

Measuring Benchmark Damages in Antitrust Litigation

Justin McCrary
University of California, Berkeley
NBER

Daniel L. Rubinfeld
University of California, Berkeley
NBER

October 2009

I. Introduction

There are a number of approaches that have been used by experts to evaluate damages in antitrust litigation. The two most common involve the use of *yardsticks* and *benchmarks*.¹ In a typical yardstick approach, one compares prices during the period in which the antitrust violation is believed to have had an effect (the “impact period”) to prices in other markets that are deemed to be reasonably comparable to the market at issue. In contrast, the benchmark approach evaluates prices only in the market at issue, comparing prices in the impact period to available prices before and/or after the alleged period of impact (the “control period”). In this paper, we offer a detailed evaluation of the benchmark approach to damages. We have found the benchmark approach to be the most commonly used damages methodology. To focus the analysis, we assume that the antitrust violation at issue involves price fixing and we also assume that the appropriate legal remedy involves *overcharges* rather than *lost profits*.²

When the time period or periods in which the alleged antitrust behavior affected prices is sufficiently long and the necessary data are available, a standard approach to the evaluation of damages involves two steps. First, one estimates a regression model that “explains” prices using only data for the control period in which the market was unimpeded. Second, the regression model is used to predict but-for prices in the

¹Alternative approaches involve variations on the yardstick approach, such as a comparison of rates of return and/or profit margins across industries.

²For a broad discussion of these alternative measures, see Hovenkamp (2005, Section 17.5(a)).

impact period.³ This “forecasting” (or “before-after”) approach relies heavily on the assumption that the regression specification adequately characterizes the nature of competition in both the impact and control periods.

An alternative approach estimates a regression model for the entire period for which data are available, and evaluates damages by looking at the statistical significance and magnitude of the coefficient on a dummy variable that distinguishes the impact period from the control period. When using this “dummy variable” approach, a secondary issue arises. Should one evaluate damages by assuming a constant price differential through the impact period (as suggested by the coefficient on the dummy variable) or should one allow for non-constant price effects of the alleged conspiracy?

In this paper, we discuss the advantages and disadvantages of each of these methodologies.⁴ In order to focus on the central methodological issues, we begin in Section II by describing the basic regression framework, assuming for simplicity that the antitrust violation involves a conspiracy to fix prices. In Section III, we discuss alternative versions of the dummy variable approach, offering in the process a suggestion as to how to compare the various methodologies. We also describe the forecasting approach and compare it to the dummy variable approach. Section IV presents four propositions that directly compare the dummy variable and forecasting approaches. The propositions tend to support the use of the dummy variable approach over the forecasting approach. However, there are particular advantages associated with the forecasting approach, and these are discussed in Section V. In Section VI, we return to the dummy variable approach, discussing some important model specification issues. In Section VII, we offer an example that illustrates the differences between the various approaches. Section VIII concludes.

II. The Basic Model

Let Y_t denote a measure of the outcome of an alleged conspiracy (e.g., price), X_t a vector of exogenous covariates not causally affected by the conspiracy (e.g., demand and cost variables), and D_t a dummy variable indicating the period of the alleged conspiracy, i.e., the impact or conspiracy period.

We assume that there is data both before and during the alleged conspiracy period. Let T_0 denote the beginning of the alleged conspiracy, so that $t = 1, 2, \dots, T_0$ corresponds to the pre-conspiracy control

³There must be sufficiently variability to allow one to appropriately account for non-collusive variables that might have affected price in the impact period.

⁴These approaches have been discussed in numerous papers. See, for example, Fisher (1980), Rubinfeld and Steiner (1983), Rubinfeld (1985), Rubinfeld (2008), and Higgins and Johnson (2003). See especially White, Marshall and Kennedy (2006); the authors strongly prefer the forecasting approach and are highly critical of the dummy-variable approach.

period and $t = T_0 + 1, T_0 + 2, \dots, T$ to the conspiracy period. We focus on the basic model

$$Y_t = \alpha + \beta' X_t + \theta D_t + \gamma' D_t X_t + \varepsilon_t \tag{1}$$

where ε_t is a mean zero residual that is independent of X_t , D_t , and $D_t X_t$.⁵ This relatively general specification takes into account the possibility that the alleged conspiracy will affect price (the outcome) directly, as given by given by θD_t (e.g., through an increase in price at each point in time in the damage period). However, it also takes into account the possibility that the effect of the conspiracy will be felt through one or more of the covariate demand and cost variables, as given by the term $\gamma' D_t X_t$. This allows the effect of X_t on Y_t to differ between the control period and the impact period. This can be a desirable feature, for example, in an industry and time period where excess profits are being dissipated over time by the process of market entry.

