THE USE OF REDUCED-FORM EQUATIONS TO ESTIMATE CARTEL PRICE EFFECTS: LESSONS FROM THE KINGSTON RETAIL GASOLINE CARTEL

Can Erutku[1]
Professor, Economics Department, York University

On March 20, 2012, Canada's Competition Bureau announced that Pioneer Energy LP, Canadian Tire Corporation, and Mr. Gas had pleaded guilty to fixing the price of retail gasoline in Kingston and Brockville (Ontario) from May to November 2007. We exploit this real-world price-fixing conspiracy to illustrate issues that affect the reliability of using a reduced-form model to estimate the price effect caused by a cartel. Those issues include: 1) the size of the sample, 2) the sensitivity to individual data points, 3) the measure of costs, 4) the problem of endogeneity, 5) the omission of explanatory variables, 6) the duration of the cartel, and 7) the matter of nonstationarity. Understanding the potential limitations of this popular method for estimating damages caused by a cartel is critical for competition lawyers and economists alike.

Le 20 mars 2012, le Bureau de la concurrence du Canada a annoncé que Pioneer Energy LP, Société Canadian Tire et Mr. Gas avaient plaidé coupable à l'accusation de fixation du prix de l'essence au détail à Kingston et à Brockville (Ontario), de mai à novembre 2007. Nous nous servons de ce complot réel de fixation des prix pour illustrer les enjeux qui touchent la fiabilité du recours à un modèle de forme réduite pour estimer l'effet d'un cartel sur les prix. Mentionnons parmi ces enjeux : 1) la taille de l'échantillon, 2) la sensibilité à chaque point de données, 3) la mesure des coûts, 4) le problème de l'endogénéité, 5) l'omission de variables explicatives, 6) la durée du cartel et 7) la question de la non-stationnarité. Comprendre les limites potentielles de cette méthode populaire d'estimation des dommages causés par un cartel est essentiel, tant pour les avocats exerçant en droit de la concurrence que pour les économistes.

1. Introduction

On March 20, 2012, Canada's Competition Bureau announced that Pioneer Energy LP, Canadian Tire Corporation, and Mr. Gas had pleaded guilty to fixing the price of retail gasoline in Kingston and Brockville (Ontario) from May to November 2007.[1] Those gasoline retailers communicated by telephone to coordinate their responses to price changes initiated by major-branded competitors. More precisely, they agreed among themselves to match major-branded competitors' price increases or decreases. During its investigation, the Competition Bureau conducted searches, seized documents (hardcopies and electronic records), intercepted phone communications, and interviewed witnesses.

We exploit this real-world price-fixing conspiracy to illustrate some common issues affecting the reliability of using a reduced-form model to estimate the price effect caused by a cartel.[2] According to Baker and Rubinfeld, the estimation of reduced-form price equations is "the most common statistical method employed in antitrust litigation..."[3] Generally speaking, a reduced-form equation explains the variation in a variable such as price by variables related to cost, demand, market structure, and dummy variables that allow the intercept to change among groups of observations. It is assumed that the equation is derived from the interaction between supply and demand and that the variables on the right-hand side of the equation are exogenous (*i.e.*, determined independently and unaffected by the dependent variable).

We also examine whether the results are robust to some issues that could arise in a courtroom, such as: 1) the size of the sample, 2) the sensitivity to individual data points, 3) the measure of costs, 4) the problem of endogeneity, 5) the omission of explanatory variables, 6) the duration of the cartel, and 7) the matter of nonstationarity. Robust results among the model's different specifications would inspire confidence about the resulting estimate of the cartel's impact. In light of the 2009 amendments, which removed the requirement for the prosecution to prove the cartel lessened competition unduly, estimation of the price impact of a cartel is no longer required for conviction. It is quite commonly used in Canadian courtrooms, however, to calculate the price impact and therefore the damages caused by the cartel. Moreover, proof of a price impact could conceivably form part of the

evidence that there was, in fact, an agreement. Understanding the potential limitations of this common econometric technique is critical for Canadian competition lawyers and economists alike.

The rest of the paper is structured as follows: Section 2 describes the structure of Kingston's retail gasoline market. Section 3 describes the data collected to implement the base empirical model, which is presented in Section 4. Section 5 discusses the regression results. Section 6 expands the analysis by covering a number of issues that can affect the results obtained with the base empirical model. Section 7 provides concluding remarks.

2. Kingston's Retail Gasoline Market[4]

The demand for gasoline in Kingston was approximately 143, 155, and 165 million litres in 2006, 2007, and 2008, respectively. During this period, the number of operating gasoline outlets varied between 33 and 38. Table 1 depicts the distribution of outlets and market shares (in terms of total sales) between Major Refiners (Esso, Shell, and Petro-Canada), Regional Refiners (Sunoco and Ultramar), and the firms that pleaded guilty to price-fixing (Pioneer, Canadian Tire, and Mr. Gas). Table 1 also illustrates the Herfindahl–Hirschman Index (HHI), which is a measure of industry concentration.[5]

The Major Refiners' market share was around 40%. Although they never operated more than 25% of outlets, the conspirators had about a third of the market (Pioneer and Canadian Tire accounted for the bulk of the conspirators' sales). Remaining sales were made through Regional Refiners or Independents.[6] Overall, the market was not highly concentrated, with an HHI below 1500.

