A structural break cartel screen for dating and detecting collusion

Carsten J. Crede
Centre for Competition Policy
School of Economics
University of East Anglia

CCP Working Paper 15-11

Abstract

In this paper, a new empirical screen for detecting cartels is developed. It can also be used to date the beginning of known conspiracies, which is often difficult in practice. Structural breaks that are induced by cartels in the data generating process (DGP) of industry prices are detected by testing reduced form price equations for structural instability. The new screen is applied to three European markets for pasta products, and it successfully reports the cartels that were present in the Italian and Spanish markets, but finds no suspicious patterns in the French market, which was not cartelised.

Contact Details:
Carsten Crede c.crede@uea.ac.uk

ISSN 1745-9648

A structural break cartel screen for dating and detecting collusion*

Carsten J. Crede

School of Economics, Centre for Competition Policy, and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich NR4 7TJ, UK

September 17, 2015

Abstract

In this paper, a new empirical screen for detecting cartels is developed. It can also be used to date the beginning of known conspiracies, which is often difficult in practice. Structural breaks that are induced by cartels in the data generating process (DGP) of industry prices are detected by testing reduced form price equations for structural instability. The new screen is applied to three European markets for pasta products, and it successfully reports the cartels that were present in the Italian and Spanish markets, but finds no suspicious patterns in the French market, which was not cartelised.

JEL Classification: D43, L41

Keywords: antitrust; cartel; detection; empirical screen

*Corresponding author: Carsten Crede (c.crede@uea.ac.uk).

I thank Rosa Abrantes-Metz, Giuliana Battisti, Farasat Bokhari, Steve Davies, Nils Gutacker, Franco Mariuzzo, Peter Ormosi, and George Papadopoulos as well as seminar participants at NIE-Doctoral Student Colloquium 2015, CCP-UEA 2014 and 2015, CLEEN 2014, and CLaSF 2014 for helpful comments. Any remaining errors are my own.

1 Introduction

Recent success of empirical methods used by competition authorities in the Netherlands, Brazil, and Mexico leading to the detection of several cartels as well as the spectacular case of the LIBOR market manipulation have increased the interest in cartel screens (Abrantes-Metz, 2014). The purpose of these empirical methods is to detect a cartel by identifying patterns in market outcomes that suggest collusion. They are meant "*not to deliver the final evidence based on which colluders will be convicted, but instead to identify markets where empirical red flags are raised and which are worth further investigations.*" (Abrantes-Metz, 2014, p.7). Rather than relying on the prosecution of failed cartels as might be the case for leniency schemes, the purpose of screens is to increase deterrence of antitrust violations by potentially targeting more stable and successful cartels (Harrington, 2007). Further, the level of deterrence that can be attained with fines is limited due to concerns over bankruptcy of convicted firms and over-enforcement. Increasing the detection probability might be necessary to prevent cartel recidivism and to further increase deterrence.

When referring to screening for collusion, it is necessary to distinguish between structural and behavioural screens. While the former refers to approaches that identify markets which are likely to be subject to cartelisation due to industry characteristics, the latter aims at detecting cartels by looking at market outcomes based on an idea how they might be affected by collusion (Abrantes-Metz, 2014; Harrington, 2007). The literature on behavioural cartel screens has grown significantly in the last decade. Most notable are the contributions of Abrantes-Metz *et al.* (2006) and Bolotova *et al.* (2008), who suggest cartel screens based on the analysis of price variance in an industry.

In this article, a new behavioural cartel screen is proposed that is based on identifying structural breaks in the data generating process (DGP) of industry prices. A DGP characterising competition is established with a reduced form price equation based on competitive periods and then used to test suspicious periods for collusion. Structural break tests are applied to check for structural instability in the DGP. The underlying idea is that a cartel alters the dynamics of industry pricing (and therefore the DGP) by raising prices arbitrarily or introducing collusive technologies to operate the cartel. In the absence of other logical explanations for the structural breaks, this raises an empirical red flag suggesting that there might be a cartel and that the market requires further investigation. In addition to cartel screening, the structural break cartel screen is suitable to date the start of a conspiracy that has already been detected. This is often an issue in antitrust litigation, when it is uncertain whether the earliest written evidence that has been found of a cartel represents the start of the conspiracy.

Application of the new screen is discussed based on three European pasta product markets, two of which were cartelised. Each market has different features, which allows to test the screen under varying circumstances: The market in Italy featured a cartel as well as a significant level of tacit collusion after cartel breakdown. The cartel in the Spanish market lasted only 3 months, which renders it difficult to detect its impact on market prices. The French market did not feature a cartel, but saw prices rise significantly during 2007 due to a strong input cost shock. Results show that the screen successfully detects the cartels in Italy and Spain, but does not report a false positive of a cartel in France.

The structure of the article is as follows. In Section 2, the relevant literature for cartel screens is summarised. The methodology is discussd in Section 3: the new screen is presented in Section 3.1, suitable structural break tests for the screen are introduced in Section 3.2, and the necessary procedual steps in the application of the new methodology is outlined in Section 3.3. The screen is applied to the three European pasta industries in Section 4: the industry is presented in Section 4.1, the determination of DGPs of competition is carried out in Section 4.2, structural break tests are carried out in Section 4.3, and the number and dates of breaks is estimated in Section 4.4. Advantages and limitations of the new cartel screen are discussed in Section 5. Section 6 concludes.

2 Literature Review

According to Abrantes-Metz (2014) there are two important rules with respect to designing and applying a cartel screen. On the one hand, the screen must be fitted to the particular industry under investigation to ensure proper identification. On the other hand, identification further depends on the quality of the data that is used for the screen. This knowledge has to be used to develop an idea how collusion or cheating might alter market outcomes in order to use statistical methods to test for significant changes in the considered market outcomes. These changes are signs of collusion. Harrington (2007) provides an extensive discussion of these *collusive markers*. For prices these include: sudden increases in list prices with a declined variation of discounts across customers, a number of substantial price increases following a trend of significant price reductions, simultaneous price increases and reduction of imports, an increase in positive correlation between different firms' pricing, a reduction of price differences between competing firms for products and ancillary services, low price variance, and prices being subject to regime switches.

The majority of behavioural screens so far has been suggested for bid-rigging conspiracies, which are now regularly used in auctions. Porter (2005) provides an extensive survey of the literature on screening for big-rigging cartels. For example, Porter and Zona (1993) find in a study of a bidding ring for highway-paving construction tasks in New York that while the lowest bid of a conspirator was most likely to be related to the fact that this firm had the lowest cost, this correlation did not exist for the higher bids of other ring members. Abrantes-Metz and Bajari (2010) and Blair and Sokal (2013) provide an overview of the different applications of screens for detecting collusion. In the last decade, researchers began to develop screens for cartels outside auctions. In particular one class of behavioural screens has received much attention recently: Variance-based screens analysing prices rely on the idea that the reduced price variance of firms across time or within geographical clusters is an indicator of collusion. Abrantes-Metz *et al.* (2006) find a significant reduction in price variance for a cartel for frozen perch in Philadelphia between 1984 and 1989 based on a comparison of the price coefficient of variation (price mean divided by its standard deviation) between collusive and competitive periods. Comparing the absence of changes in costs and significant reductions in the market price, they are also able to track the collapse of the conspiracy. They further test for geographical clusters of gasoline retail stores in Louisville, USA, based on the concept of a reduced price variance indicating collusion. In a similar study, Esposito and Ferrero (2006) find reduced price variance for two Italian cartels fixing prices for motor fuel and products sold in pharmacies. A similar but more sophisticated approach is suggested by Heijnen *et al.* (2015), who test the Netherlands' gasoline market for suspicious local clusters with reduced price variance. Bolotova *et al.* (2008) use ARCH and GARCH models to analyse price and price variance changes of the citric acid and lysine cartels. Finding strong support for the former, only mixed evidence is found for the latter. The authors provide a number of reasons that might explain the lack of robust findings regarding the price variance. Abstracting from the methodological explanations for this result, they stress that variance-based screens can fail when cartels are not all-inclusive and do not have full control over the price or abnormal supply or cost shocks affect market outcomes.[1] Blanckenburg *et al.* (2012) test whether information on the mean, kurtosis and skewness of price changes can be used as other moments of price change distributions for 11 major price-fixing cartels. They conclude that only the variance can be seen as a robust indicator for collusion. A different conclusion is reached by Hüschelrath and Veith (2013), who use sequential t-tests to test for significant changes in the mean of prices in the German cement cartel to show that with this approach the cartel could have been detected before it was uncovered by the German competition authority.

3 Methodology

3.1 The screen

As Porter (2005) points out for big rigging cartels, the collusive schemes implemented to overcome operational issues often help to discover the anticompetitive conspiracy. The existing behavioural screens in the literature exploit the existence of these collusive schemes to empirically test for specific changes in market outcomes that are assumed to result from a specific operation of the cartel. For example, the price variance-based screens assume that cartel activity reduces price variance potentially indicating collusion. This is a shortcoming of the existing behavioural screens, as they only work in situations in which the cartel show very specific behavioural pattern – the one assumed by the screen. This imposes a burden on the econometrician carrying out the analysis. As the existing behavioural screens are tailored for specific markets or collusive markers, their functionality critically depends on choosing the correct screen for the market

[1] This is not to be confused with price dispersion between firms in a point in time. Connor (2005) provides an overview of the literature on the reduced price dispersion under collusion and discusses the reasons.

under investigation.

The structural break cartel screen suggested in this article tries to overcome the problem of sensitivity of behavioural cartel screens to the underlying assumptions about cartel conduct by relying on a more general approach to cartels. It merely assumes that a cartel induces a structural break the DGP determining industry pricing over time – whether there are reductions in price variance, price wars, changes in cost pass-through, more cost increases than reductions, sudden price increases after cartel meetings, and so on does not matter. This idea is similar in spirit to the seminal work of Porter (1983), who shows how cartel activity induces regime switching measurable in market outcomes. The approach of the screen is to estimate a reduced-form equation of industry pricing for the industry under consideration and control for structural breaks induced by potential cartel activity. Reduced form price equations are an established and the most common approach to estimate cartel overcharges in antitrust litigation (Baker and Rubinfeld, 1999; Brander and Ross, 2006; Nieberding, 2006). Advantages of this approach are its simplicity to use and its limited demands towards data compared to full demand and supply systems. Usually, price changes are explained with supply and demand shifters as well as variables capturing changes in the market structure. If they capture pricing dynamics in the industry well, in the dummy variable approach an indicator variable flagging the cartel periods measures the average overcharge generated by a cartel. Alternatively, a model estimated based on data outside the cartel periods is used to predict counterfactual competitive but-for prices during the cartel periods in the cartel periods (see, e.g., Nieberding, 2006). In case of a cartel screen, whether and when a cartel might exist is unknown, such that no such cartel dummies or forecasting can be used to identify suspicious pricing.