We assume that the covariates are not causally related to the conspiracy. When the covariates *are* caused by the conspiracy, neither the forecasting approach nor the basic dummy variable approach is appropriate if applied using the model in Equation (1). When using the forecasting approach, one bases the forecast on covariates; this approach will be tainted if the covariates are caused by the conspiracy. For the dummy variable approach, the covariates are endogenous, which requires a more complex econometric model.⁶

We focus on what is to be done when the model in (1) is appropriate and there are sufficient data to apply either approach. We therefore assume that there is no causal effect of the conspiracy on the covariates in the conspiracy period. Were this not the case, other methods (e.g., instrumental variables) would be needed to estimate the model, but the discussion that follows would otherwise apply. Note, however, that assuming no *causal* relationship between the covariates and the conspiracy does not rule out the possibility that the covariates are *correlated* with the conspiracy. Indeed, we focus on the case where the covariates have different levels during pre-conspiracy period than during the conspiracy period.⁷

We will also assume, again for simplicity, that the period in which there are antitrust damages and the conspiracy period are identical. Allowing for the two to be different would add some complexity to the specification, but would not change any of the fundamental points to be made in the paper.

⁵That is, we assume that Equation (1) is the data generating process.

⁶The forecasting approach will also require a more complex model that correctly specifies the values of the taint-free covariates in the impact period.

⁷The case where the covariates have equal average levels between the pre-conspiracy period and the conspiracy period is discussed in Higgins and Johnson (2003); see their assumption 4.

The model in Equation (1) is identically equivalent to the following model of counterfactual outcomes:

$$Y_t(1) = \alpha + (\beta' + \gamma')X_t + \theta + u_t \quad (2)$$

$$Y_t(0) = \alpha + \beta'X_t + v_t \quad (3)$$

where $Y_t(1)$ is the outcome during the conspiracy period, and $Y_t(0)$ is the outcome that would occur were there no conspiracy. Under this formulation, $Y_t = D_t Y_t(1) + (1 - D_t) Y_t(0)$ and $\varepsilon_t = D_t u_t + (1 - D_t) v_t$. The formulation in Equations (2) and (3) is useful for understanding some of the conceptual points we raise, below.

III. The Dummy Variable and Forecasting Approaches

To apply the dummy variable approach, we estimate Equation (1) for the entire time period. Continuing to assume that the impact of the covariates on price is unaffected by the conspiracy, θ measures the “average” effect of the conspiracy on price per unit of time. More generally, the impact of the covariates on price may be correlated with the conspiracy, although not directly caused by it. In this more general scenario, the difference between prices with and without the conspiracy per unit of time is given by $\mathbb{E}[Y_t(1) - Y_t(0)|X_t] = \theta + \gamma'X_t$. If Q_t represents output, and all covariates are deemed to be associated with the alleged conspiracy, then aggregate overcharges are a sales-weighted sum of the difference in prices with and without the conspiracy, or

$$\widehat{OC}_1 = \sum_t D_t Q_t \{ \widehat{Y}_t(1) - \widehat{Y}_t(0) \} \equiv \widehat{\theta} \sum_t D_t Q_t + \widehat{\gamma}' \sum_t D_t Q_t X_t \quad (4)$$

where $\widehat{\theta}$ and $\widehat{\gamma}$ are regression estimates of the parameters and $\widehat{Y}_t(1) = \widehat{\alpha} + (\widehat{\beta}' + \widehat{\gamma}')X_t + \widehat{\theta}$ and $\widehat{Y}_t(0) = \widehat{\alpha} + \widehat{\beta}'X_t$.

For some purposes, it may be desirable to impose the restriction that $\gamma = 0$ (i.e., the effect of the covariates on price is the same in the impact period and the control period). In this case, we would obtain a different estimate of damages, given by

$$\widehat{OC}_2 = \widehat{\theta}_S \sum_t D_t Q_t \quad (5)$$

where $\widehat{\theta}_S$ is estimated from the “short” regression that omits the interaction between covariates and the conspiracy period dummy. Our main focus is on \widehat{OC}_1 , but we discuss \widehat{OC}_2 in Section VI.

The forecasting approach estimates the following model:

$$Y_t = \alpha + X_t\beta + e_t \quad (6)$$

for the control period $t = 1, 2, \dots, T_0$. Overcharges are then based on the difference between the predicted outcome (e.g., price) and the actual outcome, during the conspiracy period. If \widehat{Y}_t^f is the out-of-sample (damage period) price forecast, then aggregate overcharges using the forecasting method are given by:

$$\widehat{FC} = \sum_t D_t(Y_t - \widehat{Y}_t^f)Q_t \quad (7)$$

As a general rule in this context, the forecasting approach and the dummy variable approach will generate different overcharge estimates. This is because price gaps are weighted by sales here (cf., Salkever 1976). In the next section, we discuss the issues involved in choosing one approach rather than the other.

