serie ec

WP-EC 2012-10

When Size Matters: Clustering in the European Carbon Market

Fernando Palao and Ángel Pardo

Ivie

Working papers

Working papers

Working papers

Versión: diciembre 2012 / Version: December 2012

Edita / Published by:
Instituto Valenciano de Investigaciones Económicas, S.A.
C/ Guardia Civil, 22 esc. 2 1º - 46020 Valencia (Spain)

WP-EC 2012-10

When size matters: Clustering in the European carbon market[1*]

Fernando Palao and Angel Pardo[**]

Abstract

This paper documents evidence of size clustering behavior in the European Carbon Futures Market and analyzes the circumstances under which it happens. Our findings show that carbon trades are concentrated in sizes of one to five contracts and in multiples of five. We have observed the existence of price clustering of prices ending in digits 0 or 5, and we have also proved that the more clustered prices have more clustered sizes. Finally, the analysis reveals that traders use a reduced number of different trade sizes when uncertainty is high, market liquidity is poor, and the desire for opening new positions is very strong.

Keywords: clustering, size, EUA, ECX, EU ETS.

JEL classification: G12.

Resumen

Este trabajo demuestra la existencia de agrupamiento de órdenes en determinados tamaños en el Mercado Europeo de derechos de emisión y analiza los determinantes bajo los cuales se detecta dicha acumulación o *clustering*. En concreto, nuestros resultados muestran que las transacciones de contratos de futuros sobre CO_2 se concentran en tamaños de uno a cinco contratos y en múltiplos de cinco contratos. Hemos observado la acumulación de transacciones en precios terminados en los dígitos cero y cinco. Además, se ha observado que la agrupación en precios y en tamaños son fenómenos complementarios. Finalmente, el análisis de los factores clave revela que los contratos de futuros sobre CO_2 utilizan un reducido número de tamaños de transacción para simplificar el proceso negociador cuando la incertidumbre es grande, la liquidez del mercado es muy baja y cuando el deseo de entrar en el mercado de futuros es extremadamente alto.

Palabras clave: acumulación, tamaño de la transacción, EUA, ECX, EU ETS.

Clasificación JEL: G12.

[*] The authors are grateful for the financial support of the Spanish Ministry of Education and Science (project CGL2009-09604), the FEDER, and the Cátedra Finanzas Internacionales-Banco Santander.

[**] F. Palao: Trading, Middle Office Department, Repsol Group, Madrid, Spain. A. Pardo: Department of Financial Economics, Faculty of Economics, University of Valencia, Valencia, Spain. Contact author: A. Pardo, angel.pardo@uv.es.

1. Introduction

Black (1971, p.30) indicates that an asset is perfectly liquid when (i) there are always bid and ask prices for the investor who wants to trade small amounts of assets and the difference between those prices is always small; (ii) an investor can trade a large amount of the asset over a long period of time at a price not very different from the current market price; and (iii) an investor can buy or sell a large block of stock immediately, but at a premium or discount that depends on the size of the block.

Following Harris (2003, p. 399), a trader must minimize the cost of trading a given size or, similarly, maximize the size she trades at a given cost. However, the ability to trade large sizes at low costs could be hindered when the size of the orders is concentrated at specific trade sizes. This empirical fact, known in the literature as the size clustering effect, has been recently observed in foreign exchange, equity, and index futures markets.

Moulton (2005) analyzes size clustering in the foreign exchange market and shows that customers trade more precise quantities at quarter-ends because this is when investors could have a stronger desire to satisfy their quantity demands. Alexander and Peterson (2007) carry out an analysis of trade size-clustering on the NYSE and Nasdaq stock markets and observe that trade-size clustering increasingly occurs in three levels - multiples of 500, 1,000 and 5,000 shares. Furthermore, they observe that price clustering and size clustering take place at the same time. Gwilym and Meng (2010) study the effect of size clustering in the UK FTSE100 index futures market. They detect that the number of daily distinct trade sizes increases with trade frequency and with intra-day volatility. They also observe that price clustering, defined as the tendency to observe certain trade prices more frequently than others, has a trade-off with size clustering, that is to say, less clustered prices have more clustered sizes. They suggest that price clustering and size clustering are substitutes for each other and that investors place orders rounding only one of the two variables. Blau *et al*. (2012) examine trade size of short sales and non-short sales also on the NYSE and Nasdaq stock markets and find that short sales cluster less on round sizes than do non-short trades. Finally, Verousis and ap Gwilym (2012) investigate differences in trading costs between the upstairs and the downstairs market in the London Stock Exchange, and find that price and size clustering tend to occur simultaneously rather than being substitutes.

The financial literature offers some theories to explain both price and size clustering. Firstly, the price negotiation hypothesis, introduced by Ball *et al*. (1985) and by Harris (1991), indicates that the presence of uncertainly leads the traders to round both trade sizes and their equilibrium prices, with the aim of minimizing the costs of the trading process. Secondly, there are some papers that suggest that the tendency to round sizes and prices is due to trader's preferences. This is the case of different behavioral hypotheses suggested by

Wyckoff (1963), Goodhart and Curcio (1991), and Ikenberry and Weston (2007), among others, that argue that investors prefer certain numbers over others without any rational explanation. By using a rounded set of numbers, the quantity of information that has to be processed by the traders is less. Combining these hypotheses, clustering appears because traders use a restricted set of prices and trade sizes to simplify their negotiations. Therefore, the higher the market volatility and the less the trading frequency, the higher the trading costs and the higher the level of clustering.

Finally, Hodrick and Moulton (2009) examine liquidity and how it affects the behavior of uninformed traders. One of the implications of their model states that in a market with many heterogeneous uninformed investors, the number of different sizes traded increases in accordance with their desire for satisfaction. If the desire for satisfaction is very high, they choose to trade a high range of different sizes. Therefore, the degree of size clustering should be very low on days in which the desire of portfolio managers to satisfy their negotiations is very intense.