However, in case the reduced form price equation captures all demand, supply, and market characteristic factors that affect pricing, it can be used for cartel detection. Equation 1 shows a reduced form price equation that can be used to detect cartels

$$\Delta P_i = \alpha_1 \Delta C_i + \alpha_2 \Delta D_i + \alpha_3 \Delta S_i + \epsilon_t \ , \tag{1}$$

where ΔP_i denotes a price change of the industry under consideration in period i, C_i and D_i are vectors of exogenous supply and demand shifters, S_i is a vector of market characteristics, and ϵ_i denotes the model error. All variables are included in the form of the first difference to ensure stationarity and prevent spurious results.[2] As a result, the model describes short-run relationship between the price and its shifters. It features no constant, as the constant is absorbed when the first difference of the regression is taken.[3] As Baker and Rubinfeld (1999) point out, the functional form of the reduced form equation depends on the functional form of the underlying structural model of the industry. Thus, the functional form in Equation 1 should be adjusted to fit the industry, e.g., to a log-log model if there is reason to believe that

[2]In case the first difference of the included variables is nonstationary, further differencing is required until stationarity is achieved.

[3]This further removes the effects of time-invariant and unobserved variables on the price from the regression, which leads to lower requirements towards the availability of data that is necessary to estimate the regression.

this characterises the industry better. Before a cartel forms, the price changes are determined by a DGP of competition. In this situation the coefficients show the relationship between the regressors and the price changes in the industry when it is characterised by competition. A cartel that affects prices induces a structural break in the competitive DGP resulting in a bad model fit and sudden instability of coefficients during the cartel periods.

Thus, the structural break cartel screen tests whether there are significant fluctuations of model parameters in the DGP of competition as identified above over time. In case the fluctuations are significant, the industry should receive an in-depth analysis to assess whether there might be a cartel in the industry. Formally, the hypothesis H_0 and alternative hypothesis H_1 that are tested are

$$H_0 : \alpha_i = \overline{\alpha} \quad \text{and} \quad H_1 : \alpha_i \neq \overline{\alpha} \text{ for } i = 1, \ldots, T, \tag{2}$$

where α_i denotes the full vector of all coefficients at any point of time and $\overline{\alpha}$ is a vector containing the corresponding whole-sample averages of the coefficients. Note that the structural break cartel screen can detect any cartel that alters pricing in the industry. As such, it incorporates testing for many different collusive technologies. Sudden price increases, e.g., increase model residuals significantly and might lead towards arbitrary fluctuations of coefficients, if the price changes cannot be explained by the factors that usually induce price changes. Similarly, a cartel that links prices to a certain input good will result in a significant increase of the coefficient of the corresponding regressor in the regression. A reduced price variance due to collusion in turn should point to a structural break because price changes are predicted based on the competitive DGP. As such, the difference between predicted and actual price changes will increase because of the collusive price hysteresis leading to larger model residuals that induce different coefficients. Generally speaking, any successful attempt of a cartel to affect industry pricing in principle creates a structural break that can be picked up by the screen.

3.2 Structural change tests

A challenge for the structural break screen is the choice of a suitable methodology to test for structural breaks induced by cartel activity. The test should work irrespective of the number of breaks and be able to detect different types of structural breaks, e.g., changes in the level of cost pass-through or arbitrary price increases. In addition, it should tell the researcher how many structural breaks there are and determine the points in time when they occur. Further, it should allow for lagged regressors and dynamic models, which are often necessary to model industry price dynamics, e.g., when cost pass-through is not immediate or price changes are subject to hysteresis.

Dozens of structural break tests can be found in the literature, because "[...] *there are infinitely many conceivable ways of deviation from the null hypothesis of structural stability.*" (Zeileis *et al.*, 2005, p.100). Most of these structural change tests are designed against a specific H_1 hypothesis and feature the highest power against these specific alternatives rendering them too restrictive to be used for the cartel screen.[4] A suitable methodology fulfilling all the above demands is the *generalized fluctuation test framework* that "*includes formal significance tests but [...] the techniques are designed to bring out departures from constancy in a graphic way instead of parameterizing particular types of departure in advance and then developing formal significance tests intended to have high power against these particular alternatives.*" (Brown *et al.*, 1975, pp.149–150).[5] Fluctuation tests test for parameter consistency against the alternative of non-constancy, i.e., they do not rely on any assumption about the type of structural break – may it be triggered by one or several structural breaks, or different sources of departure from constancy. This renders them an ideal framework to be used for the cartel screen.

The underlying idea of fluctuation tests is to test the stability of coefficients in either widening or in moving data windows of fixed size: the sample data is decomposed into sub-samples and the regression model parameters are sequentially calculated for various sub-samples including different time periods. In case of widening data windows, a regression model is estimated, e.g., for the first 10% of the observations. Then, the next period just outside the data window is added to the window and the model parameters are re-calculated – this process is repeated until all observations are included. In case of the moving windows of fixed width, e.g., parameters are obtained sequentially for a regression including first periods 1-30, then 2-31, then 3-32, and so on. In either case, if the DGP is stable and there are no structural breaks, then the parameter estimates should not fluctuate significantly between the sub-sample estimates. However, if there is structural instability, then the parameter estimates are subject to significant fluctuations. To assess whether fluctuations in a data window are significant, an empirical fluctuation process (EFP) is calculated capturing the parameter fluctuations in the data window and is compared to a benchmark. Structural break tests provide both formal significance tests for the H_0 as well as allow for a graphical inspection of the EFPs. While the former provides an easy-to-interpret and familiar procedure to determine whether there is a structural break, the latter enables the econometrician to gain information on the dating and length of a structural break. Structural breaks induce the EFPs to systematically deviate from their mean. As such, sudden trends in the EFPs indicate structural instability in the model that are significant if the fluctuations lead the EFP to cross boundaries, which can be plotted graphically as well.

The literature on fluctuation tests can be categorised into residual-based approaches and parameter estimates-based tests. Residual-based approaches detect structural breaks by focusing on fluctuations of residuals over time, whereas parameter estimates-tests directly test all coefficients separately for fluctuations. In the following, three tests contained in the generalized

[4]Examples include tests for parameter constancy against the alternative hypotheses of a single shift (Andrews, 1993; Andrews and Ploberger, 1994), random walks (Nyblom, 1989), and unit roots (Zivot and Andrews, 1992).

[5]Kuan and Hornik (1995) show that the different fluctuation tests can be combined in the generalized fluctuation test framework.

fluctuation test framework which are particularly useful for cartel screening are presented.[6] Each test has the highest power in certain situations and all of them rely on the same assumptions. They require that heteroskedasticity and autocorrelation must not be present in the estimated regression unless heteroskedasticity and autocorrelation consistent covariance matrices are used. Further, dynamic and lagged regressors are allowed as long as they are stationary.[7]

The first test used is the residual-based OLS-CUSUM test by Ploberger and Krämer (1992) with the improved alternative boundaries suggested by Zeileis (2004). While performing worse than the OLS-MOSUM test when there are several structural breaks, it performs better than the OLS-MOSUM test when there is a single break, features good finite sample properties (in particular in dynamic models), and it well suited to detect relatively short-lasting structural instability (Chu *et al.*, 1995a). Starting from an initial data window at the beginning of the sample, the OLS-CUSUM test sequentially tests the data for structural breaks based on the cumulative sum of residuals. In period i, the model is calibrated for periods 1 until i-1 and an expected value of the dependent variable is predicted for period i. If the regressions parameters are constant across time, the residuals should fluctuate around zero. If, however, there is structural instability, the residuals will systematically increase.

Rather than relying on a growing window of observations like the OLS-CUSUM test, the residual-based OLS-MOSUM test of Chu *et al.* (1995a) is based on a window of fixed width that is sequentially "moved" through the whole sample to calculate the moving sums of OLS residuals. It reports structural instability when the sum of residuals in the data window exceeds a suitable boundary. As it is only based on a window of the sample that changes across time, it detects multiple structural breaks faster and more clearly than tests based on a growing window. The width of the window has to be chosen by specifying a bandwidth h for the test that states how many % of the data are included in the moving window.

In contrast to the residual-based OLS-CUSUM and OLS-MOSUM tests, the Moving Estimates (ME) test of Chu *et al.* (1995b) controls for the fluctuations within all regression parameters. As such, an EFP can be observed for each coefficient. This makes the ME test an attractive choice as it allows to gain inference about the source of a structural break, i.e., which coefficients are subject to structural instability. Further, the ME test is – unlike the above residual-based tests – sensitive to orthogonal shifts of the mean regressor (Chu *et al.*, 1995a; Ploberger and Krämer, 1992; Krämer *et al.*, 1988), whereas the estimates-based tests do not share this shortcoming (Zeileis *et al.*, 2005). The ME test measures the fluctuations of coefficients by comparing the moving-window estimates to the whole-sample average to pick up structural breaks. As such, it becomes apparent why the ME test has approximately constant detection probabilities across time and can detect several structural changes. As in the OLS-MOSUM test, a bandwidth h

[6] Further tests include tests based on recursive CUSUM and MOSUM processes rather than OLS-based residuals (Brown *et al.*, 1975; Chu *et al.*, 1995a), and recursive estimates processes (Ploberger *et al.*, 1989). While no test is dominated any other test under all circumstances, the OLS-based CUSUM and well as MOSUM tests as well as the Moving Estimates test introduced below in many circumstances have higher statistical power in the presence of several structural breaks (Kim, 2011).

[7] A more in-depth description of the methodology and the assumptions are presented in the Appendix.

has to be chosen to apply the test.

While the fluctuation tests above are useful for determining the existence of structural breaks in the data, the approach suggested by Bai and Perron (1998, 2003) is better suited to determine the number of breaks as well as their dates. This follows from the fact that information about the number and dates of structural breaks determined with the generalized fluctuation tests are obtained from a graphical inspection of the EFPs. This does not always allow to identify a precise date of a structural break. Bai and Perron (1998, 2003) develop a procedure to determine the number of breaks as well as the optimal breakdates within a regression. They propose an algorithm based on dynamic programming to determine the number and dates of structural breaks in the estimated regression. The structural breaks are simultaneously obtained based on the global minimisation of the sum of squared residuals (RSS). Selection of the number of breaks is carried out by sequential checks of optimal single break partitions (for details, see Bai and Perron, 2003). In essence, the approach computes the RSS for models based on a different number of segments in the data. Increasing the number of segments towards the true number of segments leads to a significant reduction in the overall RSS. However, assuming more than the true number of segments does not induce a further significant reduction in the RSS. Similar to the bandwidth in the OLS-MOSUM and ME tests, the econometrician needs so specify a minimal segment width in this test. Different information criteria are considered to assess how many segments (and therefore structural breaks) are in the data: the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and a modified Schwarz criterion (LWZ) as proposed by Liu *et al.* (1997).

