Journal of Applied Economics. Vol IX, No. 2 (Nov 2006), 361-380

ESTIMATING OVERCHARGES IN ANTITRUST CASES USING A REDUCED-FORM APPROACH: METHODS AND ISSUES

JAMES F. NIEBERDING*

LECG

Submitted February 2005; accepted July 2005

This paper presents several methods and discusses salient issues pertaining to the use of reduced-form models to estimate overcharges in antitrust matters (e.g., price-fixing) where "but-for" prices may be less than actual prices during the anticompetitive period. In particular, two common types of reduced-form estimations are discussed: the "dummy-variable approach" and the "forecasting approach". Under either methodology, an error correction model is then specified as one way to address technical problems often found in applied time-series analysis – nonstationary data and the existence of short-term and long-term dynamics.

JEL classification codes: C22, K21, L4

Key words: price-fixing; time-series, reduced-form, estimated overcharges, antitrust

I. Introduction

Along with the recent emphasis by antitrust agencies on international cartels and their anticompetitive activities such as price-fixing,[1] there appears to be

---

* 1725 Eye Street, NW, Suite 800, Washington, DC 20006. Ph: (202) 973-0526; e-mail: jnieberding@lecg.com. Jim Nieberding is a Senior Managing Economist, LECG (Washington, DC). The views expressed herein are those of the author and should not be construed as representing the position of either LECG or other experts in LECG. A version of this paper was presented at the Contemporary Issues in Law, Economics, and Antitrust session, WEA International Conference, Vancouver, BC (July 2004). I would like to thank two anonymous referees, as well as a co-editor, for very useful comments.

[1] For example, in the United States, the Department of Justice (Antitrust Division) has stated: Beginning in 1995, the Antitrust Division made the prosecution of international cartels that victimize U.S. businesses and consumers one of its highest priorities. International cartels, compared to their domestic counterparts, tend to be broader in scope, larger in terms of

renewed interest in this area by economists (e.g., Harrington 2004a, Harrington 2004b, Levenstein and Suslow 2004, Connor 2004, Connor 2001, Clarke and Evenett 2003). And because economists are trained in the study of markets, in the processes that lead to equilibrium price determination, and in the statistical methods required to analyze these issues, it is not surprising that sometimes they are asked to estimate antitrust damages in such matters.[2] Such an exercise often involves comparing prices actually paid by plaintiffs during the anticompetitive period ("cartel period") to estimates of the prices that would have prevailed in the absence of such conduct but where conditions are otherwise the same (the "but-for" or "benchmark" condition). It is this empirical estimation of unobserved but-for prices – largely ignored in the literature but addressed in this paper – which often becomes the focus of the expert economic analysis.[3]

This paper examines key issues concerning the use of either a forecasting approach or a dummy-variable approach in a reduced-form setting (widely used in antitrust litigation), as well as the various time-series phenomena often encountered (e.g., dynamics, nonstationary data).[4] With the aid of a data set, these methodologies to estimate but-for prices are examined in detail. While these time-

affected volumes of commerce, and more harmful in terms of numbers of businesses and consumers injured. Investigations have uncovered meetings of international cartels in over 100 cities and in over 35 countries, including most of the Far East and nearly every country in Western Europe.

(www.usdoj.gov/atr/public/guidelines/201436.htm#5, September 2003, visited 6/2/05).

There is now significant global consensus that cartel enforcement deserves our most focused attention…

(www.usdoj.gov/atr/public/speeches/206428.htm, November 21, 2004, visited 6/2/05).

Currently, there are approximately 50 sitting grand juries investigating suspected international cartel activity. International cartel investigations account for more than a third of the Division's grand jury investigations. The subjects and targets of the Division's international investigations are located on 6 continents and in nearly 25 different countries…Since the beginning of FY 1997, the Division has prosecuted international cartels affecting well over $10 billion in U.S. commerce. (www.usdoj.gov/atr/public/speeches/207226.htm, January 10, 2005, visited 6/2/05)

[2] Because horizontal agreements such as price fixing are per se illegal in countries like the U.S., it is in the private "follow-on" civil litigation where the type of economic evidence presented in this paper is almost exclusively found.

[3] While many economists over the years no doubt have engaged in the estimation of but-for prices in the damages phase of antitrust litigation, the details of their methods likely remain hidden in their confidential expert reports.

[4] "The most common statistical method employed in antitrust litigation involves the estimation of reduced form price equations." (Baker and Rubinfeld 1999, p. 391).

series data are "artificial", they mimic the patterns and characteristics (nonstationarity and dynamic behavior) of actual data from a price-fixing case in which the author was involved.[5]

While there are several helpful articles that provide a general overview of some issues in the estimation of but-for prices in antitrust matters (e.g., Baker and Rubinfeld 1999, Finkelstein and Levenbach 1983, Rubinfeld and Steiner 1983), their focus generally is on "higher-level" issues. This paper contributes to the literature not only by providing the necessary details of estimating but-for prices in a reduced-form setting, but also focuses on more recent time-series issues by illustrating one method that accounts for such phenomena.

II. Using reduced-form methods

Although plaintiffs' pricing history in an antitrust matter may be determined by many factors, the goal of the reduced-form analysis is to isolate the effect of the anticompetitive conduct. Defining the but-for situation that plaintiffs would have faced is key to isolating the effect of this conduct and distinguishing it from the effects of factors unrelated to, and unaffected by, the anticompetitive conduct. For example, a measure of damages typically applied in price-fixing cases is based upon an overcharge – the difference between the prices paid by plaintiffs and the but-for prices during the cartel period (Page 1996, p. 36).[6]

The ability to measure the overcharge accurately depends upon how reliably and precisely the analysis can distinguish the collusive effect on prices from other influences that are unrelated to the anticompetitive conduct. In practice, reduced-form price models are used to 1) isolate the impact of the anticompetitive conduct

[5] For confidentiality reasons, use of the actual data or identifying the matter involved is not possible. However, I believe the data used are sufficient for the purposes of this paper.

