Catching Cartels

An evaluation of using structural breaks to detect cartels in retail markets

Author: Sebastian Mårtensson

Supervisor: Håkan J. Holm

2nd year Master Thesis in Economics

Department of Economics

Lund University

June 2021

Abstract

This thesis evaluates a new method for detecting cartels based on the idea that cartels create structural breaks in their price generating process. The method is evaluated along certain criteria and is further explored using more detailed data on Spanish retail fuel prices. A novel contribution with this thesis is to present an improvement of the method. The method fails to detect cartels and fails to live up to some of the criteria presented.

Acknowledgments

I would like to thank Håkan J. Holm for his advice and guidance in writing this thesis. It would not have been possible without him and the experience would have not been as interesting. I would also like to thank Meera and Albert for giving valuable feedback and for friends and family for their encouragement.

Table of Contents

1. Introduction

Every year, cartels sap welfare from consumers, businesses, and public procurement. Each country's competition authority has a great task of stopping this, by dissuading cartel formation through fines and making prosecution easier through leniency programs. These reactive measures are one way of tackling cartels and have most likely led increased disassembling of cartels (Miller 2009). There may however still exist cartels which remain undiscovered by competition authorities. Acting as a complement to the reactionary programs is the proactive programs which typically seeks out cartels and the markets they are likely to operate in. One of these proactive programs is the use of screens which have been advertised by national competition authorities (OECD, 2013a) and by academics (Harrington, 2008). One of these screens, behavioural screens, are based on market outcomes and is the topic of thesis. More specifically, I examine if structural breaks can be caused by a transition to collusive price setting and whether this phenomenon can be used to detect cartels.

The reason why firms should cause structural breaks is based on that collusion in a market has a goal of manipulating the price, through price fixing, non-aggression agreements or standardization of business practices (Genesove & Wallace, 2001). Firms therefore affect how prices are set which differs between collusive and competitive periods. The behavioural screen which is used in this thesis uses the structural breaks test as developed by Bai and Perron (2003) which can detect unknown amounts and dates of structural breaks. The method has been suggested as a viable option for detecting and dating screens (Crede, 2019; Boswijk, Bun & Schinkel, 2017) since it is able to detect and date breaks in the price generating process and should detect when competitive and collusive pricing-schemes change.

This method is tested on Spanish diesel retail prices in the provinces of Madrid and Barcelona which saw periods of non-aggression agreements and exchange of strategic information surrounding prices. Different levels of aggregation are used, based on aggregated price data on the province and brand level for daily observations. This is done in order to explore how detailed the data needs to be in order for collusion to be detected by the structural breaks test and if more detailed data is advantageous.

I find that the method does not work very well due to low predictive power and low usability. The high degree of false negatives is believed to come from the exclusion of demand side variables and the low market share by the colluding firms. Using more detailed data makes the interpretation more complex while providing little value in predictive power.

The rest of the thesis has the following disposition. The next section describes on what criteria a good method should fulfil. The third section covers the Spanish diesel retail market and background for the sector. The fourth section then summarises previous research and explains the structural breaks test by Bai and Perron (2003). The fifth section covers the methodology, how the hypotheses are tested and what is expected to happen. The sixth section presents the results and the seventh discusses it. The eight section concludes.

2. Method evaluation

When evaluating a screen for detecting cartels, there is a discussion on what it means for it to be good or successful. Success can take different dimensions such as being the most cost-effective in finding cartels, being costly to beat and requiring a low amount of human input in terms of specifying a model or interpreting the results (Harrington, 2008). The most important requirement for a successful screen is however the ability to correctly predict the firms involved in a cartel.

The ability to correctly predict cartels is based on how well the method works and is the most crucial aspect to consider. If a screen is unable to correctly predict cartels and competitive firms then the other requirements are not relevant. The performance of a screen can be measured in terms of having a low rate of false positives and false negatives. If a cartel can easily avoid being detected, then it faces the same problem as being unable to predict cartels. A screen which is costly to beat can affect cartels in multiple ways such as reducing the internal stability and making cooperation more perilous or inhibiting the actions that a cartel can make. This can take the form of less coordination between cartel members or not as high prices. If a screen can find cartels at a high rate and they are unable to avoid it, then cost effectives becomes more important, in order to optimize the work of the competition authority.

Cost effectiveness could be seen as maximizing the number of found cartels constrained by some budget by a competition authority. While one obviously wants to correctly predict competitive and collusive firms or periods to the greatest extent, there are different costs associated with getting it wrong. Decreasing the rate of false negatives may benefit consumers as they are less affected by cartels and over a shorter period of time. Prioritizing finding cartels comes with a downside of increasing the rate of false positives and causing unnecessary investigations. Incorrectly predicting that a cartel is present, false positive, is considered to incur greater costs for competition authorities, than not finding a present cartel (Huber & Imhof, 2019). The resources spent on an investigation which leads to nothing could have been better spent on other, more probable cases or other efforts undertaken by the competition authority. If emphasising a low rate of false positives, then the rate of undiscovered cartels naturally goes up as one becomes more sceptical towards finding cartels. A low rate of false positives has been prioritized in previous research for this reason.

A screen may be used in multiple stages of a proactive program. Screens which are easy to implement can be used as the first line of investigation, continuously looking at markets. The need for a low amount of human input and easily available data is therefore necessary to make it scalable and cost efficient. Requiring a lot of human input, interpretation, and niche data, takes resources which could be better spent elsewhere.

All together, a screen for cartels should firstly, and most importantly, be effective in finding cartels and having it be costly to beat the screen. Secondly, the ability to affect the rate of false positives and false negatives and economize the screening procedure makes the competition authority more effective. Thirdly, the ability to optimize resources by requiring low human input and easily available data is needed to make the screen scalable and widely used. The results are evaluated using these requirements in the discussion.

3. Spanish fuel retailers

This section covers some background on the Spanish diesel retail sector and the data associated with this study.

3.1 Background

In 2015, 2 major Spanish diesel retail firms were sanctioned with fines for anticompetitive arrangements(Concurrencies, 2015). These included price-fixing, exchanging strategic information about upcoming price changes and other violations of Spanish competition law (Concurrencies, 2015). The 2 companies, Repsol and CEPSA, were fined for entering into a non-aggression agreement between July 2011 and August 2011 as well as exchanging strategic information in the following period. Both Repsol and CEPSA are present in all Spanish provinces and account for a large market share (OECD 2013b).

The large market share of these companies, and the market structure of the Spanish retail fuel market have been criticized as being highly prone to collusion. Prices in Spain have a history of being higher than their European counterparts and the Spanish competition authority have found cases of collusion in the past (Gonzalez & Mora, 2019). A study by Perdiguero and Jimenez (2021) found that companies deliberately and systematically lowered prices on days where price information was collected, summarized, and compared with other countries. They suspect that this was to avoid further scrutiny, as high prices would warrant further competition enhancing actions in the sector.