The approach requires assumptions similar to those required by the generalized fluctuation test framework, i.e., detrended and stationary variables are required, lagged (dependent) variables can be included, and heteroskedasticity and autocorrelation (between segments) is allowed provided that consistent estimators are used. Therefore, we can comfortably apply the approach of Bai and Perron (1998, 2003) for dating structural breaks to the same specification used for the generalized fluctuation tests. Bai and Perron (2003) report results of simulations to test which of the information criteria performs best with respect to a correct selection of the true number of structural breaks. They conclude that the AIC generally does not perform well. The BIC performs bad when there are no structural breaks present by overstating the true number of breaks in the data, i.e., they suggest breaks when there are none, but performs well in the presence of structural breaks. The LWZ in turn performs well if there is no structural break in the underlying DGP, but “very bad” (Bai and Perron, 2003, p.15) when breaks are present.

Due to the potential problems in determining the existence of structural breaks in dating approach of Bai and Perron (1998, 2003), their approach should be used in conjunction with the generalized fluctuation tests. The existence of structural break should be determined with the generalized fluctuation framework. If one or several structural breaks have been found with different tests from the framework, the approach of Bai and Perron (1998, 2003) should be used to establish the number and dates of the breaks based on the BIC criterion. In other words, the above procedure for dating structural breaks should not be used in isolation to test whether there is structural instability. This is an aspect that is of particular importance for the proposed cartel screen, in which a correct identification is crucial. This conservative and robust approach combines the strengths of both methodologies and reduces the risk of type I and II errors.

3.3 Estimation procedure and choice of specifications

In the following, the application of the the cartel screen is discussed, i.e. how the reduced form price equation has to be combined with the structural break tests to ensure correct inference of the new approach. Chu *et al.* (1996) provide a useful list of aspects with respect to factors that determine the quality of monitoring with fluctuation tests. Monitoring is not discussed here but refers to the presented tests' ability to monitor incoming new data for structural breaks on a fixed set of historical data (changing the definition of what is seen as the whole sample) that serves as a benchmark. The underlying intuition holds equally well for all historical and monitoring-based fluctuation tests. First, the precision of the fluctuation tests depends on how accurately the DGP is estimated. In other words, the higher the model fit the better is the chance to pick up a structural break. Second, chances of detecting a structural break increase with the magnitude of the structural break and the induced parameter change in the model. The larger the change in the DGP, the easier it is to distinguish observed changes from random noise. Third, shorter windows in the OLS-MOSUM and ME tests, i.e., a smaller bandwidth h, lead to a faster detection of breaks but longer windows have better finite sample properties. It is not known which bandwidth performs best. However, this indicates that it can be smaller if the data set is large, and should be increased in particularly small data sets.

To reduce the risk of type I and II errors, it is suggested to run the OLS-MOSUM and ME tests with multiple bandwidths and only conclude that breaks are present if the majority of tests point towards structural instability. Segment widths for the dynamic programming algorithm of to date breaks by Bai and Perron (1998, 2003) should be smaller than window widths of the fluctuation tests. If the distance between two breaks in the data is smaller than the minimal segment width defined for the algorithm, correct inference on the number and exact dates of the breaks is not possible. This is discussed in greater detail in Section 4.4.

As the cartel screen is based on detecting structural breaks in the DGP of industry pricing dynamics, the specification needs to contain all major, time-variant and relevant supply and demand shifters as well as industry characteristics significantly affecting industry prices to ensure proper identification. In this respect, the approach is the same as that used to estimate reduced-form price equations to determine cartel overcharges. Inclusion of all relevant time-variant price shifters is necessary to ensure that structural breaks that are detected are not caused by omitted variables. This includes substitutes to the product under investigation if they are believed to have significant effects on its price. The more of the shifters are included, the more certain can the econometrician be that structural breaks are caused by a cartel. However, the shifters need not be included when they are stable. In that case, absent significant technological change, they should not induce changes in industry prices, and are dropped from the reduced-form price equation in the transformation of the data from levels to first differences. Price shifters which have a negligible impact on prices are not essential to identification either, as they induce little to any measurable price change. Similarly, variables characterising the competitive environment do not need to be included if it not subject to significant changes in the time periods under consideration. This aspect can greatly simplify the analysis, in particular if not all data is available. As a result, it poses lower requirements towards the availability of data than for example cost-based approaches or simulations used to estimate competitive counterfactual prices for periods that are checked for collusion.

When setting up a regression model as defined in Equation 1, the specification has to be chosen carefully. It should include the relevant price shifters, and must not feature any endogenous regressors. As required by the generalized fluctuation test framework, all included regressors (and the regressand) need to be stationary. Therefore, all variables need to be controlled for stationarity with an appropriate test.[8] Further, variables that are subject to a trend need to be detrended first and then checked for stationarity. These steps are necessary to prevent spurious relationships from biasing the estimates and explosive variances to affect the statistical inference. After the above steps have been taken, an optimal specification should be chosen. In order to identify a valid DGP characterising competition in the industry, the specification should be determined based on competitive periods only. For this purpose, the suggested fluctuation test can be used by regressing the dependent variable under consideration on a constant to assess when sufficient changes in the DGP might occur. This check can be supplemented by the structural break dating methodology suggested by Bai and Perron (1998, 2003) to determine the breakdate. Further, if data on unaffected industries in other regions/countries is available, these regional benchmarks can be used to determine from which point in time on the market under consideration was subject to different pricing dynamics point to a structural break. As a cartel likely induces price increases or heavy price fluctuations triggered by price wars or cartel breakdown, periods with suspicious price movements should be excluded for determining the specification.

In order to improve the finite sample properties of the estimates given limited number of observations available in growing or moving windows used to estimate coefficients, inclusion of regressors should be parsimonious. This also prevents the risk of overfitting in the model, which might negatively affect inference in the structural break tests. The rule here is as many regressors as necessary, but as few as possible. Lagged regressors as well as lagged dependent variables are allowed and a compromise between goodness of fit and number of included lags has to be made. If seasonality is present in the DGP that has been determined, it needs to be removed from the data to ensure correct inference. While different methodologies exist to

[8]Tests for unit roots have to be chosen carefully. Structural breaks in the time series can be misinterpreted as nonstationarity by Augmented Dickey Fuller and Phillips-Perron tests. A suitable unit root test is proposed by Zivot and Andrews (1992), which tests for unit roots against the alternative of a structural break.

remove seasonality, the seasonal-trend decomposition procedure based on nonparametric loess smoothers (STL) by Cleveland *et al.* (1990) is a suitable tool for this task. A strength of this approach is that it allows for a robust estimation of seasonality that is not biased by aberrant behaviour of the data, as can follow from the effect of cartels on prices. Removing rather than modelling seasonality is necessary for the screen, as the growing or moving data windows do not allow for a robust distinction between seasonality and cartel price changes. The DGP has to be re-fitted to the deseasoned data to ensure a proper fit.

All structural break estimations are carried out in the statistical software package R with the package *strucchange* by Zeileis *et al.* (2002). In order to reduce the risk of Type I errors, a conservative approach is suggested. Results of the three fluctuation tests above are used for the analysis and results with respect to the existence and number of breaks are compared with each other. Only if all three tests report a structural break do we conclude that there is significant evidence for a structural break. Based on these results, the structural breaks identified with the generalized fluctuation tests are dated with the approach of Bai and Perron (1998, 2003).

4 Application: European pasta industries

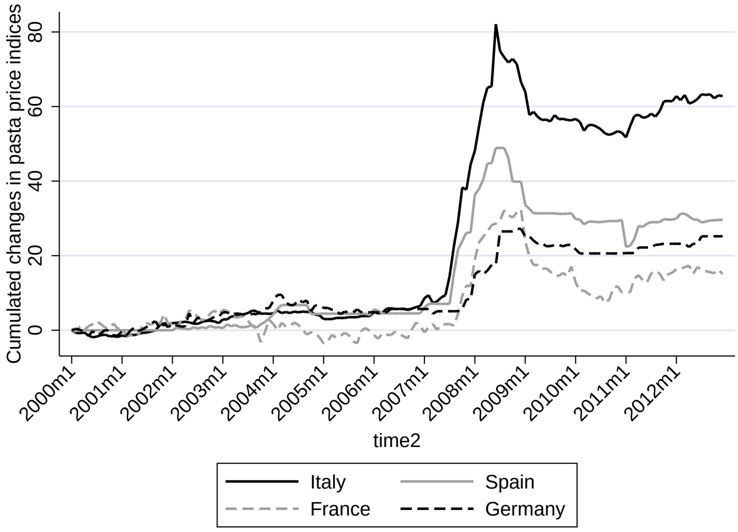
4.1 The industry

In order to test the new screen with a real world example, it is applied to the pasta industries in Italy, Spain, and France based on monthly data between 2003 and 2012. While the former two were subject to cartels between October 2006 to March 2008 and July to October 2007, respectively, the latter did not feature any cartel (Notaro, 2014; Ordóñez-de Haro and Torres, 2014). As such, we would expect a properly working cartel screen to detect cartels in Italy and Spain, but not to report structural breaks in France. Cartel formation in Italy and Spain was triggered by significant price increases in the main input good in pasta production – durum wheat – during 2006 and 2007. These markets represent useful cases to test the screen under difficult conditions: In Italy, a significant level of post-cartel tacit collusion after the cartel prevents a clear breakdown of the DGP of collusion. In Spain, the cartel lasted only for three months, providing only very few observations to detect collusive behaviour. In France, market prices rose significantly due to a strong durum wheat cost shock in the absence of a cartel. Figure 1 contains the cumulative price changes in % between 2000 and 2012 in the pasta price indices of the four largest pasta producers in Europe by order: Italy, Germany, Spain, and France (International Pasta Organisation, 2012).[9] Absent significant changes in supply, demand, and market characteristics across countries, differences in the reaction of markets to common supply shocks allow for a preliminary analysis of market prices to the cost shock.

While the prices of pasta develop similarly in the four countries before 2007, the indices rise significantly further and at an earlier date in Italy and Spain than in France and Germany.

[9]Unfortunately data on German durum wheat prices is unavailable, preventing the analysis of this market.

Figure 1: Cumulated price changes in pasta price indices



While the French and German industries suggest competitive price increases due to the common input cost shock of around 30%, prices at times rose by up to 82.1% and 48.9% in Italy and Spain. The magnitude of differences in the reactions to the increases in durum wheat prices are suspicious and would have been unlikely to be explained by the mere differences of industry characteristics between the countries. Further, the earlier increase in prices in Italy and Spain suggest that the cartels in these countries could have engaged in coordinated cost pass-through. Whether the price changes are indeed suspicious is tested with the structural break cartel screen below.

