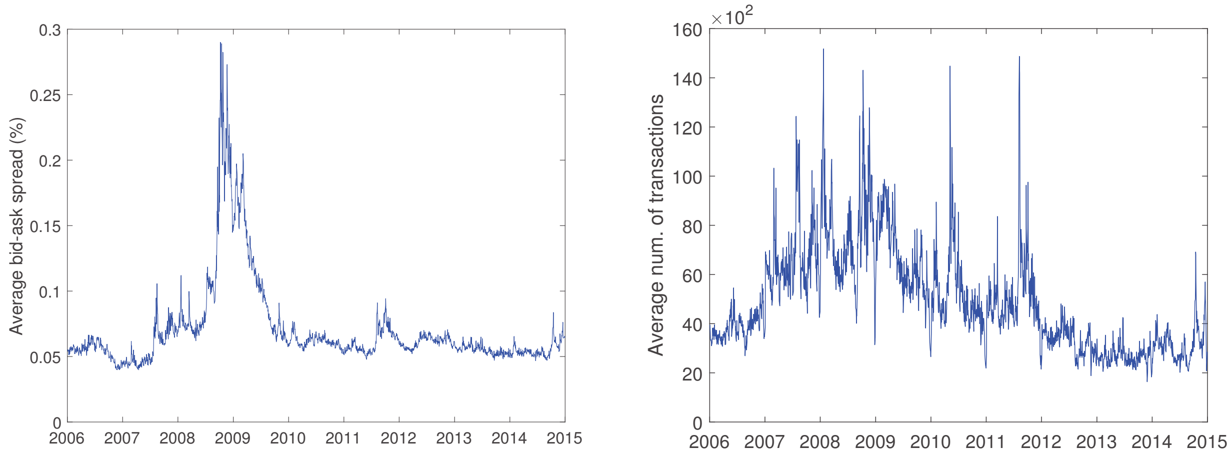

Figure 9: We report the time-series of bid-ask spread, computed by averaging over the 190 stocks (left panel), and the time-series of the number of transactions, computed by averaging over the 190 stocks (right panel).

4.4 Forecasting Value-at-Risk

We assess the effect of illiquidity on daily VaR forecasts by carrying out with actual data a test analogous to the one performed in the Monte-Carlo analysis of Section (3.2). To run this test, we choose an illiquid asset not belonging to the dataset considered so far. The reason is that this dataset is composed of frequently traded assets that feature a negligible percentage of zero returns when sampling their prices at daily frequencies. The asset considered in this analysis is Glen Burnie Bancorp (GLBZ), a stock with a market capitalization of 26.141 million dollars traded on NasdaqCM. We download its daily closing prices from Yahoo Finance in the period from 25 October 2007 to 28 September 2020. GLBZ is not traded for a significant fraction of the considered sample, as shown by its FZR measure, which is equal to 0.4693.

We estimate the discrete-continuous mixture model on the sub-sample including the first 1000 days, from 25 October 2007 to 12 October 2011. One-day-ahead VaR forecasts are computed in the remaining days by re-estimating the model on a rolling window of 1000 days. We also compute the VaR provided by Beta-t-EGARCH and GARCH using the same rolling window approach.

GLBZ				
$\hat{\omega}$	$\hat{\phi}$	$\hat{k}$	$\hat{\nu}$	$\hat{c}$
0.1714	0.8123	0.0838	4.6394	1.4962
(0.1469)	(0.1342)	(0.0246)	(1.1485)	(0.6219)

Table 6: Average of maximum likelihood estimates for GLBZ recovered by re-estimating the discrete-continuous mixture on a rolling window of 1000 days from 13 October, 2011 to 28 September, 2020. The values inside the parenthesis indicate the standard deviations of the estimated parameters.

Table (6) shows the maximum likelihood estimates of the discrete-continuous mixture, averaged over all the rolling windows on which the model is re-estimated. Given the significant amount of zero returns, the estimated threshold parameter is large and highly significant. We also note that the estimated degrees of freedom parameter is significantly lower compared to the average values found on intraday data, see Table (4). This is because the model is now re-estimated on time windows of 1000 days, which are likely to include jumps in prices.