The finding of coarse price grids, or price clustering, is common across a broad range of markets, including, among others, energy, water, foreign exchange, stock, bond futures, stock index futures, and carbon futures markets. However, as we have cited, the literature about the presence of size clustering is far less numerous.[1] This study offers the first analysis of observed patterns in European Union Allowances (EUAs) trade sizes. Specifically, the purpose of this paper is to document empirical evidence of size clustering behavior in the ECX EUA futures market and to understand under what circumstances it happens. The investigation of clustering in trade sizes could offer new insights into the liquidity of the European Carbon Futures Markets as long as its presence would be indicative of the fact that carbon traders might not negotiate their desired quantities at a given price. Additionally, the results of this study contribute to the debate by providing further evidence on whether price and size clustering are coincident or not.

The remainder of the paper is organized as follows. Section 2 describes briefly the European Carbon Market and the data used to perform this study. Section 3 analyzes the distribution of the trade sizes. Section 4 presents the findings on size clustering and its key determinants. Section 5 summarizes and concludes.

[1] See Brooks *et al*. (2013) and ap Gwilym and Meng (2010) for excellent reviews of the literature on price and size clustering, respectively.

2. Market structure and data

The European Union Emissions Trading Scheme (EU ETS) was launched in January 2005 and is, at the moment, the largest and most established regional cap-and-trade programme in the world. Established under Directive 2003/87/EC, the EU ETS limits the carbon dioxide emissions from approximately 12,000 European installations that include power generators and heavy industry. The EU ETS is operated under a cap-and-trade basis. Within this cap, companies receive emission allowances which they can sell to or buy from one another as needed. These rights to emit CO2 are known as European Union Allowances, or EUAs. An EUA unit is equal to one metric ton of carbon dioxide equivalent. At the end of each year, each company must surrender enough allowances to cover all its emissions, otherwise heavy fines are imposed. [2]

The EU ETS is organized in Phases. Phase I ran from 2005 to 2007; Phase II runs from 2008 to 2012 and coincides with the Kyoto Commitment Period; and Phase III, spanning 2013 to 2020, will cover new industries and have a prolonged compliance cycle. It will incorporate a centralized EU-wide allocation of allowances with a yearly linear decrease of the emissions cap of 1.74% per year, even beyond 2020.[3]

Several electronic markets currently offer trading on EUAs, however, the ICE ECX EUA Futures Market is considered as the benchmark as it concentrates by far the majority of the total trading volume. The ICE ECX market operates an electronic order-driven market with market makers and brokers. The daily session starts with a pre-open period of 15 minutes (from 6:45 a.m. UK local time) that finishes with a single call auction, where the opening price and the allocated volume are determined by an algorithm. During the continuous session, from 7:00 to 17:00, investors can submit limit orders, stop limit orders, market orders, and block orders. The futures contracts are traded in lots. Each lot equals 1,000 tonnes of CO_2 equivalent, that is, 1,000 EUAs. The minimum tick size was €0.05 until 27 March 2007 when it changed to €0.01. Futures contract ceases trading at 17:00 hours UK local time on the last Monday of the contract month.[4]

To carry out this study, we have chosen the complete lifespan of the ECX EUA futures contract with maturity in December 2011. This contract began to be traded on 23 March 2006, and, until its maturity on 19 December 2011, a total of 359,004 transactions took

[2] It is important to note that Australia plans to link its Emissions Trading Scheme to Europe's in 2015, cancel its floor price for CO2 permits and limit access to U.N. offsets for firms regulated by its market. The aim is to have the Australian and EU schemes fully linked from July 2018, which will constitute the union of two major emissions markets. See http://ec.europa.eu/clima/policies/ets/linking/index_en.htm for further details about this plan (last accessed on November 13, 2012).

[3] See http://www.ieta.org/overview for further details about the European Union Emissions Trading Scheme (last accessed on July 11, 2012).

[4] For further details on the EUAs futures contract, see the user guide of ECX Contracts at the www.theice.com (last accessed on July 11, 2012).

place.[5] Of note, the December 2011 ECX EUA futures contract represented the highest levels of trading ever registered on the ECX market until that date.

Specifically, for every screen trade, our database contains: the time stamp measured in GMT, the traded price in euros, the maturity of the contract, the traded volume, the daily settlement price, and the sign of the transaction (buyer or seller initiated). Following Alexander and Peterson (2007), a trade that has been buyer initiated is more likely to be followed by another buyer initiated order if the trades are rounded. Therefore, we will take into account the sign of the transaction to check if trades initiated by one of the sides could be more size clustered than trades initiated by the other side.

3. Trade size distribution

In this section, we begin by using the data on trade sizes to calculate their frequency. Figure I(a) shows the frequency of transactions with the same trade size as a function of the trade size, up to 50 futures contracts, for all the trades (359,004), independently of the sign of the transaction. Figures I(b) and I(c) present the distribution of all the buyer-initiated (175,491) and seller-initiated trades (183,513), respectively. Figure I(a) shows that the trades are concentrated in sizes of one to five contracts. We also observe spikes at five contracts and at size multiples of five that decrease steadily. Both buyer and offer initiated trades seem to be distributed in a similar way. The most common trade sizes are focused at the lowest trade size values and in multiples of five.

Table 1 presents the basic descriptive statistics of the three samples. The average trade size for the sample of all the transactions is 8.28, the minimum transaction size is one and the maximum is 900 contracts. However, the median is three, which gives an idea of the high concentration of trades around the lowest sizes. In fact, 88.16% of trades have a size lower than or equal to 20 contracts. These results are in line with those obtained by ap Gwilym and Meng (2010) for the FTSE100 futures contract. They suggest that this tendency to concentrate on small sizes could be the desire of traders to avoid trading large orders with a better-informed counterparty.