It must be remembered that simply counting the number of retailers might be a misleading reflection of the true number of competitors if there are cross-ownerships or structural links. It may very well be the case that a large number of outlets were jointly operated or managed by one or a few individuals.

In the next section, the data used to test for the cartel's price effect with the reduced-form approach is described.

Table 1
Number of Outlets, Market Shares of Gasoline Retailers, and HHI in Kingston

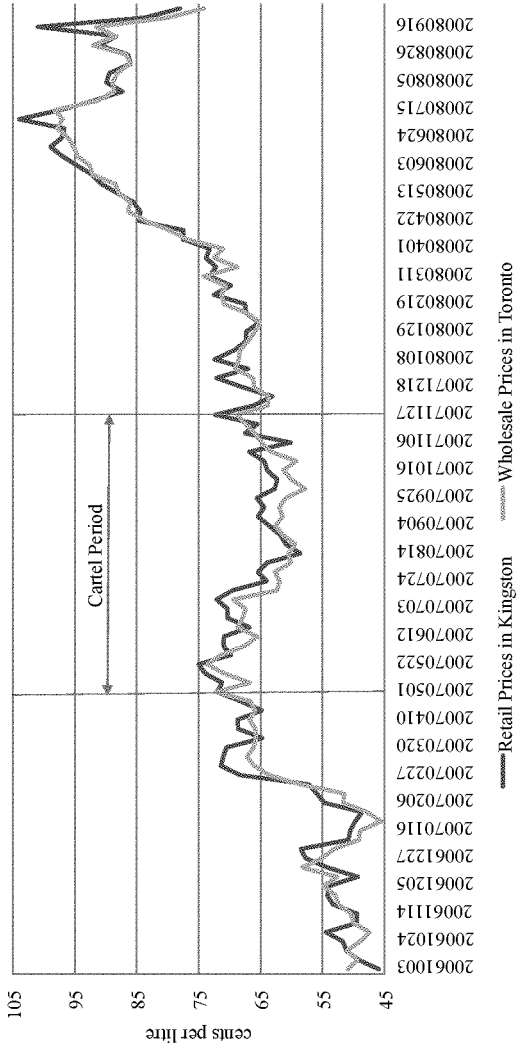
Quarter, Year	Major Refiners		Regional Refiners		Pioneer, Canadian Tire, Mr. Gas		HHI
	No. of Outlets	Market Shares (%)	No. of Outlets	Market Shares (%)	No. of Outlets	Market Shares (%)	
1st, 2006	17	41.7	9	18.1	7	34.3	1345
2nd, 2006	17	43.2	9	18.8	7	33.3	1362
3rd, 2006	16	41.5	11	21.9	7	32.0	1335
4th, 2006	16	38.9	11	24.2	7	33.1	1332
1st, 2007	15	37.3	11	25.8	7	32.9	1326
2nd, 2007	15	35.8	11	26.6	7	33.6	1337
3rd, 2007	15	37.4	11	26.7	7	32.8	1385
4th, 2007	16	37.7	10	25.2	7	34.5	1427
1st, 2008	16	34.9	10	27.0	7	35.6	1455
2nd, 2008	15	36.0	10	27.5	7	33.8	1445
3rd, 2008	16	39.6	10	26.2	7	32.0	1444
4th, 2008	16	37.6	9	26.2	7	34.0	1436

3. Data

To estimate the price effect of this not-all-inclusive cartel (it did not involve all market participants), weekly retail gasoline prices without tax (retail prices) in Kingston were obtained from MJ Ervin & Associates Inc. (MJ Ervin).[7,8] Because it is assumed that a gasoline retailer's marginal cost is the cost of purchasing wholesale gasoline, weekly wholesale gasoline prices (wholesale prices) in Toronto were also obtained from MJ Ervin.[9] All prices are expressed in cents per litre. The data cover a two-year period spanning from October 3, 2006, to September 30, 2008 (*i.e.*, from the beginning of the fourth quarter in 2006 to the end of the third quarter in 2008). This provides us with several observations either side of the cartel period.

Figure 1 depicts retail prices in Kingston and wholesale prices in Toronto. The two price series are closely aligned to one another. Table 2 presents means and standard errors for retail prices, wholesale prices, and margins (obtained by subtracting wholesale prices from retail prices) *i*) before the start of the cartel, *ii*) when the cartel was active, and *iii*) after the end of the cartel.[10]

Figure 1
Retail Prices in Kingston and Wholesale Prices in Toronto
October 3, 2006 to September 30, 2008



The increase in Kingston's retail prices throughout those periods can apparently largely be explained by the increase in wholesale prices. That being said, average retail gasoline margins are the highest when the cartel was active suggesting that more market power was being exercised during this time period.

Although it is often assumed that prices should be more stable in the presence of a cartel,[11] the relative stability (*i.e.*, the relatively low standard error) of retail prices during the cartel period is a reflection of the stability of wholesale prices between May and November 2007.