The coverage rates of the three models are reported in Table (7) together with the p-values of the Kupiec (1995) test. The null hypothesis, in this test, is that the coverage rate is equal to the nominal confidence level. The results are quite similar to those obtained in the Monte-Carlo analysis, showing that the coverage rates of the discrete-continuous mixture closely match the nominal confidence levels. The VaR forecasts of the Beta-t-EGARCH are significantly jeopardized by zero returns. As discussed in Section (3.2), this is imputable to the downward bias on the degrees of freedom parameter, which heavily downweights volatility forecasts. Given the large number of zero returns in GLBZ, the VaR provided by the Beta-t-EGARCH is virtually zero, and therefore the coverage rate reduces to the fraction of positive returns in the sample, which is 0.6996. This example is interesting because it shows that the use of the Beta-t-EGARCH to compute the VaR of a very illiquid asset can lead to a trivial result.

As found in the Monte-Carlo analysis, the GARCH overestimates the VaR at 90% and 95% confidence levels, whereas it underestimates the VaR at 99% confidence level. This result is due to a trade-off between the non-robustness of the GARCH update, which determines spuriously large volatility in the presence of jumps, and the misspecification coming from assuming a normal density, which underrates the quantiles corresponding to very large confidence levels.

	Beta-t-EGARCH	D-C mixture	GARCH
$\text{VaR}_{90\%}$	0.6996	0.9011	0.9352
	(0.0000)	(0.8660)	(0.0000)
$\text{VaR}_{95\%}$	0.6996	0.9543	0.9614
	(0.0000)	(0.3418)	(0.0098)
$\text{VaR}_{99\%}$	0.6996	0.9916	0.9840
	(0.0000)	(0.4412)	(0.0087)

Table 7: Coverage rates of out-of-sample VaR forecasts for GLBZ at 90%, 95%, 99% confidence levels. The values inside parenthesis are the p-values of the test of Kupiec (1995).

5 Conclusions

Illiquidity has deep implications for the time-series behaviour of financial securities. The model proposed in this paper accounts for the presence of zero returns in the process of generating financial data, motivating the probability mass in zero in the light of well-established theories of price formation. Our analysis shows that the use of standard volatility models can lead to trivial results when data is affected by a significant percentage of zero returns. This is especially evident in models that account for fat-tails in data. The excess kurtosis induced by zero returns heavily downweights the volatility estimates of these models, leading to unreliable VaR forecasts for illiquid assets. We show how the proposed model can be used to compute VaR forecasts by taking into account the effect of zero returns. Given the importance of VaR in the banking regulatory system, such results can have relevant policy implications for the finance industry.

The analysis also provides interesting insights into the literature on liquidity measures. It shows that using zero returns alone as a proxy for market illiquidity can lead to erroneous conclusions in periods of market distress characterized by an increase in trading activity. Our model incorporates a second relevant dimension in the problem, the volatility of the information signal. If volatility is high, the trading activity can increase as a consequence of market sentiment, even in the case in which liquidity is low. In this circumstance, the number of zero returns tends to decrease, determining spuriously large liquidity. Trading activity in our model depends on both transaction costs and volatility. By capturing the high volatility observed in these periods, the model provides a more reliable assessment of market liquidity.

6 Bibliography

Abramowitz, M. and Stegun, I. A. (1972). Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. National Bureau of Standards Applied Mathematics Series 55. Tenth Printing.

Amihud, Y. and Mendelson, H. (1987). Trading Mechanisms and Stock Returns: An Empirical Investigation. *The Journal of Finance*, 42:533–553.

Amihud, Y. and Mendelson, H. (1989). The Effects of Beta, Bid-Ask Spread, Residual Risk, and Size on Stock Returns. *The Journal of Finance*, 44:479–486.