[6] While this measure of damage may provide an estimate of that suffered by the plaintiff, it is not intended to measure additional damages such as the deadweight loss (DWL) associated with the artificially elevated prices due to the anticompetitive behavior. Broadly speaking, DWL denotes the loss in surplus by producers and consumers due to the price increase that is not re-distributed to some agent in the economy. It may arise because as a result of the price increase, some buyers switch from the now relatively more expensive item to substitutes which presumably were less desirable prior to the price increase. This constitutes a welfare loss for buyers. Also, if production expands for those products that are considered substitutes to the now higher-priced item, and presumably because such production would not have been profit maximizing or efficient absent the price increase, it represents a production-related DWL.

on the price under study after controlling for influences unrelated to such conduct (the "dummy variable approach"), or 2) forecast but-for prices during the cartel period based upon relationships and trends using data outside the cartel period (the "forecasting approach"). These approaches are discussed in more detail below.[7]

A. The underlying structural model

To isolate the cartel and non-cartel influences on price, economists may quantify the interaction of buyers and sellers through the explicit construction of behavioral supply and demand relations in the context of a structural model as specified by economic theory pertinent to the industry under study. For example, suppose that demand and supply can be represented by the two estimating equations:

$$Q_t^d = \alpha_0 + \alpha_1 P_t + \alpha_2 D_t, \tag{1}$$

$$Q_t^s = \beta_0 + \beta_1 P_t + \beta_2 S_t + \beta_3 M_t,$$

where Q_t^d and Q_t^s represent the quantity demanded and supplied, respectively, at time t; P_t is market price, and D_t, S_t, and M_t represent exogenous demand and supply influences (e.g., S_t may represent an exogenous input cost; M_t represents the quantity of imports of a competing product). Since equilibrium price and quantity are jointly determined in the above system, proper econometric techniques (e.g., 3SLS) need to be employed assuming the equations are identified. The expected signs of the parameters attached to D_t, S_t, and M_t will depend on the theoretical reason for their inclusion in the structural model. For example, if S_t represents an exogenous "cost shifter," then a reasonable expectation would be

[7] Whether or not regression analysis can be used reliably to estimate overcharges in a specific antitrust matter depends upon the circumstances of each case. An econometric approach to measuring overcharges is only one recognized method of damages estimation, and the chosen method to estimate but-for prices naturally will depend on the specific set of facts about the industry under study, the availability and quality of data, and the nature of the hypotheses being tested. For the sake of simplicity, in this paper it is assumed that defendants' liability has been established prior to the damages analysis, that the task is to determine an overcharge due to horizontal price-fixing using a reduced-form price equation, that the plaintiffs represent direct purchasers, and that the benchmark and cartel periods have been established.

that b_2 ought to be negative reflecting that as this cost increases, Q_t^s will decrease (all else equal).[8]

B. The reduced-form model

By using the equilibrium condition that supply equals demand, equilibrium price (P_t) can be solved for using the equations in (1),

$$P_t = \delta_0 + \delta_1 S_t + \delta_2 D_t + \delta_3 M_t \tag{2}$$

This yields a reduced-form equation that relates equilibrium price to exogenous supply and demand factors. When interpreting reduced-form regression coefficients estimated from a model like (2), one needs to remember that these parameters are a function of those from the underlying structural model. This can be seen by solving for P_t from the equations in (1) which can be written as:

$$P_t = \left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}\right) + \left(\frac{\beta_2}{\alpha_1 - \beta_1}\right) S_t - \left(\frac{\alpha_2}{\alpha_1 - \beta_1}\right) D_t + \left(\frac{\beta_3}{\alpha_1 - \beta_1}\right) M_t \tag{3}$$

Equation (3) is identical to the reduced-form price equation in (2) assuming that:

$$\delta_0 = \left(\frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1}\right) \quad \delta_1 = \left(\frac{\beta_2}{\alpha_1 - \beta_1}\right) \quad \delta_2 = \left(\frac{-\alpha_2}{\alpha_1 - \beta_1}\right) \quad \delta_3 = \left(\frac{\beta_3}{\alpha_1 - \beta_1}\right); \tag{4}$$

$$(\alpha_1 - \beta_1) \neq 0$$

Therefore, while the expected sign of a reduced-form regression coefficient is informed by economic theory, it also is dependent on the underlying structural

[8] Because input costs can be a significant component of price, their variation can be expected to influence price changes. The cost variables typically included are exogenous input prices for the product whose price is under study. If, however, the firms are vertically integrated into the production of such upstream inputs making their prices likely not "arm's length", then suitable proxies must be found.

regression coefficients (some of which may have opposing effects on equilibrium price).[9]

Likewise, the reduced-form quantity equation, $Q_t = f(S_t, D_t, M_t)$, can be obtained from (1),

$$Q_t = \left(\frac{\alpha_1\beta_0 - \beta_1\alpha_0}{\alpha_1 - \beta_1}\right) + \left(\frac{\beta_2\alpha_1}{\alpha_1 - \beta_1}\right)S_t - \left(\frac{\beta_1\alpha_2}{\alpha_1 - \beta_1}\right)D_t + \left(\frac{\beta_3\alpha_1}{\alpha_1 - \beta_1}\right)M_t, \quad (5)$$

which is identical to the reduced-form quantity equation:

$$Q_t = \delta_4 + \delta_5 S_t + \delta_6 D_t + \delta_7 M_t, \quad (6)$$

assuming that:

$$\delta_4 = \left(\frac{\alpha_1\beta_0 - \beta_1\alpha_0}{\alpha_1 - \beta_1}\right), \ \delta_5 = \left(\frac{\beta_2\alpha_1}{\alpha_1 - \beta_1}\right), \ \delta_6 = \left(\frac{-\beta_1\alpha_2}{\alpha_1 - \beta_1}\right), \ \delta_7 = \left(\frac{\beta_3\alpha_1}{\alpha_1 - \beta_1}\right), \quad (7)$$