Table 2
Mean and Standard Error for Retail Prices and Margins in Kingston and Wholesale Prices in Toronto

Time Period	Retail Prices in Kingston (cents per litre)	Wholesale Prices in Toronto (cents per litre)	Retail Margins in Kingston (cents per litre)
Before the Cartel (Oct. 3, 2006, to Apr. 24, 2007)			
Mean	57.19	55.80	1.40
Standard Error	8.12	7.33	2.92
During the Cartel (May 1, 2007, to Nov. 27, 2007)			
Mean	66.84	64.67	2.17
Standard Error	4.32	4.31	2.59
After the Cartel (Dec. 4, 2007, to Sep. 30, 2008)			
Mean	81.51	80.41	1.10
Standard Error	11.78	11.40	2.48

Now that we have illustrated the common evolution of retail and wholesale prices, along with retail margins, through descriptive statistics, we turn our attention to the empirical model that will be used to

measure the effect of the cartel on retail prices and to discuss potential limitations of regression analysis in competition policy cases.

4. Empirical Model

The following reduced-form equation is estimated to assess the cartel's impact on retail prices

$$P_{K,t} = \alpha + \beta_0 W_{T,t} + \beta_1 W_{T,t-1} + \gamma \mathbf{Q}_t + \lambda Trend_t + \delta Cartel_t + \varepsilon_t. \quad (1)$$

$P_{K,t}$ measures retail prices in Kingston during period t; $W_{T,t}$ and $W_{T,t-1}$ represent Toronto's wholesale prices (with the subscript T denoting Toronto) in period t (*i.e.*, current wholesale prices) and in period t-1 (*i.e.*, wholesale prices lagged by one week), respectively, to control for changes in retail prices that are explained by changes in costs;[12] $\mathbf{Q}_t$ is a vector of quarterly fixed effects intended to capture unobserved seasonal variations that may affect the demand for retail gasoline; $Trend_t$ takes into account the long-term movement in prices; $Cartel_t$ is a dummy variable taking a value of one when the cartel was active and a value of zero otherwise; and ε_t stands for the error term. The variable $Cartel_t$ isolates the conspiracy's impact controlling for demand and supply factors unrelated to the price-fixing conduct.

In accordance to the Competition Bureau's backgrounder,[13] it is assumed that the cartel was active from May 2007 to November 2007. Thus, the variable $Cartel_t$ takes a value of one during this time period and zero otherwise. Equation (1) is estimated using Ordinary Least Square (OLS) with standard errors being adjusted for heteroskedasticity using White's (1980) method.[14]

The Greek letters in Equation (1) represent the parameters to be estimated (one parameter for each variable) with α being referred to as the *Constant* in the Tables below.

5. Empirical Results

Table 3 presents regression results from Equation (1). For brevity, coefficients and standard errors for the quarterly dummies and the trend are not reported. Those variables, as well as the constant, are not statistically significant at usual levels of confidence with the exception of the dummy variable for the fourth quarter.

Table 3[15]
OLS Regression Results from Equation (1)
October 3, 2006, to September 30, 2008

Variables	Coefficients (Standard Errors)	t-statistic
Constant	1.31 (2.05)	0.64
$W_{T,t}$	0.71 [a] (0.08)	8.88
$W_{T,t-1}$	0.31 [a] (0.08)	3.88
$Cartel_t$	1.20 [b] (0.67)	1.79
R^2	0.97	
Durbin-Watson Statistics	1.90	
No. of Observation	105	

[a] Statistically significant at the 1% level.
[b] Statistically significant at the 10% level.

Wholesale prices, current and lagged by one week, are statistically significant at the 1% level (*i.e.*, we expect the result to be correct 99 times out of 100, which is a very high level of confidence). Suppose, for instance, that both current and lagged wholesale prices are equal to 65 cents per litre. Then, estimated retail prices in Kingston before and after the cartel would be 66.3 cents per litre (0.71*65 + 0.31*65). When the cartel was active, however, the estimated retail prices would be 67.5 cents per litre (66.3 + 1.20) since the statistically significant coefficient of the variable $Cartel_t$ is equal to 1.20. This suggests that the presence of the price-fixing cartel lead retail prices in Kingston to be higher than in its absence (after controlling for current and lagged wholesale prices and fixed effects).

The plaintiff could use this empirical model's results to require some compensation through a class-action suit, for example. Indeed, the plaintiff's hypothesis would be that the presence of the cartel resulted in higher prices ($\delta > 0$ in Equation (1)). But the defendant could argue that results are not robust and do not reflect the cartel's influence on retail prices. The defendant's hypothesis would be that the cartel had no effect on prices ($\delta = 0$ in Equation (1)). This claim is highlighted in the next section.

6. Empirical Issues

Like any piece of evidence, a regression analysis can affect the outcome of an antitrust case. As such, a testifying economic expert should disclose the assumptions underlying the empirical model as well as how the results are sensitive to the methodology and the data used to perform the analysis. This, in turn, should help the court in determining the appropriate weight that should be placed on the empirical evidence presented during the proceeding. Therefore, we illustrate in this section the results' sensitivity when modifying how Equation (1) is estimated. Seven issues are investigated starting with the size of the sample.

Sample size. Suppose that the sample's size is increased so that the data cover a period spanning from one year before the start of the cartel to one year after the end of the cartel. Hence, consider data that extend from May 2, 2006, to November 25, 2008 (30 observations are added to the initial data set). Table 4 presents regression results from Equation (1) with the extended data set.