Andersen, T. G., Dobrev, D., and Schaumburg, E. (2012). Jump-Robust Volatility Estimation using Nearest Neighbor Truncation. *Journal of Econometrics*, 169:75–93.

Bandi, F. M., Kolokolov, A., Pirino, D., and Renò, R. (2020). Zeros. *Management Science*, 66:3466–3479.

Bandi, F. M., Pirino, D., and Renò, R. (2017). EXcess Idle Time. *Econometrica*, 85:1793–1846.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2009). Realized Kernels in Practice: Trades and Quotes. *The Econometrics Journal*, 12:C1–C32.

Bekaert, G., Harvey, C. R., and Lundblad, C. (2007). Liquidity and Expected Returns: Lessons from Emerging Markets. *The Review of Financial Studies*, 20:1783–1831.

Blasques, F., Gorgi, P., and Koopman, S. (2020). Missing Observations in Observation-Driven Time Series Models. *Journal of Econometrics, Forthcoming*.

Blasques, F., Koopman, S. J., and Lucas, A. (2015). Information-Theoretic Optimality of Observation-Driven Time Series Models for Continuous Responses. *Biometrika*, 102:325–343.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327.

Buccheri, G., Bormetti, G., Corsi, F., and Lillo, F. (2020). A Score-Driven Conditional Correlation Model for Noisy and Asynchronous Data: An Application to High-Frequency Covariance Dynamics. *Journal of Business & Economic Statistics*, 0:1–17.

Catania, L., Mari, R. D., and de Magistris, P. S. (2020). Dynamic Discrete Mixtures for High-Frequency Prices. *Journal of Business & Economic Statistics*, 0:1–19.

Creal, D., Koopman, S. J., and Lucas, A. (2011). A Dynamic Multivariate Heavy-Tailed Model for Time-Varying Volatilities and Correlations. *Journal of Business & Economic Statistics*, 29:552–563.

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics*, 28:777–795.

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

Glosten, L. R. and Milgrom, P. R. (1985). Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. *Journal of Financial Economics*, 14:71–100.

Goyenko, R. Y., Holden, C. W., and Trzcinka, C. A. (2009). Do Liquidity Measures Measure Liquidity? *Journal of Financial Economics*, 92:153–181.

Griffin, J. M., Kelly, P. J., and Nardari, F. (2010). Do Market Efficiency Measures Yield Correct Inferences? A Comparison of Developed and Emerging Markets. *The Review of Financial Studies*, 23:3225–3277.

Gryglewicz, S. (2011). A Theory of Corporate Financial Decisions with Liquidity and Solvency Concerns. *Journal of Financial Economics*, 99:365–384.

Harvey, A. and Ito, R. (2020). Modeling Time Series when Some Observations are Zero. *Journal of Econometrics*, 214:33–45.

Harvey, A. and Lange, R.-J. (2017). Volatility Modeling with a Generalized t Distribution. *Journal of Time Series Analysis*, 38:175–190.

Harvey, A. and Liao, Y. (2019). Dynamic Tobit models. Cambridge Working Papers in Economics 1913, Faculty of Economics, University of Cambridge.

Harvey, A. C. (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series*. Econometric Society Monographs. Cambridge University Press.

Hasbrouck, J. (1999). The Dynamics of Discrete Bid and Ask Quotes. *The Journal of Finance*, 54:2109–2142.

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.

Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives*, 3:73–84.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, 53:1315–1335.

Lesmond, D. A. (2005). Liquidity of Emerging Markets. *Journal of Financial Economics*, 77:411–452.

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.

Linton, O. and Wu, J. (2020). A Coupled Component DCS-EGARCH Model for Intraday and Overnight Volatility. *Journal of Econometrics*, 217:176–201.

Lo, A. W., Petrov, C., and Wierzbicki, M. (2003). It's 11 PM - Do You Know Where Your Liquidity Is? The Mean-Variance-Liquidity Frontier. *Journal of Investment Management*, 1:55–93.

Longstaff, F. A. (2009). Portfolio Claustrophobia: Asset Pricing in Markets with Illiquid Assets. *The American Economic Review*, 99:1119–1144.