3.2 Data

Data relating to those areas affected by their collusion were gathered from a previous study examining asymmetric fuel prices responses (Balaguer & Ripollés, 2016) and consists of a station index, time index, date, daily price with and without taxes, what brand of station as well as coordinates and postal codes of that particular station. The data covers the two provinces of Madrid and Barcelona from 10th of September 2010 to 26th of February 2013 and 10th of June 2010 to 25th of November 2012 respectively. In addition to the data from Balaguer & Ripollés (2016), prices of Brent crude oil in the EU were collected from the Federal Reserve of St. Louis and converted to euros.

Table 1 shows all brands above 3% market share in Madrid. Market share is calculated as the share of stations under each brand. The collusive firms, REPSOL and CEPSA have the two highest shares, amounting to 49% in Madrid. REPSOL and CEPSA have a far lower combined market share of 33.2% in Barcelona.

Table 1 – Madrid Summary statistics

Brand	Mean Price	Average Price	Market Share	Nr Stations
REPSOL	1.277	0.01	33.5%	221
Other	1.269	0.01	19.4%	128
BP	1.276	0.01	8.3%	55
CEPSA	1.276	0.01	16.1%	106
CAMPSA	1.276	0.01	9.5%	63
SHELL	1.272	0.01	6.4%	42
GALP	1.277	0.009	6.8%	45
Total	1.277	0.01	100%	660

Table 2 – Barcelona Summary statistics

Group	Mean Price	Average Price	Market Share	Nr Stations
OTHER	1.277	0.012	40.7%	294
REPSOL	1.291	0.012	23%	166
BP	1.295	0.011	4.3%	31
CAMPSA	1.289	0.012	4.8%	35
PETROCAT	1.284	0.012	4.3%	31
GALP	1.29	0.011	7.9%	57
CEPSA	1.289	0.011	10.2%	74
MEROIL	1.279	0.011	4.7%	34
Total	1.277277	0.012	100%	722

4. Previous research and theory

This section surveys previous research on cartel pricing and builds a theory for why structural test should be able to detect anticompetitive behaviour. This section also provides an overview of how collusion in a homogenous product, such as diesel, can be modelled and what is expected to happen upon collusion. It also provides an overview of screens and the test proposed by Bai and Perron (2003).

4.1 The effect of collusion on price

When modelling competition in oligopolies, there are two major models one can use depending on the market. In markets where firms compete in quantity, the ability to quickly change the amount of sold goods can be described as Cournot competition. In markets where firms compete in prices, such as fuel where output easily follows from the set price, firms can flexibly change their price to earn the highest profit. In this Bertrand competition setting, where goods are homogenous, whichever firm sets a price slightly lower than its competitors serves all of the customers. A race to the bottom then ensues in the Bertrand Nash equilibrium where each firm sets their price equal to the marginal cost, earning zero profit. This Bertrand setting is not very realistic as firms can still compete in price but without earning zero profit. The spatial Bertrand solves this problem by introducing a dimensional aspect.

In the spatial Bertrand model, each firm is placed along a road where they sell their good. This can be likened to the solid lines in figure 1 in which 3 firms are placed along the road on the horizontal axis. The firms charge a price, represented on the vertical axis which is the same as its competitors. Consumers are evenly spread along the horizontal line and experience a disutility of traveling from their position towards the placement of the closest firm. The disutility of traveling is represented by the slanted lines and the final price a consumer pays consists of the firm price and the disutility of traveling (Pepall, Richards & Norman, 2014). Each individual firm can then be thought of as monopolist within their own geographical area where the edge of their market is defined by which price they and their neighbour competitor set. The spatial Bertrand model can be used to explain how partial collusion may be profitable as well as how it can lead to price leadership in a spatial Bertrand market.