$(\alpha_1 - \beta_1) \neq 0.$

The two reduced-form equations given by (3) and (5) show explicitly how the two endogenous variables (P_t and Q_t) are jointly determined, and how changes in the exogenous variables affect the endogenous ones. If one were to estimate the two-equation system given by (2) and (6), then at least theoretically and under certain circumstances, one could solve the system of eight simultaneous equations given in (4) and (7) and "recover" the seven implied structural coefficients in (1) from the estimates of the reduced-form coefficients in (2) and (6).[10] For example, solving these eight equations (i.e., δ_0 - δ_7) and assuming $(\alpha_1 - \beta_1) \neq 0$ yields:

[9] Moreover, to the extent that the independent variables in a reduced-form model like (2) exhibit multicollinearity, it can be difficult to precisely estimate the separate impact of each on the dependent variable.

[10] That there are eight equations but only seven unknowns in this example stems from the fact that the supply equation in (1) is exactly identified while the demand equation technically is over-identified. Mathematica software was used to solve this system.

$$\alpha_0 = \delta_4 - \frac{\delta_0 \delta_7}{\delta_3}, \quad \alpha_1 = \frac{\delta_7}{\delta_3}, \quad \alpha_2 = \delta_6 - \frac{\delta_2 \delta_7}{\delta_3}, \tag{8}$$

$$\beta_0 = \delta_4 - \frac{\delta_0 \delta_6}{\delta_2}, \quad \beta_1 = \frac{\delta_6}{\delta_2}, \quad \beta_2 = \delta_5 - \frac{\delta_1 \delta_6}{\delta_2}, \quad \beta_3 = \delta_7 - \frac{\delta_3 \delta_6}{\delta_2}.$$

In practice, however, perhaps because it might be easier to obtain pricing data than quantity data, it is more typical to see the economic expert in antitrust matters estimate only the reduced-form price equation such as (2).[11] That is, data on prices and the independent variables are used to estimate values for the reduced-form regression parameters. Then, by using the actual values of the independent variables and the reduced-form regression coefficients, predicted values for prices – as well as but-for prices – can be generated for the computation of overcharges. This is the approach utilized in this paper using a "double-log" model for the reduced-form price equation.[12]

There likely are many factors that influence the specific values of the price under study as time unfolds. The estimation of the reduced-form equation proceeds as if data for all the relevant factors are included and available. In practice, this usually is not the case; some relevant variables are likely to be omitted perhaps due to the unavailability of reliable data. These omissions may or may not be significant. For example, they may not be significant if, in fact, the values and influences of these variables were similar in the cartel and the benchmark periods.[13]

III. Two methods of using reduced-form models to estimate overcharges

Two common methods of using a reduced-form model to estimate overcharges

[11] See footnote 4.

[12] In a double-log model, estimated regression coefficients are elasticities. Moreover, taking natural logs of the data typically narrows the range of the variables which make any statistical analysis less sensitive to outlying observations.

[13] The consequence of omitting an independent variable that is correlated with included independent variables (known as "omitted variable bias") is that all estimated regression coefficients from the under-specified model will be biased, inconsistent, and perhaps inefficient. As a result, their standard errors are likely to give misleading conclusions about the statistical significance of these estimated coefficients. If the omitted variable is not correlated with the included explanatory variables, the slope coefficients from the under-specified model will be unbiased estimators of the true parameters. For a more detailed treatment of this topic, see Wooldridge (2000, pp. 87-100).

are the "dummy variable (DV) approach" and the "forecasting approach." The DV approach is characterized by Rubinfeld and Steiner (1983):

> In it, one would estimate the econometric model for all time periods for which there was data, both conspiratorial and nonconspiratorial. However, this estimating model would include one additional variable, a dummy variable equal to 1 when the cartel was in effect, and 0 otherwise. The coefficient on the cartel dummy would then indicate whether the average price during the cartel's existence was significantly different and, in particular, higher than the average price during the competitive period.[14]

Under this approach, it is common practice to use one or more "intercept" dummy variables during the cartel period, the coefficient(s) of which measure the average difference in price between the benchmark and cartel periods after controlling for other factors thought to influence price, yet unrelated to the cartel. This dummy variable – if found to be significant – can be thought of as representing a "shift" in the price line under study rather than causing a change in its shape (or slope).[15] However, as discussed by Finkelstein and Levenbach (1983), if prices are affected in more varying and complex ways, the use of one scale dummy variable for the entire cartel period (which assumes that the cartel added a fixed dollar or percentage amount to price during the conspiracy) may be too simple an approach.[16]

Under the forecasting approach, reduced-form price models are estimated using data from periods unaffected by the anticompetitive conduct. Then, using the regression coefficients along with the observed values of the independent variables during the cartel period, the regression equation is used to forecast prices that would have prevailed but-for the cartel.[17] The overcharge then is calculated as the

---

[14] Rubinfeld and Steiner (1983, p. 123, fn. 77).

[15] A slope dummy variable may be used to investigate whether the relationship between price and an independent variable is different in the benchmark and cartel period. The analysis here is confined to one scale dummy variable.

[16] Finkelstein and Levenbach (1983, p. 156). These authors also point out that under the DV approach, to the extent that the cartel dummy variable is correlated with other independent variables in the reduced-form equation, then applying OLS will allocate explanatory power among the independent variables (including the dummy variable) differently than would a forecasting approach.

[17] As stated by Rubinfeld (1985, p. 1087), "in a price-fixing case, for example, a regression model that explains price in a period of nonviolation can be used to predict what the price would have been during the period of violation."

difference between the actual and but-for prices. This approach requires that after excluding data from the cartel period from the estimation sample, enough data exist in the benchmark period so that a forecasting equation reliably and precisely can be estimated.