Lucas, A., Opschoor, A., and Schaumburg, J. (2016). Accounting for Missing Values in Score-Driven Time-Varying Parameter Models. *Economics Letters*, 148:96–98.

Mei, J., Scheinkman, J. A., and Xiong, W. (2009). Speculative Trading and Stock Prices: Evidence from Chinese A-B Share Premia. *Annals of Economics and Finance*, 10:225–255.

Naes, R., Skjeltorp, J. A., and Ødegaard, B. A. (2011). Stock Market Liquidity and the Business Cycle. *The Journal of Finance*, 66:139–176.

Nelson, D. B. (1992). Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model. *Journal of Econometrics*, 52:61–90.

Nelson, D. B. and Foster, D. P. (1994). Asymptotic Filtering Theory for Univariate ARCH Models. *Econometrica*, 62:1–41.

Nelson, D. B. and Foster, D. P. (1995). Filtering and Forecasting with Misspecified ARCH Models II: Making the Right Forecast with the Wrong Model. *Journal of Econometrics*, 67:303–335.

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

Sandmann, G. and Koopman, S. J. (1998). Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood. *Journal of Econometrics*, 87:271–301.

Stoll, H. R. (1978a). The Pricing of Security Dealer Services: An Empirical Study of Nasdaq Stocks. *The Journal of Finance*, 33:1153–1172.

Stoll, H. R. (1978b). The Supply of Dealer Services in Securities Markets. *The Journal of Finance*, 33:1133–1151.

Stoll, H. R. (2000). Friction. *The Journal of Finance*, 55:1479–1514.

Sucarrat, G. and Grønneberg, S. (2020). Risk Estimation with a Time-Varying Probability of Zero Returns. *Journal of Financial Econometrics, Forthcoming*.

APPENDIX

A Conditional variance of observed returns

Since the discrete-continues mixture is symmetric around $y_t = 0$, it has conditional mean equal to zero. The conditional variance reduces to the second moment of the conditional distribution:

$$\operatorname{Var}\left[y_t|\mathcal{F}_{t-1},\sigma_t\right] = 2\int_0^{+\infty} y_t^2 p_Y(y_t;\mathcal{F}_{t-1},\sigma_t)dy_t = 2\int_c^{+\infty}(x_t-c)^2 p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t, \tag{15}$$

with $\sigma_t = e^{\lambda_t}$. Carrying out the square, Equation (15) becomes:

$$\begin{aligned}\operatorname{Var}\left[y_t|\mathcal{F}_{t-1},\sigma_t\right] = 2\Bigg[&\int_c^{+\infty} x_t^2 p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t + c^2\int_c^{+\infty} p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t \\ &- 2c\int_c^{+\infty} x_t p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t\Bigg].\end{aligned} \tag{16}$$

The second integral is given by $\int_c^{+\infty} p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t = F_X(-c;\mathcal{F}_{t-1},\sigma_t)$. Let us compute the first integral:

$$\int_c^{+\infty} x_t^2 p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t = \int_c^{+\infty} x_t^2\cdot\frac{1}{\sqrt{\nu}\sigma_t B(1/2,\nu/2)}\left(1+\frac{x_t^2}{\nu\sigma_t^2}\right)^{-\frac{\nu+1}{2}}dx_t. \tag{17}$$

Making a change of variable to $t_t = \frac{x_t}{\sqrt{\nu}\sigma_t}$, $m = \frac{\nu-1}{2}$ and $dt = \frac{1}{\sqrt{\nu}\sigma_t}dx$, we get:

$$\int_c^{+\infty} x_t^2\cdot\frac{1}{\sqrt{\nu}\sigma_t B(1/2,\nu/2)}\left(1+\frac{x_t^2}{\nu\sigma_t^2}\right)^{-\frac{\nu+1}{2}}dx = \frac{\nu\sigma_t^2}{B(1/2,m-\frac{1}{2})}\int_{\frac{c}{\sqrt{\nu}\sigma_t}}^{+\infty}\frac{t_t^2}{(1+t_t^2)^m}dt_t. \tag{18}$$