Wholesale prices, current and lagged by one week, remain statistically significant at the 1% level of confidence. The coefficient of the variable $Cartel_t$ becomes, however, not statistically different from zero. Assume once again that both current and lagged wholesale prices are equal to 65 cents per litre. Then, estimated retail prices in Kingston would be 65 cents per litre (0.58*65 + 0.42*65) before, during, and after the cartel. Use of the larger data set suggests that the presence of the price-fixing cartel had no significant effect on retail prices in Kingston.

The larger sample size increases the precision in the estimated coefficients. For example, the standard error (a measure of the observation's variation from the mean) of the variable $Cartel_t$ decreases from 0.67 in Table 2 to 0.56 in Table 3. This illustrates the danger of producing Type I errors (i.e., finding a price effect when there is none) when using a short data set.

That being said, using a longer time series is not a panacea and we continue the analysis with the shorter time series. Two reasons explain

Table 4
OLS Regression Results from Equation (1)
May 2, 2006, to November 25, 2008

Variables	Coefficients (Standard Errors)	t-statistic
Constant	2.37 (1.61)	1.47
$W_{T,t}$	0.58 [a] (0.07)	8.29
$W_{T,t-1}$	0.42 [a] (0.07)	6.00
$Cartel_t$	0.68 (0.56)	1.21
R^2	0.96	
Durbin-Watson Statistic	1.93	
No. of Observation	135	

[a] Statistically significant at the 1% level.

this decision. First, we want to show how each issue, taken individually, can affect the results generated with the base model. While we have the opportunity to add observations to the initial database, this is not a luxury that applies to all antitrust cases. Second, it is very difficult to define what is the appropriate length of the time series. Older pre-cartel data may not be that relevant in yielding more accurate estimates of the cartel's effect if, for example, past prices were largely affected by some unobservable variables (perhaps a different market structure characterized by the presence of a vigorous competitor). Post-cartel data can also be problematic. For instance, there might be some residual collusion after the collapse of the cartel, whereby firms replaced explicit collusion with tacit collusion.[16]

Sensitivity to data points. When using the OLS estimation procedure, it is assumed that the distribution of the error terms (which capture the effects of omitted variables) is normal (*i.e.*, bell shaped). When the error terms do not follow a normal distribution, the t-statistic (which is the ratio of the estimated coefficient to its standard error) may overstate or understate the significance level of a coefficient (*i.e.*, the degree of confidence one should place in a coefficient).

Related to this non-normality issue is the sensitivity of the regression results to individual data points. For instance, some observations may take very high or very low values. It is important to consider how those extreme observations can influence results. Looking at Figure 1, there seem to be a few large variations of retail and wholesale prices in February 2007, July 2008, and September 2008. As such, we create a dummy variable that takes a value of one when retail price variations exceed the mean (in absolute value) of those fluctuations plus or minus three standard errors, and a value of zero otherwise. This variable, Big_t, is added to the right-hand side of Equation (1), which is estimated with OLS.[17] The results of this regression are presented in Table 5.

Table 5
OLS Regression Results from Equation (1)
Controlling for Large Retail Price Fluctuations

Variables	Coefficients (Standard Errors)	t-statistic
Constant	1.52 (2.06)	0.74
$W_{T,t}$	0.72 [a] (0.09)	8.00
$W_{T,t-1}$	0.30 [a] (0.09)	3.33
$Cartel_t$	1.37 [b] (0.66)	2.08
Big_t	4.56 [b] (2.30)	1.98
R^2	0.97	
Durbin-Watson Statistics	2.00	
No. of Observation	105	

[a] Statistically significant at the 1% level.
[b] Statistically significant at the 5% level.

The coefficient of the variable Big_t is statistically significant at the 5% level. The addition of the variable increases the significance threshold of the coefficient $Cartel_t$ from 10% in Table 2 to 5% in Table 5. This reinforces the confidence that the cartel had a positive impact on retail prices in Kingston when compared to the results presented in Table 3.

Measure of cost. There might be some uncertainty as to the exact measure of costs. It is well known that the cost of wholesale gasoline represents a major fraction of the price of retail gasoline. But there might be some debate about the source of supply. For instance, suppose that Maitland's wholesale prices are used to measure retailers' cost in Kingston rather than Toronto's wholesale prices (Maitland is located about 80 km away from Kingston compared to around 180 km for Toronto). With the same sample size as the one used in Section 4, Table 6 presents regression results from Equation (1) (the subscript *M* associated with wholesale prices stands for Maitland).

Table 6
OLS Regression Results from Equation (1)
Wholesale Prices in Maitland as a Measure of Cost

Variables	Coefficients (Standard Errors)	t-statistic
Constant	1.14 (2.25)	0.51
$W_{M,t}$	0.79 [a] (0.10)	7.90
$W_{M,t-1}$	0.27 [a] (0.09)	10.78
$Cartel_t$	-0.17 (0.69)	-0.25
R^2	0.96	
Durbin-Watson Statistics	1.77	
No. of Observation	105	

[a] Statistically significant at the 1% level.

The null hypothesis that the coefficient of the variable $Cartel_t$ is equal to zero cannot be rejected. Assuming that current and lagged wholesale prices are both equal to 65 cents per litre, retail prices in Kingston are estimated to be equal to 68.9 cents per litre (0.79*65 + 0.27*65) over the entire data range. This suggests that the presence of the price-fixing cartel had no effect on retail prices in Kingston.

