

Visible and Invisible Forces: What Drives the Intensity of Trading in the European Carbon Market?

IORDANIS KALAITZOGLOU

*Audencia Business School**

This study models the trading intensity in European Allowances (EUA) futures contracts in the European Climate Exchange (ECX) using various specifications and investigates the forecasting ability of observable versus unobservable factors. This set up tests empirically the impact of the evolving market structure through regulatory updates, captured by trading volume fluctuations, to the intensity of trading in the European Carbon market. The findings suggest that observable market characteristics capture better the dynamics of trading intensity than their latent counterparts, which implies that regulatory changes that enhance transparency would also improve market efficiency.

Keywords: Carbon market, Duration modelling Ultra-High Frequency data.

JEL classification: C30, C41, G14

1 Introduction

This study investigates explicitly the modelling of trading intensity in the last two phases of regulatory amendments in the European Carbon market, prior to the introduction of MiFID II (January 2018). The primary objective is to identify whether observable factors, such as trading volume (Dufour and Engle, 2000) and trading frequency (e.g., Engle and Russell, 1998), or unobservable factors, captured by distributions with progressively nesting properties, contribute more to trading intensity clustering (e.g., Kalaitzoglou and Ibrahim, 2013a). According to the empirical findings it is the observable factors that exhibit a higher explanatory power of the data generation process of trading intensity, suggesting that greater market transparency – the main objective of MiFID II – would ameliorate the predictive power of liquidity models and thus, it would contribute through liquidity to higher market efficiency (Kalaitzoglou and Ibrahim, 2015) and better price discovery (Medina et al. 2014).¹

* Email:

© 2018 Iordanis Kalaitzoglou . Licensed under the Creative Commons Attribution - Noncommercial 3.0 Licence (<http://creativecommons.org/licenses/by-nc/3.0/>). Available at <http://rofea.org>.

¹ For more information about MiFID II please refer to the European Commission updated rules, https://ec.europa.eu/info/business-economy-euro/banking-and-finance/financial-markets/securities-markets/investment-services-and-regulated-markets-markets-financial-instruments-directive-mifid_en, as well as

In more detail, the European Carbon market, known as the “European Union Greenhouse Gas Emission Trading System” (EU ETS), is the regulatory response of the European Union to the Kyoto Protocol (1997) and creates a new “commodity” market that has become the most important mechanism in reducing emissions.² The EU ETS exhibits some distinctive features that are associated with its relative illiquidity. It is a “cap and trade” market, where a certain number of allowances (i.e., EUAs) is allocated to only specific industrial sectors and each company can trade its surplus/deficit, while it allows Over-The-Counter (OTC) EUA holders to enter and register their positions. Consequently, the overall quantity of available EUAs is politically influenced and their trading is affected both by the organised venues, as well as by the OTC market. Recently, the introduction of the MiFID II regulation, with its suggested increased transparency, is expected to affect market liquidity (e.g., Rannou and Barneto, 2016) and through it market efficiency (Kalaitzoglou and Ibrahim, 2015) and price discovery (Medina et al. 2014). Consequently, an empirical investigation of intraday trading patterns could provide a further insight on the expected impact of increased transparency on liquidity.

Several studies, in early stages of the market, have discussed several stylized factors and provide a general description of the trading mechanism (e.g., Mansanet-Bataller and Pardo, 2008; Chevallier, 2009), as well as of price evolution (e.g., Christiansen et al., 2005; Uhrig-Homburg and Wagner, 2006; Mansanet-Bataller et al., 2007), mainly on a daily frequency (e.g., Daskalakis et al., 2009). However, other studies (e.g., Benz and Trück, 2009) examine the same issue on an intraday basis and report the existence of trading patterns (e.g., Kalaitzoglou and Ibrahim, 2013b) that can be employed to identify the existence of different groups of traders (e.g., Kalaitzoglou and Ibrahim, 2013a), as well as their interaction with price discovery (e.g., Bredin et al., 2014), trading costs (e.g., Ibrahim and Kalaitzoglou, 2016) and order to trade ratios (e.g., Galariotis et al., 2018). All these suggest that liquidity, especially on that level, has a significant impact on market quality in the form of frictions (e.g., Medina et al., 2014) and overall efficiency (e.g., Ibikunle et al., 2016). Consequently, the identification of observable or latent factors that might

on the report from the Commission to the European parliament and the council, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52015DC0576>.

² The majority of countries (Annex I) have ratified a treaty – “The Kyoto Protocol” – aiming at their emissions’ reduction. The agreement defines three periods for countries to meet their targets. Phase I (2005-2007) is the pilot period, Phase II (2008-2012) is the commitment period and Phase III (2013-2020) is the post commitment period for re-evaluation and further adjustments. There are three “flexibility mechanisms” for achieving emission targets: (i) Joint Implementation (JI), (ii) Clean Development Mechanism (CDM) and (iii) Emission Trading Scheme (ETS). Each one provides units – JI provides Emission Reduction Units (ERUs), CDM provides Certified Emission Reduction Units (CERs), while Allocated Allowance Units (AAUs) are allocated according to national allocation plans –that can be used for compliance.

improve liquidity forecasting ability, would improve market quality and help the market achieve its emission reduction targets.

Unlike when using daily data, an intraday analysis using Ultra-High-Frequency (UHF) data poses some challenges with respect to the arrival time of events, which is irregularly spaced and might depend on various factors. According to Engle and Russell (1998), the time between two consecutive trades, namely duration, might convey price relevant information, as it measures economic and not calendar time. They model the Data Generation Process (DGP) of durations as a dependent point process, with an autoregressive model named Autoregressive-Conditional-Duration (ACD) model, in which the conditional mean varies over time as a function of past durations, while an associated distribution with positive support must be specified. Many studies challenge the simplicity of the conditional mean specification (e.g., Dufour and Engle, 2000; Zhang et al., 2001), as being inadequate in capturing non-linear asymmetries, and of the conditional density function (e.g., Grammig and Maurer, 2000; Hujer and Vuletić, 2007), as insufficient in capturing duration's higher moments. Combining the two empirical challenges, Bauwens et al. (2004) suggest that the ACD framework could be used to distinguish between the impact of observable (i.e., conditional mean specification) and unobservable (i.e., conditional density specification) factors on fitting of the model and, consequently, on the drivers of the intraday liquidity formation.

Towards that direction, several studies try to identify observable factors that might affect intraday liquidity, with volume, duration and price change being the most widely discussed observable factors. In the context of the EU ETS, several studies (e.g., Bredin et al., 2014; Ibrahim and Kalaitzoglou, 2016) examine the interaction between duration, volume and price and report a positive correlation between the intensity of trading and price change. Kalaitzoglou and Ibrahim (2015) unify volume and duration into a new, rescaled time variable, proxy for trading intensity and report that deviations from the conditionally expected levels are associated with information inflow and the induction of uncertainty in the price discovery process. Following the propositions of Easley and O'Hara (1992), Kalaitzoglou and Ibrahim (2013a) investigate this further and report a link between these deviations and the existence of different types of market participants with respect to their association to information. Informed trading cannot be identified, but it could be inferred from the trading activity. This is usually done by assuming that the DGP of trading intensity consists of distinct groups of traders. Modelling-wise, this is usually modelled with an increasing flexibility of the conditional density function (e.g., Hujer and Vuletic, 2007; Kalaitzoglou and Ibrahim, 2013; Rannou and Barneto, 2016). Therefore, the empirical practice so far considers observable factors affecting the conditionally expected trading intensity (i.e., conditional mean specification), while latent factors are assumed to be manifested in the distribution of the trading intensity (i.e., captured by a more flexible conditional density specification).

This is the primary concern of this study. An empirical investigation of the drivers of intraday liquidity, in the form of duration and volume, is conducted, aiming at identifying the main determinants of the trading intensity process, with respect to observable or unobservable market factors. This is done by testing the fitting of different specifications of trading intensity using the Autoregressive-Conditional-Weighted-Duration (ACWD) framework (Kalaitzoglou and Ibrahim, 2015). This framework describes the DGP of trading intensity, modelled as a rescaled point process, combining both duration and volume into one variable.

The empirical findings suggest that the trading activity in the EU ETS can be better described by enhanced conditional mean specifications, rather than by more flexible conditional density specification. This implies that trading history, which is observable, is an important determinant of subsequent trading activity. In contrast, inferred information about trading activity, such as the existence of better informed agents, usually captured by the distributional assumptions, does improve the fitting of the models, but not to the extent of the conditional mean specification improvements. Consequently, latent-unobservable-inferred factors, have a lesser contribution to determining subsequent trading activity. This finding has a direct regulatory implication, in favour of MiFID II, suggesting that greater transparency, which would enhance the quantity and quality of available information, can lead to a better understanding of the fundamentals that drive trading activity and thus, to a market with less frictions. This can lead to a better balance between market innovation and liquidity needs. According to Viswanathan (2010), this would provide the foundations of more accurate pricing and would help the market to serve its purpose of reducing emissions. A more precise modelling of trading activity could also improve market making practices in EU ETS. Market makers can better manage risk if they know how long (i.e., expected duration) they are going to be exposed to it, and this would result in narrower spreads. Limit order traders can also benefit from a better trading intensity model, since they can develop better trading strategies by managing the time dimension of “execution” risk.

The remainder of the paper is organized as follows. The next section presents the methodology and a description of the data collection and manipulation. The next section discusses the empirical findings, and the last section concludes.

2 Methodology and Data Sample

2.1 The ACDW Model

Engle and Russell (1998) propose the Autoregressive Conditional Duration (ACD) models for modelling the inter-trade intervals, durations $x_i = t_i - t_{i-1}$ (where t_i is the time stamp of a trade), and thus, liquidity, as a dependent stochastic process. Following Kalaitzoglou and Ibrahim

(2015), this is complemented by a measure of relative volume (number of contracts, $v \sim N(\bar{v}, \sigma_v)$), $K(u_i) = \exp\left(\frac{-u_i}{2}\right)$, where $u_i = (v_i - \bar{v})/\sigma_v$, which is used to rescale time and form a new proxy for trading intensity, $S_i = x_i * K(u_i)$. The new variable, i.e., trading intensity S_i is then modelled following the ACWD framework:

$$S_i = \theta_i \varepsilon_i \quad (1)$$

$$\theta_i = \theta(S_{i-1}, \dots, S_1; \varphi_1) \quad (2)$$

$$\varepsilon_i \sim i. i. d. \text{ with density } f(\varepsilon_i; \varphi_2) \quad (3)$$

where, θ_i is the expected trading intensity, conditioned on past trading intensity, ε_i is the standardized volume weighted durations, and φ_1 and φ_2 are vectors of parameters. Both the conditional mean, $\theta(S_{i-1}, \dots, S_1; \varphi_1)$, as well as the conditional density function of S , $f(S_i)$, can be modelled using various specifications in order to capture various market stylized factors.