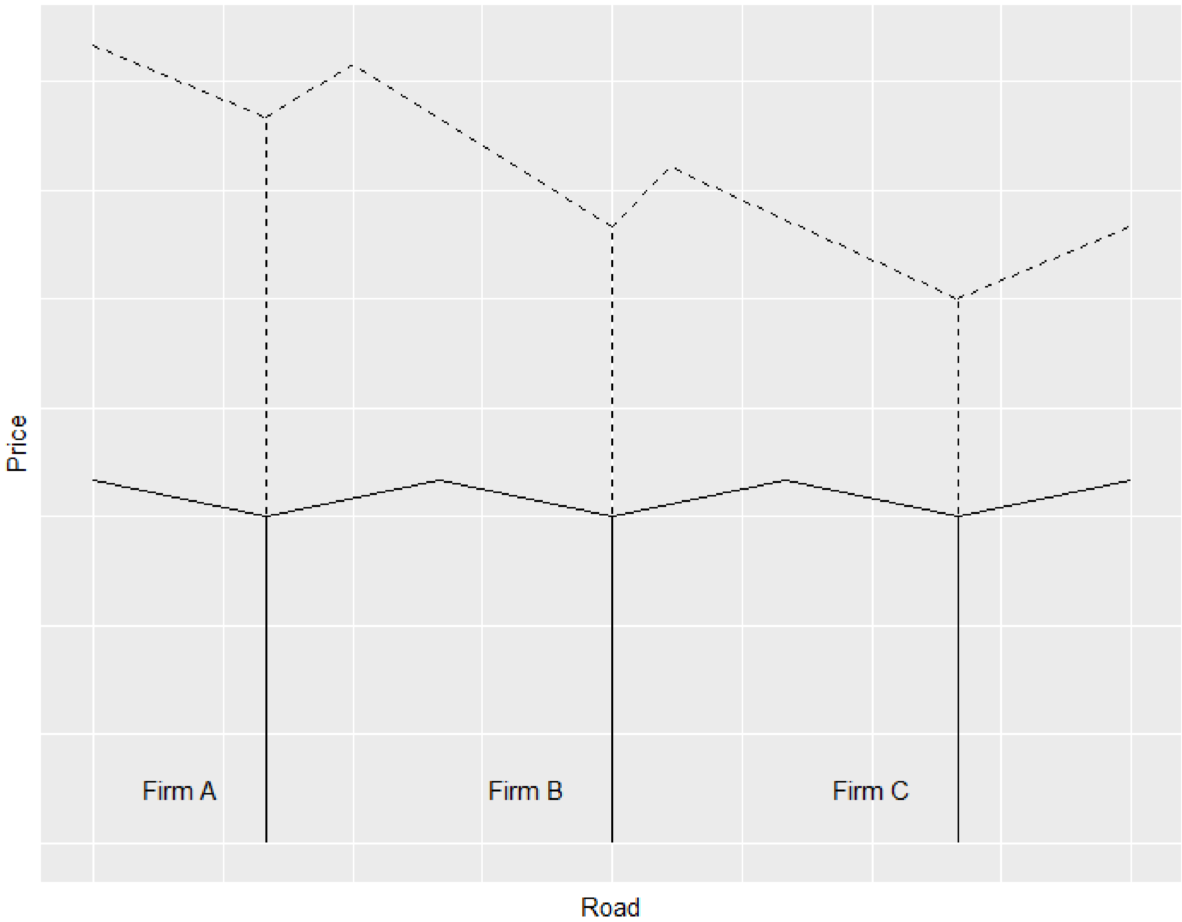

Figure 1 – Spatial Bertrand with 3 firms

When firms form a cartel and start colluding, they do so to increase their profit. One of the most typical way of colluding is by acting as a monopolist, increasing the price and thereby sharing the market. While a cartel consisting of all firms on the market has a greater ability to raise prices, even a partial cartel can impact the price. This can be seen by the dashed lines in figure 1, where firm A and B started to collude and maximize their joint profits. When the cartel increases prices, they also increase the price of other firms which are not part of the cartel since the prices act as strategic complements. Firms which collude may therefore exhibit price leadership, in which competing firms are able to raise their prices because the price leader did so.

An increase in price is one sign of collusion and is a typical sign of cartels. Other markers of collusion also indicate that price variance is lower for collusive firms from a theoretical and empirical perspective. (Blanckenburg, Geist & Kholodilin, 2011; Bolotova, Connor & Miller, 2008). Higher margins may also be a sign of collusion. Harrington and Chen (2006) provide theoretical basis for why cartels price differently and how the price vary less when firms also attempt to avoid being detected by competition authorities. They find that cartels have two phases of pricing behaviour. Upon formation, the cartel slowly increases prices to avoid suspicion as a sudden increase in prices may look like coordination and trigger an investigation by competition authorities. After the transitionary phase comes the stationary phase in which a cartel ceases to increase prices and instead maintains the joint profit maximising price. The stationary phase is also characterised by lower variance which arises because cost shocks are not passed through completely, although they don’t provide an explanation for why that phenomenon happens.

All together, a cartelised fuel industry should experience lower price variance, higher price and higher price margin for the colluding firms compared to non-collusive firms. The firms that are not colluding should have a similar pattern of higher price and margin, but not as much as the firms of the cartel. Based on these stylized facts, different screens have been designed in order to capture the effect on price variance, the distribution of prices and how cartels cause a break in the price generating process.

4.2 Behavioural Screens

Crede (2019) uses structural break tests to identify possible cartels and their starting dates in the Italian, French and Spanish pasta industries using monthly national prices. The pasta cartels had high market share in both cases but lasted different amount of time. The Italian cartel lasted for almost 2 years while the Spanish cartel lasted for only 3 months. The tests successfully detect cartels in the Spanish and Italian markets and does not provide a false positive for the French market. A cartel is said to be plausible, according to Crede, if structural breaks are dated at the same time, across different minimum segment sizes. The article serves as the main inspiration for this thesis and in particular the use of the structural break test by Bai and Perron (2003).

The same methodology is used by Boswijk et al (2017) who apply the structural breaks test to the Sodium Chlorate cartel. The cartel supplied 90% of European consumers over a period of 12 years. They found that dates aligned with structural breaks matched those found by the European commission in 2008. In addition to showing how the method can be used in a screening sense, they also show how cartel misdating may underestimate cartel overcharges, by lowering the estimated price absent collusion.

Previous research and application of other behavioural screens not related to structural screens have mostly focused on detecting bid rigging conspiracies (Porter, 2005; Harrington 2008). Huber & Imhof (2019) examines statistical moments of each tender to detect bids that were rigged and finds that the variance is lower within tenders and that the bids are more skewed than during competition. A central assumption of their study is that the way bids are generated differs between the competitive and the bid-rigging tenders. Abrantes, Froeb, Geweke & Taylor (2006) also finds that the price variance of frozen perch filet sold to the US government decreases after the collapse of a cartel, indicating that low variance could be associated with collusion.

Screens for retail markets have also used statistical moments to detect collusion, such as Blanckenburg, Geist & Kholodilin (2011) who used it on 11 German industries. They found that variance was the only robust indicator of collusion whilst skewness and kurtosis were less reliable. Silveira, Vasconcelos, Resende & Cajueiro (2021) examines Brazilian gas stations and finds a significant difference in how the moments are distributed between competitive and cartel markets. They additionally train a machine learning model to detect suspicious markets based on these moments, which accurately predict 84% of the cases.

Bolotova, Connor & Miller (2008) use a GARCH and ARCH model to analyse the price variance in 2 cartels in citric acid and lysine acid. They find that variance is not necessarily higher during the cartels for citric acid which they believe may be due to increase international competition. They find however that the variance was lower for lysine acid.

4.3 Structural tests

Structural breaks in time series data are dates at which the model before and after the break are different. The most well-known structural break test is that by Chow (1960) which requires the date of a hypothesised structural break to be specified. Different coefficients are estimated between each structural break and the Residual Sum of Squares (RSS) are summated. This unrestricted RSS is used together with the restricted RSS, which have the same coefficients throughout the entire sample, to create a F-statistic. Under the null, the F-statistic follows an F-distribution which if rejected implies that the unrestricted RSS have better explanatory power and that the coefficients do change between structural breaks.

Bai and Perron (2003) introduces an efficient way of finding multiple structural breaks when the dates and the number of breaks is unknown. Their paper relies on the authors' previous study from 1998 which develop the asymptotics of their test and is similar to the Chow test in many ways. The test works by estimating the coefficients and RSS for each possible dates where there is some minimum distance (h) between possible structural breaks. The date which minimizes the RSS is then selected as the best possible date which explains the model and where a structural break is likely to have happened. This procedure is then done for some maximum number of breaks (m) where each minimized RSS is saved to later determine the amount of breaks which are present.

In order to estimate the number of breaks, the Bayesian Information Criterion (BIC) is used as suggested by Bai and Perron (2003) who rely on Yao (1998). The number of breaks are the ones which optimise BIC. When using an information criterion, the penalty of introducing additional coefficients, or breaks in this case, must be compensated by an improved model fit. If breaks are included that do not improve the model sufficiently enough, then the breaks should not be included as they don't provide enough explainability.

5. Method

This section explains how structural breaks are tested and how the different aggregation levels of the data are used.