Notaro (2014) discusses the Italian pasta cartel in detail and estimates the average overcharge to be around 11% of the competitive price. The cartel covered 76% of the market (90% if private labels produced by the cartelists are included). Market coverage of the Spanish cartel was less complete with a joint market share of 55-60% including private labels produced by the cartel members (Comisión Nacional de la Competencia, 2009). Between May 2006 and May 2008, the price charged by the producers to retailers in Italy increased by about 51.8% (Italian Competition Authority, 2009). However, Notaro (2014) shows with a dynamic treatment effects analysis that the observed cost changes including the significant price increases for durum wheat could not explain the price increases. In addition, he argues that the market was characterised by significant levels of tacit collusion after breakdown. This can easily be seen in Figure 1, as the prices in Italy remain significantly above the prices of the other European countries. Notaro (2014) reports that domestic and import durum, as well as labour and energy costs represented 73% and 77% of the total direct costs and 54% and 60% of the total costs of the Italian pasta industry in 2006 and 2007, respectively. The strong dependence of pasta prices on few input goods renders the industry suitable for an analysis with the proposed cartel screen.

4.2 Specification of the DGPs of competition

In the first analytical step, a representative model describing the DGPs of competition in the markets has to be established based on periods that are assumed to be characterised by competition. The strong dependence of pasta prices on few input goods renders the industry suitable for an analysis with the proposed cartel screen. In the following, five cost shifters are used together with a demand shifter for the screen. All variables are used as first differences to achieve stationarity. Denote *Domestic durum* and *International durum* as the domestic and the international durum price indices. The domestic durum wheat prices capture the majority of actual costs for the main input good in pasta production, and the prices usually are fixed in contracts between suppliers and customers several months in advance. The international durum prices measure both import prices as well as expected price changes of durum wheat. Further, let *Labour costs*, *Energy*, *Energy sq.*, and *Borrowing costs*, denote the Italian price indices for industry labour and (squared) energy costs, and costs of borrowing capital for companies. *Expenditure* is included as a demand shifter and measures household expenditure on goods to approximate demand fluctuations. Table 4 in the Appendix contains a description of all variables used in the analysis as well as the data sources. The variables are detrended if necessary and tested for stationarity: augmented Dickey-Fuller and Phillips-Perron tests for unit roots confirm that all variables are stationary at a 1% significance level. A potential caveat of the analysis below is the absence of variables that control for market characteristics due to a lack of data. Thus, the implicit assumption that has to be imposed is that market characteristics in the industries under consideration were roughly stable across time.

The optimal specification for the DGP that characterises competition in the industry has to be determined based on unsuspicious, non-collusive periods only. Here, an inspection of the pasta industry prices in Figure 1 suggests that prices tend to increase significantly during 2007, suggesting that these periods should be excluded to establish the specification. If no data is available for cross-country comparisons or no apparent suspicious periods arise from a graphical inspection, identification of potentially collusive periods can further be achieved by regressing price changes on a constant and testing for parameter instability with the structural break tests. Indeed, such tests report structural breaks indicating changes in the average price changes in the mid of 2007 for all three countries. Therefore, a conservative approach is taken and only periods before 2007 are considered for the computation of the DGPs of competition.[10]

The pasta industry is not characterized by contemporaneous cost pass-through, as input prices are often determined by fixed contracts several months in advance. Further, the industries are characterised by different market characteristics and dominated by different firms. As such, there will be divergent lag structures across countries to capture the DGP of competition. A parsimonious lag structure varying across the countries is used that provides in each case the

[10]For Italy, only the periods prior to the cartelisation of the industry in October 2006 are used. Yet, this additional exclusion restriction has no effects on results, as the cartel did not influence market prices before June 2007, as will be shown below.

best fit for the observed pasta pricing dynamics in the pre-cartel periods. While not all supply and demand shifters are always significant, they are included due to their theoretical importance for prices as well as the fact that they might become significant determinants of price changes outside the time frame used to determine the DGP of competition to prevent potential omitted variable bias. No lagged dependent variable is included in any of the regressions, as there is neither theoretical reason to do so nor observable patters of price hysteresis in the pasta industry. The specifications of the different DGPs capturing competition can be found in Table 1.

Table 1: Specification of DGPs of competition

	Italy	Spain	France
Domestic durum	0.017^{**}	0.047^{***}	0.093^{***}
	(0.008)	(0.012)	(0.032)
	4	3	4
International durum	0.007^{**}	0.009^{**}	0.013^{**}
	(0.003)	(0.003)	(0.006)
	3	3	0
	-	-	0.021^{***}
			(0.007)
			4
Labour costs	0.012	0.069	2.791^{*}
	(0.011)	(0.053)	(1.425)
	0	2	5
Energy costs	0.175^{***}	0.394^{***}	0.331^{***}
	(0.044)	(0.061)	(0.103)
	3	1	6
Energy costs sq.	-	-	-0.209^{***}
			(0.045)
			6
Borrowing costs	-1.114	-0.911	3.967^{***}
	(0.692)	(0.551)	(1.340)
	2	4	1
Household expenditure	0.000	0.140^{*}	0.414
	(0.000)	(0.083)	(0.412)
	4	6	2
Adj. R^2	0.373	0.486	0.523
Observations	42	43	46

*Notes: The dependent variables are the domestic producer prices for pasta products. All variables are included as first differences. The first row for each variable features the coefficient, with ***, **, and * showing significance of the coefficient at a 1%, 5%, and 10% significance level. The second row contains the (robust) standard errors in brackets. The third row indicates which lag of the variable is used with 0 denoting the contemporaneuous value*

In line with theory, coefficients of all significant cost and demand shifters are positive. The results are robust to different functional forms. Note that the estimates of Table 1 are not used for the analysis. They are merely used to establish the lag structure that is later tested for stability with the structural break tests and for various specification tests. Heteroskedasticity is suspected for the Spanish market, such that heteroskedasticity and autocorrelation-robust standard errors are reported throughout the analysis for the market. The Italian and French markets show no signs of either heteroskedasticity or autocorrelation as reported by appropriate tests, such that OLS residuals are used for these markets throughout the analysis. The careful use of heteroskedasticity and autocorrelation consistent covariance matrices is necessary, as samples in the growing or moving windows in the fluctuation tests typically are small. This ensures that poor asymptotics in small samples do not lead to significant biases in the covariance matrices leading to wrong inference (Angrist and Pischke, 2008; Wooldridge, 2012). Seasonality is detected in regressions for the French pasta industry, whereas no seasonality is present in the Italian and Spanish markets. Therefore, the seasonality in the French data is removed with the nonparametric STL deseasoning approach by Cleveland *et al.* (1990) using the robust option to prevent aberrant behaviour to affect the results. The deseasoned time series of French pasta prices is used throughout the analysis. Ramsey's RESET test results for all regressions indicate that the linear functional form is correct (Ramsey, 1969): in case of the French industry, energy costs have to be quadratic such that the functional form of the specification is valid.

4.3 Determining the existence of structural breaks

In the second analytical step, the OLS-CUSUM, OLS-MOSUM and ME tests are applied to the specifications proposed in Table 1 to test for structural breaks in the DGPs identified above. As such, the regressions are tested for stability throughout all available time periods between 2003 and 2012. Table 2 reports the p-values for the tested H_0 hypotheses of stability of the regression models throughout time. As both OLS-MOSUM and ME tests are based on a moving window, the window width h in terms of % with respect to the whole sample has to be specified. The choice of h is no obvious task and requires the researcher to make a choice. As pointed out in Section 3.3, shorter windows increase the sensitivity of the tests to fluctuations of coefficients, and allows to better detect periods of short instability in the data. Yet, overly small window widths also increase the risk of false positives. Therefore, four window widths $h = [0.15, 0.2, 0.25, 0.3]$ are tested and results across the different tests are compared with each other. To reduce the risk of Type I and Type II errors, we will only conclude that there is a structural break if evidence for it arises consistently in the majority of tests.

The results show that there is strong evidence for structural breaks in the DGP of pasta price changes both in Italy and Spain. Only for large window widths, the ME test fails to detect breaks in the Italian market. This might occur if the industry is characterised by a distinct DGP for few periods only, such that the coefficients in the window do not fully capture the changed DGP. Yet, given the strong and largely robust results, the structural break cartel screen suggests that DGPs changed significantly both in Italy and Spain, which creates serious doubt about the state of competition in the markets. In other words, price changes are governed by significantly different processes at different points in time in these markets. A different conclusion arises for the French market, in which there was no cartel. Indeed, despite the strong