2.1.1 Conditional Mean Specifications

In order to account for the impact of observable trading history on expected trading intensity, the following enhanced version of the original ARMA-like specification is introduced:

$$\theta_i = \omega + \sum_{j=1}^m (a_j + \zeta_j * I_{i-j}) S_{i-j} + \sum_{j=1}^q \beta_j \theta_{i-j} \quad (4)$$

which considers the impact of lagged, observable variables I on expected trading intensity. Considering previous literature that recognizes the asymmetric impact of volume (e.g., Kalaitzoglou and Ibrahim, 2013b; Rannou and Barneto, 2016) on subsequent trading activity, volume is introduced as a determinant of subsequent trading intensity.³ Previous transactions of differing volume revise the coefficient of the past values of trading intensity, a . Compared to the linear model, a represents the “average” impact of previous transactions, while ζ is adjusts its

³ This is not an identification statement, but it serves the purpose of investigating whether observable variables have a greater forecasting ability than unobservable, summarized in the conditional density. Other variables could also be introduced alongside or instead of volume, but the objective of this study is to investigate which part of the modelling, i.e., the conditional mean specification that captures observable information or the conditional density specification that captures inferred/unobservable information, contributes more to describing trading intensity.

value according to the magnitude of the volume of lagged transactions. When $\zeta > 0$ ($\zeta < 0$) market activity slows down (accelerates) with higher volume, in the form of longer (shorter) expected duration and/or lower (larger) expected volume. Various economic explanations, mainly related to information and liquidity, might be proposed to link the sign of ζ to an economic context. This is relevant to the modelling of trading intensity in the sense that, given the depth and the size of the market, large trades are expected to influence the floor trading. A positive ζ could possibly mean that when large trades appear in the market, market members might perceive this as information inflow and might abstain from trading in the fear of losing money by trading with informed traders. Another explanation could be that larger trades might exhaust liquidity, at least on the best bid-ask level and therefore, trading activity might need some time to recover. Whatever the reason, this empirical specification should be able to capture the impact of observable market factors on subsequent market activity.

In addition, in order to address potential non-linearities, a Box-Cox transformation (Hautsch, 2004), i.e., BCACWD, is also introduced:

$$\ln \theta_i = \omega + \sum_{j=1}^m (a_j + \zeta_j * I_{i-j}) * (\varepsilon_{i-j})^\delta + \sum_{j=1}^q \beta_j \ln \theta_{i-j} \quad (5)$$

In this model, two types of non-linearities are introduced. First, it is a logarithmic transformation of the expected trading intensity, which places a greater weight on low values of Θ , which capture a high expected trading activity. This is also measured in relative terms, i.e., ε , with respect to the actual duration, S . Second, the impact of these relative innovations might be asymmetric, of higher or lower order δ , determined by the data. In fact, if $\delta < 0$, larger shocks have a lower impact on expectations and this implies a tendency to return to a “normal” stage of market activity, e.g., mean reversion, after stages of excessive activity or inactivity. In contrast when $\delta > 0$, larger shocks have a stronger impact on subsequent trading intensity, which implies a tendency to prolong trading intensity episodes. This specification, could potentially capture a more complex structure of trading intensity.

2.1.2 Conditional Density Specifications

Previous literature (e.g., Hujer and Vuletic, 2007; Kalaitzoglou and Ibrahim, 2013b) suggests that the impact of latent variables can be captured by the conditional density specification, assuming that it is manifested by the trading activity of different groups of traders. Because the trading intensity is a positive-support variable, it is customary for $f(\varepsilon_i)$ and consequently $f(S_i) = \theta_i f(\varepsilon_i)$, to be modelled using density functions belonging to the family of generalized Exponential distributions. The conditional density specifications employed in this study are:

Exponential (E) Distribution:

$$f(S_i | S_{i-1}, \dots, S_1; \tau) = 1/\theta_i \exp(-S_i/\theta_i) \quad (6)$$

Weibull (W) distribution:

$$f(S_i | S_{i-1}, \dots, S_1; \tau) = \gamma/S_i [S_i \Gamma(1 + 1/\gamma)/\theta_i]^\gamma \exp(-[S_i \Gamma(1 + 1/\gamma)/\theta_i]^\gamma) \quad (7)$$

where, $\Gamma(\cdot)$ is the gamma function of the distribution parameter $\gamma, \gamma > 0$. When $\gamma = 1$, W nests E. More specifically, E is a decreasing function, while W assigns higher (lower) probability in extreme values (very short or very long durations) when $\gamma < 1$ ($\gamma > 1$). In addition, in order to account for a mixture of market participants, assuming that their trading is captured by a different distribution, a mixture of Weibull distributions (S) is also employed:

$$f(S_i | J_i; \tau) = (\gamma_i/S_i) [S_i \Gamma(1 + 1/\gamma_i)/\theta_i]^{\gamma_i} \exp(-[S_i \Gamma(1 + 1/\gamma_i)/\theta_i]^{\gamma_i}) \quad (8)$$

The overall shape parameter of W, $\gamma_i = \gamma(J_i; \tau) = \gamma_1 + (\gamma_2 - \gamma_1) * G_1(J_i; g_1, j_1) + (\gamma_3 - \gamma_2) * G_2(J_i; g_2, j_2)$, is a function of a threshold variable, J_i , represented by S , and a vector of parameter coefficients $\tau = (\gamma_1, \gamma_2, \gamma_3, g_1, g_2, j_1, j_2)$. γ_i , is the weighted average of γ_1, γ_2 and γ_3 , i.e., the shape parameters in the respective three regimes determined by the threshold variable J_i , and the weights are determined by the smooth transition functions, $G_k(S_i; g_k, j_k) = (1 + \exp\{-g_k * (J_i - j_k)\})^{-1}$, with smoothness parameters $g_k, k = 1, 2$.

2.1.3 Estimation and Performance

The estimation method used is maximum likelihood using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) (1970) algorithm with numerical derivatives. In addition, in order to deal with heteroskedasticity of known form, the models have been estimated using robust errors that account for it.⁴ In addition, the different specifications are fitted using the Box-Jenkins methodology.⁵ Then the performance of the models is tested, by examining the in-sample “forecasts”, as another “goodness of fit” test. The “short-term” (1-step) forecasts are generated

⁴ The market gained complexity and liquidity gradually. Duration has been decreased over the years, while volume has been increased. Consequently, the values of the trading intensity measure we use are consistently lower over the years, introducing heteroskedasticity of known form.

⁵ The null hypothesis is that there is no remaining autocorrelation after 15 lags. The Q-stat (15) test statistic is distributed as a χ^2_{15} with a 5% critical value (for the Q-stats(15)) of 25.

according to the conditional mean specification, while “long-term” (10-step) forecasts are computed according to Dufour and Engle (2000).⁶

Their performance is evaluated, first, with the correlation (CORR) between the actual and the forecasted values of trading intensity, as well as by using a comparing the percentiles of their respective distributions (Q-Q plots). These are measures of the ability of the models to forecast large/small values when they are observed. Practically, they capture the ability of the models to forecast the level of liquidity. That could be very important for active traders, for dealers and for policy makers since liquidity is an important component of the price/cost of an asset. The second measure of performance is:

$$UNL = MSE / \bar{x}, \text{ where } MSE = 1 / (N_T + 1) \sum_{i=R}^{T-s} (x_{i+s} - \hat{x}_{i+s,i})^2 \quad (9)$$

where, UNL is the Unitized Loss, which measures the average squared error for every unit of forecasted values. MSE is the mean squared error, while \bar{x} is the realised trading intensity, \hat{x} is the forecasted trading intensity, and N_T is the number of observations. This measure captures the proportional error for every unit of forecasted value and it is important for active traders, such as market makers, who condition their signed price setting on expected incoming liquidity on each side of the spread.

Finally, previous literature (e.g., White, 2000; Hansen, 2005) considers performance measures that focus on the first moment of forecasts, such as the above, to be biased because they ignore their distributional assumptions. This is a relevant critique to the forecasting exercise above because it employs many competing models, which might consistently over- or under-forecast the actual series. For these purpose, an additional two tests are employed in order to investigate whether a model or a family of models exhibits a statistically significant improvement in overall performance. The first test is the Reality Check (RC) introduced by White (2000) and the second is the Superior Predictive Ability (SPA) test of Hansen (2005). Both are consider the differential function:

⁶ The long-term forecasts of the linear ACWD are given by: $E_i[x_{i+s}] = \omega \frac{1-(\alpha+\beta)^{s-1}}{1-(\alpha+\beta)} + (\alpha + \beta)^{s-1}\psi_{t+1}$. When OTC transactions are taken into account (ACWD-OTC) the following formula is used instead. $E_i[x_{i+s}] = \omega \frac{1-((\alpha+\zeta+D_{i-1})+\beta)^{s-1}}{1-((\alpha+\zeta+D_{i-1})+\beta)} + ((\alpha + \zeta * D_{i-1}) + \beta)^{s-1}\psi_{t+1}$, where $\hat{\omega}$, $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\zeta}$ are the estimated parameters of the respective ACWD model. The long term (10-step) forecasts for non-linear models are given by: $E_i[x_{i+s}] = \xi_1 \cdot \xi_2 \cdot \dots \cdot \xi_{s-1} \left[\exp \left\{ \omega \frac{1-\beta^{s-1}}{1-\beta} + \beta^{s-1} \ln(\psi_{t+1}) \right\} \right]$, where $\hat{\omega}$, $\hat{\alpha}$ and $\hat{\beta}$ are the estimated parameters from the ACD models and $s = 10$ is the forecasting horizon, and $\xi_m = E_{i+s-m-1}[\exp\{\beta^{m-1}g(\varepsilon_{i+s-m})\}]$ are parameters that are used to correct biased forecasts and can be estimated using the following sample moment; $\hat{\xi}_m = 1/R \sum_{i=1}^R \exp\{\beta^{m-1}g(\varepsilon_{i+s-m})\}$, where $m=1, \dots, s-1$ and $g(\varepsilon_i) = (\alpha + \zeta * D_{i-1})\varepsilon_i^{\delta'}$, where $\delta' = \delta_0 + \delta_1 * D_{i-1}$.