Endogeneity. In Section 4, Equation (1) was estimated with the assumption that all variables on the right hand side were exogenous.

However, there is the potential issue of endogeneity of wholesale prices. The error term ε_t in Equation (1) captures factors that are unobservable to the econometrician but that can affect retail prices. Those factors might also be correlated with wholesale prices (either in Toronto or Maitland). Local shocks to the demand of retail gasoline might affect wholesale prices in Toronto or Maitland if they are somewhat distinct from national prices.[18]

Table 7
Instrumental Variable Results from Equation (1)
October 3, 2006, to September 30, 2008

Variables	Coefficients (Standard Errors)	t-statistic
Constant	2.24 (2.50)	0.90
$W_{T,t}$	0.82 [a] (0.13)	6.31
$W_{T,t-1}$	0.17 (0.13)	1.31
$Cartel_t$	1.06 (0.69)	1.54
Durbin-Watson Statistics	1.88	
No. of Observation	105	

[a] Statistically significant at the 1% level.

To handle this issue, the potentially endogenous variable (wholesale prices in Toronto) is regressed on all the other variables on the right hand side of Equation (1) and one instrument.[19] An instrument is a variable (we take wholesale prices in Vancouver) correlated with the potentially endogenous variable but not correlated with the error term ε_t in Equation (1). The fitted values of this first regression are then used in place of the potentially endogenous variable in Equation (1). The results of this instrumental variable procedure are presented in Table 7. The t-statistic associated with the variable $Cartel_t$ suggests that the cartel did not have a statistically significant impact on Kingston retail gasoline prices.

Omitted variables. In a similar vein, many factors can influence retail prices in Kingston during the time period covered by the analysis. The

reduced-form method assumes that the data for all the relevant factors are included in the estimated equation. In practice, however, some relevant variables might be omitted. The result is that estimated coefficients along with their standard errors might be inaccurate. This, in turn, may lead the econometrician to identify incorrectly a causality (a price effect coming from the cartel's presence) when none is there. Stated differently, relying on the reduced-form approach to test for the cartel's effects is to make the strong assumption that nothing else material happened to influence positively or negatively retail prices during the time period covered by the analysis.

Table 8
OLS Regression Results from Equation (1)
HHI_t Added as an Explanatory Variable

Variables	Coefficients (Standard Errors)	t-statistic
Constant	-5.15 (19.43)	-0.27
$W_{T,t}$	0.72 [a] (0.08)	9.00
$W_{T,t-1}$	0.31 [a] (0.08)	3.88
$Cartel_t$	1.15 (0.74)	1.55
HHI_t	0.005 (0.01)	0.50
R^2	0.97	
Durbin-Watson Statistics	1.90	
No. of Observation	105	

[a] Statistically significant at the 1% level.

For instance, consider the inclusion of the HHI on the right-hand side of Equation (1). Table 1 illustrates that the HHI increased during the period considered by the analysis and a higher level of concentration in a market is often associated with higher prices.[20] Table 8 shows that the variable $Cartel_t$ loses its explanatory power after the addition of the variable measuring the concentration level in the market even though its estimated coefficient is not statistically significant.

Table 9
OLS Regression Results from Equation (1)
Cartel Duration: January 2007 to November 2007

Variables	Coefficients (Standard Errors)	t-statistic
Constant	-0.66 (2.39)	0.28
$W_{T,t}$	0.71 [a] (0.08)	8.88
$W_{T,t-1}$	0.31 [a] (0.08)	3.88
$Cartel_t$	1.79 [b] (0.71)	2.52
R^2	0.97	
Durbin-Watson Statistics	1.99	
No. of Observation	105	

[a] Statistically significant at the 1% level.
[b] Statistically significant at the 5% level.

Cartel duration. It was assumed that the price-fixing cartel was active between May 2007 and November 2007. This assumption is based on the Competition Bureau's backgrounder following the guilty pleas.[21] However, the cartel may have been active before May 2007 and could have continued its activities after November 2007. As such, a sensitivity analysis was conducted by varying the possible duration of the cartel. For instance, we allowed the cartel to be active from January, February, March, and April 2007 until December 2007. For brevity, Table 9 only reports results when the variable $Cartel_t$ takes a value of one between January and November 2007 and zero otherwise.[22] The variable $Cartel_t$ remains statistically significant suggesting that the price-fixing cartel might have been active before May 2007.

Nonstationarity. Time-series data, such as the ones used here, are subject to the presence of nonstationarity. While a complete discussion of nonstationarity falls outside the scope of this paper, suffice it to say that there is the potential of spurious correlation: a relationship between two variables may be wrongly inferred simply because they each follow a similar trend over time. When this is the case, an

Table 10
OLS Regression Results from Equation (2)
Error Correction Model

Variables	Coefficients (Standard Errors)	t-statistic
$\Delta P_{K,t}$	0.02 (0.05)	0.40
$\Delta W_{T,t}$	0.71 [a] (0.08)	8.88
z_{t-1}	-0.99 [a] (0.08)	-12.38
$Cartel_t$	0.91 (0.57)	1.60
R^2	0.65	
Durbin-Watson Statistics	2.03	
No. of Observation	104	