Figures 1(b) and 1(c) provide us with a first glance at the distribution of trade size depending on its sign. There appear to be little difference in the pattern of trade sizes between buyer- and seller-initiated trades. Next, we have formally tested the equality of means, medians and variances of both distributions with the parametric Anova *F*-test, the non-parametric Kruskal-Wallis test and the Brown-Forsythe's test, respectively. The results are displayed in Table 1.

[5] It is very unlikely that the change in the minimum tick size in March 2007 affects our results, given that only 74 out of 359,004 transactions took place before that change.

Figure 1. Frequency of the trades with the same trade size

Figure 1(a)

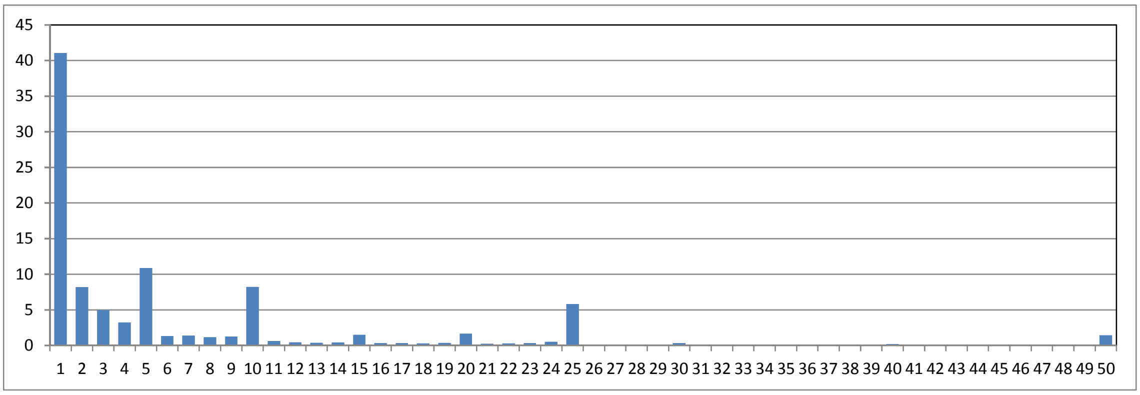


Figure 1(b)

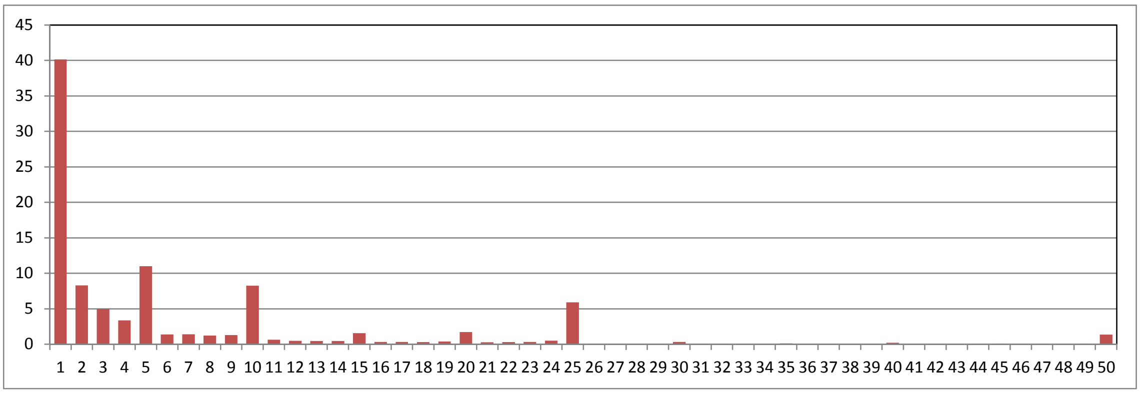
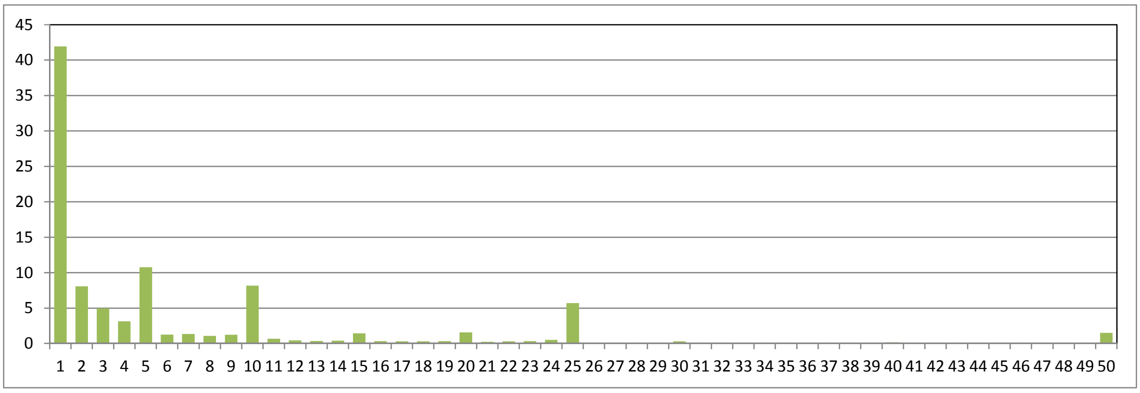


Figure 1(c)



Note: These figures show the frequency of the trades with the same trade size expressed in percentage (a) for all the trades, (b) for buyer-initiated trades, and (c) for seller-initiated trades for the ECX EUA futures contract with maturity in December 2011. A total of 359,004 transactions took place, 175,491 buyer-initiated and 183,513 seller-initiated trades.