$$d_{k,i} = UNL_{0,i} - UNL_{k,i} \quad (10)$$

which $d_{k,n} > 0$ if model k has a better performance than a benchmark model 0, at transaction time $i = 1, \dots, n$. The two statistics examine whether one or a group of models exhibits a performance that is statistically different from 0. White (2000) suggests using the differential functions themselves, $\bar{d}_k = \frac{1}{n} \sum_{i=1}^n d_{k,i}$, while Hansen (2005) proposes the following standardized version:

$$RC_n = \max_{1, \dots, k} \{ \sqrt{n} \bar{d}_1, \dots, \sqrt{n} \bar{d}_k \} \quad (11)$$

$$SPA_n = \max \left[\max_{1, \dots, k} \left\{ \frac{\sqrt{n} \bar{d}_k}{\text{var}(\sqrt{n} \bar{d}_k)} \right\}, 0 \right] \quad (12)$$

Then the moments of the distributions of RC and SPA are estimated by using the stationary bootstrap of Politis and Romano (1994), by drawing $b = 1, \dots, B$ samples and re-estimating empirically equations (12) and (13).⁷ Then the rejection probabilities, are computed as:

$$p_{RC} = \sum_{b=1}^B \frac{1_{\{RC_n > RC_{b,n}\}}}{B} \text{ or } p_{SPA} = \sum_{b=1}^B \frac{1_{\{SPA_n > SPA_{b,n}\}}}{B} \quad (13)$$

where the null hypothesis of $E(RC_n) = E(\max_{1, \dots, k} \{ \sqrt{n} \bar{d}_1, \dots, \sqrt{n} \bar{d}_k \}) \leq 0$ is rejected for high values of p .

2.2 Sample

The data sample consists of all transactions and best-quote revisions of EUA Futures contracts with December maturity, traded on the European Climate Exchange (ECX), extracted from the Thomson Reuters Tick History (TRTH) database. Due to lack of prior information the dataset covers a period from 29/09/2008 until 30/04/2016. The information collected is the time stamp at millisecond accuracy, the price and the volume of each transaction. According to market microstructure practice the following data manipulation is applied.

First, observations with no or erroneous inputs, such as a 0 price, and/or out-of-trading-hours are omitted. Second, in order to consider the main price discovery process only the contracts with the highest trading volume are considered. The roll over dates from one maturity to the next are chosen using the following criterion: when the new contract exhibits higher daily volume for at

⁷ For estimating $\text{var}(\sqrt{n} \bar{d}_k)$, Hansen (2005) suggests various assumptions for the null distribution of SPA_n . Because of the nesting property of the models employed here, the expected value of SPA_n is considered to be 0. The moments of the statistics are based on a number of 1,000 samples, with a sample rate of 10% and dependence parameter of 25%.

least two consecutive days. The rollover days are 09/12/2008, 04/12/2009, 16/12/2010 and 14/12/2011 for phase II and 13/12/2013, 12/12/2014, 11/12/2015 for phase III. This produces two datasets, namely ECX II and ECX III, with 1,206,114 and 1,517,191 transactions, respectively. Third, trades with identical price, time stamp and trade direction are considered as one segmented trade and are aggregated. This thinning process results in two datasets of 864,368 and 835,289 transactions, for ECX II and ECX III, respectively. Fourth, the trade initiation is computed according to the ‘EMO’ rule (Ellis et al., 2000) and assigns a +1 (-1) to buyer- (seller)-initiated trades. Finally, the remaining durations are diurnally adjusted according to Dufour and Engle (2000).⁸

Table 1 presents the descriptive statistics of the two rolling series; ECX II and ECX III. The first notable observation is that the main asset traded in each phase is the contracts with the maturity in the same phase. Until the end of 2012, the number of trades in the ECX II series is far superior to ECX III. In addition, the number of observations before 2012 for ECX III accounts for about 10% of the total trades, but with an increasing number every year. Consequently, the price discovery for EUA Futures is expected to occur in the contracts with the closest maturity. The large difference in overall liquidity between the two series is also observed in the decreasing duration over the trading years. In both phases, the duration in the early trading period is significantly higher compared to the main trading period of each asset. For example, the average duration in 2010 for ECX III is 2,753 seconds and it drops to 35.61 seconds in 2013 – comparable to 40.74 seconds in 2012 for ECX II – when Phase III contracts become the main asset. The transaction size does not change significantly, so the differences are observed due to a higher number of transactions.

⁸ First, each trading day has been divided into five time intervals, each two hours long. The nodes (time benchmarks) used were 09:00:00 (32400), 11:00:00 (39600), 13:00:00 (46800), 15:00:00 (54000), 17:00:00 (61200). The values in parentheses are seconds after midnight. In order to avoid problems with convergence to infinity when the log-transformation is used, a unit, i.e., 1, is added to all raw durations. Then duration (dt) was regressed on the following time function, in order to obtain the expected duration, $E(d_i|f(t))$. $f(t) = \beta_0 + \sum_{h=1}^3 \sum_{j=1}^5 (\beta_j(t - n_j)^h)$, where $j=1\dots 5$ and stands for the five nodes used and $h=1\dots 3$. In addition, n_j ’s are five dummy variables which are calculated as: $n_j = \begin{cases} t - k_j, & \text{when } k_{j-1} < t < k_j \\ 0, & \text{elsewhere} \end{cases}$. Then, after estimating β_j ’s, the expected durations ($E(d_i|f(t))$) are calculated. Finally, using the following ratio, the durations are normalized (diurnally adjusted), $x_i = d_i/E(x_i|f(t))$, where, x_i stands for the diurnally adjusted durations, d_i stands for the raw durations and $E(x_i|f(t))$ is the expected duration derived by regressing the raw durations on the cubic spline $f(t)$.

KALAITZOGLOU Intensity of Trading in the European Carbon Market

Table 1: Descriptive Statistics

		Duration	Volume	Price	Price Chg	Duration	Volume	Price	Price Chg
Total ECXII #864,368 ECXIII #835,289	Avg	44.90	10.09	12.27	0.00	68.04	15.51	6.16	0.00
	Med	6.51	4.00	13.12	0.00	5.46	5.00	5.83	0.00
	Max	24878	2000	24.39	0.95	35742	5000	20.22	0.64
	Min	0.00	1.00	5.61	-1.09	0.00	1.00	2.46	-0.95
	Std	129.26	21.81	3.56	0.02	430.22	44.58	2.08	0.01
2008 ECX II #36,645	Avg	64.27	4.76	18.92	0.00				
	Med	8.20	2.00	18.25	0.00				
	Max	23901	216	24.39	0.92				
	Min	0.00	1.00	13.50	-1.09				
	Std	223.24	7.37	2.79	0.03				
2009 ECX II #202,288	Avg	45.39	6.03	13.24	0.00				
	Med	6.06	2.00	13.54	0.00				
	Max	24878	478	16.04	0.95				
	Min	0.00	1.00	8.05	-0.95				
	Std	134.29	9.99	1.63	0.02				
2010 ECX II #190,012 ECX III #3,108	Avg	48.31	10.73	14.56	0.00	2572.65	19.90	16.57	0.00
	Med	7.03	4.00	14.80	0.00	174.05	10.00	16.63	0.00
	Max	19584	995	16.73	0.27	35742	995	19.19	0.64
	Min	0.00	1.00	12.25	-0.26	0.00	1.00	14.53	-0.59
	Std	137.95	21.30	1.02	0.01	5254.79	40.81	0.91	0.07
2011 ECX II #216,302 ECX III #23,912	Avg	42.36	11.56	13.03	0.00	383.07	14.28	13.60	0.00
	Med	6.55	5.00	12.91	0.00	25.95	5.00	13.73	0.00
	Max	18002	689	18.18	0.51	24783	995	20.22	0.61
	Min	0.00	1.00	6.73	-0.88	0.00	1.00	7.26	-0.79
	Std	116.73	22.34	2.99	0.01	1069.67	28.19	3.37	0.03
2012 ECX II #219,121 ECX III #58,973	Avg	40.74	12.72	7.53	0.00	156.28	17.13	7.88	0.00
	Med	6.16	5.00	7.51	0.00	18.60	5.00	7.83	0.00
	Max	3744	2000	9.63	0.24	17448	2000	10.45	0.39
	Min	0.00	1.00	5.61	-0.21	0.00	1.00	5.93	-0.32
	Std	104.29	29.31	0.76	0.01	378.51	47.92	0.85	0.02
2013 ECX III #258,745	Avg					35.61	14.81	4.46	0.00
	Med					4.60	4.00	4.47	0.00
	Max					11755	5000	6.84	0.43
	Min					0.00	1.00	2.46	-0.95
	Std					104.29	44.83	0.71	0.01
2014 ECX III #241,635	Avg					38.16	15.56	5.90	0.00
	Med					4.78	5.00	5.93	0.00
	Max					5102	5000	7.57	0.43
	Min					0.00	1.00	3.71	-0.42
	Std					109.11	47.09	0.69	0.01
2015 ECX III #166,877	Avg					55.44	16.75	7.56	0.00
	Med					5.99	5.00	7.49	0.00
	Max					5088	3685	8.71	0.15
	Min					0.00	1.00	6.28	-0.18
	Std					151.36	47.38	0.58	0.01
2016 ECX III #82,039	Avg					33.80	14.11	5.68	0.00
	Med					5.29	5.00	5.37	0.00
	Max					2130	2000	8.33	0.17
	Min					0.00	1.00	4.62	-0.17
	Std					81.74	29.08	0.89	0.01

Note: Table presents the descriptive statistics of Duration (sec), volume (#contracts), Price (£) and Price Change (€) for the full sample and for each year over the sample period. This study uses the same sample as in Galariotis et al. (2018) and this table is an extract from this study. Source: Galariotis et al. (2018) paper.

In addition, prices consistently decrease over the years due to macroeconomic factors (e.g., Chevallier, 2009), and this pattern is also followed by price volatility. Average prices move from 18.92 in 2008 to 5.68 in 2016, while price change volatility decreases from 0.03 in ECX II and 0.07 in ECX III to 0.01, when the contracts are most actively traded. They too suggest that the main price discovery occurs when the contracts with the closest maturity are traded.