[a] Statistically significant at the 1% level.

independent variable (*e.g.*, wholesale prices) may appear to be more significant than it actually is in explaining variations in the dependent variable (*e.g.*, retail prices). A standard technique to deal with nonstationarity is to take first differences (i.e., $\Delta x_t = x_t - x_{t-1}$) and use the newly created variables in the regression model. This can also be accompanied with an error correction term (ECT) to handle the long-term relationship between two price series that could spread into their short-run dynamics. We, therefore, implement the following regression model[23]

$$\Delta P_{K,t} = \Sigma_{l=1}^{l=L_P} \alpha_{t-l} \Delta P_{K,t-l} + \Sigma_{l=0}^{l=L_W} \beta_{t-l} \Delta W_{T,t-l} + \gamma \mathbf{Q}_t + \delta Cartel_t + \lambda \underbrace{\left(P_{K,t-1} - \varphi_0 - \varphi_1 W_{T,t-1}\right)}_{z_{t-1}} + \varepsilon_t \qquad (2)$$

where $\Delta P_{K,t} = P_{K,t} - P_{K,t-1}$ and $\Delta W_{T,t} = W_{T,t} - W_{T,t-1}$ are the first differences in Kingston's retail prices and Toronto's wholesale prices, respectively. Equation (2) includes current and lagged variations in wholesale prices and lagged fluctuations in retail prices to capture the transmission of shocks.[24] The term in parenthesis in Equation (2) is the ECT, which captures the long-run relationship between retail and wholesale prices.[25] The other variables are defined as before. Results obtained from estimating Equation (2) are presented in Table 10.[26] The variable $Cartel_t$ is non-significant at standard level of confidence suggesting that the conspiracy had no impact on retail prices.

The findings of this section, which are summarized in Table 11, make it difficult to reach a definitive conclusion as to the cartel's impact on retail prices in Kingston. The empirical results are not robust to the size of the sample, the specification of the equation, or the estimation procedure.

Table 11
Summary of Findings

Model	The cartel had a positive impact on retail prices ($\delta > 0$)	The cartel did not have a positive impact on retail prices ($\delta = 0$)
Base Model (Table 3)	¶	
Larger Sample (Table 4)		¶
Data Sensitivity (Table 5)	¶	
Measure of Cost (Table 6)		¶
Endogeneity (Table 7)		¶
Omitted Variable (Table 8)		¶
Duration of the Cartel (Table 9)	¶	
Error Correction Model (Table 10)		¶

That being said, the reader should not infer from the large number of tick marks in the right-hand side column of Table 11 that we have shown that the conspiracy had no effect. Our goal here is rather to point out that there are many economic and statistical issues that need to be taken into account before any estimation can be considered robust.

Measuring precisely the price effect of the cartel would imply knowing which statistical model best fits the structural and behavioral features of the Kingston retail gasoline market around the time the cartel was active. This is a difficult task. However, knowledge of the

industry can help guide the econometrician in determining the specification of the regression model. For instance, it could be possible to use documentary evidence to determine exactly when the cartel was active or to identify with certainty the source of supply. Hence, fact gathering about the industry and regression analysis should be seen as complementary tools in measuring the price effect of the cartel.

7. Conclusion

Effective March 2010, Section 45 (the conspiracy provision) of the *Competition Act* makes horizontal agreements to fix prices, allocate markets, or restrict output *per se* illegal.[27] As such, the need to show beyond a reasonable doubt that competition is unduly lessened because of an agreement to secure a conviction is no longer required.[28] This makes it more difficult for conspirators to escape conviction by providing statistical evidence (such as those presented in Section 6) refuting the plaintiff's claims of an undue lessening of competition.

The results presented in Section 6 and the amendment of Section 45 do not imply, however, that regression analysis should be dismissed. It is a powerful tool in measuring competitive effects (such as damages in cartels) as long as the empirical model is appropriate, the underlying assumptions are respected, and the data are reliable.[29] Since econometric analysis is now commonly used in antitrust cases and proceedings and is subject to cross-examination, it is important for antitrust practitioners to understand why an empirical model is adopted and how sensitive it is to minor or major changes. This would avoid both embarrassment and a loss of credibility in the courtroom.

When results are robust throughout sensitivity analysis, the courts should be confident that the empirical evidence presented during the proceedings deserves to be considered. Otherwise, as is the case here, less weight should be put on the empirical evidence. In any event, regression analysis can be a helpful tool when used correctly and rigorously in merger, abuse of dominance, and cartel cases. Hence, an adjudicative body such as the Competition Tribunal should ensure that its lay members are able to understand and evaluate econometric techniques and results.[30]

Bibliography

Rosa M Abrantes-Metz et al, "A Variance Screen for Collusion" (2006) 24 International Journal of Industrial Organization 467.

Eva Audy & Can Erutku, "Price Tests to Define Markets: An Application to Wholesale Gasoline in Canada" (2005) 5 Journal of Industry, Competition and Trade 137.

Jonathan B Baker & Daniel L Rubinfeld, "Empirical Method in Antitrust Litigation: Review and Critique" (1999) 1 Am L & Econ Rev 386.

Yuliya Bolotova, John M Connor & Douglas J Miller, "The Impact of Collusion on Price Behavior: Empirical Results from Two Recent Cases" (2008) 26 Industrial Journal of Industrial Organization 1290.