The principal difference between the forecasting approach and the DV approach is that the latter uses data for the entire time period as the estimation sample while a forecasting model uses data only from the benchmark period. However, these two approaches are related. As noted by Higgins and Johnson (2003), the forecasting approach is equivalent to the DV approach when the dummy variables meant to investigate the presence or absence of overcharges during the cartel period are defined at the same frequency as the data in the estimation sample. For example, when the underlying data are monthly (or, say, quarterly), the DV approach is equivalent to the forecasting approach when a separate dummy variable is included for every month (or quarter) during the cartel period.[18]

There are several assumptions implicitly being made when one approach is selected over the other. In using the forecasting approach, one assumes that the same relationship among price and the independent variables that exists in the benchmark period also holds in the cartel period. However, if this relationship does appreciably change during these two periods, the forecast model may not reliably predict but-for prices vis-à-vis an approach that accounts for this. Under the DV approach, one also must maintain the assumption that the same relationship between price and the independent variables exists in the benchmark and cartel period. If the influence of supply and demand factors affects equilibrium price differently in these two periods, then the DV approach may not reliably measure overcharges vis-à-vis an approach that accounts for such a change.[19]

With the forecasting approach and assuming that enough data exist pre- and post-cartel, one needs to investigate the appropriateness of pooling both sets of prices vis-à-vis using only pre-cartel prices for forecasting into the cartel period, or using only post-cartel prices for backcasting. If market realities and price-setting behavior differ before and after the cartel, this needs to be reflected in the analysis if the factors driving such differences are identifiable and able to be controlled for. If not, care needs to be taken that exclusion of one of the periods will not introduce

[18] This equivalency occurs because each observation during the cartel period effectively is "dummied out" thereby removing it from the analysis. Such a result need not occur if, in using the DV approach, quarterly or annual dummy variables are introduced during the cartel period when all data are, say, monthly.

[19] As noted by Rubinfeld and Steiner (1983, p. 123 fn. 77):

bias into the estimated but-for prices (Harrington 2004b). Moreover, since the regression equation under the forecasting approach is estimated only with data outside the cartel period, a good fit within the benchmark period does not necessarily mean that the regression equation will forecast well outside the benchmark period.

Similar issues arise under the DV approach. For example, such models often uses data "before and after" the cartel period. In doing so, the inclusion or exclusion of post-cartel price data sometimes becomes a point of contention among the parties. The debate often revolves around whether the effects of the anticompetitive conduct have "lingered" beyond the purported end of the cartel thereby throwing into question the appropriateness of including such data in the benchmark.[20]

In some instances, direct evidence of cartel behavior can be tested against pricing data.[21] With the DV approach, however, if the beginning and end of the cartel period is not known prior to estimation (e.g., there is no *prima facie* evidence for this from an earlier guilty plea), then this likely will become a point of contention among the parties. Harrington (2004b) notes:

> A second problem is that it is more difficult to identify when the cartel started than when it ended. If one presumes that the beginning of an investigation caused collusion to stop then the beginning of the post-cartel period may be relatively easy to identify. However, the end of the pre-cartel period is typically not so straightforward.

[T]here is one important limiting assumption made [in the DV approach]. That assumption is that the overall behavior of the regression model, that is, the behavior of supply and demand, can be modeled in precisely the same way during both the conspiratorial and nonconspiratorial periods…In contrast, the described forecasting approach requires no such assumption…

A similar sentiment can be found in Rubinfeld (1985, pp. 1090-1091); and Fisher (1980, pp. 727-728). However, Higgins and Johnson (2003) argue (and theoretically prove) that under certain assumptions and in certain instances, the DV approach that does not control for the different effects that independent variables have on the dependent variable between the benchmark and cartel periods (when in fact they exist), will result in an unbiased estimated overcharge that is more precisely estimated than if a more complex model with "slope-changing" dummy variables were estimated.

[20] In practice, the inclusion or exclusion of post-conspiracy data may have a significant effect on estimated overcharges. As discussed by Harrington (2004b) who formally explores issues related to the lingering effects of cartels, if elevated prices linger after the cartel period, the inclusion of such data in the estimation sample necessarily will overstate but-for prices thereby underestimating overcharges.

[21] If there is evidence that such an agreement was reached on a certain date, the resulting price elevation attributable to the anticompetitive conduct should be apparent in an analysis of firms' transactions data after that date.

These and other relevant issues ought to be investigated for their applicability in particular instances to know which approach – or if an alternate methodology – may be more reasonable to estimate overcharges. Under either the forecasting or DV approach, however, another more basic consideration is that if an appropriately specified reduced-form model is to be used to test for and measure an average overcharge, it must first have been established that the average is a reasonable representation of the overcharge suffered by all plaintiffs. That is, using either methodology, it may be appropriate to investigate whether or not it is more reasonable to calculate plaintiff-specific but-for prices rather than relying on a "one size fits all" approach.[22]

IV. An illustrative example

This section illustrates in more detail the methodologies discussed above. Figure 1 illustrates the time path of the price under study in both natural logs and levels over the entire time period, or 84 observations (e.g., seven years of monthly data; 21 years of quarterly data). Assume that the cartel operates in periods 45 through 64, that equilibrium price is a function of a supply-side factor (S_t) and a demand-side factor (D_t) – both expected to be positively related to equilibrium price – as well as of the quantity of imports of a competing product (M_t) which is substitutable in consumption to the product under study (i.e., M_t and P_t are inversely related, all else equal).[23]

A. The DV and forecasting approach

Under this methodology, Dum_t equals 1 during the cartel period, and 0 otherwise. Equation (9) reports the OLS estimation of (2) using all 84 observations; but-for prices during the cartel period are estimated by setting Dum_t equal to 0:

---

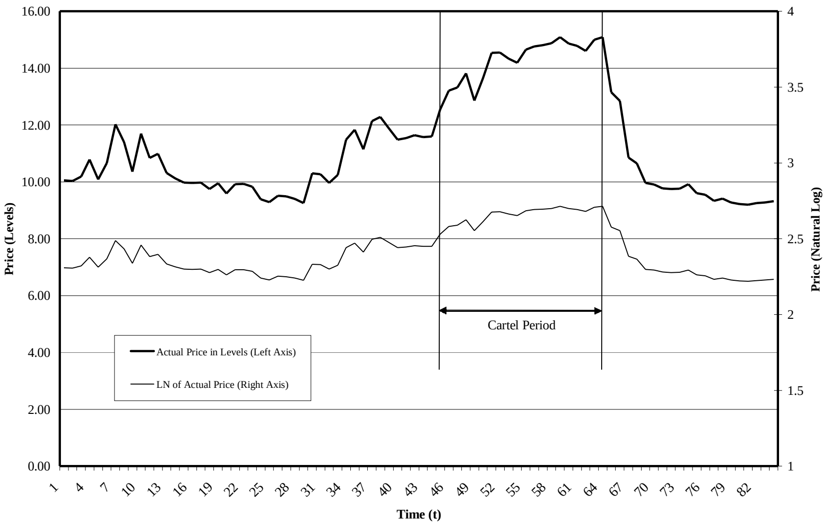
[22] One may want to investigate whether overcharge estimates are significantly sensitive to the method employed. If both the forecasting and DV approach yield similar estimated overcharges, then one might be reassured that overcharges are robustly estimated in that they are not being driven by the method employed.

[23] For the purpose of this paper, potential endogeneity among independent variables and equilibrium price in the reduced-form model is ignored (e.g., while M_t may be responsive to and jointly determined by P_t, such potential simultaneity is assumed not to be present). In practice, one can apply formal tests of exogeneity to variables suspected of being endogenous. If found, and assuming that sufficient instruments exist for the endogenous variable, econometric techniques such as 3SLS can be used to produce consistent parameter estimates.

$$\ln \hat{P}_t \;=\; \underset{(2.70)}{2.489 + 0.145 \ln S_t} + \underset{(2.87)}{0.134 \ln D_t} - \underset{(-4.07)}{0.103 \ln M_t} + \underset{(5.62)}{0.137 Dum_t} \qquad (9)$$

The coefficient and t-statistic on Dum_t indicates that prices were elevated significantly during the cartel period relative to the benchmark period. Under this approach, the overcharge during the cartel period equals 14.73%.[24]

Figure 1. Actual price (Cartel period: t = 45 through 64)



In some instances, there may be evidence that the price under study does not adjust instantaneously to changes in market conditions because, even after controlling for contemporaneous and exogenous market factors, current price (P_t) may still depend on past prices (e.g., P_{t-1}, $P_{t-2,}$ etc.). One method of capturing such

[24] This is calculated as the percent change between the predicted and but-for price for each time period during the cartel period which, under this specification, is equivalent to $\mathbf{e}^{a} - 1$ where a represents the estimated coefficient on Dum_t. With a positive and significant coefficient for Dum_t in models such as (9), a log-log DV model will result in a constant overcharge percentage whereas a model in levels would result in a constant overcharge in dollars. For example, estimating (9) in levels results in a constant overcharge of $2.06 for each period during the cartel period. This results in an average percentage overcharge during the cartel period of 14.46%.

price dynamics is to allow for the presence of a lagged dependent variable. For example, incorporating P_{t-1} into (9) and estimating yields:

$$\ln \hat{P}_t = 0.742 + 0.580 \ln P_{t-1} + 0.090 \ln S_t + 0.093 \ln D_t - 0.040 \ln M_t + 0.052 Dum_t \quad (10)$$

(12.38) (2.91) (3.37) (-2.58) (3.31)

Here, P_{t-1} is significantly correlated to current price P_t, while its inclusion does not significantly perturb the sign and significance of the other included independent variables compared to (9). The average overcharge during the cartel period from (10) equals 12.09%.[25]

In (10), the percentage overcharge (in levels) for each period during the cartel is ($e^{\text{factor}} - 1$) where "factor" equals a for the first period of the cartel; $\alpha \cdot (1+\beta)$ for the second; $\alpha \cdot (1+\beta+\beta^2)$ the third, etc.; a is the coefficient on Dum_t and b is the coefficient on P_{t-1}.[26] Equivalently, the percentage overcharge (in levels) for each period during the cartel is $e^{\alpha}\left(\hat{P}_{-1} / P_{-1}^{BF}\right)^{\beta} - 1$, where $\hat{P}_{-1}$ is the lagged predicted price, and P_{-1}^{BF} is the lagged but-for price. "Linger" – the elevation in the predicted price vis-à-vis the but-for price post-cartel – decays according to the expression $\left(\hat{P}_{-1}/P_{-1}^{BF}\right)^{\beta} - 1$, and will persist if the coefficient on lagged price is large.

The expressions in the preceding paragraph can be derived as follows. Consider the general reduced-form price equations:

$$\ln \hat{P} = \boldsymbol{\lambda}\mathbf{X} + \alpha Dum + \beta \ln \hat{P}_{-1}, \quad (11)$$

$$\ln P^{bf} = \boldsymbol{\lambda}\mathbf{X} + \beta \ln P_{-1}^{bf}, \quad (12)$$

where $\hat{P}$ is the predicted price from (11), P^{bf} is the but-for price from (12), **X** is a matrix of exogenous independent variables including a constant, l is a vector of parameters, *Dum* is the cartel dummy variable equal to 1 in (11) and 0 in (12), $\hat{P}_{-1}$ is the lagged predicted price, and P_{-1}^{bf} is the lagged but-for price. To obtain the

---

[25] When generating the predicted and but-for prices from (10), lagged predicted and but-for prices rather than lagged actual prices should be used in dynamically simulating the time-path of these prices.