Combined, the descriptive statistics suggest that the overall liquidity of the market is a very important in price discovery. These unconditional, basic statistics suggest that information about the carbon market is resolved into the closest maturity asset and this is reflected on liquidity levels.

Consequently, liquidity fluctuations in these assets could be associated with patterns of observable or unobservable factors and their modelling could further assist in improving market efficiency in the form of a better price discovery due to identifiable patterns. Some partial evidence of that is the impact of the improved overall liquidity over the years on uncertainty. Price change volatility, i.e., uncertainty, decreases over the years alongside an increase in overall liquidity. Consequently, liquidity contributes to efficiency and therefore, a proper modelling of liquidity should be able to contribute to lower uncertainty and thus, better price discovery. The objective of this paper is aligned with this observation and it tries to identify which aspects of the trading activity, observable trading history or inferred signals from observing past trading history, contribute more in liquidity modelling.

3 Empirical Results

Estimation results are presented in Tables 2a and 2b for ECX phase II (ECX II) and ECX phase III (ECX III), respectively. With respect to modelling, almost all parameter estimates are statistically significant. In particular, estimates of α and β for ACWD in all four tables add up to less than, but close to, one. This indicates stationarity with high persistence, underlying autoregressive dynamics, and hence the appropriateness of the ARMA/ACD modelling framework, for trading intensity. Thus, intensity shocks have prolonged subsequent effects, in the sense that a shock does not die out quickly, but it persists over time.

The BCACWD specification also investigates persistence, but of higher order due to the implied non-linearities. δ is consistently less than one, in all estimations and this shows that larger shocks do not have a prolonged effect. Therefore, large illiquidity deviations in the form of lower actual market activity than expected, i.e., $\varepsilon > 1$, should not be expected to last for very long, since its impact on subsequent trading activity is asymmetrically less pronounced. This is an indication of a mean reversion process. In the opposite case, when the actual trading activity is more intense than expected, i.e., $\varepsilon < 1$, the impact of high liquidity shocks seems to be prolonged, due to an implied shorter expected trading intensity due to $\delta < 1$. This, asymmetric effect suggests

that slow markets are not expected to last for long, but an increased liquidity shock is expected to create an “episode” of more intense trading. This might be due to increase association with information (e.g., Kalaitzoglou and Ibrahim, 2013a, 2013b) and is consistent with prior literature. For example, Easley and O’Hara (1992) suggest that informed trading is manifested into higher trading intensity due to exploitable arbitrage opportunities, which should be expected to last until information is price resolved. Engle (2000) reports that there is also a significant price impact mainly due to increased frequency of trading, i.e., lower duration, but recent literature report that volume might also have a comparable impact (e.g., Bredin et al., 2014; Ibrahim and Kalaitzoglou, 2016).

Another main result relates to the question whether other observable factors, such as prior volume, or any other variable of interest, spur a different, than average, level of trading intensity for transactions that immediately follow. The hypothesis is that an increased volume might be associated with the arrival of price unresolved information and therefore it might cause a more intense subsequent trading. In parallel, given the illiquid character of the market, a high volume transaction, especially when it exceeds the available quantity on the opposite side of the spread, it might consume fast available liquidity and therefore, it might slow down temporarily the trading activity. The estimate of the coefficient for volume is consistently positive, i.e., $0 < x$ activity. The estimate of the coefficient for volume is consistently positive, i.e., $0 < \zeta < 1$, and suggests that larger volume leads to slower subsequent trading, in the form of increased duration, reduced volume or both.⁹ This is closer to the second hypothesis and partially explains the observations derived from the basic statistics. According to the basic statistics, price discovery improves with higher liquidity. The estimates of ζ suggest that this might be due to a significant adverse impact of requested volume on subsequent available liquidity. This should be expected to be less pronounced as overall liquidity improves. This is indeed the case with the estimates of ζ in ECX III being consistently lower than in ECX II.

Furthermore, the importance of volume, and extrapolating on this idea of other observable factors, on subsequent trading activity is also captured by the improvement in the values of the log-likelihood function, reported in the middle panel in each table. There is an improvement in the log-likelihood function and the associated BIC between the ACWD-volume and ACWD. This means that, according to log-likelihood, the ACWD-volume is a better model than the ACWD – a result that supports the hypothesis that volume is a significant determinant of the DGP of trading intensity and therefore, it improves the fitting of the model.

⁹ The variable used to proxy trading intensity is the volume weighted durations. Lower values (short duration or large volumes) indicate an active market. On the contrary, higher values refer to a decreased liquidity. Therefore, the positive ζ indicates a higher Θ , and thus the market is expected to be less liquid.

Table 2a: Estimation Results ECX II

	ACWD			ACWD-volume			BCACWD-volume		
	E	W	STM	E	W	STM	E	W	STM
ω	0.0216 (3.88)	0.0353 (7.06)	0.0681 (9.91)	0.0242 (5.12)	0.0364 (9.32)	0.0765 (16.91)	-0.1738 (-5.44)	-0.3676 (-9.44)	-1.4281 (-6.67)
α	0.0962 (5.43)	0.1480 (9.87)	0.2661 (13.45)	0.0766 (6.37)	0.1098 (10.81)	0.1671 (18.34)	0.2050 (5.25)	0.4201 (9.31)	1.5474 (6.97)
ζ				0.0403 (5.79)	0.0682 (10.75)	0.1217 (17.25)	0.0391 (5.38)	0.0672 (8.95)	0.1999 (21.73)
δ							0.3827 (15.96)	0.3236 (19.08)	0.1047 (7.72)
β	0.8840 (40.37)	0.8149 (42.89)	0.7014 (33.39)	0.8832 (50.27)	0.8220 (54.36)	0.7111 (54.24)	0.9693 (57.69)	0.9392 (46.77)	0.8632 (30.02)
γ (or γ_1)		0.6325 (25.10)	0.5693 (44.41)		0.6340 (31.06)	0.5740 (38.34)		0.6339 (26.63)	0.5711 (46.77)
γ_2			4.3632 (17.85)			4.1097 (40.47)			3.9931 (22.90)
γ_3			1.5326 (14.26)			0.9604 (14.12)			1.2185 (16.21)
j_1			0.6958 (8.25)			0.6874 (9.25)			0.6874 (8.93)
j_2			0.7521 (5.81)			1.1914 (7.24)			1.6405 (7.61)
g_1			1.1601 (6.15)			0.8698 (8.43)			1.1615 (9.54)
g_2			0.1856 (2.70)			0.2308 (2.78)			0.1934 (9.32)
$-L$	64541	43090	40438	64144	42858	40026	63701	42388	38481
BIC	1.5156	1.0124	0.9504	1.5065	1.0072	0.9409	1.4964	0.9964	0.9044
H(0)									
$\zeta=0$				33.49 (0.00)	115.62 (0.00)	297.40 (0.00)	28.93 (0.00)	80.17 (0.00)	472.03 (0.00)
$\delta=1$							292.56 (0.00)	1115.5 (0.00)	2080.7 (0.00)
$\gamma_1=1$		2987.18 (0.00)	133.35 (0.00)		138.63 (0.00)	161.28 (0.00)		2526.46 (0.00)	154.65 (0.00)
$\gamma_2=1$			156.64 (0.00)			253.53 (0.00)			168.56 (0.00)
$\gamma_3=1$			4.84 (0.05)			1.26 (0.11)			3.55 (0.06)

Note: Table 2a consists of three panels. The first presents the estimation results for the linear ARMA specification (ACWD), the volume augmented ARMA specification presented in equation (4) (ACWD-volume) and its non-linear extension (BCACWD-volume), using the Exponential (E), the Weibull (W) or a mixture of Weibull (STM) distributions. The values in (:) are t -statistics. The second panel presents the Bayesian information Criterion (BIC) and the log-likelihood function value (L). The last panel presents the hypothesis testing of the estimated values of the parameters δ and γ , with values in (:) being the p -values.

Table 2b: Estimation Results ECX III

	ACWD			ACWD-volume			BCACWD-volume		
	E	W	STM	E	W	STM	E	W	STM
ω	0.0468 (10.34)	0.0700 (11.30)	0.1233 (18.31)	0.0459 (8.12)	0.0675 (11.89)	0.1265 (20.91)	-0.4228 (-8.60)	-0.8570 (-10.23)	-0.3417 (-28.23)
α	0.1363 (16.87)	0.1855 (18.66)	0.2726 (27.02)	0.1302 (14.38)	0.1689 (18.89)	0.2251 (26.57)	0.5334 (9.26)	0.9928 (11.14)	0.3944 (26.03)
ζ				0.0235 (4.35)	0.0680 (19.79)	0.1332 (20.04)	0.0215 (2.79)	0.0189 (2.82)	0.0644 (9.87)
δ							0.2861 (12.91)	0.2022 (12.19)	0.1174 (6.95)
θ	0.8283 (68.24)	0.7508 (51.52)	0.6375 (46.20)	0.8306 (63.05)	0.7568 (56.34)	0.6417 (51.39)	0.9151 (64.79)	0.8741 (32.93)	0.6699 (53.82)
γ (or γ_1)		0.6372 (32.74)	0.5480 (61.71)		0.6370 (44.60)	0.5427 (49.51)		0.6415 (34.49)	0.5478 (61.79)
γ_2			3.8028 (16.19)			2.6642 (17.53)			3.8397 (19.29)
γ_3			1.3262 (13.38)			1.0878 (13.84)			1.3692 (14.93)
j_1			0.8892 (21.84)			0.6601 (37.92)			0.8919 (21.43)
j_2			1.0287 (7.71)			0.9022 (9.20)			1.0840 (10.40)
g_1			1.3293 (10.96)			1.2690 (10.83)			1.3406 (10.74)
g_2			0.1850 (13.04)			0.2445 (12.39)			0.1839 (13.19)
$-L$	83456	61053	56294	83418	60983	56041	81895	59996	56354
BIC	1.8293	1.3384	1.2343	1.8286	1.3370	1.2287	1.7953	1.3157	1.2358
$H(0)$									
$\zeta=0$				33.49 (0.00)	115.62 (0.00)	297.40 (0.00)	28.93 (0.00)	80.17 (0.00)	472.03 (0.00)
$\delta=1$							292.56 (0.00)	1115.5 (0.00)	2080.7 (0.00)
$\gamma_1=1$		2987.18 (0.00)	133.35 (0.00)		138.63 (0.00)	161.28 (0.00)		2526.46 (0.00)	154.65 (0.00)
$\gamma_2=1$			156.64 (0.00)			253.53 (0.00)			168.56 (0.00)
$\gamma_3=1$			4.84 (0.05)			1.26 (0.11)			3.55 (0.06)

Note: Table 2b consists of three panels. The first presents the estimation results for the linear ARMA specification (ACWD), the volume augmented ARMA specification presented in equation (4) (ACWD-volume) and its non-linear extension (BCACWD-volume), using the Exponential (E), the Weibull (W) or a mixture of Weibull (STM) distributions. The values in (:) are t-statistics. The second panel presents the Bayesian information Criterion (BIC) and the log-likelihood function value (L). The last panel presents the hypothesis testing of the estimated values of the parameters δ and γ , with values in (:) being the p-values.