David A Dickey & Wayne A Fuller, "Likelihood Ratio Statistics for Autoregressive: Time Series with a Unit Root"(1981) 49 Econometrica 1057.

James Durbin & Goeffrey S Watson, "Testing for Serial Correlation in Least-Squares Regression, I"(1950) 37 Biometrika 409.

Robert F Engle & Clive W. J. Granger, "Co-Integration and Error Correction: Representation, Estimation, and Testing" (1987) 55 Econometrica 251.

MO Finkelstein & H Levenbach, "Regression Estimates of Damages in Price-Fixing Cases" (1983) 46 Law & Contemp Probs 145.

Joseph E Harrington Jr, "Post-Cartel Pricing During Litigation" (2004) 52 Journal of Industrial Economics 517.

Jerry A Hausman, "Specification Tests in Econometrics" (1978) 46 Econometrica 1251.

James F Nieberding, "Estimating Overcharges in Antitrust Cases Using a Reduced-Form Approach: Methods and Issues" (2006) 9 Journal of Applied Economics 361.

Daniel L. Rubinfeld "Econometrics in the Courtroom" (1985) 85 Colum L Rev1048.

Gideon E Schwartz, "Estimating the dimension of a model" (1978) 6 Annals of Statistics 461.

Anindya Sen, "Higher Prices at Canadian Gas Pumps: International Crude Oil

Prices or Local Market Concentration? An empirical Investigation"(2003) 25 Energy Economics 269.

Halbert White, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity"(1980) 48 Econometrica 817.

Jeffrey M Wooldridge, *Introductory Econometrics: A Modern Approach, 4th ed* (Mason, Ohio: South-Western CENGAGE Learning, 2009).

Endnotes

[1] See "Gasoline Companies Plead Guilty to Price-Fixing in Kingston and Brockville, Ontario" (20 March 2012), online: Competition Bureau <http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/03448.html>.
[2] Nieberding (2006) also discusses issues in using a reduced-form model when estimating overcharges in antitrust cases. However, the data he uses are artificial, *i.e.*, do not come from a real-world case, and he discusses only a few of the issues that commonly arise.
[3] Jonathan B Baker & Daniel L Rubinfeld, "Empirical Method in Antitrust Litigation: Review and Critique" (1999) 1 Am L & Econ Rev 386 at 391. Many examples can be found in Finkelstein and Levenbach (1983).
[4] The data used in this section were obtained from Kent Marketing Services Limited.
[5] The Herfindahl–Hirschman Index is calculated at the brand level, not at the outlet level, thus over-estimating the concentration level in the market. It is calculated by summing the squared market share of each individual brand.
[6] The Independents we are referring to here are Olco, FS, Gas, and Sunys. These firms operated five outlets at the beginning of 2006. At the end of 2008, only Olco was still in the market with two outlets.
[7] Gasoline refers to regular gasoline, which represents more than 90% of the sale of all gasoline types (regular, mid-grade, premium) made in Kingston between 2006 and 2008.
[8] MJ Ervin collects retail gasoline prices every Tuesday morning at 10:00 am local time. As such, retail gasoline prices are a snapshot of those prices on that particular time and day and do not represent a weekly average price. This means, in turn, that the data used here can allow a comparison of retail gasoline prices before, during, and after the cartel on Tuesdays at around 10:00 am.
[9] There are no gasoline wholesalers in Kingston and the Toronto data are used because of the close proximity between the two cities.
[10] Margins can be seen as an indicator of market power. The higher the margins, the more firms have been able to raise their price above marginal cost, *i.e.*, competitive levels.

[11] See Abrantes-Metz et al (2006) for a discussion and an application to a retail gasoline market in the United States.

[12] The inclusion of lags for wholesale prices captures the possibility that retail gasoline prices do not adjust instantaneously to changes in wholesale prices. The number of lags is chosen to minimize the Bayesien Information Criterion (Schwartz, 1978). In a nutshell, the Bayesien Information Criterion (BIC) establishes a criterion to select the appropriate model. It is possible to improve how a model fits the data by adding variables to the right hand side of an equation to be estimated, even though these variables might be irrelevant. To avoid this over-fitting problem, the BIC introduces a penalty for the number of parameters to be estimated. According to the BIC, only current wholesale prices and those lagged by one period should be included in Equation (1).

[13] *Supra* note 1.

[14] The coefficients obtained with the OLS estimation procedure minimize the sum of the squared deviations of the predicted values of the model (the right-hand side of the regression equation evaluated with the estimated coefficients, or estimators, as weights for each variable) from the actual observations. OLS is commonly used in antitrust and other fields because it is a simple and intuitive method to estimate the relationship between two or more variables. It also produces estimators that have desirable properties: unbiasedness and consistency. An estimator is unbiased when its expected value is equal to the true value (*i.e.*, OLS estimations give, on average, the true value of the estimated coefficient(s)). An estimator is consistent when it tends toward the true value of the parameter as the sample available for estimation gets larger. Of course, OLS estimations are not free of limitations (some of them like endogeneity and misspecification, *e.g.*, omitted variables, are discussed in Section 6). The reader is referred to Wooldridge (2009) for a thorough discussion of OLS.