Defining $k = \frac{\nu\sigma_t^2}{B(1/2,m-\frac{1}{2})}$ and making the change of variable $b_{t,t} = \frac{t_t^2}{(1+t_t^2)}$, so that $\frac{1}{1+t_t^2} = 1-b_{t,t}$, $t_t = \sqrt{\frac{b_{t,t}}{1-b_{t,t}}}$ and $db_{t,t} = \frac{2t_t}{(1+t_t^2)^2}dt_t$, we obtain that Equation (18) becomes:

$$\frac{k}{2}\int_{b_{c,t}}^1(1-b_{t,t})^{m-2}\left(\sqrt{\frac{b_{t,t}}{1-b_{t,t}}}\right)db_{t,t} = \frac{k}{2}\int_{b_{c,t}}^1(1-b_{t,t})^{m-1-3/2}b_{t,t}^{1-1/2}db_{t,t}, \tag{19}$$

with $b_{c,t} = \frac{c^2/\nu\sigma_t^2}{1+c^2/\nu\sigma_t^2}$. Therefore, knowing that:

$$\int_0^1(1-b_{t,t})^{m-1-3/2}b_{t,t}^{1-1/2}db_{t,t} = B(1+1/2,\nu/2-1), \tag{20}$$

we can re-write Equation (19) as follows:

$$\frac{k}{2}\int_{b_{c,t}}^1(1-b_{t,t})^{m-1-3/2}b_{t,t}^{1-1/2}db_{t,t} = \frac{k}{2}B(1+1/2,\nu/2-1) - \frac{k}{2}\int_0^{b_{c,t}}(1-b_{t,t})^{m-1-3/2}b_{t,t}^{1-1/2}db_{t,t}, \tag{21}$$

where $\int_0^{b_{c,t}}(1-b_{t,t})^{m-1-3/2}b_{t,t}^{1-1/2}db_{t,t} = B(b_{c,t};1+1/2,\nu/2-1)$ is the incomplete beta function. Thus, the first integral in Equation (16) is given by:

$$\begin{aligned}\int_c^{+\infty} x_t^2 p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t &= \frac{k}{2}B(1+1/2,\nu/2-1) - \frac{k}{2}B(b_{c,t};1+1/2,\nu/2-1)\\ &= \frac{\nu\sigma_t^2 B(1+1/2,\nu/2-1)}{2B(1/2,m-\frac{1}{2})} - \frac{\nu\sigma_t^2 B(b_{c,t};1+1/2,\nu/2-1)}{2B(1/2,m-\frac{1}{2})}\\ &= \frac{\nu}{2(\nu-2)}\sigma_t^2 - \frac{\nu\sigma_t^2 B(b_{c,t};1+1/2,\nu/2-1)}{2B(1/2,\nu/2)}.\end{aligned} \tag{22}$$

We now compute the last integral in Equation (16):

$$-2c\int_c^{+\infty} x_t p_X(x_t;\mathcal{F}_{t-1},\sigma_t)dx_t = -c\int_c^{+\infty} 2x_t\cdot\frac{1}{\sqrt{\nu}\sigma_t B(1/2,\nu/2)}\left(1+\frac{x_t^2}{\nu\sigma_t^2}\right)^{-\frac{\nu+1}{2}}dx_t. \tag{23}$$

This is an elementary integral that can be computed immediately:

$$\begin{aligned}-\frac{c\sqrt{\nu}\sigma}{B(1/2,\nu/2)}\int_c^{+\infty}\frac{2x}{\nu\sigma_t^2}\left(1+\frac{x^2}{\nu\sigma_t^2}\right)^{-\frac{\nu+1}{2}}dx_t &= \frac{c\sqrt{\nu}\sigma}{B(1/2,\nu/2)}\left.\frac{\left(1+\frac{x_t^2}{\nu\sigma_t^2}\right)^{-\frac{(\nu-1)}{2}}}{\frac{(\nu-1)}{2}}\right|_c^{+\infty}\\ &= -\frac{c\sqrt{\nu}\sigma}{B(1/2,\nu/2)}\frac{\left(1+\frac{c^2}{\nu\sigma_t^2}\right)^{-\frac{(\nu-1)}{2}}}{\frac{(\nu-1)}{2}}.\end{aligned} \tag{24}$$