5.1 Levels of aggregation

The thesis uses 2 different levels of aggregation in order to investigate the data requirement needed for the test and to contrast it with Crede (2019) and Boswijk et al(2017) who used mean market prices. The 2 different levels of aggregation are also used to explore if it is a viable way to construct the screen. The first level of aggregation is on province wide level which only uses the mean price in each region. As seen from Table 1, the market share for the colluding firms is quite sizeable as measured by the amount of stations. In Barcelona, REPSOL and CEPSA cover 23% and 10% of the market whilst in Madrid they cover 33% and 16% respectively. It is therefore reasonable that they have a significant impact on the price as outlined in the spatial Bertrand, although maybe not as much as if they were covering even more of their markets. The possible problem of market share is mitigated by using more detailed price information on the other level of aggregation.

The next level of aggregation is on brand-level and uses the mean price of each brand for each day. Brands which have market shares below 3% are categorized into “Other” which contains all smaller brands. Using brand specific prices has multiple advantages as the test can be applied to each brand separately where they control 100% of the price. Concerns for how large the companies need to be to have a sizeable effect on the price is therefore lessened since if the firm start colluding this should carry over fully into the mean brand price. Applying the test on brand level also has the advantage of being able to pinpoint which brands the test was triggered for, which is not the case when using province-level data. Another reason for using brand-level data is to be able to separate the effects of collusion with that of a demand shock.

To use 2 different levels of aggregation is a novel approach which contrasts the different levels of aggregation and how viable each level is in a screen. The novel approach also explores how brand specific prices could be used for cartel detection. Using brand-level data also makes it viable on cartels where they don’t cover a majority of the market and allows for pinpointing of which cartels that appear suspect.

5.2 Selecting a model

The process of model selection is done by finding a good fit for each level of data and province so that there is not a high degree of multicollinearity and where the coefficients for each lag are mostly significant. Each regression is tested for unit root and cointegration when applicable in order to take advantage of the superconsistency from cointegrated variables (Enders, 2015). This results in quicker convergence of the coefficients and lower SSR due to being estimated at levels which should improve the break detection ability.

The variables included in the regressions and the tests are now discussed. Price, being the most important variable to this study is assumed to be dependent on itself. Given the time series nature of the data, this is an unsurprising statement as prices should be affected by its own previous values due to price stickiness. When stations enter their collusive phase, this dependence on its own previous values should decrease. In other words, the coefficient of lagged y in an AR(1) process will be smaller during the cartel phase.

$$p_t = \rho p_{t-1} + \epsilon_t \tag{1}$$

$$\sigma_p^2 = \frac{\sigma_\epsilon^2}{1-\rho^2} \tag{2}$$

Given an AR(1) process as seen in equation (1) , the variance equation for the independent variable, p, is given by equation (2). Since the variance of price, σ_p^2, is lower during the cartel phase, this implies that ρ should be smaller (Abrantes, Froeb, Geweke & Taylor, 2006; Huber & Imhof, 2019; Silveira, Vasconcelos, Resende & Cajueiro, 2021; Silveira, Vasconcelos, Bogossian & Neto, 2021). An assumption for how ρ affects the variance of price is then that the variance of the residuals is unchanged.

The European price of Brent Crude oil is the second variable to be included and is an important cost variable. Since the main cost of fuel retailers is the price of the input good, including oil price accounts for large changes in retail fuel prices (Gonzalez & Moral 2019) and so supply shocks are captured. The coefficient from oil ought to be smaller during cartels, following Harrington and Chen (2006) in which they show that cost pass throughs are lower during collusion.

While there exists good daily data on cost variables, factors affecting demand are less available at that high frequency. Household income and expenditures are not updated each day, week or even month which makes their inclusion difficult to implement. This thesis excludes demand side variables which can be rationalised by highlighting that Crede (2019) and Boswijk, et al (2018) find that these are not statistically significant in their regressions and in the following discussion.

By excluding demand side variables, the risk is that changes in demand causes structural breaks and makes it appear as if a cartel has been formed. This can be mitigated in two ways depending on the level of aggregation. If province-level aggregation is used and a structural break is caused by a demand shock, then proxies for demand can be used to explain the structural break. If a structural break occurs and can not be linked to a demand shock, then collusion seems more probable.

Using brand-level aggregation can indicate which brands that are colluding and structural breaks between brands can be compared. This works well for demand shocks since a demand shock will cause structural breaks for all brands around the same time. If all brands show structural breaks close to each other, then two possibilities exist. The first possibility is that all brands started colluding at the same time and the second possibility is that the structural break arose from a demand shock. By matching the structural breaks to known demand shocks, it can be determined if collusion or a change in demand caused it.

When only some brands collude, structural breaks for those brands should be found close to each other. Brands that are not colluding, may however have structural breaks very close to the colluding brands since they act as price followers. By following the cartel's price, non-colluding brands increase their price and create a structural break in their data. In order to avoid confusing cartels and non-colluding brands, the brands with the first structural breaks close to each other should be identified. Given that the colluding brands act as price leaders, the first structural breaks belong to the colluding brands.

This naturally raises the question of how close brands' structural breaks should be to each other to warrant a comparison. This is analysed using sets of common structural breaks. The sets of common structural breaks encompass 2 weeks for which there are the following reasons. Firstly, having a longer period naturally means that more structural breaks and more sets of common structural breaks are analysed, adding to the complexity of the method, and including unrelated structural breaks. The second reason is to combat noise in the data which can shift the date of the structural break so it's not aligned with the start and end of the non-aggression agreement. Having a period smaller than 2 weeks may therefore leave out structural breaks that happened close to each other.

Different minimum segment sizes, which is the shortest possible time between structural breaks, are used to balance the positive and negative aspects associated with the different lengths. Smaller minimum segment sizes have a faster break detection ability since there is less possible time between structural breaks, but this may result in worse finite sample properties due to the low number of observations. Having greater minimum segment sizes runt into the opposite problem where there are better finite sample properties but structural breaks that happen close to each other can not be identified. Based on this, 4 minimum segment sizes of 0.3, 0.2, 0.1 and 0.05 are used. A cartel is said to be detected if the same dates appear for 3 or more different minimum segment sizes. This is similar to Crede (2019) and avoids a high rate of false positive, which is preferred over a high rate of false negatives.

To consider the case that CEPSA and REPSOL were colluding before the official start of July 2011, 2 months before that will be highlighted in the discussion. 2 months after the end of august 2011 are also analysed to consider that brands continued colluding longer than was found by the courts.