Table 2: Break test p-values

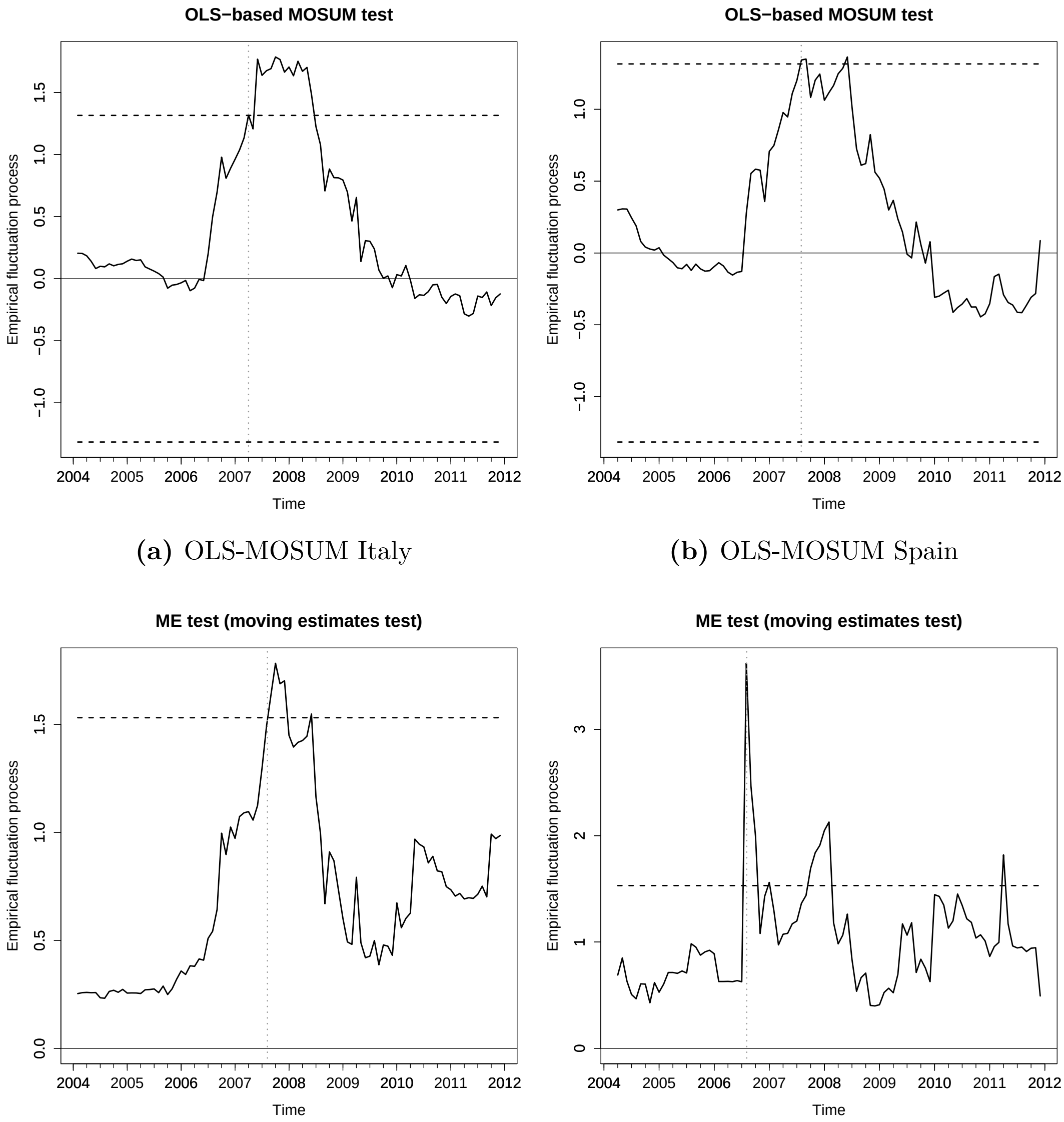
	h	Italy	Spain	France
OLS-CUSUM	-	0.000	0.015	0.755
OLS-MOSUM	0.15	0.010	0.014	0.088
	0.20	0.010	0.036	0.043
	0.25	0.010	0.050	0.123
	0.30	0.010	0.052	0.151
ME	0.15	0.010	0.010	0.034
	0.20	0.010	0.010	0.276
	0.25	0.089	0.010	0.235
	0.30	0.160	0.010	0.212

and sudden price increases during 2007, most of the tests report that there is no structural break in the data. Note that there are Type I errors for small window widths for OLS-MOSUM and ME tests pointing towards false positives for the French market. Nevertheless, given that the majority of results fails to reject the H_0 hypothesis of stability of the coefficient over time, it can be concluded that that the DGP is stable across time in France. Put differently, price changes in the French pasta industry can be explained throughout time by established patterns of cost pass-through and adaptation to changes in demand. This result shows the advantage of including price shifters in the estimated regressions rather than relying on an analysis of the dependent variable time series alone. If it is not controlled for the price shifters, a structural break in the French pasta price would have incorrectly be picked up by the OLS-MOSUM and ME tests when the price changes are only regressed on a constant. This shows that the inclusion of price shifters significantly improves identification by relaxing the assumption about constancy of all supply and demand factors.

In the third analytical step, the number and the dates of the breaks in the Italian and French markets are determined. A graphical inspection of the EFPs of the residuals in the different tests can be used to gain an impression of the structural instability in the models. The EFPs of the OLS-MOSUM and ME tests for the Italian and Spanish markets are plotted (black lines) together with the boundaries of the corresponding 95% confidence intervals (horizontal grey dashed lines) in Figure 2.

The DGP of pasta prices in Italy is stable until mid 2006 when the EFP suddenly increases significantly and crossing the boundary in early 2007. Both OLS-MOSUM and ME tests point towards a significant instability of the DGP between early 2007 and the mid of 2008, after which the industry reverts to a stable DGP. Yet, as the higher EFP in the ME test after the mid of 2008 shows, the regression model fitted to the DGP of competition provides a worse and less stable fit after cartel breakdown, possibly due to post-cartel tacit collusion. A similar picture arises in the Spanish market. However, note that the patterns in the ME test for the Spanish market should be interpreted with care, as the covariance matrices used to correct for potential heteroskedasticity and autocorrelation might be unreliable given the small sample sizes of the

Figure 2: OLS-MOSUM and ME EFPs for Italy and Spain



(a) OLS-MOSUM Italy

(b) OLS-MOSUM Spain

(c) ME Italy

(d) ME Spain

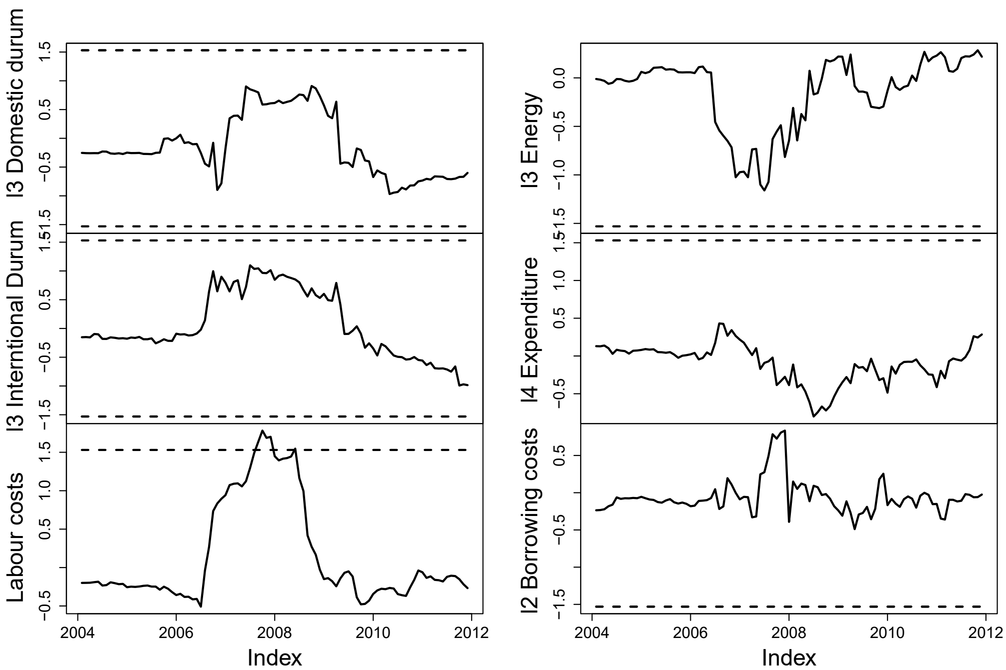
moving windows.

As the ME test is a parameter-based test, we can also look at the EFPs of the model coefficients for, e.g., the Italian market, which are plotted in Figure 3.[11] This allows to gain insights on the sources of instability in the model. Two potential patterns can be expected. On the one hand the cartel could engage in arbitrary price changes not linked to price developments of price shifters. As the estimated regression has no intercept, such a collusive behaviour results in instability of all coefficients, and (arbitrary) significant fluctuations in one or more coefficients. On the other hand, if the cartel links price changes to the development of specific price shifters, a strong and persistent change for that price shifter would be expected, while the other coefficients are not subject to strong and arbitrary fluctuations. The Figure shows that all coefficients are subject to instability starting in early 2007, when the cartel started to affect market prices.

[11]Results of the ME test with respect to the coefficients for the Spanish industry can be found in the Appendix.

However, only the fluctuations in the Labour cost coefficient is significant, despite the fact that – unlike durum wheat – it was not subject to significant price changes. Taken together, this suggests that the structure of cost pass-through was not significantly altered by the cartel, but that the it engaged in arbitrary price increases.

Figure 3: ME test results for coefficient stability in Italy



4.4 Dating the breaks

In the fourth and last step of the analysis, the dynamic programming algorithm of Bai and Perron (1998, 2003) is used to determine the number and dates of the structural breaks in the Italian and Spanish markets. For this purpose, the specifications established in Section 4.2 are again applied to the whole sample, and the optimal partition into different segments is carried out by the algorithm. Similar to the choice of window widths for the OLS-MOSUM and OLS-CUSUM tests, a decision with respect to the minimal segment width has to be chosen for the dynamic programming algorithm. Segment widths should not be chosen too large, as this could negatively affect identification: if the segment width is larger than the distance between two structural breaks, the breaks cannot be dated correctly. Therefore, segment widths of $s = [0.07, 0.08, 0.1]$ are used for dating the structural breaks.[12] Results are reported in Table 3.

The results show that the dating algorithm detects three structural breaks for the Italian industry, and either three or four breaks for the Spanish industry. For Italy, the algorithm detects a first break induced by cartel formation in June 2007. A comparison of this result with

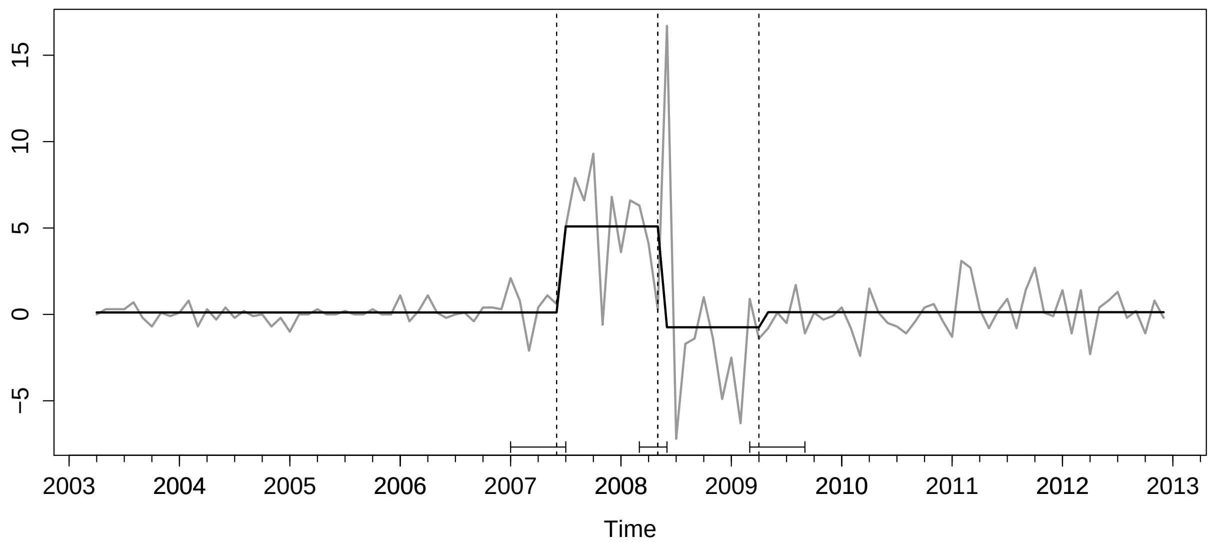
[12]A smaller segment width potentially increases the number of structural breaks reported by the dynamic programming algorithm of Bai and Perron (1998, 2003). However, as the identification of the existence of structural breaks is carried out with the fluctuation tests, this poses no significant problem for the structural break cartel screen.