The null hypotheses of equality of means, medians and variances are rejected at the 1% level. Firstly, the mean of the distribution of the buyer initiated trade sizes (=8.466) is statistically significantly different from the mean of seller initiated trade sizes (=8.112). Secondly, the Kruskal-Wallis test rejects the equality in medians. While the buyer median is three contracts, the offer median is equal to two. Thirdly, the Brown-Forsythe variance test reveals that the standard deviations of buyer (=18.651) and seller (=16.556) trade sizes are also statistically different. Finally, the non-parametric Wilcoxon/Mann-Whitney statistic (=9.444) confirms that the two series have different distributions at the 1% level. Therefore, the results of all these tests indicate that the distributions of trade sizes are conditioned by the sign of the order.

Following Alexander and Peterson (2007), in order to formally test if size clustering in all the samples is significant, we conduct a linear regression analysis:

$$\begin{aligned} ln\,Perc.\ Size_i \ &= \alpha_i + \beta_5 D5_i + \beta_{10} D10_i + \beta_{15} D15_i + \beta_{20} D20_i \\ &+ \beta_{25} D25_i + \beta_{upper\ 25} DM5_i + \beta_{Ln\ Size,\ i} LnSize_i + \varepsilon_i \end{aligned} \tag{1}$$

where the dependent variable is the natural log of the percentage of trades that occur at size *i* and we include as independent variables some dummy variables that will capture whether the trade size sample is affected by the round numbers, in particular as the 88.16% percent of trade sizes occurs in the range defined between one and twenty-five contracts, we include the dummy variables that will detect if the trade size is equal to 5, 10, 15, 20 or 25 contracts. Besides, we adapt from Blau *et al*. (2012) the variable *upper 25* which identifies trade sizes which are multiples of five and bigger than twenty-five. Finally, to control how the size clustering is affected by the size of the transaction, we include the variable $LnSize_i$ which is the natural logarithm of trade size *i* measured in number of contracts.

Table 2 shows the results of the round trade sizes regression analysis carried out using both ordinary least squares and the Newey and West correction that accounts for heteroskedasticity and serial correlation. In all the cases the adjusted-R squared exceed the level of 96% and all the coefficients are statistically significant at the 1% level. For the three cases, the dummy variables are positively related with the dependent variable, as we expected from the Figure 1 and from the results obtained by the previous literature. Regarding the variable *LnSize*, we find that its coefficient is negative and significant, meaning that the bigger the size of the transaction the lower the frequency of trades with such size.

Table 1. Descriptive statistics of trade sizes

	All trades	Buyer	Seller	Test
Mean	8.285	8.466	8.112	F- test: 36.324*
Median	3	3	2	KW- test: 89.194*
Std. Deviation	17.612	18.651	16.556	BF- test: 25.304*
Minimum	1	1	1	
Maximum	900	900	518	
Observations	359,004	175,491	183,513	

Note: This table shows the descriptive statistics of all the trade sizes for all the trades, for buyer-initiated trades, and seller-initiated trades. The sample period goes from 2006 to 2011. The *F*-test stands for the *F* statistic that tests the null hypothesis of equality of means of trade sizes. KW-test is the Kruskal-Wallis statistic that tests the null hypothesis of equality of medians of trade sizes. The BF-test is the Brown–Forsythe's statistic that tests the null hypothesis of equality of variances. *Denotes statistical significance at the 1% level.

Table 2. Round numbers analysis regression

$$\begin{aligned} ln\,Perc.\ Size_i \ &= \alpha_i + \beta_5 D5_i + \beta_{10} D10_i + \beta_{15} D15_i \\ &+\beta_{20} D20_i + \beta_{25} D25_i + \beta_{upper\ 25} DM5_i \\ &+\beta_{Ln\ Size,\ i} LnSize_i + \varepsilon_i \end{aligned}$$

	Full sample	Buyer	Offer
α	3.656	3.635	3.673
β_5	1.548	1.538	1.543
β_{10}	2.475	2.447	2.492
β_{15}	1.492	1.487	1.483
β_{20}	2.079	2.070	2.071
β_{25}	3.733	3.697	3.751
$\beta_{upper\ 25}$	2.489	2.395	2.568
β_{LnSize}	-1.748	-1.725	-1.765
Adj R^2	0.965	0.965	0.965

Note: This table shows the results of a regression analysis in order to test how round numbers affect trade sizes. $LnPerc.Size_i$ is the natural logarithm of the percentage of trade size *i*. $D5_i$, $D10_i$, $D15_i$, $D20_i$ and $D25_i$ are five dummy variables which take value 1 if the trade sizes *i* are equal to 5, 10, 15, 20 and 25, respectively, and 0 otherwise. $DM5_i$ takes value 1 if the trade sizes *i* is multiple of 5 *upper 25* and 0 otherwise. $LnSize_i$ is the natural logarithm of trade size *i* measured in number of contracts. This regression analysis has been applied to the full sample, buyer and offer sample. All the coefficients are significant at the 1% level.

4. Size clustering

4.1 Univariate analysis

Previous empirical evidence has created uncertainty regarding whether price and size clustering are complementary or substitutes. Alexander and Peterson (2007) and Verousis and ap Gwilym (2012) find that price and size clustering tend to occur simultaneously in stock markets. On the contrary, studies such as ap Gwilym and Meng (2010), for index futures markets, and Blau *et al*. (2012), for NYSE short sales, observe that less clustered prices have more clustered sizes, implying that price and size resolution may be substitutes. Given that Palao and Pardo (2012) show the existence of price clustering in December 2010 ECX EUA futures contract at prices ending in digits 0 or 5, we have also tested for its presence in the December 2011 futures contract, with the purpose of studying possible links between price and size clustering in the European Carbon Market.