5.3 Expectation

It is expected that the test detects breaks for REPSOL and CEPSA in the beginning of July 2011 and end of August 2011 as that is when the non-aggression agreement began and ended. It is also expected that other brands have structural breaks which come after REPSOL and CEPSA in July 2011 and august 2011, due to being price followers. Structural breaks from the non-colluding brands are expected to appear inside the range of common structural breaks. A similar expectation is that structural breaks are dated at the province-level for both Madrid and Barcelona in the beginning of July 2011 and end of August 2011. The third expectation is that demand shocks will cause structural breaks for all brands within a set of common structural breaks.

6. Results

This section displays and comments the results from the structural breaks test at the different levels and provides unit root and cointegration test to make sure the inference is valid.

6.1 Province

As can be seen from Table 3, unit root is present for Barcelona and Madrid and the price of crude oil for both the unit root tests. The variables are cointegrated in Barcelona but not so in Madrid according to table 4. This allows us to treat the cointegrated variables in levels and is represented as in equation 3 in order to take advantage of the super consistency of Oil. Additional lags of Price and Oil were not considered as they proved to be highly multicollinear. Since there was no cointegration found for Madrid, the same specification as in equation 3 was used except that all variables were differenced, resulting in equation 4.

Table 3 - Unit root test for Provinces

Barcelona Price			Madrid Price		
Name	Statistic	p-value	Name	Statistic	p-value
ADF-level	1.563	0.50	ADF-level	-1.637	0.43
ADF-diff	-18.495	0.00	ADF-diff	-19.718	0.00
Kpss-level	11.138	0.00	Kpss-level	10.5	0.00
Kpss-diff	0.1243	> 0.1	Kpss-diff	0.127	> 0.1
Oil Price			ADF is the Augmented Dickey Fuller test for which the null hypothesis is unit root. Kpss is the Kwiatkowski–Phillips–Schmidt–Shin where the null hypothesis is that the data is stationary, ie it has no unit root. Level and diff represent the data in level form or differenced, respectively.		
Name	Statistic	p-value			
ADF-lvl	0.5684	0.38			
ADF-diff	-30.236	0.00			
Kpss-lvl	9.1506	0.00			
Kpss-diff	9.15	0.00			

Table 4 - Engel Granger cointegration test of OilPrice and Price		
Province	Test-statistic	p-value
Barcelona	-4.481	< 0.01
Madrid	-1.932	> 0.1
Engel Granger cointegration test has null hypothesis of no cointegration		

$$price_t = \rho * price_{t-1} + \beta * Oil_t + \epsilon_t \tag{3}$$

$$\Delta price_t = \rho * \Delta price_{t-1} + \beta * \Delta Oil_t + \epsilon_t \tag{4}$$

Table 5 - Structural breaks – Province level				
h (#days)	0.05 (45)	0.1 (90)	0.2 (180)	0.3 (270)
Dates - Barcelona	-	-	-	-
Dates – Madrid	2010-10-29	2010-12-19	2011-03-18	2011-06-28
The table shows the dates which the structural breaks algorithm found a structural break. h represents the minimum segment size as a fraction of the sample period being used. #days represents the minimum segment size in terms of number of days.				

As can be seen from table 5, no structural breaks are detected at any of the minimum segment sizes for Barcelona. There is only one date which is sufficiently close to date the beginning and end of the non-aggression agreement which is the minimum segment size of 0.3.

6.2 Brand

Brand level data is explored in this section where each price is the mean of each brands' price. This should cause more structural breaks to be detected, as the brands have more influence on their brand specific price.

Table 6 – Panel Unit root test for Brands

Barcelona Price			Madrid Price		
Name	Statistic	p-value	Name	Statistic	p-value
LLC	-26.39	0	LLC	-25.78	0
Breitung	18.7577	0	Breitung	-17.44	0
LLC represent the Levin-Lin-Chu test statistic for panel unit root. The null hypothesis for both tests is that the data has unit root and the alternative is that all panels are stationary.					

The brand specific unit roots in table 6 indicate that neither province have unit roots and so one can treat them as stationary. The specification applied to the brand data is seen in equation 5 where an additional lag has been added compared to equation 3 and 4. Regressions of 3 or more brand specific price lags were not indicative that the third lag contributed much to the model as they were usually insignificant. This is similar to Boswijk et al (2017) who also use 2 price lags.

$$price_t = \rho_1 price_{t-1} + \rho_2 price_{t-2} + \beta_1 * Oil_t + \epsilon_t \qquad (5)$$

Table 7 - Barcelona Province, structural breaks on Brand level

h(#days)	Brands							
0.3(270)	BP	CAMPSA	CEPSA	GALP	MEROIL	OTHER	PETROCAT	REPSOL
			2011-03-07	**2011-03-10**	**2011-03-07**		**2011-03-15**	2011-04-12
	2011-05-13	2011-05-19				2011-05-10		
	2012-02-09	2012-02-23	**2012-02-10**	**2012-02-08**	**2012-02-10**	2012-02-06	**2012-02-10**	2012-02-07
0.2(180)	BP	CAMPSA	CEPSA	GALP	MEROIL	OTHER	PETROCAT	REPSOL
	2011-01-16	**2011-01-02**	2011-05-28	**2011-01-13**	**2011-01-15**	**2011-01-13**	**2011-01-03**	2011-01-22
		2011-07-03			2011-07-14			
	2011-08-06		2012-04-20	**2011-07-29**		2011-11-06	**2011-08-01**	**2011-08-07**
	2012-02-09	2012-01-26		2012-02-08	2012-02-10	2012-05-10	2012-02-10	2012-02-07
0.1(90)	BP	CAMPSA	CEPSA	GALP	MEROIL	OTHER	PETROCAT	REPSOL
	2010-09-13	2010-09-18	2010-09-25	2010-09-10	2010-09-08	2010-09-09	2010-09-14	2010-09-27
		2010-12-25		**2010-12-14**	**2010-12-17**	**2010-12-18**		
	2011-01-16							2011-01-22
			2011-03-07				**2011-03-09**	
		2011-03-31		2011-03-31	2011-03-31	2011-03-20		
	2011-05-26		**2011-06-06**			**2011-06-18**	**2011-06-13**	**2011-06-02**
	2011-08-31	2011-07-02	2011-09-10	2011-09-01	2011-08-15	2011-09-18	2011-11-24	2011-09-04
	2012-02-09	2011-10-09	2011-12-10	**2012-02-08**	2011-11-26	2011-12-28	**2012-02-22**	**2012-02-07**
		2012-01-09	2012-03-15		2012-04-22	2012-04-04		
		2012-05-26	2012-06-23		2012-08-03		2012-05-23	
							2012-08-25	

Dates that are **bold** or underlined are highlighted to show sets of common structural breaks. Common sets of structural breaks are on the same row. Bold and underlined dates are discussed separately from each other.