Table 3: Dates of structural breaks

s	Italy			Spain			
0.07	6/2007	5/2008	1/2009	4/2007	12/2007	10/2008	5/2011
0.08	6/2007	5/2008	2/2009	3/2007	12/2007	10/2008	5/2011
0.10	6/2007	5/2008	4/2009	7/2007	-	10/2008	5/2011

the cumulated price changes as shown in Figure 1 suggests that this break coincides with the first period when the cartel that came into existence in October 2006 had a first visible impact on industry prices. Further, as can be seen in the Figure, industry prices collapsed by about 20% after cartel detection in March 2008. The dating algorithm dates the breakdown of collusive prices to June 2008. The third reported break varies by segment width and captures the point in time in which the industry entered into a state characterised by tacit collusion, as became evident by a comparison to cumulated price changes in the other European markets after cartel breakdown. A visualisation for the breaks identified for the Italian market for $s = 0.1$ can be found in Figure 4, where the price changes over time are shown by the grey line, and the mean in each of distinct phases of the cartel are shown by the black line. The vertical dashed lines show the detected structural breaks, and the small horizontal black lines at the bottom of the break lines represent the 95% confidence intervals for the breaks. Thus, the cartel screen works well for the Italian market, in which it detects the cartel 9 months before the Italian Competition Authority.

Figure 4: Identified structural breaks in the Italian pasta industry



The choice of the segment width parameter can affect the dating of a break when the distance between two breaks is smaller than the minimal segment width. This becomes evident in particular for the Spanish market. The dating algorithm reports three or four structural breaks in Spanish pasta prices. The short length of the Spanish pasta cartel – it lasted only between July and September 2007 – causes problems for the dating algorithm, as the structural instability is much shorter than the minimum segment width of the algorithm. Compared to the Italian market, this leads to less stable break dates across the different segment widths in the Spanish markets. Therefore, neither the establishment nor the breakdown of the Spanish pasta cartel after detection can be dated with precision for $s = 0.07$ and $s = 0.08$ leading to some arbitrary results for the exact break dates. This highlights a caveat of the methodology used to date the structural breaks. Only for $s = 0.1$ does the reported structural break with the start of the cartel, but no break is reported for its breakdown. However, two breaks can be dated consistently, which are further apart from each other. The first of these breaks is found in October 2008, when Spanish pasta prices stop to decline and converge towards the prices in the non-cartelised French and German markets. Similar to Italy, this can be interpreted as the start of potential post-cartel tacit collusion in the Spanish market. The second break is found in June 2011, when a noticeable price decline not existent in the other markets occurs, after which the price variance in the Spanish market increases. This points towards a collapse of tacit collusion in the market ending the resulting price hysteresis and bringing price variance more in line with the French market.

5 Discussion

In the following, advantages and limitations are discussed. With respect to limitations in scope, first and foremost – as all behavioural cartel screens – its application assumes that a cartel has significant control over the market price. In other words, it can only detect a cartel that alters the DGP of industry prices. Similarly, when used to date the beginning of a conspiracy, it can only detect a cartel once it has a significant effect on prices. Second, determining the number of breaks and in particular the correct dates can be difficult when the break dates are not far apart. In these cases the minimum segment length required by the chosen specification might be larger than the distance between the breaks resulting in wrong or missing break dates. Third, in line with some of the other behavioural screens, it cannot distinguish between explicit and tacit collusion. Fourth, the screen assumes that at least some data is available about price shifters, and that unobserved but important variables are not subject to significant instability across time. This assessment on the likelihood of significant changes in unobserved variables is necessary to attribute potential structural breaks in the DGP to cartelisation rather than to substantial and persistent demand and supply shocks or major changes in the industry. However, note that the requirements with respect to observable supply shifters are not as high as in the estimation of cartel overcharges based on cost information. What is required is not the construction of the exact cost structure of the market, but to capture important input cost changes affecting prices in an unbiased way. This implies that significant levels of (unobserved) time-invariant fixed costs are not problematic, as they do not lead to fluctuations of prices (see, e.g., Nieberding, 2006).

The cartel screen suggested in this article has a number of advantages compared to other behavioural cartel screens in the literature. The biggest advantage is that it is not based on any particular collusive marker for identification of cartel activity – it only requires the cartel to affect industry prices. As such, it can detect different collusive technologies which reduces the risk of type II errors compared to other cartel screens. In addition, the approach works equally well for both stable and unstable cartels. While the former are likely to be characterised by a reduced price variance during collusive periods, the latter might be subject to frequent price wars and cartel recidivism. This follows from the fact the proposed structural break tests can deal equally well for a single or multiple structural breaks. Further, the identification strategy is not based on a manual (arbitrary) choice of time periods that is to be formally tested for collusion. Therefore, it avoids the problem faced, e.g., in the ARCH-based price variance test of Bolotova *et al.* (2008), in which the results of formal hypothesis tests depend on the exact definition of time periods to be tested for collusion. It is the cartel screen that informs the econometrician about the likely start (and end) of collusion. This renders the approach ideal not just for testing the start of detected cartels, but also for a proactive search for cartels. Finally, it is not as prone to manipulation as the established cartel screens. It is often argued by critics of cartel screens that they are useless because cartels will react to screening by developing mechanisms that trick the screens. For example, a cartel that knows that the market is screened with a price-variance based cartel screen could easily avoid detection by engaging in arbitrary price changes to keep the price variance stable over time. This problem does not exist for the proposed structural break cartel screen. Any artificial cartel price increase *does* alter the DGP. The only way for a cartel to avoid detection would be to engage in a large number of very small price increases that would not be detected as significant alterations of the DGP but being hidden in the margin of error. However, this has destabilising effects on collusion, as can easily be shown in infinitely repeated games.

The structural break cartel screen should be seen as a useful tool for dating and detecting cartels. While a conservative econometric approach, e.g., by requiring different structural break tests to unanimously report a break or across different specifications, can reduce the risk of type I errors, a change in the DGP should not be seen as definite evidence of a cartel. Instead, a significant structural break calls for a more in-depth analysis of the industry. Only if the supplementary analysis of the market shows that the change in industry pricing cannot be attributed to other factors not considered in the econometric model can the screen be seen to provide substantial evidence for collusion. Further, it should be seen as complementary to the other behavioural cartel screens. One the one hand, it is not always possible to obtain all data necessary for a particular type of screen. On the other hand, in some situations behavioural screens which are specifically tailored for a market might provide more compelling evidence based on a technically simple analysis.

6 Conclusion

In this article a new empirical methodology to date the beginning and end of cartels and detect previously unknown cartels is proposed and tested. Based on the idea that cartels change the data generating process (DGP) of industry prices compared to competitive periods, an approach is presented that detects collusion by testing the stability of this DGP with structural break tests. Different structural break tests suitable for this task are discussed and those from the generalized fluctuation test framework are found to be the most useful for determining the existence structural breaks in an industry. The approach by Bai and Perron (1998, 2003) is suggested as a useful complement to provide exact estimates of both the number and dates of structural breaks. The new structural break cartel screen is applied to the three European pasta industries: the markets in Italy and Spain were cartelised, while no cartel was present in France. The screen successfully detects the Italian pasta cartel 9 months before the Italian Competition Authority learned about the cartel. Despite its shirt length of three months only, the Spanish cartel is detected as well. However, the screen cannot precisely dating of the beginning and end of the cartel by the screen due to its short length. Although the pasta prices were characterised by a sudden and significant price increase by about 30%, the screen does not wrongly report the existence of a cartel in the French market. This results follows from the fact that – unlike in the Italian and Spanish industries – the price increase could reasonably well be explained by changes in the input costs.

The results show that the development of cartel screening is a promising field of research that should receive more work. Recent developments in the behavioural cartel screen literature provide the opportunity to further strengthen antitrust enforcement. Given the new technology at hand, competition authorities could increase the proactive search for cartels to increase the risk of detection for cartels. This renders collusion less viable and attractive to cartels, and could complement other efforts to increase deterrence that are not based on the problematic unilateral increases of cartel fines.

References

Abrantes-Metz, R. M. (2014). Recent Successes of Screens for Conspiracies and Manipulations: Why Are There Still Skeptics? *Antitrust Chronicle*, **10** (2).

— and Bajari, P. (2010). A Symposium on Cartel Sanctions: Screens for Conspiracies and Their Multiple Applications. *Competition Policy International*, **6** (2), 129–253.

—, Froeb, L. M., Geweke, J. and Taylor, C. T. (2006). A variance screen for collusion. *International Journal of Industrial Organization*, **24** (3), 467–486.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, **61** (4), 821–856.

— and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, **62** (6), 1383–1414.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion.* Princeton: Princeton University Press.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, **66** (1), 47–78.

— and — (2003). Critical values for multiple structural change tests. *The Econometrics Journal*, **6** (1), 72–78.

Baker, J. B. and Rubinfeld, D. L. (1999). Empirical methods in antitrust litigation: review and critique. *American Law and Economics Review*, **1** (1), 386–435.

Blair, R. D. and Sokal, D. D. (eds.) (2013). *Oxford Handbook on International Antitrust Economics.* Oxford: Oxford University Press.

Blanckenburg, K., Geist, A. and Kholodilin, K. A. (2012). The Influence of Collusion on Price Changes: New Evidence from Major Cartel Cases. *German Economic Review*, **13** (3), 245–256.

Bolotova, Y., Connor, J. M. and Miller, D. J. (2008). The impact of collusion on price behavior: Empirical results from two recent cases. *International Journal of Industrial Organization*, **26** (6), 1290–1307.

Brander, J. A. and Ross, T. W. (2006). Estimating damages from price-fixing. *Canadian Class Action Review*, **3** (1), 335–369.

Brown, R. L., Durbin, J. and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society*, **37** (2), 149–192.

Chu, C.-S. J., Hornik, K. and KAUN, C.-M. (1995a). MOSUM tests for parameter constancy. *Biometrika*, **82** (3), 603–617.

—, — and Kuan, C.-M. (1995b). The moving-estimates test for parameter stability. *Econometric Theory*, **11** (4), 699–720.

—, Stinchcombe, M. and White, H. (1996). Monitoring Structural Change. *Econometrica*, **64** (5), 1045–1065.

Cleveland, R. B., Cleveland, W. S., McRae, J. E. and Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, **6** (1), 3–73.

Comisión Nacional de la Competencia (2009). Resolución Expte., S/0053/08, FIAB Y ASOCIADOS Y CEOPAN.