First of all, to investigate the presence of price clustering, we focus on the distribution of the last decimal of the transaction price, in particular the frequency distribution of prices among x.x0 and x.x9. We analyze the price clustering as the frequency of the number of transactions occurring at each digit. Furthermore, following Brooks *et al*. (2013), we have also studied the price clustering as the frequency of the total amount of contracts traded at each digit.

The price clustering has been analyzed for the sample of all the transactions, for buyer-initiated trades, and for seller-initiated trades. Table 3 shows that the most clustered digit for the three samples is 0 followed by 5. This issue is observed at each digit both in the number of trades and in the sum of contracts. It is notable that when we take into account the total number of contracts traded, the percentage observed for trade prices at x.x0 and x.x5 is higher. This suggests that customers not only trade more frequently at digits 0 and 5 but also, when they trade at these digits, they place a higher amount of contracts than in the rest of the cases.

Additionally, we have applied the Goodness of Fit Chi-squared statistic, shown in Panel B, to test the null hypothesis of no difference between the observed distribution and the expected distribution. The Goodness of Fit Chi-squared statistic, showed in Panel B as *GOF,* is defined as:

$$GOF = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{N-1}$$

where O_i is the observed frequency of the last digit; E_i is the expected frequency under a uniform distribution, and *GOF* is the distributed Chi-square with *N-1* degrees of freedom under standard conditions. In all the cases, the tests reject the null hypothesis at the 1% level, confirming the presence of price clustering both in the number of trades and in the sum of contracts in the December 2011 ECX EUA futures contract at prices ending in digits 0 and 5.

Table 3. Price clustering

Panel A. Distribution of last digit of the price and the amount of associated contracts

Pricing grid	All sample		Buyer		Offer	
	% Trades	% Contracts	% Trades	% Contracts	% Trades	% Contracts
x.x0	15.03	18.65	14.82	17.97	15.23	19.33
x.x1	8.17	7.49	9.16	8.35	7.22	6.63
x.x2	9.00	8.22	9.35	8.66	8.67	7.77
x.x3	9.14	8.48	9.36	8.75	8.93	8.20
x.x4	8.73	7.97	7.99	7.21	9.43	8.73
x.x5	13.58	15.54	13.20	15.27	13.93	15.81
x.x6	8.84	8.01	9.52	8.87	8.18	7.15
x.x7	9.03	8.60	9.14	9.14	8.92	8.06
x.x8	9.53	8.97	9.44	8.64	9.62	9.29
x.x9	8.96	8.07	8.01	7.12	9.86	9.02
Total	359004	2974389	175491	1485754	183513	1488635
x.x0 and x.x5	28.61	34.19	28.02	33.24	29.17	35.14

Panel B. Clustering tests

	All sample		Buyer		Offer	
	% Trades	% Contracts	% Trades	% Contracts	% Trades	% Contracts
GOF	17379.34	393088.78	7771.94	174294.15	10724.64	231185.30
p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Note: Panel A shows the frequency of the number of transactions occurring at each digit (% Trades) and the frequency of the total amount of contracts traded at each digit (% Contracts) expressed in percentage for all the trades, for buyer-initiated trades, and for seller-initiated trades for the ECX EUA futures contract with maturity in December 2011 traded at the 1 cent interval. Panel B presents the Goodness of Fit Chi-squared statistic (*GOF*) and its p-value that tests the null hypothesis of no difference between the observed distribution and the expected distribution.

Next, we define different variables and perform different tests in order to check for the presence of size clustering in carbon markets. We follow the methodology proposed both by Moulton (2005) and by ap Gwilym and Meng (2010). Specifically, we define the variable *Size* as the daily number of different trade sizes; *Count* as the daily trading frequency, and *Volume* as the daily volume.

A simulated example of the daily trading activity in a fictitious market in Table 4 will help to clarify these variables. Panel A presents all the intraday trades for two consecutive days. Panel B shows how these trades are classified according to different subsamples. In our case, we perform the analysis for the full sample and two subsamples that takes into account prices that end in digit 0, in digit 5, in digits 0 or 5, and in digits different from 0 and 5. Finally, Panel C shows the percentages for each subsample.

For example, for Day 1, there are three transactions recorded (see Panel A), the variable *Count* indicates three transactions, while two trades of size one and one trade of size ten constitute two *Sizes* on the same day (see Panel B). Panel C deserves special attention, because the proportion of size one (2/5) is bigger in the sample of prices ended in 0 than in the whole sample (3/10) as happens for size one. It is explained because, for each sample, we only consider the sum of the different sizes corresponding to each sample. The full sample has ten trades while the subsample of digit-0 has five. For this reason, additively cannot be supposed when comparing the different samples of *Size*.

The reason for employing the daily number of distinct trade sizes (the variable *Sizes*) instead of the variable trade sizes is because we are interested in analyzing the amplitude of the range of the trade sizes and not the frequency of the observations for each trade size. Proceeding in this way, we avoid that a trading day with a high number of trades could determine the size clustering level. For instance, we do not mind if the trade size quantity equal to one repeats 20 times, though we do mind if an investor can trade at such quantity.

Next, we perform different tests in order to check size clustering in carbon trades. We examine all trade sizes of all intraday screen transactions of the December 2011 ECX EUA futures contract and we calculate the daily variables *Size*, *Count* and *Volume*. Table 5 presents some descriptive statistics and some tests about the three trading-related variables for the different samples considered. Panel A and B show the *size, count* and *volume* measures for prices ending in 0 or 5 and for prices ending in digits different from 0 or 5, respectively. Furthermore, for each panel, we show descriptive statistics for all the transactions, for buyer-initiated trades, and for seller-initiated trades. Panel C and D present the p-value of the Kruskal-Wallis statistic that tests the null hypothesis of equality in the medians for different samples analyzed.