Another main result relates to the level of generality required of the underlying distribution in order to take into account necessary features of the data's higher moments. Estimates of the coefficients γ and γ_k are statistically significant in all tables. Moreover, hypotheses tests reported in the bottom panels, reveal that $\gamma = 1$ can all be rejected apart from the low trading intensity regime, identified in the STM-(BC)ACWD. This shows that the Weibull (W) distribution is more appropriate than its reduced form of Exponential (E) distribution. More importantly, the optimum log-likelihood function value (L) reported in the middle panel of the tables is significantly (statistically) greater for models estimated with the mixture of Weibull distributions (STM) than with the W distribution than with the E distribution, in this order.¹⁰ In the Weibull models, always is $0 < \gamma < 1$ in a very narrow range. This indicates a monotonically decreasing hazard function. That means that the probability of a transaction to occur decreases with time. This is consistent with the findings in the non-linear version of the conditional mean specification, episodes of more intense trading. Here, the probability of a transaction decreases with time, indicating that it is more likely for transactions to be observed together with short time intervals in between. The probability to observe another transaction when time lapses decreases, suggesting that transactions might be clustered together with a persistent shorter trading intensity.

Next, the various combinations of models and distributions are examined, using the measures of in sample fit and forecasting accuracy. The empirical findings are presented in the three sections, a., b. and c., of Table 3. The major outcome is that a higher complexity, either in the form of a more sophisticated specification of the conditional mean or the conditional density is rewarded with better fitting and higher forecasting accuracy. However, a secondary finding, which is also the main proposition of this study is that, although the density specification improves performance, it is the conditional mean that determines the ranking of the model.

More specifically, the values of Q-stat(15) are all insignificant when volume is included as an explanatory variable and indicate the non-linear BCACWD-volume as the best fitting model. This shows that the contribution of volume to subsequent trading is very important, and

¹⁰ The nesting property of W to E facilitates testing using the Likelihood Ratio, which is distributed as chi-squared with a degree of freedom equal to the difference in the number of parameters between pairs of models. For example, in Table 2 the optimal Log-likelihood function values, L, are -43089.8 for W and -40438.0 for STM. These values give a Likelihood Ratio test statistic of 5303.46, which is far larger than the 5% critical value of 3.84. This indicates that the ACWD model with the mixture of Weibull distributions is significantly better (statistically and on the basis of likelihood) than with the same model with the Weibull distribution. The same is true in all tables for the ACWD-volume, which is a linear model, and the BCACWD-volume, which is a non-linear model. Thus, the STM provides a statistically better estimation in both phases and across linear and non-linear models.

Strictly speaking, the Likelihood Ratio cannot be used to test whether the non-linear BCACWD-volume fits better than the ACWD-volume because the former does not directly nest the latter. However, comparison is conducted later on the basis of in-sample forecast performance.

extrapolating on this idea, the contribution of observable variables and the patterns that they might create capture better the dynamics of the DGP of trading activity. In contrast, when volume is not considered some autocorrelation remains unaccounted for by the model. This ranking of the performance of the models considered is also consistent in all other measures in both 1-step and 10-step forecasts. More specifically, the empirical findings presented in Table 3a show that The UNL is consistently lower in models that consider observable factors as an integral part of the modelling process and this is consistent across samples and forecasting length. According to Tables 3b and 3c, the cross model differences in performance are significant and therefore, a more appropriate modelling of the conditional mean specification can contribute to forecasting accuracy in a statistically significant manner.

Furthermore, although the performance measures employed so far focus on the “average” performance of each model, they seem to point towards the right direction with respect to the superiority of the predicting ability of each model, even if the full distributional properties, e.g., the higher moments, of the forecasts are taken into consideration. This concern was raised in the literature (e.g., White, 2000) with respect to cross-model comparisons based on the same datasets and thus, of biased estimates due to data snooping. To this extend, the last panel of Table 3a presents the results of the RC and SPA tests, which test whether the “presumed” superior predicting ability is due to data snooping or because a model, or a family of models predicts better the original series than a benchmark model. As an additional level of robustness check, these tests are run with each one of the models as the benchmark model, alternatively. The findings suggest that the rejection probabilities significantly increase when the simple linear specification (ACWD) is employed as a benchmark, suggesting that it is really likely that another specification is superior in its ability to predict the original series. The rejection probabilities decrease significantly when the volume-enhanced specification is employed (ACWD-volume) and they reach their lowest level when the non-linear specification is used as a benchmark. This finding highlights that the probability of finding a better model (from the ones employed) in forecasting better the original series is at its lowest level when the non-linear specification is employed. This confirms that that a more sophisticated modelling of the conditional mean, explicitly modelling observable factors (e.g., volume) or patterns (non-linearity) can significantly (e.g., Tables 3b and 3c) improve liquidity modelling.

In addition, focusing on the density functions, increased flexibility is also rewarded with better ranking. STM performs better than W, which is better than E. However, the main improvement in fitting and forecasting comes from the mean equation specification. This is apparent in all tests employed so far, presented in Table 3. A more sophisticated distribution should be able to capture in a more flexible way the empirical distribution of trading intensity, because it should be more flexible, due to the additional parameters considered, in capturing the

Table 3a: Performance Evaluation

			Q-stat (15)					
			Rank	ECX II	P-value	Rank	ECX III	P-value
ACWD	E		9	71.40	(0.00)	9	69.78	(0.00)
	W		8	67.12	(0.00)	8	58.25	(0.00)
	STM		7	17.79	(0.27)	7	41.54	(0.00)
ACWD-volume	E		6	12.81	(0.62)	6	6.49	(0.97)
	W		4	2.57	(1.00)	4	5.77	(0.98)
	STM		2	0.42	(1.00)	2	4.99	(0.99)
BCACWD-Volume	E		5	8.23	(0.91)	5	6.26	(0.98)
	W		3	1.22	(1.00)	3	5.37	(0.99)
	STM		1	0.40	(1.00)	1	4.59	(1.00)

			UNL				CORR					
			Rank	ECX II	Rank	ECX III	Rank	ECX II	P-value	Rank	ECX III	P-value
1-step Forecasts	ACWD	E	5	9.1653	8	7.5821	8	0.0889	(0.01)	6	0.2169	(0.00)
		W	7	9.6919	9	7.7088	5	0.0924	(0.00)	8	0.2153	(0.00)
		STM	9	10.5178	7	7.5573	3	0.0954	(0.00)	5	0.2175	(0.00)
	ACWD-volume	E	6	9.2627	3	6.5539	7	0.0890	(0.00)	3	0.2179	(0.00)
		W	8	10.3353	6	6.9941	6	0.0921	(0.00)	4	0.2175	(0.00)
		STM	2	6.1559	5	6.8021	4	0.0953	(0.00)	9	0.2103	(0.00)
	BC-ACWD-Volume	E	3	6.4320	2	6.4694	9	0.0881	(0.01)	7	0.2166	(0.00)
		W	4	7.0481	4	6.7937	2	0.1001	(0.00)	2	0.2182	(0.00)
		STM	1	5.9759	1	6.4358	1	0.1132	(0.00)	1	0.2182	(0.00)
10-step Forecasts	ACWD	E	9	11.8853	9	25.7027	9	0.0127	(0.36)	9	-0.0025	(0.40)
		W	8	8.1354	8	22.7088	8	0.0169	(0.34)	8	0.0013	(0.40)
		STM	4	4.4104	7	18.8300	7	0.0187	(0.33)	7	0.0058	(0.39)
	ACWD-volume	E	6	5.6463	5	5.6463	6	0.0619	(0.05)	5	0.2031	(0.00)
		W	7	6.4553	6	6.4553	5	0.1003	(0.00)	6	0.1437	(0.00)
		STM	5	5.4895	4	5.4895	4	0.1110	(0.00)	4	0.2578	(0.00)
	BC-ACWD-Volume	E	3	2.8811	2	4.9802	3	0.2737	(0.00)	2	0.3131	(0.00)
		W	2	2.7701	3	5.0156	2	0.2915	(0.00)	3	0.2872	(0.00)
		STM	1	2.4146	1	4.3988	1	0.4078	(0.00)	1	0.5278	(0.00)

			AWD			ACWD-volume			BCACWD-volume		
			E	W	STM	E	W	STM	E	W	STM
ECX II	1-step	RC	0.77	0.76	0.87	0.65	0.81	0.27	0.26	0.31	0.15
		SPA	0.78	0.76	0.87	0.70	0.83	0.27	0.25	0.33	0.15
	10-step	RC	0.92	0.71	0.36	0.41	0.47	0.06	0.08	0.05	0.04
		SPA	0.94	0.72	0.36	0.40	0.45	0.07	0.10	0.05	0.03
ECX III	1-step	RC	0.89	0.90	0.89	0.37	0.48	0.47	0.19	0.19	0.18
		SPA	0.90	0.85	0.85	0.36	0.49	0.46	0.17	0.20	0.19
	10-step	RC	0.99	0.96	0.96	0.46	0.35	0.34	0.20	0.27	0.16
		SPA	0.99	0.95	0.93	0.47	0.33	0.34	0.22	0.25	0.12

Note: Table 3a consists of three panels. The first panel presents the in-sample ranking of the different specifications, presented in equations (4)-(8), according to the Q-statistic at 15 lags, with the initial hypothesis that there is no autocorrelation in the first 15 lags. The second panel presents the ranking of the models according to their 1-step and 10-step forecasting accuracy according to the unitized loss (UNL) and correlation (CORR), the values in (:) are t-statistics of H(0): CORR=0 performance measures. The last panel presents the rejection probabilities of the Reality Check (RC) and Superior Predictive ability (SPA) tests, Eq. (13), considering the model in each column as the benchmark model.