[15] The t-statistic is used in hypothesis testing. Here, we want to test the null hypothesis that an estimated coefficient is equal to zero. As such, the t-statistic is computed as the ratio of the estimated coefficient over its standard error. A significance level (usually 1%, 5%, or 10%) is chosen as a rejection rule of the null hypothesis. Suppose we decide on a significance level of 5%. This implies that we are willing to mistakenly reject 5% of the time the null hypothesis when it is true (*i.e.*, 5% of the time we conclude that an estimated coefficient is statistically different than zero when it is in fact equal to zero). A more stringent (tolerant) level of significance is 1% (10%). As a rule of thumb, a t-statistic of 2.00 is used for a level of significance of 5%. The R^2 of a regression measures the fraction of the variation in the dependent variable that is explained by the dependent variables. When the R^2 is close to 1 (0), the OLS regression provides a good (poor) fit to the data. That being said, it is recognized that not too much weight should be put on the size of the R^2 in evaluating the goodness-of-fit of a regression equation. The Durbin-Watson statistic (Durbin and Watson, 1951) tests for serial

correlation of order 1 in the error terms, *i.e.*, detects whether the error term in one period are correlated directly with errors in the following period. When serial correlation is present, the standard errors of the estimated coefficients are smaller than their true values. This leads to the conclusion that the estimated coefficients are more precise than they actually are. In other words, there will be a tendency to reject the null hypothesis that an estimated coefficient is equal to zero when it should not be rejected. As a rule of thumb, a Durbin-Watson statistic close to 2 indicates that there is not serial correlation of order 1 in the error terms.

[16] See Harrington (2004).

[17] Big_t takes a value of one on February 27, 2007, September 16, 2008, and September 23, 2008 (those three observations fall outside the period during which the cartel is assumed to be active). If Big_t takes a value of one when retail price variations exceed the mean (in absolute value) of those fluctuations plus or minus two standard errors, then its coefficient becomes statistically equal to zero. Wooldridge (2009) at 301 provides a discussion on outliers and suggests that "...OLS results should probably be reported with and without outlying observations in cases where one or several data points substantially change the results."

[18] Audy and Erutku (2005) argue that relevant geographic markets for wholesale gasoline can be larger than cities but smaller than East and West Canada.

[19] The Hausman (1978) procedure is used to test for the endogeneity of current and lagged wholesale prices in Toronto. The procedure suggests that current (lagged) wholesale prices are exogenous (endogenous). However, they are jointly exogenous according to a F-test. Hence, the F-test determines here that the two null hypothesis that $W_{T,t}$ is exogenous and that $W_{T,t-1}$ is exogenous both hold simultaneously.

[20] See Sen (2003) for an application to Canadian retail gasoline markets.

[21] *Supra* note 1.

[22] Results are similar when we let the cartel's activities start closer to May 2007. The ending date of the conspiracy is also debatable. But we make the assumption that the Competition Bureau conducted searches in November 2007. Bolotova et al (2008) also make the assumption that searches by an antitrust authority correspond to the ending date of a conspiracy. Obviously, one could argue that the cartel continued after November 2007 questioning both the results and the appropriateness of including the post-November 2007 data (see *supra* note 13). That being said, the variable $Cartel_t$ remains statistically significant when all post-November 2007 data are removed from the sample.

[23] It must be noted, however, that residuals obtained in estimating Equation (1) are stationary, *i.e.*, $P_{K,t}$ and $W_{T,t}$ are co-integrated, according to a Dickey-Fuller (1981) test.

[24] The number of lags is chosen to minimize the Bayesien Information Criterion (see *supra* note 12).

[25] As proposed by Engle and Granger (1987), the ECT is estimated first with OLS by running the regression $P_{K,t-1} = \alpha_0 + \alpha_1 W_{T,t-1} + \mu_{t-1}$ and residuals are then inserted directly into Equation (2). The coefficient on the term $P_{K,t-1} - \alpha_0 - \alpha_1 W_{T,t-1}$ should be negative. When retail prices are above (below) their equilibrium level, *i.e.*, $P_{K,t-1} - \alpha_0 - \alpha_1 W_{T,t-1} > (<) 0$, retail prices they should fall (rise).

[26] Coefficients and standard errors for the quarterly dummies are not reported for brevity as they are not statistically significant.

[27] *Competition Act*, RSC 1985, c C-34, s 45.

[28] The amendment eliminates the need to quantify the impact on competition. For instance, empirical results presented in Table 5 suggest that retail prices could have been up to 1.37 cents per litre higher because of the cartel. This represents a 2% price increase (1.37 cents per litre divided by 66.84 cents per litre, which is the average retail price when the cartel was active – see Table 1). Statistical tests show that the cartel could not have had an impact greater than 3%. The question as to whether a 2% or 3% price increase constitutes an undue lessening of competition is now removed from the enforcement of Section 45.

[29] Consider the issue of data reliability. If we had known that the cartel would have taken place in Kingston during a specific period of time, we could have collected a database richer than the one we used (*supra* note 8). But we did not know. And this is precisely one of the limitations of using econometrics in the courtroom. Data are not always perfect and reliable. Econometricians can do the best they can with the data they have and judges need to determine the weight they want to put on the statistical evidence presented during the proceedings.

[30] Rubinfeld (1985) makes this argument.