Table 4. Example of trading-related variables

Panel A: Trade Volume Information on two days

Day 1		Day 2	
Price	**Contracts**	**Price**	**Contracts**
15.05	10	15.02	10
15.00	1	15.03	2
15.01	1	15.01	3
		15.00	3
		16.00	18
		15.00	3
		15.00	1

Panel B: Trade volume classification for the three variables

Day 1	Sizes	Count	Volume
Full	2	3	12
0 digit	1	1	1
5 digit	1	1	10
0 and 5	2	2	11
Different	1	1	1
Day 2	**Sizes**	**Count**	**Volume**
Full	5	7	40
0 digit	3	4	25
5 digit	-	-	-
0 and 5	3	4	25
Different	3	3	15

Panel C: Trade size percentage by sample

Sizes	Full	0 digit	5 digit	0 and 5	Different
1	3/10	2/5	-	2/6	1/4
2	1/10	-	-	-	1/4
3	3/10	2/5	-	2/6	1/4
10	2/10	-	1	1/6	1/4
18	1/10	1/5	-	1/6	-
Count	10	5	1	6	4

Note: This table shows the classification of the trades according to the variables used in the size analysis. Panel A provides an example of the trade negotiation on two days. Panel B shows how these transactions are distributed according to the distinct trade sizes (*Sizes)*, the frequency of observations (*Count)* and the total volume of contracts traded *(Volume)* for the full sample, for trades where the last decimal is 0, for trades where the last decimal is 5, for trades where the last decimal is 0 or 5, and trades whose last decimal is different from 0 and 5. Panel Care 0 or 5, and trades whose last decimal is different from 0 and 5. Panel C shows the percentage of trade sizes over the two days for each sample.

Table 5. Descriptive statistics of daily trades distribution

Panel A: Prices ending in digits 0 or 5

	All transactions			Buyer initiated			Offer initiated		
	Size	**Count**	**Volume**	**Size**	**Count**	**Volume**	**Size**	**Count**	**Volume**
Mean	13.38	106.65	1056.06	10.35	56.85	570.97	10.31	60.07	587.09
Median	8,00	24,00	180,00	6,00	14,00	113,00	6,00	14,00	99,00
Min	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Max	86,00	1383,00	21124,00	65,00	636,00	11945,00	61,00	747,00	9179,00

Panel B: Prices ending in digits different from 0 and 5

	All transactions			Buyer initiated			Offer initiated		
	Size	**Count**	**Volume**	**Size**	**Count**	**Volume**	**Size**	**Count**	**Volume**
Mean	18.43	266.15	2032.61	19.59	140.66	1104.53	19.59	143.64	1066.89
Median	14,00	63,00	486,00	6,00	34,00	284,00	6,00	36,00	272,00
Min	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Max	81,00	2369,00	24426,00	94,00	1133,00	12040,00	95,00	1236,00	12386,00

Panel C: Kruskal-Wallis test between prices ending in digits 0 or 5 and prices ending in digits different from 0 and 5

	All transactions			Buyer initiated			Offer initiated		
	Size	**Count**	**Volume**	**Size**	**Count**	**Volume**	**Size**	**Count**	**Volume**
0&5 vs different	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Panel D: Kruskal-Wallis test among different samples

	All vs Buyer			All vs Offer			Buyer vs Offer		
	Size	**Count**	**Volume**	**Size**	**Count**	**Volume**	**Size**	**Count**	**Volume**
Digits 0&5	0.001	0.001	0.001	0.001	0.001	0.001	0.764	0.999	0.867
Digits different from 0 and 5	0.309	0.001	0.001	0.281	0.001	0.001	0.957	0.223	0.708

Note: *Size* refers to the daily number of distinct trade sizes; *Count* refers to the frequency of observations; *Volume* is the daily volume of contracts traded. Panel A and B show the *size, count* and *volume* measures for prices ending in 0 or 5 and for prices ending in digits different from 0 and 5, respectively. For each panel, we present descriptive statistics for all the transactions, for buyer-initiated trades, and for seller-initiated trades. Panel C and D show the p-value of the Kruskal-Wallis statistic that tests the null hypothesis of equality in the medians for the different samples compared.

Panel A of Table 5 shows that the mean and median of the daily number of different trade sizes is statistically lower for prices ending in 0 or 5 (13.38 and 8) than for prices ending in digits different from 0 or 5 (18.43 and 14). Panel C confirms that this difference is statistically significant at the 1% level. Therefore, more clustered prices have more clustered sizes, suggesting that price and size clustering are complementary. The frequency of observations and the daily volume of contracts traded are significantly higher for prices ending in digits different from 0 and 5. This is an expected result given that we are summing the total trading volume for each category. However, the average trade size for the digits ending in 0 or 5 is 9.9 versus 7.63 for the rest of the prices. When we divide the sample between buyer- and offer-initiated trades, we cannot reject the equality between the medians for the two samples. Therefore, the sign of the order affects the trade size of the order (see Table 1) but it does not influence the variable *Size* (see Panel D of Table 4). This is the reason why, from now on, the analysis will be focused only on the sample composed of all the transactions.

4.2 Multivariate analysis

Finally, based on previous empirical evidence obtained for other assets, a multivariate analysis is carried out to determine the key factors which affect size clustering in carbon prices. To study possible links between price and size clustering, and following ap Gwilym and Meng (2010), we have split the data set into two parts in order to capture any differences between observations with prices ending in x.x0 and x.x5 and those with prices ending in the remaining digits. To do this, we have defined *Dt* as a dummy variable that takes the value 1 for observations where prices end in digits different from 0 or 5, and 0 otherwise. Observations are indexed by t (t = 2032, across 1016 trading days).