Structural breaks for Barcelona on the brand level are shown in table 7 and table 8. Examining table 7, there appears to be many sets of common structural breaks for each minimum segment size. A trend is however that the number of structural breaks increase as minimum segment sizes increase. There exists one period where multiple sets of common structural breaks span across different minimum segment sizes, which is the beginning of February 2012. The brands who continue to have structural breaks as minimum segment sizes change are BP, GALP and PETROCAT.

The brands who have sets of common structural breaks in the beginning of the non-aggression agreement are CAMPSA and MEROIL for a minimum segment size of 0.2 Sets of common structural breaks in the end of august 2011 are found for BP, CEPSA, GALP and REPSOL for a minimum segment size of 0.1 and MEROIL and OTHER, when using a minimum segment size of 0.05, as can be seen in Table 8.

Common sets of structural breaks within 2 months of the start of the non-aggression agreement include BP, CAMPSA and OTHER for a minimum segment size of 0.3 which have their set of common structural breaks in the beginning of May 2011. BP, CAMPSA, CEPSA and GALP also have a common set of structural breaks in the middle of May for a minimum segment size of 0.05.

Additionally, BP, CAMPSA, CEPSA and PETROCAT have a set of common structural breaks in the beginning of October 2011, which correspond to 1 month after the end of the official non-aggression agreement. BP and CAMPSA are the two brands which all have common structural breaks within 2 months of the official non-aggression agreement.

Table 8 - Barcelona Province, structural breaks on Brand level for minimum segment size 0.05

BP	CAMPSA	CEPSA	GALP	MEROIL	OTHER	PETROCAT	REPSOL
2010-07-29			**2010-07-25**	**2010-08-07**		**2010-07-25**	**2010-07-25**
	<u>2010-08-29</u>	<u>2010-08-30</u>	2010-09-26		<u>2010-08-21</u>	2010-09-18	
2010-10-03				**2010-10-17**			**2010-10-03**
<u>2010-12-15</u>	<u>2010-12-05</u>	<u>2010-12-02</u>	<u>2010-12-11</u>	2011-01-21			2010-11-24
					2010-10-29	**2010-11-09**	
					<u>2011-01-13</u>	<u>2011-01-19</u>	<u>2011-01-13</u>
2011-03-13	**2011-03-09**	**2011-03-07**	**2011-03-07**			**2011-03-20**	
				<u>2011-03-31</u>	<u>2011-04-08</u>		<u>2011-04-09</u>
2011-05-26	**2011-05-15**	**2011-05-17**	**2011-05-17**				
				<u>2011-06-04</u>		<u>2011-06-01</u>	<u>2011-06-11</u>
		2011-07-07			**2011-07-01**		
	<u>2011-08-10</u>		<u>2011-07-31</u>	<u>2011-07-30</u>	2011-09-13		
2011-08-19						**2011-08-16**	**2011-08-29**
		2011-08-24	<u>2011-09-18</u>	<u>2011-09-13</u>			
				2011-12-03	**2011-12-03**		
<u>2011-10-06</u>	<u>2011-10-12</u>	<u>2011-10-10</u>			2012-02-24	<u>2011-10-04</u>	
2011-11-28	**2011-11-28**	2012-01-08	**2011-11-18**		2012-05-11	**2011-11-24**	**2011-11-26**
<u>2012-02-06</u>	<u>2012-02-07</u>	2012-03-15	<u>2012-02-06</u>	2012-01-25		<u>2012-02-10</u>	<u>2012-02-06</u>
2012-04-12	**2012-04-14**	2012-05-26	**2012-04-12**	**2012-04-03**		2012-04-25	**2012-04-01**
<u>2012-06-22</u>	<u>2012-06-09</u>						<u>2012-06-08</u>
			2012-06-28	2012-06-15	**2012-07-10**	**2012-07-13**	
	<u>2012-07-31</u>	<u>2012-07-30</u>					<u>2012-07-27</u>
2012-09-01			**2012-09-12**	2012-08-03	**2012-09-09**		
	<u>2012-09-25</u>	<u>2012-09-27</u>		<u>2012-09-22</u>			<u>2012-09-18</u>

Dates that are **bold** or <u>underlined</u> are highlighted to show sets of common structural breaks. Common sets of structural breaks are on the same row. Bold and underlined dates are discussed separately from each other.

Table 9 - Madrid Province, structural breaks on Brand level

h(#days)	Brands						
0.3(270)	BP	CAMPSA	CEPSA	GALP	Other	REPSOL	SHELL
	2011-06-06	**2011-06-15**				**2011-06-08**	**2011-06-12**
			2011-07-14	2011-07-26	2011-07-17		
	2012-05-11	**2012-05-08**	**2012-04-26**	**2012-05-08**	**2012-05-01**	**2012-04-25**	2012-05-19
0.2(180)	BP	CAMPSA	CEPSA	GALP	Other	REPSOL	SHELL
	2011-05-03	2011-05-26	**2011-04-20**	2011-07-28	**2011-05-06**	**2011-05-05**	**2011-05-06**
	2011-11-10	2012-05-01	2011-10-19	2012-06-11	2012-02-27	2011-11-08	2011-11-11
0.1(90)	BP	CAMPSA	CEPSA	GALP	Other	REPSOL	SHELL
	2010-12-08	2011-02-06		2010-12-11	2010-12-10	2010-12-30	2010-12-10
	2011-03-29	2011-05-12	2011-04-20	2011-03-11			**2011-04-02**
					2011-05-06	2011-05-05	
	2011-06-28			**2011-06-21**	2011-10-17		
		2011-08-14	2011-08-09			2011-08-16	2011-08-17
	2011-10-02	**2011-11-22**	**2011-11-07**	2011-09-22		2011-12-02	**2011-11-17**
	2012-01-05			2011-12-23			
			2012-02-14		**2012-01-29**	**2012-03-01**	**2012-02-15**
	2012-04-08			2012-03-25	2012-04-30	2012-06-05	
		2012-05-08	**2012-05-15**				**2012-05-19**
		2012-08-08	2012-08-17				2012-10-26
	2012-07-07			**2012-06-24**			
	2012-10-06	2012-11-10	2012-11-27	2012-10-06	2012-09-24		

Dates that are **bold** or underlined are highlighted to show sets of common structural breaks. Common sets of structural breaks are on the same row. Bold and underlined dates are discussed separately from each other.

Table 9 and 10 show structural breaks for Madrid at different minimum segment sizes. Sets of common structural breaks that correspond to the beginning of the non-aggression agreement are found for the segment sizes of 0.1 and 0.05. For the minimum segment size of 0.1, BP and GALP have a set of common structural breaks at the end of June 2011 and very close to the beginning of July. CAMPSA, CEPSA, GALP and SHELL also have a set of common structural breaks for a minimum segment size of 0.05. The brands who have a set of common structural break change drastically between the two different minimum segment sizes. Only GALP have a common structural break in the beginning of July, for both a minimum segment size of 0.05 and 0.1.