Connor, J. M. (2005). Collusion and price dispersion. *Applied Economics Letters*, **12** (6), 335–338.

Esposito, F. M. and Ferrero, M. (2006). Variance screens for detecting collusion: an application to two cartel cases in Italy. *Mimeo.*

Harrington, J. E. (2007). Behavioural Screening and the Detection of Cartels. In C.-D. Ehlermann and I. Atanasiu (eds.), *European Competition Law Review 2006: Enforcement of Prohibition of Cartels*, Oxford/Portland, Oregon: Hart Publishing.

Heijnen, P., Haan, M. A. and Soetevent, A. R. (2015). Screening for collusion: a spatial statistics approach. *Journal of Economic Geography*, **15** (2), 417–448.

Hüschelrath, K. and Veith, T. (2013). Cartel Detection in Procurement Markets. *Managerial and Decision Economics*, **35** (6), 404–422.

International Pasta Organisation (2012). Annual Report 2012. *Available online at http://www.internationalpasta.org/resources/report/IPOreport2012.pdf, retrieved 19/07/2015.*

Italian Competition Authority (2009). Desicion of the Autorità Garante della Concorrenza e del Mercato regarding UNIPI – Unione Industriali Pastai Italiani e Union Alimentari – Unione Nazionale della Piccola e Media Industria Alimentare. *Available online at http://www.governo.it/backoffice/allegati/42113-5213.pdf, retrieved 17/05/2014.*

Kim, J.-H. (2011). Comparison of Structural Change Tests in Linear Regression Models. *Korean Journal of Applied Statistics*, **24** (6), 1197–1211.

Krämer, W., Ploberger, W. and Alt, R. (1988). Testing for structural change in dynamic models. *Econometrica*, **56** (6), 1355–1369.

Kuan, C.-M. and Hornik, K. (1995). The generalized fluctuation test: A unifying view. *Econometric Reviews*, **14** (2), 135–161.

Leisch, F., Hornik, K. and Kuan, C.-M. (2000). Monitoring structural changes with the generalized fluctuation test. *Econometric Theory*, **16** (6), 835–854.

Liu, J., Wu, S. and Zidek, J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, **7** (2), 497–525.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Berlin: Springer Science & Business Media.

Nieberding, J. F. (2006). Estimating overcharges in antitrust cases using a reduced-form approach: Methods and issues. *Journal of Applied Economics*, **9** (2), 361–380.

Notaro, G. (2014). Methods for quantifying cartel damages: The pasta cartel in Italy. *Journal of Competition Law and Economics*, **10** (1), 87–106.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, **84** (405), 223–230.

Ordóñez-de Haro, J. M. and Torres, J. L. (2014). Price hysteresis after antitrust enforcement: Evidence from spanish food markets. *Journal of Competition Law and Economics*, **10** (1), 217–256.

Ploberger, W. and Krämer, W. (1992). The CUSUM test with OLS residuals. *Econometrica*, **60** (2), 271–285.

—, Krämer, W. and Kontrus, K. (1989). A new test for structural stability in the linear regression model. *Journal of Econometrics*, **40** (2), 307–318.

Porter, R. H. (1983). A Study of Cartel Stability: The Joint Executive Committee, 1880-1886. *The Bell Journal of Economics*, **14** (2), 301–314.

— (2005). Detecting Collusion. *Review of Industrial Organization*, **26** (2), 147–167.

— and Zona, J. D. (1993). Detection of Bid Rigging in Procurement Auctions. *Journal of Political Economy*, **101** (3), 518–538.

Ramsey, J. B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 350–371.

Sen, P. K. (1980). Asymptotic theory of some tests for a possible change in the regression slope occurring at an unknown time point. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **52** (2), 203–218.

Wooldridge, J. (2012). *Introductory econometrics: A modern approach.* Mason, Ohio: South-Western, Cengage Learning.

Zeileis, A. (2004). Alternative boundaries for CUSUM tests. *Statistical Papers*, **45** (1), 123–131.

— (2005). A Unified Approach to Structural Change Tests Based on ML Scores, F Statistics, and OLS Residuals. *Econometric Reviews*, **24** (4), 445–466.

—, Kleiber, C., Krämer, W. and Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, **44** (1), 109–123.

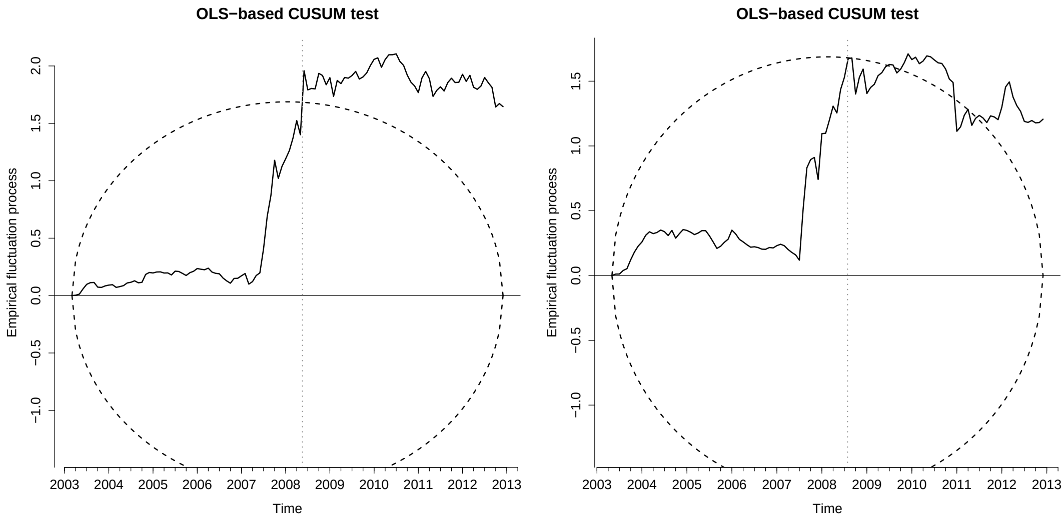
—, Leisch, F., Hornik, K. and Kleiber, C. (2002). strucchange. An R package for testing for structural change in linear regression models. *Journal of Statistical Software*, **7** (2), 1–38.

—, —, Kleiber, C. and Hornik, K. (2005). Monitoring structural change in dynamic econometric models. *Journal of Applied Econometrics*, **20** (1), 99–121.

Zivot, E. and Andrews, D. W. K. (1992). Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics*, **10** (3), 251–270.

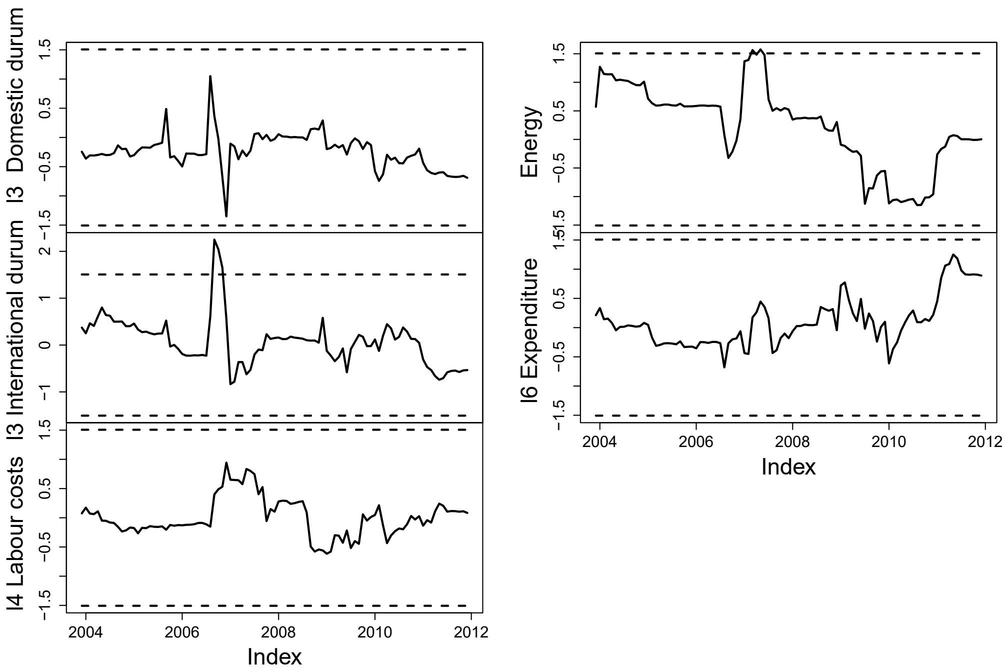
Appendix

Figure 5: OLS-CUSUM EFPs of Italy and Spain



(a) OLS-CUSUM IT **(b)** OLS-CUSUM ES

Figure 6: ME test results for coefficient stability in Spain



The gerneralized fluctuation test framework

This Appendix is not intended to be part of the article but provides reference with respect to the generalized fluctuation test methodology for those unfamiliar with it.

In order to simplify notation in the discussion of the structural break methodology used for the cartel screen, express the reduced form price equation in Equation 1 as

$$\Delta P_i = \Delta x_i^\top \beta_i + \Delta\epsilon_i \quad \text{for } i = 1, \ldots, T, \tag{3}$$

where i denotes time, ΔP_i is a $k \times 1$ vector of price changes in the industry, $x_i = (1, \Delta x_{i2}, \ldots, \Delta x_{ik})^\top$ is a $k \times 1$ vector of exogenous regressors, β_i is a $k \times 1$ vector of coefficients, and $\Delta\epsilon_i$ is a $k \times 1$ vector of model errors. In order to simplify notation I will suppress Δ and assume that P_i, x_i, and ϵ_i denote the first difference of the variables contained in these vectors. Technically speaking, all tests in the generalized fluctuation test framework are based on the two assumptions:

$$A1 : \{\epsilon_i\} \text{ is a homoskedastic martingale difference sequence with respect to } \mathcal{F}^i, \tag{4}$$
$$\text{which is the } \sigma\text{-field that is generated by } \{(x_{t+1}, \epsilon_t), t \leq i\} \text{ such that } E\left[\epsilon_i^2 | \mathcal{F}^i\right] = \sigma^2$$

$$A2 : \{x_i\} \text{ is such that } \lim_{T\to\infty} \frac{1}{T}\sum_{i=1}^{T} E|x_i|^{2+\delta} < \infty \text{ for a } \delta > 0, \tag{5}$$
$$\text{and } Q_{[Tt]} = \frac{1}{[Tt]}\sum_{i=1}^{[Tt]} x_i x_i^\top \xrightarrow{P} Q,$$