Substituting Equation (24) and the result of Equation (22) into Equation (16), we retrieve the conditional variance of y_t as follows:

$$\begin{aligned}\mathrm{Var}[y_t|\mathcal{F}_{t-1},\sigma_t] &= \frac{\nu}{(\nu-2)}\sigma_t^2 - \frac{\nu\sigma_t^2 B(b_{c,t};1+1/2,\nu/2-1)}{B(1/2,\nu/2)} + 2c^2F_X(-c;\mathcal{F}_{t-1},\sigma_t)\\ &\quad -2\frac{c\sqrt{\nu}\sigma}{B(1/2,\nu/2)}\frac{\left(1+\frac{c^2}{\nu\sigma_t^2}\right)^{-\frac{(\nu-1)}{2}}}{\frac{(\nu-1)}{2}},\end{aligned} \tag{25}$$

where $\frac{\nu}{(\nu-2)}\sigma_t^2 = \mathrm{Var}[x_t|\mathcal{F}_{t-1},\sigma_t]$ is the conditional variance of x_t. By noting that $F_X(-c;\mathcal{F}_{t-1},\sigma_t) = 1/2 - 1/2\beta(b_{c,t};1/2,\nu/2)$, we obtain the expression in Proposition 2.1.
Observe now that $b_{c,t}\to 1$ when $c\to+\infty$, and thus $\frac{\nu\sigma_t^2 B(b_{c,t};1+1/2,\nu/2-1)}{B(1/2,\nu/2)} \to \frac{\nu}{\nu-2}\sigma_t^2$. Observe also that, since $F_X(-c;\mathcal{F}_{t-1},\sigma_t)$ asymptotically behaves as $\frac{1}{c^\nu}$, for $\nu>2$ we get $2c^2F_X(-c;\mathcal{F}_{t-1},\sigma_t)\to 0$ as c diverges. Since also the last term in the equation above goes to zero as c diverges, we obtain $\mathrm{Var}[y_t|\mathcal{F}_{t-1},\sigma_t]\to 0$ for $c\to+\infty$.

B Score of the discrete-continuous mixture

The score of the discrete-continuous mixture in Equation (4) is:

$$u_t = (1-I_{y_t=0})\frac{\partial \ln p_X(x_t;\mathcal{F}_{t-1},\lambda_t)}{\partial\lambda_t} + I_{y_t=0}\frac{\partial\ln[F_X(c;\mathcal{F}_{t-1},\lambda_t) - F_X(-c;\mathcal{F}_{t-1},\lambda_t)]}{\partial\lambda_t}. \tag{26}$$

To prove the result, we start by computing the first term in the right hand side of the equation. Consider first the conditional log-likelihood:

$$\ln p_X\left(x_t; \mathcal{F}_{t-1}, \lambda_t\right) = \ln\left(\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\right) - \lambda_t - \frac{(\nu+1)}{2}\ln\left(1+\frac{x_t^2}{\nu e^{2\lambda_t}}\right). \tag{27}$$