Table 3b.2: 2-Tail Tests of Cross-Model Performance Comparison-UNL

			ACWD			ACWD-volume			BCACWD-volume		
			E	W	STM	E	W	STM	E	W	STM
ECX II	ACWD	E	-	<i>7.76</i>	<i>7.76</i>	<i>-62.97</i>	<i>-36.01</i>	<i>-47.77</i>	<i>-68.14</i>	<i>-48.28</i>	<i>-70.20</i>
			-	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>
		W	32.25	-	-9.28	-70.72	-43.77	-55.52	-75.90	-56.04	-77.96
		(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	STM	32.25	50.57	-	-61.44	-34.49	-46.25	-66.62	-46.76	-68.68	
		(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	ACWD-volume	E	5.96	-26.29	-76.86	-	26.95	15.20	-5.17	14.69	-7.23
			(0.00)	(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
		W	71.65	39.40	-11.17	65.69	-	-11.76	-32.13	-12.27	-34.19
	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
STM	-184.29	-216.54	-267.11	-190.25	-255.93	-	-20.37	-0.51	-22.43		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	(0.61)	(0.00)		
BCACWD-vol	E	-167.38	-199.63	-250.20	-173.34	-239.03	16.91	-	19.86	-2.06	
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	(0.04)	
	W	-129.65	-161.90	-212.48	-135.62	-201.30	54.63	37.72	-	-21.92	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	
STM	-195.31	-227.56	-278.13	-201.28	-266.96	-11.03	-27.93	-65.66	-		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	
ECX III	ACWD	E	-	<i>-183.34</i>	<i>-183.34</i>	<i>-1228.20</i>	<i>-1178.66</i>	<i>-1237.80</i>	<i>-1268.99</i>	<i>-1266.82</i>	<i>-1304.59</i>
			-	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>	<i>(0.00)</i>
		W	-229.63	-	-237.53	-1044.86	-995.32	-1054.46	-1085.65	-1083.48	-1121.25
		(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	STM	-229.63	-228.11	-	-807.33	-757.80	-816.94	-848.13	-845.96	-883.73	
		(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	ACWD-volume	E	-382.06	-152.42	75.68	-	49.54	-9.60	-40.79	-38.62	-76.39
			(0.00)	(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
		W	-332.52	-102.88	125.22	49.54	-	-59.14	-90.33	-88.16	-125.93
		(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
	STM	-391.66	-162.03	66.08	-9.60	-59.14	-	-31.19	-29.02	-66.79	
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	(0.00)	(0.00)	(0.00)
BCACWD-volume	E	-551.39	-321.76	-93.65	-169.33	-218.87	-159.73	-	2.17	-35.60	
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.03)	(0.00)	
	W	-558.19	-328.55	-100.45	-176.13	-225.67	-166.53	-6.80	-	-37.77	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	(0.00)	
STM	-579.96	-350.33	-122.22	-197.90	-247.44	-188.30	-28.57	-21.77	-		
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	-	(0.00)	

Note: Table 3b presents the empirical results of cross-model comparison 2-tail t-tests for both samples employed and across models, presented in equations (4)-(8). The lower triangular matrix presents the results for 1-step forecasts ($H(0): UNL_{row}=UNL_{column}$), while the upper triangular matrix, in *italics*, presents the results for 10-step forecasts ($H(0): UNL_{column}=UNL_{row}$). The values in (:) are the associated p-values.

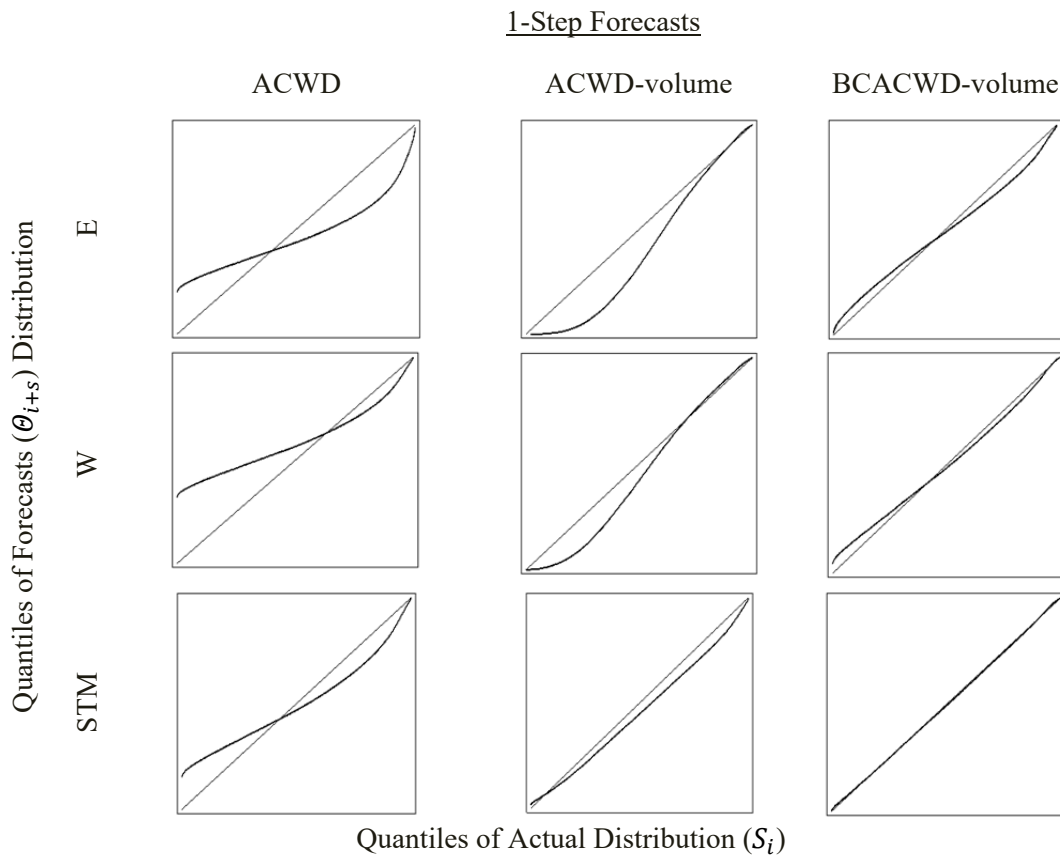
Table 3c: 2-Tail Tests of Cross-Model Performance Comparison-CORR

			ACWD			ACWD-volume			BCACWD-volume		
			E	W	STM	E	W	STM	E	W	STM
ECX II	ACWD	E	–	-5.20 (0.00)	-5.20 (0.00)	3.11 (0.00)	1.94 (0.05)	-20.98 (0.00)	-0.86 (0.39)	4.06 (0.00)	4.16 (0.00)
		W	11.13 (0.00)	–	7.09 (0.00)	8.31 (0.00)	7.14 (0.00)	-15.78 (0.00)	4.34 (0.00)	9.26 (0.00)	9.36 (0.00)
		STM	11.13 (0.00)	9.32 (0.00)	–	1.21 (0.22)	0.05 (0.96)	-22.87 (0.00)	-2.75 (0.01)	2.17 (0.03)	2.26 (0.02)
	ACWD-vol	E	0.13 (0.90)	-11.00 (0.00)	-20.32 (0.00)	–	-1.17 (0.24)	-24.08 (0.00)	-3.97 (0.00)	0.96 (0.34)	1.05 (0.29)
		W	10.15 (0.00)	-0.98 (0.33)	-10.29 (0.00)	10.02 (0.00)	–	-22.92 (0.00)	-2.80 (0.01)	2.12 (0.03)	2.22 (0.03)
		STM	20.13 (0.00)	9.01 (0.00)	-0.31 (0.76)	20.01 (0.00)	9.98 (0.00)	–	20.11 (0.00)	25.04 (0.00)	25.13 (0.00)
	BCACWD-vol	E	-2.64 (0.01)	-13.77 (0.00)	-23.08 (0.00)	-2.77 (0.01)	-12.79 (0.00)	-22.77 (0.00)	–	4.93 (0.00)	5.02 (0.00)
		W	35.28 (0.00)	24.15 (0.00)	14.83 (0.00)	35.15 (0.00)	25.13 (0.00)	15.14 (0.00)	37.92 (0.00)	–	0.09 (0.93)
		STM	76.95 (0.00)	65.82 (0.00)	56.50 (0.00)	76.82 (0.00)	66.80 (0.00)	56.81 (0.00)	79.59 (0.00)	41.67 (0.00)	–
ECX III	ACWD	E	–	11.87 (0.00)	11.87 (0.00)	650.01 (0.00)	462.21 (0.00)	823.11 (0.00)	998.12 (0.00)	916.18 (0.00)	1676.89 (0.00)
		W	13.41 (0.00)	–	14.50 (0.00)	638.13 (0.00)	450.33 (0.00)	811.24 (0.00)	986.25 (0.00)	904.31 (0.00)	1665.02 (0.00)
		STM	13.41 (0.00)	5.42 (0.00)	–	623.63 (0.00)	435.83 (0.00)	796.74 (0.00)	971.75 (0.00)	889.81 (0.00)	1650.52 (0.00)
	ACWD-volume	E	155.70 (0.00)	142.30 (0.00)	136.87 (0.00)	–	-187.80 (0.00)	173.10 (0.00)	348.11 (0.00)	266.17 (0.00)	1026.89 (0.00)
		W	277.14 (0.00)	263.73 (0.00)	258.31 (0.00)	121.44 (0.00)	–	360.90 (0.00)	535.91 (0.00)	453.97 (0.00)	1214.69 (0.00)
		STM	310.83 (0.00)	297.42 (0.00)	292.00 (0.00)	155.13 (0.00)	33.69 (0.00)	–	175.01 (0.00)	93.07 (0.00)	853.78 (0.00)
	BCACWD-volume	E	825.22 (0.00)	811.82 (0.00)	806.40 (0.00)	669.52 (0.00)	548.09 (0.00)	514.39 (0.00)	–	-81.94 (0.00)	678.77 (0.00)
		W	881.49 (0.00)	868.09 (0.00)	862.67 (0.00)	725.79 (0.00)	604.36 (0.00)	570.67 (0.00)	56.27 (0.00)	–	760.71 (0.00)
		STM	1249.52 (0.00)	1236.11 (0.00)	1230.69 (0.00)	1093.82 (0.00)	972.38 (0.00)	938.69 (0.00)	424.30 (0.00)	368.03 (0.00)	–

Note: Table 3b presents the empirical results of cross-model comparison 2-tail t-tests for both samples employed and across models, presented in equations (4)-(8). The lower triangular matrix presents the results for 1-step forecasts (H(0): CORRrow=CORRcolumn), while the upper triangular matrix, in italics, presents the results for 10-step forecasts (H(0): CORRcolumn=CORRrow). The values in (:) are the associated p-values.