where Q is a non-stochastic, positive definite matrix, $c < t \leq 1$ with $c > 0$, and $[Tt]$ is the integer part of Tt (Kuan and Hornik, 1995, p.145). Assumption A1 implies that residuals follow a homoskedastic martingale difference sequence such that $E[\epsilon_i\epsilon_{i-k}] = 0 \; \forall \; k > 0$, i.e., the residuals are uncorrelated (Lütkepohl, 2005, p.689). It follows that heteroskedasticity and autocorrelation must either not be present, or that heteroskedasticity and autocorrelation consistent covariance matrices need to be used. Assumption A2 implies that dynamic and lagged regressors are allowed as long as they are stationary (Zeileis *et al.*, 2005, p.102; Zeileis *et al.*, 2003, p.111). If these two assumptions are fulfilled, a functional central limit theorem (FCLT) is satisfied for

$$\left(\frac{1}{\hat{\sigma}_T\sqrt{T}} Q_t^{-1/2} \sum_{i=1}^{Tt} x_i\epsilon_i, 0 \leq t \leq 1\right) \Rightarrow \mathbf{W}, \tag{6}$$

where $\mathbf{W}$ denotes a standard Wiener process with n dimensions, and $\hat{\sigma}$ and $Q_{(t)} = T^{-1}\sum_{i=1}^{T} x_i x_i^\top$ are consistent estimators of σ and Q (Kuan and Hornik, 1995, p.147). The different fluctuation tests construct empirical processes based on two components: one element that is approximately a straight line and another element that satisfies a FCLT as outlined above. After elimination of the straight line component, the resulting empirical fluctuation process (EFP) that results converges to and therefore follows the FCLT if parameters are constant over time. If, however, there are structural breaks, the empirical process will fluctuate such that the FCLT does not fully hold. The constancy of parameters is tested with a functional measuring the fluctuations in the EFP: if fluctuations are too large, the null hypothesis of no structural breaks is rejected.

The OLS-CUSUM test

The first test used is the OLS-CUSUM test by Ploberger and Krämer (1992), which is based on the Recursive-CUSUM test by Brown *et al.* (1975), but relies on OLS residuals instead of recursive residuals. While not uniformly superior to the Recursive-CUSUM test, the OLS-CUSUM test performs better in most situations, in particular if a structural break appears late in the sample (Zeileis, 2004; Ploberger and Krämer, 1992). The alternative boundaries suggested by Zeileis (2004) should be used in the application to increase the power of the test. While performing worse than the OLS-MOSUM test when there are several structural breaks, it performs better than the OLS-MOSUM test when there is a single break, features good finite sample properties (in particular in dynamic models), and its performance is not as dependent on the alternative hypothesis being true as the structural break tests based on F-tests by Andrews (1993) or Andrews and Ploberger (1994). Further, it is well suited to detect relatively short-lasting structural instability (Chu *et al.*, 1995a, pp.610-611).

Starting from an initial data window at the beginning of the sample, the historical OLS-CUSUM test sequentially tests the data for structural breaks based on the cumulative sum of residuals. In period t, the model is calibrated for periods 1 until t-1 and an expected value of the dependent variable is predicted for period t. If the regressions parameters are constant across time, the residuals should fluctuate around zero. If, however, there is structural instability, the residuals will systematically increase. The OLS-CUSUM test statistic is based on the accumulation of these residuals, such that the structural instability results in the test statistic to systemically drift away from zero. If this drift crosses a boundary determining whether fluctuations are (in)significant, the test finds evidence for structural breaks. Formally, the underlying empirical process of the test can be characterised as

$$W_T^0(t) = \frac{1}{\hat{\sigma}\sqrt{T}} \sum_{i=1}^{[Tt]} \hat{\epsilon}_t \quad (0 \leq t \leq 1), \tag{7}$$

where $[Tt]$ is the integer part of Tt. Therefore, the cumulative sum of OLS residuals $\hat{\epsilon}_t$ based on $i = 1, \ldots, Tt$ (where t increases by 1 in each sequential step) is scaled by the estimated error standard deviation $\hat{\sigma}$ (Ploberger and Krämer, 1992, p.274). The adaption of the OLS-MOSUM test to monitoring as introduced by Leisch *et al.* (2000) and extended to dynamic models by Zeileis (2005) works in a similar way. For monitoring, the regression coefficients are estimated once and the resulting historical model is used to predict residuals for the monitoring period.

The OLS-MOSUM test

Rather than relying on a growing window of observations like the OLS-CUSUM test, the OLS-MOSUM test is based on a window of fixed width Th that is sequentially “moved” through the whole sample to calculate the moving sums of OLS residuals. It reports structural instability when the sum of residuals in the data window exceeds a suitable boundary. As it is only based on a window of the sample that changes across time, it detects multiple structural breaks faster and more clearly than tests based on a growing window (Chu *et al.*, 1995a, p.603). For example, Chu *et al.* (1995a, pp.610-611) show that the OLS-MOSUM tests has a higher power than F-based tests or CUSUM tests when there are two structural breaks. The OLS-MOSUM test is chosen here as it often dominates the Recursive-MOSUM test. Further, it outperforms F-tests in many circumstances, in particular if they rely on wrong alternative assumptions (Chu *et al.*, 1995a, p.612). The corresponding empirical process can be expressed as

$$M_T^0(t|h) = \frac{1}{\hat{\sigma}\sqrt{T}} \left(\sum_{i=[N_T t]+1}^{[N_T t]+[Th]} \hat{\epsilon}_i \right) \quad (0 \leq t \leq 1-h) \tag{8}$$

where $N_T = (T - [Th])/(1-h)$, h is a parameter denoting the bandwidth (in %) of a data window containing a sub-sample of all available time periods, such that $[Th]$ represents the integer number of observations within the window as introduced by Chu *et al.* (1995a, p.607) but loosely following the notation in Zeileis *et al.* (2005, p.104). The OLS-MOSUM is extended to monitoring by obtaining the moving sum of residuals in the moving data window in the monitoring periods based on the coefficients that have been estimated once based on the historical periods (Zeileis, 2005; Leisch *et al.*, 2000).

The Moving Estimates test

In contrast to the OLS-CUSUM and OLS-MOSUM tests, which are based on the fluctuation of the residuals, the Moving Estimates (ME) test of Chu *et al.* (1995b) extends the analysis to the fluctuations within all regression parameters. As such, an EFP can be observed for each coefficient. This makes the ME test an attractive choice, because it allows to gain inference about the source of a structural break, i.e., which coefficients in the model are subject to structural instability. Compared to the residual-based fluctuation tests above, the ME test has another advantage. The residual-based tests are not sensitive to orthogonal shifts of the mean regressor (Chu *et al.*, 1995a; Ploberger and Krämer, 1992; Krämer *et al.*, 1988), whereas the estimates-based tests do not share this shortcoming (Zeileis *et al.*, 2005, p.111). Another estimates-based test is the Recursive Estimates (RE) test of Sen (1980) and (Ploberger *et al.*, 1989). Here, the ME test is preferred, as unlike the RE test it provides non-parametric estimates of the corresponding mean functions (Kuan and Hornik, 1995, p.136). Further, it usually has higher power than the RE test when there are multiple structural breaks (Chu *et al.*, 1995b, pp.713-714). In addition, it is better suited for monitoring than the RE test, which performs well for early structural breaks in the monitoring period but has significant detection delays if the break occurs later on (Zeileis *et al.*, 2005, p.109). The empirical process of the ME test proposed by (Chu *et al.*, 1995b, p.703) but again following the notation in Zeileis *et al.* (2005, p.104) can be written as

$$Z_T\,(t|h) = \frac{[Th]}{\hat{\sigma}\sqrt{T}} \sqrt{Q_{(T)}} \left(\hat{\beta}^{([Tt]-[Th]+1,[Th])} - \hat{\beta}^{(T)} \right) \quad (0 \leq t \leq 1-h), \tag{9}$$

where again $Q_{(T)} = T^{-1} \sum_{i=1}^{T} x_i x_i^\top$. $\hat{\beta}^{([Tt]-[Th]+1,[Th])}$ denotes the coefficient estimates from the moving window, whereas $\hat{\beta}^{(T)}$ is the whole-sample estimate. Thus, the ME test measures the fluctuations of coefficients by comparing the moving-window estimates to the whole-sample average to pick up structural breaks. As such, it becomes apparent why the ME test has approximately constant detection probabilities across time and can detect several structural changes. The ME test is extended to monitoring by obtaining the estimate of $\hat{\beta}^{(T)}$ from the historical periods and the estimates of the moving window from the monitoring periods (Zeileis, 2005; Leisch *et al.*, 2000).

Table 4: Data Sources and Definition of Variables

Variable/Country	**Source**	**Description**
Pasta price index		
IT	ISTAT	Producer price index for industrial products: manufacture of macaroni, noodles, couscous, base year 2005.
ES	INE	Producer price index for industrial products: manufacture of macaroni, noodles, couscous and similar farinaceous products, base year 2010.
FR	INSEE	Producer price index for industrial products: pasta products, base year 2010.
DE	DESTATIS	Producer price index for industrial products: manufacture of macaroni, noodles, and similar farinaceous products, base year 2010.
Domestic durum		
IT, ES, FR	EC	Durum wheat price in the domestic market, in €/tonne. This variable is country-specific.
International durum		
IT, ES, FR	INSEE	International price for imported durum wheat, in US cents / bushel of 60 pounds.
Labour costs		
IT	ISTAT	Labour costs per full time equivalent unit in the manufacturing industry, base year 2010.
ES	INE	Total wage cost per effective hour of work in the manufacturing industry, base year 2012.
FR	INSEE	Labour cost index including wages and payroll taxes in the manufacturing industry, base year 2012.
Energy		
IT	ISTAT	Producer price index for electricity, gas, steam and air conditioning supply, base year 2010.
ES	INE	Industrial price index for electric power generation, transmission and distribution, base year 2010.
FR	INSEE	Industrial market price index for energy, base year 2005.
Borrowing costs		
IT, FR	ECB	Monthly average interest rate of borrowing capital for non-financial cooperations with new businesses coverage. This variable is country-specific.
ES	INE	Average interest rates for loans over 1 million € to non-financial corporations.
Household expenditure		
IT	ISTAT	Quarterly national expenditure of households and non-profit institutions serving households, base year 2010.
ES	INE	Quarterly final household consumption expenditure, base year 2010.
FR	INSEE	Monthly household consumption expenditure on goods, base year 2005.

Notes: IT, ES, FR, and DE denote the variables for the Italian, Spanish, French, and German markets. DESTATIS denotes the German Statistical Office, EC the Europan Commission, ECB the European Central Bank, INE the Spanish Stastical Office, INSEE the French Statistical Office, and ISTAT the Italian Statistical Office. Variables are selected according to being the best available proxy variable for the price shifters between 2003 and 2012.