[26] Note that when $b = 0$, this expression for calculating overcharges collapses to $e^a - 1$, identical to the expression to calculate overcharges in the case without P_{t-1} as an independent variable.

difference in the predicted and but-for price for each period during the cartel, subtract (12) from (11). Doing this for the first several periods in the cartel yields:

$$\left(\ln \hat{P} - \ln P^{bf}\right)_{45} = \alpha, \tag{13}$$

$$\left(\ln \hat{P} - \ln P^{bf}\right)_{46} = \alpha + \beta\left(\ln \hat{P}_{-1} - \ln P_{-1}^{bf}\right) = \alpha(1+\beta), \tag{14}$$

$$\left(\ln \hat{P} - \ln P^{bf}\right)_{47} = \alpha + \beta\left(\ln \hat{P}_{-1} - \ln P_{-1}^{bf}\right) = \alpha(1+\beta+\beta^2), \tag{15}$$

and so on. Alternatively, to obtain the expression for the overcharge for each period during the cartel (%DP), subtract (12) from (11):

$$\ln \hat{P} - \ln P^{bf} = \alpha + \beta\left(\ln \hat{P}_{-1} - \ln P_{-1}^{bf}\right) \tag{16}$$

By algebraically rearranging (16) and exploiting the properties of logarithms, this can be written as:

$$\left(\frac{P}{P_{bf}} - 1\right) = \Delta\%P = e^{\alpha}\left(\hat{P}_{-1} \big/ P_{-1}^{bf}\right)^{\beta} - 1 \tag{17}$$

With dynamic reduced-form models like (10), the implied long-run equilibrium price (P_e) can be estimated by setting $P_e = P_t = P_{t-1}$ and solving for P_e. That is,

$$\ln P_e = \frac{\delta_0 + \delta_1 \ln S_t + \delta_2 \ln D_t + \delta_3 \ln M_t + \alpha Dum_t}{1-\beta} \tag{18}$$

The expression $\frac{\Delta \ln P_e}{\Delta Dum_t} = \frac{\alpha}{1-\beta}$ gives the long-run effect of the cartel on equilibrium price, or approximately 12.38% ($\alpha = 0.052$; $\beta = 0.58$). This is not much greater than the average overcharge during the cartel period (12.09%) due to the fact that the but-for price quickly returns to the predicted price after the cartel period ends.

Under a forecasting approach – with cartel period data omitted from the estimation sample – using the price data in Figure 1 and the same specifications as above, the average overcharge equals 15.27% (without P_{t-1} as a regressor) and 11.53% (including P_{t-1} as a regressor). These estimates are not too different from that estimated under the DV approach (14.73% and 12.09%, respectively).[27]

B. Using an error correction model to address additional time-series issues

Because time-series data frequently are used in econometric analyses related to antitrust issues, this section introduces one method to account for the presence of nonstationarity in the data. While the source of potential spuriousness discussed here is due to the time-series properties of the data under study (e.g., all variables may be similarly trending so that any relationship between them may simply be because each is growing over time), there are other situations in antitrust (i.e., related to model specification) that could introduce spurious findings.[28]

It is standard practice to test individual time-series for stationarity prior to inclusion in a regression model. While a complete discussion of stationarity, the related concept of cointegration, and the formal statistical tests used to diagnose such phenomena is not necessary for the aims of this paper, suffice it to say that regressions containing nonstationary time series may produce spurious results (e.g., variables are related only through their correlation with an omitted variable). The concern in using nonstationary data is that a spurious regression may have been performed which generally overstates how good is a model's "fit", as well as

[27] In comparison to the "dynamic" DV approach (i.e., with P_{t-1} as a regressor), a "static" forecasting model may result in linger if the actual price exceeds the predicted price after the cartel period. Similar to the approach under the dynamic DV method, when generating but-for prices under the static forecasting approach, lagged predicted and but-for prices rather than lagged actual prices should be used in generating the time-path of these prices.

[28] For example, statistical tests of associations among regional prices in the determination of the extent of a geographic market may be driven in part by the effect of common factors or influences. If so, it is possible that a finding of correlation (or cointegration) among regional prices does not necessarily show regions are linked in a global market, but might reflect parallel movements of key input costs in each region over time. Suppose it is the case that there is a high correlation between prices of retail gasoline in two distant cities. However, this does not necessarily indicate that they are in the same retail geographic market. Rather, prices in both cities may respond at least in part to similar fluctuations in the price of crude oil. Statistical models in such instances need to take account of this.

makes independent variables appear to be more significant than they actually are in explaining variation in the dependent variable.[29]

A standard technique to rid a time series of nonstationarity should it be found to exist is to take first differences (i.e., $\Delta \ln P_t = \ln P_t - \ln P_{t-1}$) and, if these are stationary, to use them in the regression model.[30] Moreover, a common econometric technique – an "error correction model" (ECM) – may be used in the context of reduced-form price equations to address not only the nonstationarity of the data (because it uses data that are expressed as first differences), but also to allow for short-run dynamics (by incorporating lags of the variables), as well as the long-term relationship between the equilibrium price being modeled and its actual value (by including an "error correction term").

Define $\ln P_t^e$ as the equilibrium price at time *t* being modeled in the reduced-form price model; $\ln P_t$ represents the actual price; and $Gap_t = (\ln P_t - \ln P_t^e)$ represents the difference between the actual price and its equilibrium value. This "disequilibrium gap" represents the fact that at any time period *t*, it can be expected that the actual and equilibrium values of $\ln P_t$ will differ for factors unrelated to collusion. Given this, an ECM can be written as:

$$\Delta \ln P_t = \eta(\Delta \ln P_{t-1}) + \gamma(\ln P_{t-1} - \ln P_{t-1}^e), \qquad (19)$$

which expresses the current change in price as a function of the previous period's change in price and the previous period's disequilibrium gap – also called the error correction term – with $\gamma < 0$.[31] For example, if $(\ln P_{t-1} > \ln P_{t-1}^e)$ then price in the previous period has overshot its equilibrium value, but because $\gamma < 0$, the term

[29] For an introduction to these topics as used in antitrust, see, e.g., Rubin (2004) and Forni (2004). Technically, a stochastic process which generates a time series is said to be stationary if its mean is constant over time, and exhibits a constant and finite variance so that "shocks" only have transitory effects. Time series data that violate these stationarity conditions are termed "nonstationary" and often exhibits an upward or downward trend without reversion to a constant mean, and any "shock" will have a lasting impact on future values of this series. Two or more nonstationary series are cointegrated if they have a linear combination that is stationary. If the nonstationary time series to be used in a regression are cointegrated, then regression results containing these series may be meaningful (and not spurious).