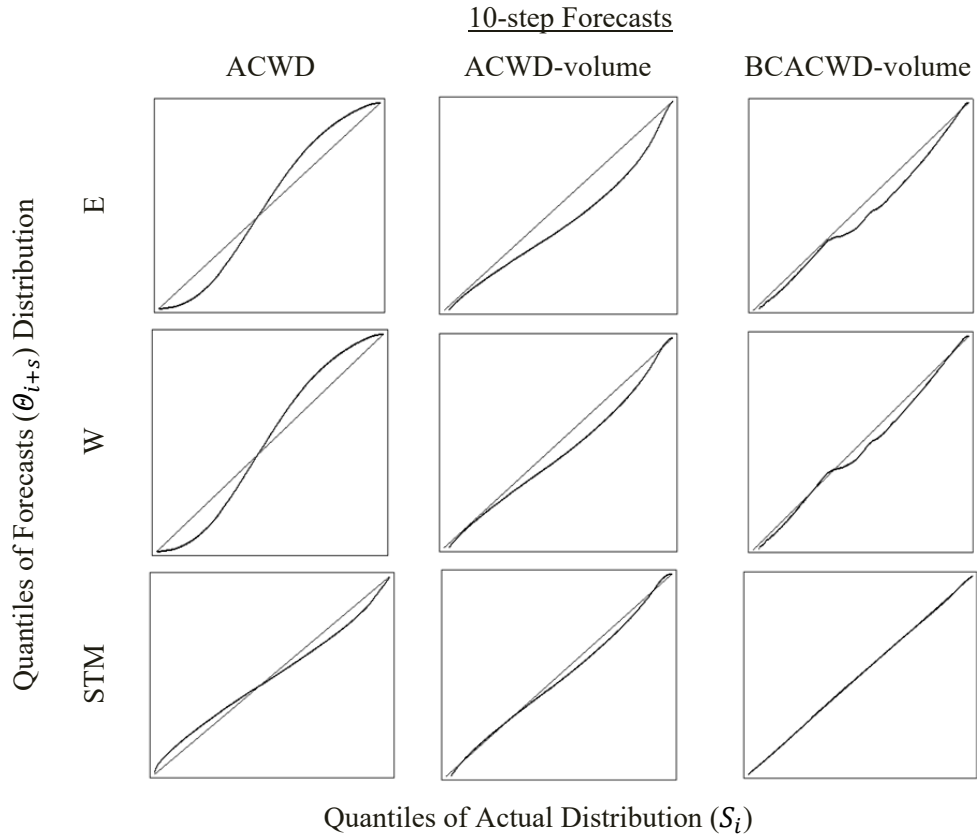
existence of observations that might arise due to unobservable factors, evidently not included in the conditional mean specification. Therefore, an increasing level of sophistication in the conditional density function should be a better proxy for latent or “invisible” factors in modelling, assuming that they will be manifested in the empirical distribution. According to Table 3, higher degrees of complexity indeed improve the fitting and forecasting of the model. Q(15), UNL and CORR are significantly (e.g., Tables 3b and 3c) improved progressively from E to W and to STM, with the latter exhibiting the best fitting in each conditional mean

Figure 1a. Q-Q Plots ECX II



Note: Figure 1a presents the Q-Q plots of the empirical distribution versus the distribution of forecasts generated by each model for the ECX II sample. This figure consists of 9 graphs for 1-step forecasts and another 9 graphs for the 10-step forecasts. Each graph compares the quantiles of the empirical distribution on the bottom axis to the quantiles of the distributions of the forecasts generated by the models in equations (4)-(8). The straight line (45°) represents a perfect fit.

Figure 1a continued. Q-Q Plots ECX II



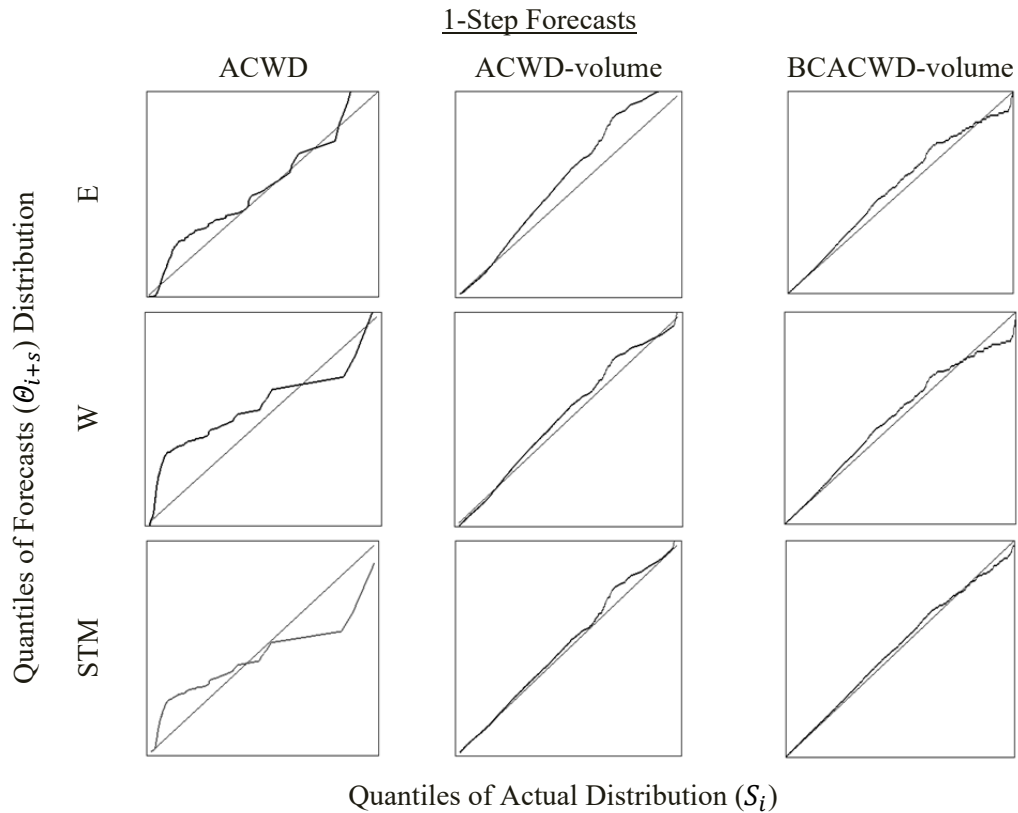
specification. However, it is the conditional mean specification that warrants a better performance and not the distributional assumption. For example, the STM distribution is undeniably a better fit across models and datasets, but the STM-ACWD does not fit and forecast better than a E-BCACWD-volume specification, suggesting that the model enhancements that come from modelling observable factors are superior to those that come from a better proxy for latent factors.

This is also confirmed in Figure 1, which compares the distribution of the actual trading intensity versus the distribution of the forecasts generated by each model. This is another full-distribution robustness check of the empirical findings that also considers whether a model systematically over- (above the straight line on the top quartiles or below the line on the bottom quartiles of the empirical distribution) or under-(below the straight line on the top quartiles or above the line on the bottom quartiles of the empirical distribution) forecasts the original series. Again, an enhanced conditional density function improves the fitting of the model (Q-Q plot line is closer to the straight line), but the main improvement comes from the conditional mean

specification. This also confirms Bauwens et al. (2004), in the sense that inferred latent signals summarized in a density function might not be sufficient in describing the DGP duration (or weighted duration), while variations of the conditional mean specification should contribute more.

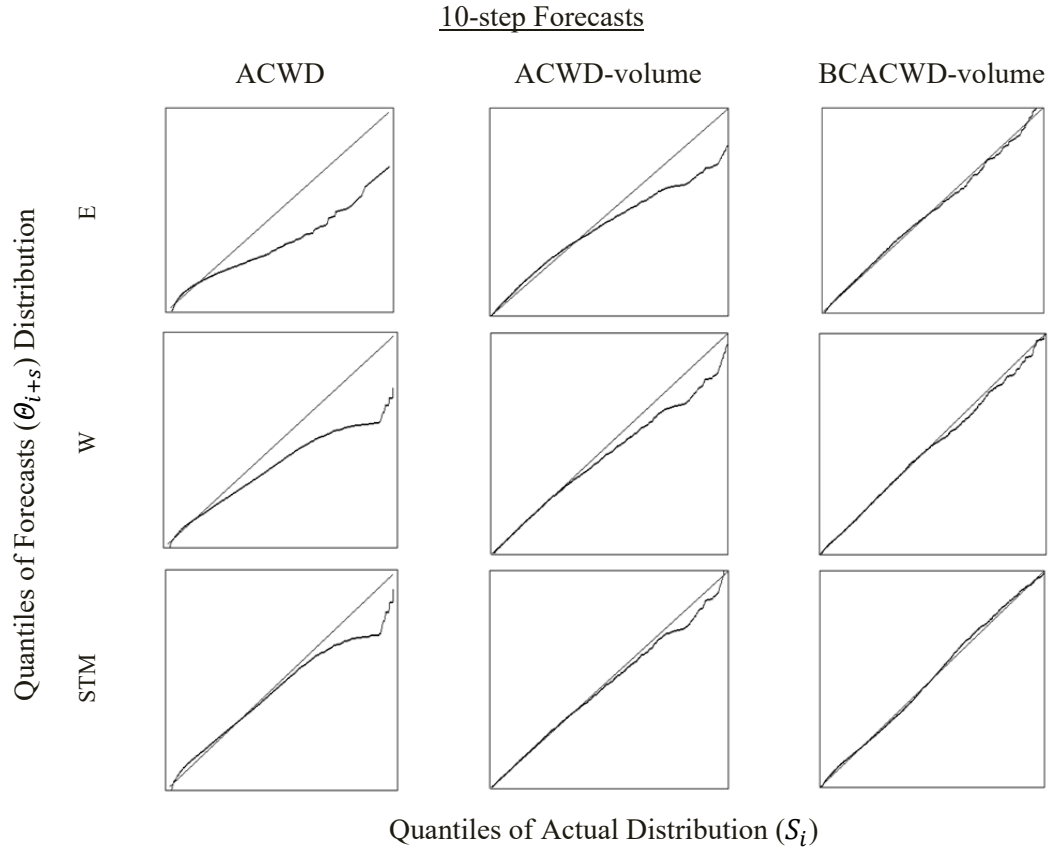
This finding is of utmost importance for liquidity modelling in the Carbon market and constitutes the major contribution of the current study. The implications of this result is that a more sophisticated specification of the conditional mean (of observable factors that can be considered in the structure of deterministic function that describes the DGP of trading activity) carry more weight in describing how the market activity evolves than a more sophisticated

Figure 1b. Q-Q Plots ECX III



Note: Figure 1b presents the Q-Q plots of the empirical distribution versus the distribution of forecasts generated by each model for the ECX III sample. This figure consists of 9 graphs for 1-step forecasts and another 9 graphs for the 10-step forecasts. Each graph compares the quantiles of the empirical distribution on the bottom axis to the quantiles of the distributions of the forecasts generated by the models in equations (4)-(8). The straight line (45°) represents a perfect fit.