The following model has been estimated using both ordinary least squares and the Newey and West correction that accounts for heteroskedasticity and serial correlation problems:

$$Size_t = \alpha_t + \beta_1\sigma_t + \beta_2 Count_t + \beta_3 Trade\ Size_t + \beta_4 R3_{H,t} + \beta_5 R3_{M,t}$$
$$+\beta_6 R3_{L,t} + \beta_7 D_t + \beta_8 D_t\sigma_t + \beta_9 D_t Count_t + \beta_{10} D_t Trade\ Size_t$$
$$+\beta_{11} D_t R3_{H,t} + \beta_{12} D_t R3_{M,t} + \beta_{13} D_t R3_{L,t} + \epsilon_t \qquad (2)$$

The dependent variable that represents the level of size clustering is the variable $Size_t$ which refers to the daily number of distinct trade sizes. The higher the measure, the lower the degree of size clustering. σ_t stands for an estimation of the intraday volatility that has been calculated following the measure proposed by Parkinson (1980):

$$\sigma_t = \frac{1}{4log2}(logH_t - logL_t)^2$$

where H_t is the highest and L_t are the lowest traded prices on day t. We will use volatility as a proxy of uncertainty. According to Harris (1991), more information arrival implies more volatility and a wider range of trade sizes. Moulton (2005) observes higher volatility associated with more sizes traded in the majority of the currencies she analyzed and ap Gwylim and Meng (2010) also observe this relationship for the FTSE100 Index futures contract. $Count_t$ is the number of daily trades for each sample. Following Moulton (2005), there cannot be more sizes than trades in a day. Furthermore, Alexander and Peterson (2007) shows that trade-size rounding tends to increase when trading activity is abnormally heavy. Therefore, the number of trades should be positively related to the number of distinct trade sizes. $Trade\ Size_t$ is calculated as the daily average trade size, i.e. the sum of the total amount of the trades divided by the number of the total transactions on such day. As we have seen, our preliminary results suggest that the average trade size is higher for the most clustered prices, and by introducing this variable into the regression, we can test whether the daily average trade size influences the range of the different trade sizes.

Finally, motivated by the theoretical paper by Hodrick and Moulton (2009), we have introduced three dummy variables. Their paper examines liquidity and how it affects the behavior of portfolio managers. One of the implications of their model is that in a market with many heterogeneous uninformed investors, an asset will trade at more distinct quantities when investors have a stronger desire to satisfy their exogenous demands, where ‘‘at more distinct quantities’’ refers to more variation in the quantities ($Size_t$) traded, not necessarily more trades or more total volume. Assuming this theory, the degree of size clustering on days with extreme desire would be negatively linked with the desire of uninformed investors (portfolio managers) to satisfy their negotiations.

We apply the $R3_t$ measure proposed by Lucia and Pardo (2010) as a proxy to study the behavior of the portfolio manager activity in the European Carbon Market. This measure is defined as the ratio between the change in the open interest and the daily trading volume over a day t. The ratio has no dimension, and can take any value ranging from –1 to +1. A positive (negative) number indicates that the number of open (closed) positions is greater than the number of closed (open) positions. After calculating the ratio for all the trading days, we have constructed three variables. $R3_{Ht}$, $R3_{Mt}$ and $R3_{Lt}$ which take value 1 when $R3_t$ is in the intervals [0.95, 1], [-0.025, 0.025] and [-1, -0.95], respectively. The first dummy variable indicates days in which the opening of new positions outnumbers by far the closing of positions; the second variable identifies those days with an abnormal number of intraday traders, while the last variable takes into account days in which the traders are massively closing positions.

The results of the regression analysis are presented in Table 6 and show a high explanatory power, given that the adjusted R^2 is 82.7%. After controlling for all the possible determinants of size clustering, the dummy variable for prices ending in digits different from 0 or 5 is positive and statistically different from zero at the 1% level, indicating that there is a higher range of different size quantities in such prices.

Therefore, as Table 5 suggests, less clustered prices have less clustered sizes. We find that volatility in clustered prices is negatively related with the dependent variable at the 5% level, which means that when uncertainty increases, carbon traders prefer to trade at a small range of sizes. No significant link is observed between average trade size and size clustering. However, there is a positive and significant relationship between the daily number of transactions and the daily number of distinct trade sizes that is not counterbalanced for prices ending in digits different from 0 or 5. The overall result of these findings supports both the price negotiation and the behavioral hypotheses.

Finally, it is important to note the results obtained when we observe how size clustering behaves under different investor decision scenarios. The number of trade sizes is not affected by intraday trading activity. However, the coefficient of the dummy variable that represents massive opening positions ($R3_{H,t}$) is negative, while the variable representing massive closing positions ($R3_{L,t}$) is positive, both at the 1% level. The same pattern is observed for the coefficients of the interaction variables. This means that on days when carbon traders have an extreme desire to open new positions, they concentrate the size of their trades, but when their desire is great for cancelling old positions, they prefer to use a wide range of sizes. Therefore, this result backs the theory by Hodrick and Moulton (2009) which states that, in a market with many heterogeneous uninformed investors, the number of different sizes traded increases with their desire for satisfaction.