[30] When data are in logs, a first difference approximates the percentage change in P_t (i.e., the price measured in levels).

[31] See Wooldridge (2000, pp. 591-593) for a discussion of ECMs and their general specification.

$\gamma(Gap_{t-1})$ in (19) will be negative and operate to decrease the current price back towards equilibrium. Likewise, if $(\ln P_{t\text{-}1} < \ln P^{e}_{t\text{-}1})$, because $\gamma < 0$, the term $\gamma(Gap_{t-1})$ will be positive and work to increase current price towards equilibrium.

The reduced-form price equation lagged one period can be written as:

$$\ln P^{e}_{t\text{-}1} = \delta_0 + \delta_1 \ln S_{t\text{-}1} + \delta_2 \ln D_{t\text{-}1} + \delta_3 \ln M_{t\text{-}1}, \tag{20}$$

which, when substituted into (19), yields:

$$\begin{aligned} \Delta \ln P_t &= \eta(\Delta \ln P_{t-1}) + \gamma \ln P_{t\text{-}1} - \gamma\delta_0 - \gamma\delta_1 \ln S_{t\text{-}1} - \gamma\delta_2 \ln D_{t\text{-}1} - \gamma\delta_3 \ln M_{t\text{-}1}, \\ \Delta \ln P_t &= \mu_0 + \mu_1(\Delta \ln P_{t-1}) + \mu_2 \ln P_{t\text{-}1} + \mu_3 \ln S_{t\text{-}1} + \mu_4 \ln D_{t\text{-}1} + \mu_5 \ln M_{t\text{-}1}, \end{aligned} \tag{21}$$

where $\mu_0 = -\gamma\delta_0$; $\mu_1 = \eta$; $\mu_2 = \gamma$; $\mu_3 = -\gamma\delta_1$; $\mu_4 = -\gamma\delta_2$; and $\mu_5 = -\gamma\delta_3$. Therefore, under the DV approach, the ECM to be estimated becomes:

$$\begin{aligned} \Delta \ln P_t &= \mu_0 + \mu_1 \Delta \ln P_{t-1} + \mu_2 \ln P_{t\text{-}1} + \mu_3 \ln S_{t\text{-}1} \\ &+ \mu_4 \ln D_{t\text{-}1} + \mu_5 \ln M_{t\text{-}1} + \mu_6 Dum_t \end{aligned} \tag{22}$$

Equation (23) reports the regression equation based upon the estimation of (22):

$$\begin{aligned} \Delta \ln P_t = 0.955 &- \underset{(-1.84)}{0.177} \Delta \ln P_{t-1} - \underset{(-5.81)}{0.441} \ln P_{t\text{-}1} + \underset{(3.25)}{0.125} \ln S_{t\text{-}1} + \\ &+ \underset{(1.07)}{0.039} \ln D_{t\text{-}1} - \underset{(-1.62)}{0.032} \ln M_{t\text{-}1} + \underset{(4.68)}{0.084} Dum_t \end{aligned} \tag{23}$$

Here, the estimated coefficient on Dum_t is positive and statistically significant indicating that the price increases during the cartel period are higher than that predicted by the other variables in the model. Figure 2 illustrates the corresponding predicted and but-for prices (in levels) from (23), and the but-for prices are estimated by setting Dum_t equal to 0 for the cartel period. Under this approach, the average overcharge during the cartel period equals 19.11%.

Following the same procedure as above, one can use (23) to find the long-run effect of the anticompetitive behavior on equilibrium price, lnP_e. That is, by setting the changes in price equal to 0 (i.e., $0 = \Delta \ln P_t = \Delta \ln P_{t-1}$), setting ln$P_e$ = lnP_{t-1}, and solving for lnP_e, the expression $\frac{\Delta \ln P_e}{\Delta Dum_t} = \frac{\mu_6}{\mu_2}$ represents the ratio of the coefficient for the dummy variable to that for the lagged price, or 19.05% (using $\mu_2 = 0.441$; $\mu_6 = 0.084$)[32] which is similar to the estimated average overcharge of 19.11%.

Figure 2. Predicted and but-for price from ECM — DV approach (Cartel period: t = 45 through 64)