Sets of common structural breaks close to the end of the non-aggression agreement are found for two minimum segment sizes. CAMPSA, CEPSA, REPSOL and SHELL have common structural breaks in the middle of August 2011 for the minimum segment size of 0.1. This is similar to BP, CAMPSA, SHELL and other brands which also have a set of common structural breaks in the middle of august 2011 for a minimum segment size of 0.05. CAMPSA and SHELL are the only brands which are in both sets of common structural breaks.

Madrid has the same set of common structural breaks in February 2012, but not across as many different minimum segment sizes. Barcelona have sets of common structural breaks across all 4 different minimum segment sizes whilst Madrid only has it for 0.1 and 0.05. No other sets of common structural breaks in Madrid, for 3 or more different minimum segment sizes, are found.

Sets of common structural breaks 2 months before the official start are found for minimum segment sizes of 0.3, 0.2 and 0.1. BP, CAMPSA, REPSOL and SHELL have a set of common structural breaks one month before the official non-aggression agreement, when looking at a minimum segment size of 0.3. For a minimum segment size of 0.2, BP, CEPSA, REPSOL, SHELL and other brands have a set of common structural breaks 2 months before the official period. REPSOL and other brands have a set of common structural breaks 2 months before the official non-aggression agreement for the minimum segment size of 0.1. A set of common structural breaks appear in the middle of October for BP, CEPSA and REPSOL for a minimum segment size of 0.05

Table 10 - Madrid Province, structural breaks on Brand level for minimum segment size 0.05

BP	CAMPSA	CEPSA	GALP	Other	REPSOL	SHELL
2010-11-18		**2010-10-31**	**2010-10-24**	2010-11-07	**2010-10-26**	**2010-10-25**
	<u>2010-12-03</u>		<u>2010-12-11</u>	2011-01-20		2011-02-01
		2010-12-30			**2010-12-30**	
<u>2011-03-29</u>	<u>2011-03-22</u>	<u>2011-03-31</u>	2011-02-05	<u>2011-03-25</u>	<u>2011-03-28</u>	2011-04-02
2011-06-10			2011-04-14			
	<u>2011-06-16</u>	<u>2011-06-24</u>	<u>2011-06-27</u>	2011-06-09	2011-05-21	<u>2011-06-28</u>
2011-08-24	**2011-08-14**	2011-09-04	2011-09-22	**2011-08-21**	2011-08-07	**2011-08-12**
<u>2011-10-13</u>		<u>2011-10-19</u>			<u>2011-10-16</u>	
	2011-11-08		2011-12-21	**2011-11-13**		2011-09-29
<u>2011-12-12</u>	<u>2011-12-25</u>	2012-01-07	2012-02-18			2011-12-02
2012-01-26				**2012-01-29**	**2012-01-12**	**2012-01-17**
	2012-02-08	<u>2012-03-03</u>			<u>2012-02-26</u>	
2012-04-12	2012-05-01	**2012-04-19**	**2012-04-18**	**2012-04-05**	2012-05-16	**2012-04-10**
<u>2012-06-21</u>	2012-07-16	<u>2012-06-18</u>	<u>2012-06-30</u>	<u>2012-06-21</u>	2012-07-01	<u>2012-06-23</u>
2012-08-31		**2012-08-17**		**2012-08-15**		2012-09-06
	<u>2012-09-28</u>		<u>2012-09-19</u>		<u>2012-09-28</u>	
		2012-10-02		**2012-10-01**		
<u>2012-10-25</u>						<u>2012-10-26</u>
	2012-12-11		**2012-12-08**		**2012-12-18**	
2013-01-09		<u>2012-12-10</u>		<u>2012-12-14</u>		

Dates that are **bold** or <u>underlined</u> are highlighted to show sets of common structural breaks. Common sets of structural breaks are on the same row. Bold and underlined dates are discussed separately from each other.

7. Discussion

Based on the results and with the purpose of evaluating if the structural breaks test is viable as a screen, the results are now discussed.

When examining the province-level data, one wouldn't be able to say that there was a cartel present in either Madrid or Barcelona. Barcelona has no structural breaks and Madrid has structural breaks which did not stay the same when changing to another minimum segment size. The structural breaks in Madrid, instead, appear as close to the beginning of the sample period, as is allowed by the minimum segment size.

The structural breaks on brand level did not indicate that collusion was present either. In order to say that collusion was present, 3 sets of common structural breaks at the same date across different minimum segment sizes had to be identified. This does not appear except February 2012 for Barcelona for all 4 minimum segment sizes. These 4 sets of common structural break are not associated with any period in the cartel's arrangement and appears after the end of the non-aggression agreement. The date does however correspond to a demand shock, in which Spain entered its second quarter of negative growth and therefore the start of its second recession CNN (2012). The demand shock can be seen in Madrid as well but only appears for the segment sizes of 0.1 and 0.05.

The brands that have structural breaks at the beginning of the non-aggression agreement change between different minimum segment sizes, and only one of the colluding brands have a structural break at the same time as other companies in Madrid, namely CEPSA. CEPSA shares this set of common structural breaks with CAMPSA, GALP and SHELL for the minimum segment size of 0.05. It was expected that REPSOL and CEPSA would have sets of common structural breaks in July 2011, but this does not happen. The brand which has 2 sets of common structural breaks across different minimum segment sizes is instead GALP for 0.1 and 0.05.

The end of the non-aggression agreement in August 2011 is not identified either. Madrid has no sets of common structural breaks at all for that date whilst Barcelona has 2 different sets. REPSOL and CEPSA have a set of common structural breaks together with BP and GALP for a minimum segment size of 0.1. MEROIL and OTHER have a set of common structural breaks for 0.05.

To consider that REPSOL and CEPSA had a longer non-aggression agreement than they were fined for, structural breaks 2 months before and after the agreement are considered. The brands may have colluded some time before the official period although there was not enough evidence of it to convict them. To consider this possibility, sets of common structural breaks are considered between the beginning of May and end of October 2011. In Barcelona, CEPSA have a common structural break with BP, CAMPSA and GALP in the middle of May for a minimum segment size of 0.05. REPSOL does not share that set of common structural breaks. Examining a minimum segment size of 0.2 for Madrid, a set of common structural breaks is found for BP, CEPSA, REPSOL and SHELL in May. There are no sets of common structural breaks 2 months after the end of the non-aggression agreement for Madrid.