Taking the derivative with respect to λ_t gives:

$$\frac{\partial \ln p_X\left(x_t; \mathcal{F}_{t-1}, \lambda_t\right)}{\partial \lambda_t} = [(\nu+1)b_t - 1], \tag{28}$$

where

$$b_t = \frac{x_t^2/\left(\nu e^{2\lambda_t}\right)}{x_t^2/\left(\nu e^{2\lambda_t}\right)+1} = \frac{z_t}{z_t+1}, \quad b_{c,t} \leq b_t \leq 1, \tag{29}$$

is distributed as a beta with parameters $1/2$, $\nu/2$, $z_{c,t} = c^2/\left(\nu e^{2\lambda_t}\right)$ and $b_{c,t}$ is defined in Section (2.2). We now compute the second term in the right hand side of Equation (26). Let's remind from Section (2.2) that $F_X(c; \mathcal{F}_{t-1}, \lambda_t) - F_X(-c; \mathcal{F}_{t-1}, \lambda_t) = \beta\left(b_{c,t}; 1/2, \nu/2\right)$, where $\beta(\cdot; \cdot, \cdot)$ is the incomplete regularized beta function. Thus:

$$\begin{aligned}
\frac{\partial \ln \beta(b_{c,t}; 1/2, \nu/2)}{\partial \lambda_t} &= \frac{\partial \ln \beta(b_{c,t}; 1/2, \nu/2)}{\partial b_{c,t}} \frac{\partial b_{c,t}}{\partial \lambda_t} \\
&= \frac{1}{\beta(b_{c,t}; 1/2, \nu/2)} \frac{\partial \beta(b_{c,t}; 1/2, \nu/2)}{\partial b_{c,t}} \frac{\partial b_{c,t}}{\partial z_{c,t}} \frac{\partial z_{c,t}}{\partial \lambda_t} \\
&= -\frac{2 p_B(b_{c,t}; 1/2, \nu/2) b_{c,t}\left(1-b_{c,t}\right)}{\beta\left(b_{c,t}; 1/2, \nu/2\right)} = -\frac{2 b_{c,t}^{1/2}(1-b_{c,t})^{\nu/2}}{\beta(b_{c,t}; 1/2, \nu/2) B(1/2, \nu/2)},
\end{aligned} \tag{30}$$

where $p_B(b_{c,t}; 1/2, \nu/2)$ is the PDF of a beta distribution with parameters $1/2$, $\nu/2$. Summing Equations (28) and (30), we obtain the expression in Equation (10).

C Proof that the score has zero expectation

We can write the expected value of Equation (10) as:

$$\begin{aligned}
\mathrm{E}(u_t) &= \mathrm{E}\left[(1-I_{y_t=0})[(\nu+1)b_t - 1] - I_{y_t=0}\left(\frac{2 b_{c,t}^{1/2}(1-b_{c,t})^{\nu/2}}{\beta(b_{c,t}; 1/2, \nu/2) B(1/2, \nu/2)}\right)\right] \\
&= \int_{b_{c,t}}^{1} \left[(\nu+1)b_t - 1\right] p_B(b_t; 1/2, \nu/2) db_t - \beta\left(b_{c,t}; 1/2, \nu/2\right) \frac{2 b_{c,t}^{1/2}\left(1-b_{c,t}\right)^{\nu/2}}{\beta\left(b_{c,t}; 1/2, \nu/2\right), B(1/2, \nu/2)}
\end{aligned} \tag{31}$$

The integral is between $b_{c,t}$ and 1 because of the restriction $b_{c,t} \leq b_t \leq 1$, which holds for $y_t \neq 0$. The expectation of the second term gives the mass of the censored region, which is

$\beta\left(b_{c,t};1/2,\nu/2\right)$. We can re-write Equation (31) as:

$$\mathrm{E}(u_t) = (\nu+1)\int_{b_{c,t}}^{1} b_t p_B(b_t;1/2,\nu/2)db_t - \int_{b_{c,t}}^{1} p_B(b_t;1/2,\nu/2)db_t - \frac{2b_{c,t}^{1/2}\left(1-b_c\right)^{\nu/2}}{B(1/2,\nu/2)}. \tag{32}$$