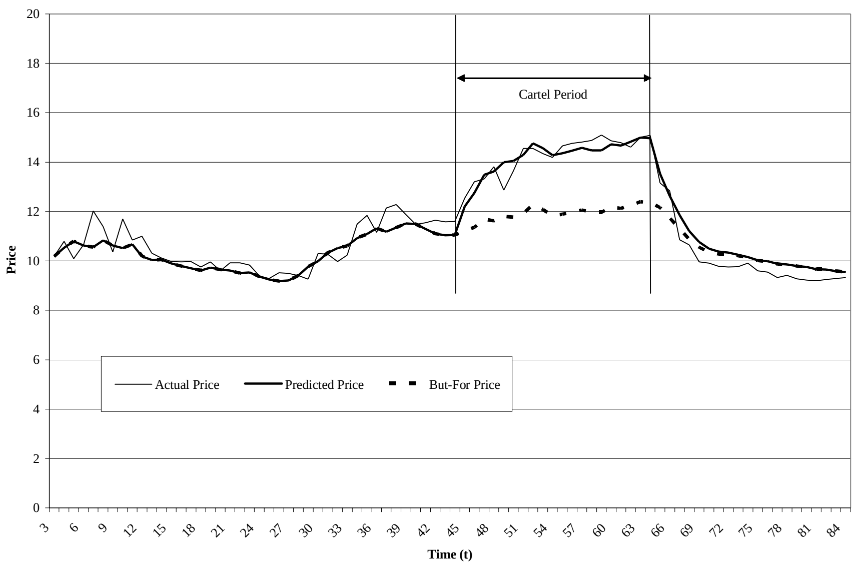
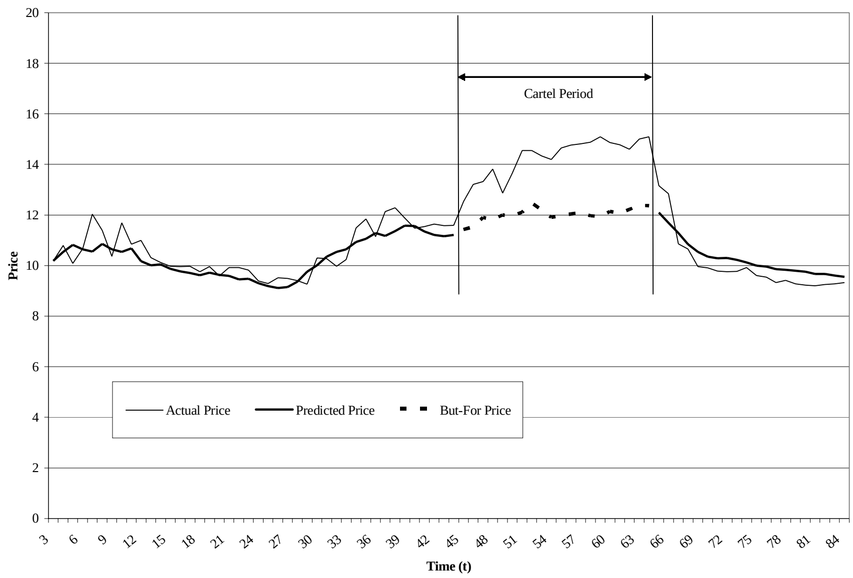


Figure 3 illustrates the predicted and but-for prices (in levels) under a forecasting approach using the specification in (23). Here, the average overcharge for the cartel period equals 18.69%.

[32] Note that in the ECM, $m_2 = g < 0$ by construction, so that the estimated coefficient on lnP_{t-1}, m_2, equals - (-0.441), or +0.441.

Figure 3. Predicted and but-for price from ECM — forecasting approach (Cartel period: t = 45 through 64)



V. Conclusion

This paper illustrates the workings of a reduced-form price model to estimate but-for prices in an antitrust setting under both a dummy-variable approach and a forecasting approach. Salient issues are discussed concerning each method. Because pricing data frequently are plagued with time-series phenomena such as nonstationarity and dynamic behavior, an error correction model is specified as one way to account for such issues in the context of a reduced-form pricing model. While reduced-form models in antitrust litigation often appear to be *ad hoc* with respect to how these various time-series issues are handled, the ECM approach offers one systematic way to address them in the computation of overcharges in antitrust matters.

References

Baker, Jonathan B. and Daniel L. Rubinfeld (1999), "Empirical methods in antitrust litigation: Review and critique", *American Law and Economics Review* **1**: 386-435.

Clarke, Julian L. and Simon J. Evenett (2003), "The deterrent effects of national anticartel laws: Evidence from the International Vitamins Cartel", *Antitrust Bulletin* **48**: 689-726.

Connor, John M. (2001), *Global Price Fixing: Our Customers are the Enemy*, Boston, MA, Kluwer Academic.

Connor, John M. (2004), "Global antitrust prosecutions of modern international cartels", *Journal of Industry, Competition and Trade* **4**: 239-267.

Finkelstein, Michael O. and Hans Levenbach (1983), "Regression estimates of damages in price-fixing cases", *Law and Contemporary Problems* **46**: 145-169.

Fisher, Franklin (1980), "Multiple regression in legal proceedings", *Columbia Law Review* **80**: 702-736.

Forni, Mario (2004), "Using stationarity tests in antitrust market definition", *American Law and Economics Review* **6**: 441-464.

Harrington, John. E., Jr. (2004a), "Cartel pricing dynamics in the presence of an antitrust authority", *RAND Journal of Economics* **35**: 652-674.

Harrington, John. E., Jr. (2004b), "Post-cartel pricing during litigation," *Journal of Industrial Economics* **52**: 517-533.

Higgins, Richard S. and Paul A. Johnson (2003), "The mean effect of structural change on the dependent variable is accurately measured by the intercept change alone," *Economics Letters* **80**: 255-259.

Levenstein, Margaret C. and Valerie Y. Suslow (2004), "What determines cartel success?", in Peter Z. Grossman, ed., *How Cartels Endure and How They Fail: Studies of Industrial Collusion*, Cheltenham, UK, Edward Elgar.

Page, William H., ed. (1996), *Proving Antitrust Damages: Legal and Economic Issues*, Chicago, Antitrust Law Section of the American Bar Association.

Rubin, Jonathan L. (2004), "Cointegration and antitrust: A primer", *Economics Committee Newsletter (American Bar Association)* **40**: 10-19.

Rubinfeld, Daniel L. (1985) "Econometrics in the courtroom", *Columbia Law Review* **85**: 1048-1097.

Rubinfeld, Daniel L. (2000), "Reference guide on multiple regression", *Reference Manual on Scientific Evidence*, 2nd ed., Washington, DC, Federal Judicial Center.

Rubinfeld, Daniel L. and Peter O. Steiner (1983), "Quantitative methods in antitrust litigation", *Law and Contemporary Problems* **46**: 69-144.

Wooldridge, Jeffrey M. (2000), *Introductory Econometrics: A Modern Approach*, Mason, OH, South-Western College Publishing.