The method of using structural breaks failed when applied to this dataset. Periods which saw collusion and the brands colluding were not identified in either province. The difference in combined market shares for CEPSA and REPSOL between Madrid and Barcelona did not play a part either. The method did not fulfill the main requirement of a screen, which is its ability to correctly predict cartels and non-colluding periods.

The second requirement from the section on how to evaluate a method, is that a screen should be costly to beat. This can not be answered by the results and would require other types of studies. The second requirement is also not relevant due to the low predictive power of the screen. The third and more interesting requirement is how the threshold for categorising cartels and non-collusive firms can be changed. The proposed method here, and the two previous studies, used the number of different minimum segment sizes required to change the threshold.

The fourth requirement of scalability needed the screen to be scalable and require a low amount of human input. In this respect, the screen failed for the brand-level screen. Analysing the structural breaks at brand-level was complicated and time-consuming to do and as is hinted at by the size of the tables. Examining province-level data was however far easier and more time efficient.

Reasons for why the method failed in terms of its predictive ability may depend on 3 factors. Firstly, the joint market share of REPSOL and CEPSA did not exceed 50%. The market shares of the 2 cartels examined by Crede (2019) spanned 55% to 90% and Boswijk et al. (2017) had market shares of 90% on their single cartel. Crede was not able to identify the cartel which spanned 55% of the market. Around 50% might therefore be a lower limit on how well a screen works.

A second possible reason for why the method failed is due to the exclusion of demand side variables, which appeared to be important given that the demand shock in February 2012 caused structural breaks. Demand shocks may however not explain all the structural breaks that are seen in table 7 through 10 as that would imply that a model is viable for only a few months.

8. Conclusion

This thesis has examined a new method of cartel detection. Using a screen based on structural breaks, the method is found to perform poorly based on its predictive power when using aggregated data. A novel approach in this thesis is to use data specific to each major brand. This was done to explore how more detailed data can be used to improve the structural breaks screen. Using data specific to brands resulted in results that showed low predictive power, usability and scalability.

This study contributes to previous research on behavioural screens of cartels in multiple ways. The first way is by implementing a variation of the structural breaks test used in previous research. The study also highlights the importance of market shares of the colluding firms and its adverse effects on being able to detect them. My thesis is also the only study using the method of structural breaks, which finds that the screen performs poorly. My study also sets up a clear and prioritised list of what is expected from a good screen which is lacking from a majority of previous studies.

Future research should take advantage of the used dataset to evaluate methods which have been successful on data where cartels have correctly found. Future research should also further explore the use of structural breaks and other screens for detecting cartels.

Cartels continue to sap welfare from consumers, and it is the mission of competition authorities and academics to continue dismantling them. Leniency programs have had great success so far, but other methods should be developed to further reduce the effect of cartels and improve the welfare of consumers.

9. References

Abrantes R., Froeb L. , Geweke J. & Taylor T. C. (2006) A variance screen for collusion, *International Journal of Industrial Organization*, vol. 24, issue 3

Bai J. & Perron P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica*, vol. 66, no. 1

Bai J. & Perron P. (2003). Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics*, Vol 18, pp 1-22

Balaguer J. & Ripollés J. (2016) Building a panel data set on fuel stations located in the Spanish regional areas of Madrid and Barcelona, *Data in Brief*, vol. 7

Blanckenburg K. , Kholodilin A. K. & Geist A. (2011) The Influence of Collusion on Price Changes: New Evidence from Major Cartel Cases, *German Economic Review*, vol. 13, issue 3

Boltova Y., Connor J. & Miller J. D. (2008). The impact of collusion on price behavior: Empirical results from two recent cases, *International Journal of Industrial Organization*, vol. 26, issue 6

Boswijk H. P., Bun J. G. M. & Schinkel P. M. (2017). Cartel dating, *Journal of Applied Econometrics*

Chow C. G. (1960) Tests of Equality Between Sets of Coefficients in Two Linear Regressions, *Econometrica*, vol. 28, No. 3

Concurrencies (2015). THE SPANISH COMPETITION AUTHORITY FINES VARIOUS PETROL PRODUCT OPERATOR FOR ANTI-COMPETITIVE AGREEMENTS. Available online: https://www.concurrences.com/en/bulletin/news-issues/february-2015/The-Spanish-Competition-Authority-76769 [Accessed online 12 May 2021]

Crede J. C. (2019). A Structural Break Cartel Screen for Dating and Detecting Collusion, *Review of Industrial Organization*, 54:543–574

Enders W. (2015). Applied Econometric Time Series, USA: John Wiley & Sons, Inc

Genesove, D. & Wallace P.M. (2001). Rules, Communication, and Collusion: Narrative Evidence from the Sugar Institute Case. *American Economic Review*, Vol 91, No 3

Gonzalez X. & Moral J. M. (2019) Effects of antitrust prosecution on retail fuel prices, *International Journal of Industrial Organization*, vol. 67

Harrington E. J. Jr & Chen J. (2006). Cartel pricing dynamics with cost variability and endogenous buyer detection, International Journal of Industrial Organization, vol. 25, issue 6

Harrington E. J. Jr. (2008). Detecting Cartels, in Paolo Buccirossi (eds), *Handbook of Antitrust Economics*, MIT Press

Huber M. & Imhof D. (2019). Machine learning with screens for detecting bid-rigging cartels, *International Journal of Industrial Organization*, 65:277-301

Isidore C. (2012). Spain falls into recession, CNN 30 April, Available online: https://money.cnn.com/2012/04/30/news/economy/spain-recession/index.htm [Accessed 13 May 2021]

Nathan H. Miller (2009). Strategic Leniency and Cartel Enforcement, *The American Economic Review*, Vol 99, No 3

OECD (2013a). Ex officio cartel investigations and the use of screens to detect cartels

OECD (2013b). Competition in Road Fuel

Pepall L., Richards D. & Norman G. (2014) Industrial Organization: Contemporary Theory and Empirical Applications, USA: John Wiley & Sons

Perdiguero J. & Jiménez L. J. (2021). Price coordination in the Spanish oil market: The monday effect, *Energy Policy*, vol. 149(C)

R. H. Porter (2005). Detecting Collusion, *Review of Industrial Organization*, 26:147–167

Silveira D., Vasconcelos S., Bogossian P. & Neto J. (2021). Cartel screening in the Brazilian fuel retail market, *EconomiA*, vol. 22, issue. 1

Silveira D., Vasconcelos, S., Resende, M. and Cajueiro, D. O. (2021) Won't Get Fooled Again: A Supervised Machine Learning Approach for Screening Gasoline Cartels, *CESifo Working Paper Series*, no. 8835

Yao Y-C. 1988. Estimating the number of change-points via Schwarz' criterion. *Statistics and Probability Letters*, vol. 6. Issue 3