Since $b_t p_B(b_t;1/2,\nu/2) = \frac{B(1/2+1,\nu/2)}{B(1/2,\nu/2)} p_B(b_t;1/2+1,\nu/2) = \frac{1}{\nu+1} p_B(b_t;1/2+1,\nu/2)$, Equation (32) can also be expressed as:

$$\mathrm{E}(u_t) = \int_{b_{c,t}}^{1} p_B(b_t;1/2+1,\nu/2)db_t - \int_{b_{c,t}}^{1} p_B(b_t;1/2,\nu/2)db_t - \frac{2b_{c,t}^{1/2}\left(1-b_{c,t}\right)^{\nu/2}}{B(1/2,\nu/2)}. \tag{33}$$

The first two terms in Equation (33) are the complementary cumulative distribution functions of a beta distribution with parameters $3/2$, $\nu/2$, and $1/2$, $\nu/2$ respectively. Thus:

$$\begin{aligned}\mathrm{E}(u_t) &= \left[1-\beta(b_{c,t},1/2+1,\nu/2)\right] - \left[1-\beta(b_{c,t};1/2,\nu/2)\right] - \frac{2b_{c,t}^{1/2}\left(1-b_{c,t}\right)^{\nu/2}}{B(1/2,\nu/2)} \\ &= -\beta(b_{c,t},1/2+1,\nu/2) + \beta(b_{c,t};1/2,\nu/2) - \frac{2b_{c,t}^{1/2}\left(1-b_{c,t}\right)^{\nu/2}}{B(1/2,\nu/2)}.\end{aligned} \tag{34}$$

Using identity (26.5.16) in Abramowitz and Stegun (1972), namely $-\beta(b_{c,t},1/2+1,\nu/2) + \beta(b_{c,t};1/2,\nu/2) = \frac{2b_{c,t}^{1/2}(1-b_{c,t})^{\nu/2}}{B(1/2,\nu/2)}$, we end up with:

$$\mathrm{E}(u_t) = \frac{2b_{c,t}^{1/2}\left(1-b_{c,t}\right)^{\nu/2}}{B(1/2,\nu/2)} - \frac{2b_{c,t}^{1/2}\left(1-b_{c,t}\right)^{\nu/2}}{B(1/2,\nu/2)} = 0. \tag{35}$$

RECENT PUBLICATIONS BY *CEIS Tor Vergata*

Corruption Bias and Information: A Study in the Lab
Germana Corrado, Luisa Corrado and Francesca Marazzi
CEIS Research Paper, 505, January 2021

The Future of the Elderly Population Health Status: Filling a Knowledge Gap
Vincenzo Atella, Federico Belotti, Kim Daejung, Dana Goldman, Tadeja Gracner, Andrea Piano Mortari and Bryan Tysinger
CEIS Research Paper, 504, December 2020

The Macroeconomic Effects of Aerospace Shocks
Luisa Corrado, Stefano Grassi and Edgar Silgado-Gómez
CEIS Research Paper, 503, November 2020

On Cointegration for Processes Integrated at Different Frequencies
Tomás del Barrio Castro, Gianluca Cubadda and Denise R. Osborn
CEIS Research Paper, 502, September 2020

Climate Actions and Stranded Assets: The Role of Financial Regulation and Monetary Policy
Francesca Diluiso, Barbara Annicchiarico, Matthias Kalkuhl and Jan C. Minx
CEIS Research Paper, 501, July 2020

Pre-selection Methods for Cointegration-based Pairs Trading
Marianna Brunetti and Roberta De Luca
CEIS Research Paper, 500, June 2020

Poisson Search
Francesco De Sinopoli, Leo Ferraris and Claudia Meroni
CEIS Research Paper, 499, June 2020

ESG Investment Funds: A Chance To Reduce Systemic Risk
Roy Cerqueti, Rocco Ciciretti, Ambrogio Dalò and Marco Nicolosi
CEIS Research Paper, 498, June 2020

The Legacy of Literacy: Evidence from Italian Regions
Roberto Basile, Carlo Ciccarelli and Peter Groote
CEIS Research Paper, 497, June 2020

A Test of Sufficient Condition for Infinite-step Granger Noncausality in Infinite Order Vector Autoregressive Process
Umberto Triacca, Olivier Damette and Alessandro Giovannelli
CEIS Research Paper, 496, June 2020