Table 6. Determinants of size clustering

$$Size_t = \alpha_t + \beta_1\sigma_t + \beta_2 Count_t + \beta_3 Trade\ Size_t + \beta_4 R3_{H,t} + \beta_5 R3_{M,t} \\ +\beta_6 R3_{L,t} + \beta_7 D_t + \beta_8 D_t\sigma_t + \beta_9 D_t Count_t + \beta_{10} D_t Trade\ Size_t \\ +\beta_{11} D_t R3_{H,t} + \beta_{12} D_t R3_{M,t} + \beta_{13} D_t R3_{L,t} + \epsilon_t$$

	Coefficient	Std. Error	t-Statistic	P-value
α	7.006	0.502	13.965	0.000
σ_t	-73.776	32.996	-2.236	0.025
$Count_t$	0.069	0.002	29.589	0.000
$Trade\ Size_t$	-0.001	0.001	-1.043	0.297
$R3_{H,t}$	-4.978	0.486	-10.240	0.000
$R3_{M,t}$	0.709	0.661	1.072	0.284
$R3_{L,t}$	0.241	0.023	10.570	0.000
D_t	3.255	0.438	7.430	0.000
D_t x σ_t	8.785	30.007	0.293	0.770
D_t x $Count_t$	-0.035	0.002	-20.791	0.000
D_t x $Trade\ Size_t$	0.000	0.001	-0.092	0.926
D_t x $R3_{H,t}$	-2.737	0.682	-4.013	0.000
D_t x $R3_{M,t}$	-0.159	0.782	-0.203	0.839
D_t x $R3_{L,t}$	3.894	0.312	12.477	0.000
R-squared	0.828	Adjusted R-squared		0.827
F-statistic	698.15	Prob (F-statistic)		0.000

Note: $Size_t$ refers to the daily number of distinct trade sizes. σ_t is the daily volatility. $Count_t$ is the number of trades per day. $Trade\ size_t$ indicates the daily average trade size. $R3_H$, $R3_M$ and $R3_L$ are dummy variable that take value 1 when *R3* is in the intervals [0.95, 1], [-0.025, 0.025] and [-1, -0.95], respectively. *D* is a dummy variable equal to 1 if contract prices end in a price different from 0 or 5, and 0 otherwise.

5. Conclusions

This study investigates the presence and the key determinants of the size clustering in the ICE ECX futures market taking into account intraday transactions data. We have found evidence of a tendency for carbon trades to cluster in small sizes and in round numbers multiples of five contracts. The finding of the evidence of this effect in carbon trades implies that carbon market participants may not be able to trade the desired quantity easily.

We have confirmed the existence of price clustering in the December 2011 ECX EUA futures contract at prices ending in digits 0 or 5, both in the number of trades and in the sum of contracts, and we have also proved that more clustered prices have more clustered sizes, suggesting that price and size resolution in the European Carbon Market are complementary. Finally, the analysis of the key determinants of size clustering suggests that carbon traders use a reduced number of different trade sizes to simplify their trading process when uncertainty is high, market liquidity is poor, and the desire for opening new positions is very strong. We interpret all these findings as being supportive of both the price negotiation and the behavioral hypotheses.

BIBLIOGRAPHY

Alexander, G.J., and M.A. Peterson, 2007, An analysis of trade-size clustering and its relation to stealth trading, *Journal of Financial Economics* 84, 435–471.

ap Gwilym, O., and L. Meng, 2010, Size Clustering in the FTSE100 Index Futures Market, *Journal of Futures Markets* 30, 432–443.

Ball, C.A., W.A. Torous and A.E. Tschoegl, 1985, The degree of price resolution: the case of the gold market, *Journal of Futures Markets* 5, 29-43.

Black, F., 1971, Toward a fully automated stock exchange, *Financial Analysts Journal* 27, 29–35.

Blau, B.M., B.F. Van Ness and R.A. Van Ness, 2012, Trade-size and price clustering. The case of short sales and the suspension of price tests, *The Journal of Financial Research* XXXV, 159–182.

Brooks, R., E. Harris and Y. Joymungul, 2013, Price clustering in Australian water markets, *Applied Economics* 45, 677-685.

Goodhart, C, and R. Curcio, 1991, The clustering of bid-ask prices and the spread in the foreign exchange market, Working Paper, (London School of Economics Financial Markets Group Discussion Paper 110*)*.

Harris, L., 1991, Stock price clustering and discreteness, *Review of Financial Studies* 4, 389–415.

Harris, L., 2003, *Trading and exchanges: Market microstructure for practitioners* (Oxford University Press, New York).

Hodrick, L.S., and P., Moulton, 2009. Liquidity: Considerations of a Portfolio Manager, *Financial Management,* Spring, 59-74.

Ikenberry, D.L., and J.P. Weston, 2007, Clustering in US stock prices after decimalization, *European Financial Management* 14, 30-54.

Lucia, J.J., and A. Pardo, 2010, On measuring speculative and hedging activities in futures markets from volume and open interest data, *Applied Economics* 42, 1549-1557.

Moulton, P., 2005, You can't always get what you want: Trade-size clustering and quantity choice in liquidity, *Journal of Financial Economics* 78, 89–119.

Palao, F., and A. Pardo, 2012, Assessing price clustering in European Carbon Markets, *Applied Energy* 92, 51–56.

Parkinson, M., 1980, The extreme value method for estimating the variance of the rate of return, *Journal of Business* 53, 61–65.

Verousis, T. and O. ap Gwilym, 2012, Trade Size Clustering and the Cost of Trading at the London Stock Exchange, *International Review of Financial Analysis,* doi: 10.1016/j.irfa.2012.08.007.

Wyckoff, P., 1963, *Psychology of Stock Market Timing* (Prentice Hall, Englewood Cliffs, NJ).

Ivie
Guardia Civil, 22 - Esc. 2, 1º
46020 Valencia - Spain
Phone: +34 963 190 050
Fax: +34 963 190 055

Website: http://www.ivie.es
E-mail: publicaciones@ivie.es