Figure 1b continued. Q-Q Plots ECX III



specification of the density function (of unobservable factors that describe the range and the frequency of the outcomes of the decisions (transactions) of various types of market participants and how they are captured by the market-wide frequency distribution of trading activity). Consequently, any microstructure aspect that could increase the transparency of trading, would also increase the quantity and the quality of available observable factors and, therefore, it would also improve liquidity modelling. According to prior literature, this would have a beneficial impact on market efficiency in the form of lower frictions (e.g., Medina et al., 2014) and faster uncertainty resolution (e.g., Kalaitzoglou and Ibrahim, 2015).

This becomes increasingly important with the introduction of MiFID II, which aims at increasing transparency in the derivatives market, including EUA futures too. According to the findings presented in this study, the anticipated impact of this legislation in the carbon market, with respect to liquidity, is that it would improve the quantity and quality of available information, which could then be used to develop better liquidity forecasting models. This would

evidently improve liquidity based strategies – necessary for market making activities – which in turn would reduce execution risk – necessary for developing an optimal order-type selection strategy (e.g., Kalaitzoglou and Ibrahim, 2018; Galariotis et al., 2018) – and ultimately would narrow spreads – a sign of improved market efficiency and reduced trading frictions.

4 Conclusion

The EU ETS has evolved significantly since its inception in April 2005 from a rather illiquid trading environment to a reference point for carbon pricing in the future. Especially, the European Climate Exchange (ECX) hosted by the Intercontinental Exchange (ICE), which clears the majority of EUA futures transactions. The importance of this market in achieving European and in extension global emission targets is also important, but several market frictions, either in the form of macroeconomic implications (e.g., Daskalakis et al., 2009) or market structure issues (e.g., Chevallier, 2009), have posed significant challenges to the efficient functioning of the market. One of the main factors that has been found to determine the price discovery in the EU ETS is liquidity and its evolution (e.g., Bredin et al., 2014; Medina et al., 2014; Kalaitzoglou and Ibrahim, 2015). To that extent several studies consider its interactions with other factors, especially pricing, but no prior study has investigated whether it is observable or inferred information that determines the evolution of trading activity.

The current study focuses exactly on this idea and extends prior work done on liquidity modelling in the European Carbon market (e.g., Kalaitzoglou and Ibrahim, 2015). More precisely, the main objective of this paper is to empirically evaluate which factors are more significant in modelling liquidity. The direct distinction is between observable and inferred factors and their contribution to the fitting and forecasting of the different models is evaluated. This is done following recent advancements in duration modelling (e.g., Engle and Russell, 1998) and the debate about whether the specification of the conditional mean or the conditional density better describes the DGP of duration (e.g., Bauwens et al., 2004). In the current study the conditional mean specification is associated with the deterministic pattern of duration and thus, with the observable factors that can explain the evolution of trading activity. In parallel, drawing on previous literature (e.g., Hujer and Vuletic, 2007), the conditional density function is associated with the manifestation of the trading activity of different types of market participants and therefore it is used as a proxy for inferred-unobservable information.

The empirical findings suggest that both the conditional mean and density specifications improve the performance of the models, but it is the conditional mean specification that determines the overall performance. With an eye on MiFID II, the implications of this finding are increasingly important. It suggests that observable factors have a higher predicting power on subsequent liquidity and therefore they are its main determinants. Unobserved/inferred factors, summarized in the conditional density, have only a marginal contribution to fitting and

forecasting. Consequently, any market structure amendment that increases transparency, as MiFID II intends to do, would increase the quantity and the quality of the available observable factors and therefore, it would improve liquidity modelling. This in turn could improve market making practices, could enhance liquidity motivated trading strategies, could result in narrower spreads and in general, it could improve market efficiency. According to Viswanathan (2010), who argues that regulation needs to ensure viability and efficiency, by sustaining liquidity in functional levels, while allowing for trading innovation due to information asymmetry, this would assist the EU ETS in achieving its emission reduction targets.

References

- Bauwens, L., Giot, P., Grammig, J., and Veredas, D. (2004), A Comparison of Financial Duration Models via Density Forecasts. *International Journal of Forecasting*, 20(4), 589-609.
- Benz, E., and Trück, S. (2009), Modeling the Price Dynamics of CO2 Emission Allowances. *Energy Economics*, 31(1), 4-15.
- Bredin, D., Hyde, S., and Muckley, C. (2014), A microstructure analysis of the carbon finance market. *International Review of Financial Analysis* 34, 222-234.
- Broyden, C. G. (1970), The convergence of a class of double-rank minimization algorithms 1. General considerations. *IMA Journal of Applied Mathematics*, 6(1), 76-90.
- Chevallier, J. (2009), Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Economics*, 31(4), 614-625.
- Christiansen, A. C., Arvanitakis, A., Tangen, K., and Hasselknippe, H. (2005), Price determinants in the EU emissions trading scheme. *Climate Policy*, 5(1), 15-30.
- Daskalakis, G., Psychoyios, D., and Markellos, R. N. (2009), Modeling CO2 emission allowance prices and derivatives: Evidence from the European trading scheme. *Journal of Banking and Finance*, 33(7), 1230-1241.
- Dufour, A. and R. F. Engle (2000), Time and the Impact of a Trade. *Journal of Finance*, 55(6), 2467-2498.
- Easley, D., and O'Hara, M. (1992), Time and the process of security price adjustment. *Journal of Finance*, 47(2), 577-605.
- Ellis, K., Michaely, R., and O'Hara, M. (2000), The accuracy of trade classification rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis* 35(4), 529-551.
- Engle, R. F. (2000), The econometrics of ultra-high-frequency data, *Econometrica*. 68(1), 1-22.
- Engle, R. F. and J. R. Russell (1998), Autoregressive conditional duration: A new model for irregularly spaced transaction data, *Econometrica*. 66(5), 1127-1162.
- Fletcher, R. (1970), A new approach to variable metric algorithms. *The Computer Journal*, 13(3), 317-322.

- Galariotis, E., Kalaitzoglou, I. A., Kosmidou, K., Papaefimiou, S., and Spyrou, S. (2018), Could Market Making be Profitable in the European Carbon Market? *The Energy Journal*, (Forthcoming).
- Goldfarb, D. (1970), A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24(109), 23-26.
- Grammig, J. and Maurer, K. O. (2000), Non - monotonic hazard functions and the autoregressive conditional duration model. *The Econometrics Journal*, 3(1), 16-38.
- Hansen, P. R. (2005), A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23(4), 365-380.
- Hautsch, N. (2004), *Autoregressive Conditional Duration Models In Modelling Irregularly Spaced Financial Data*. Springer, Berlin, Heidelberg.
- Hujer, R., and Vuletić, S. (2007), Econometric analysis of financial trade processes by discrete mixture duration models. *Journal of Economic Dynamics and Control*, 31(2), 635-667.
- Ibikunle, G., Gregoriou, A., and Pandit, N. R. (2016), Price impact of block trades: the curious case of downstairs trading in the EU emissions futures market. *The European Journal of Finance*, 22(2), 120-142.
- Ibrahim, M. B., and Kalaitzoglou, I. A. (2016), Why do carbon prices and price volatility change? *Journal of Banking & Finance*, 63, 76-94.
- Kalaitzoglou, I. A., and Ibrahim, M. B. (2013a), Trading patterns in the European carbon market: The role of trading intensity and OTC transactions. *The Quarterly Review of Economics and Finance*, 53(4), 402-416.
- Kalaitzoglou, I. A., and Ibrahim, M. B. (2013b), Does order flow in the European Carbon Futures Market reveal information? *Journal of Financial Markets*, 16(3), 604-635.
- Kalaitzoglou, I. A., and Ibrahim, M. B. (2015), Liquidity and resolution of uncertainty in the European carbon futures market. *International Review of Financial Analysis*, 37, 89-102.
- Kalaitzoglou, I. A., and Ibrahim, M. B. (2018), Optimal Conditions for Order Type Selection. Available at SSRN: <https://ssrn.com/abstract=3077907>
- Mansanet-Bataller, M., Pardo, A., and Valor, E. (2007), CO₂ prices, energy and weather. *The Energy Journal*, 28(3), 73-92.
- Mansanet-Bataller, M. and Pardo, A. (2008), What you should know about carbon markets. *Energies*, 1(3), 120-153.
- Medina, V., Pardo, A., and Pascual R. (2014), The timeline of trading frictions in the European carbon market. *Energy Economics*, 42, 378-394.
- Politis, D. N., and Romano, J. P. (1994), The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303-1313.
- Rannou, Y., and Barneto, P. (2016), Futures trading with information asymmetry and OTC predominance: Another look at the volume/volatility relations in the European carbon markets. *Energy Economics*, 53, 159-174.

- Shanno, D. F. (1970), Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation*, 24(111), 647-656.
- Uhrig-Homburg, M., & Wagner, M. W. (2006), Market Dynamics and Derivative Instruments in the EU Emissions Trading Scheme-an Early Market Perspective. *Energy Environ*, 19(5), 635-655.
- Viswanathan, S. (2010), Developing an Efficient Carbon Emissions Allowance Market. Duke University., mimeo.
- White, H. (2000), A reality check for data snooping. *Econometrica*, 68(5), 1097-1126.
- Zhang, M. Y., Russell, J. R., and Tsay, R. S. (2001), A nonlinear autoregressive conditional duration model with applications to financial transaction data. *Journal of Econometrics*, 104(1), 179-207.