IV. When Do the Approaches Differ?

Equation (1) was introduced as a description of the true relationship between the outcome, the conspiracy period, and the covariates. A related interpretation of Equation (1) is as an in-sample decomposition of prices into predicted and unexplained components. Specifically, we have

$$Y_t = \widehat{\alpha} + \widehat{\beta}'X_t + \widehat{\theta}D_t + \widehat{\gamma}'D_tX_t + \widehat{\varepsilon}_t \quad (8)$$

where $\widehat{\varepsilon}_t$ is a fitted price residual which has zero correlation with the regression prediction, in the sample at hand, by construction.

We can use this decomposition to connect the forecasting and the dummy variable approaches. We describe four important results. All proofs are contained in the appendix.

Proposition 1: When quantity varies over the conspiracy period, the forecasting and dummy variable approaches will differ, depending on whether or not quantity is correlated with the fitted residual during the conspiracy period. Formally,

$$\widehat{FC} = \widehat{OC}_1 + \sum_t D_tQ_t\widehat{\varepsilon}_t \quad (9)$$

Proposition 2 (Salkever 1976): When quantity is constant over the conspiracy period, the forecasting and dummy variable approaches yield numerically identical overcharge estimates.

Proposition 3: Assuming that quantity is population uncorrelated with omitted factors during the con-

spiracy period, the dummy variable and forecasting approaches share a probability limit. However, the forecasting approach produces a more variable estimate of damages than the dummy variable approach.

Proposition 1 establishes that the difference between the forecasting and dummy variable approaches hinges on whether the quantity of sales would affect price in the regression model. Classical demand theory would suggest that when price is unexpectedly high (i.e., when ε_t is high) that quantity is likely to be low. Hence, it is reasonable to presume that the forecasting estimate of overcharges will negatively biased relative to the dummy variable estimate of overcharges. However, there is little to be gained by making such a presumption, since it is straightforward to estimate damages using both approaches, and to assess the difference. This is appropriate because classical demand theory cannot be expected to successfully predict the sign of the difference when there is a regression misspecification.

Proposition 2 follows from the formal result in Proposition 1. If quantity is constant, then the difference between the two approaches is proportional to $\sum_t D_t \hat{\varepsilon}_t$, which is zero by construction, as noted by Salkever (1976). Proposition 3 also follows from the formal result in Proposition 1. Unless quantity is constant over the conspiracy period, the term $\sum_t D_t Q_t \hat{\varepsilon}_t$ will add noise to the forecasting estimate that is not present with the dummy variable approach.

V. Why Forecast?

Given that the forecasting approach may under some circumstances generate biased and inefficient damage estimates, when should the approach be used? In this section, we discuss the potential advantages of the forecasting approach.

Advocates of the forecasting approach sometimes use sophisticated model selection procedures to choose the regression model. One motivation for this approach is that the model selection process is based purely on data *prior* to the conspiracy period and will therefore not be corrupted by any effects that the conspiracy might have had on the covariates in the conspiracy period. There is an important benefit associated with this approach, but there is a further drawback. The benefit is that an appropriate model searching methodology minimizes the scope for “overfitting” and “cherrypicking”.⁸ If data *during* the conspiracy period are used to choose the regression model, then there is a risk that the model will produce a biased damages estimate, which is inappropriate. For example, it is always possible to use an in-sample model selection procedure to produce a damages estimate of zero, just by adding a sufficient number of irrelevant covariates so that the model fully explains prices in the conspiracy period (“overfitting”).

⁸This is a benefit emphasized by White et al. (2006).

Choosing the model that is most beneficial to a particular position (“cherrypicking”) can also be a problem. If an expert knows the damages estimate that is beneficial to the client, there is a risk that in-sample model selection could be tantamount to choosing the model that generates a damages estimate that is most preferred. A forecasting approach that is based on an appropriate model selection methodology serves as a good disciplining device.⁹

The drawback of using data prior to the conspiracy period to select the model is that it may be too disciplining. In particular, the use of only pre-conspiracy data prevents the expert from selecting a model using all of his or her knowledge of the economics of the problem. Particularly in dynamic markets, the relationship between covariates and prices may be so rapidly evolving that the pre-conspiracy period will not be an especially good guide to model selection for the conspiracy period. In such a setting, prior knowledge may be of great value and the expert may want to use such knowledge. Suppose, for example, that the market at issue is a highly innovative one in which new technologies are developed on average every two years, and also that the rate of innovation is growing over time. Suppose also that the conspiracy period is four years long. Then, the forecasting approach is likely to underestimate the extent to which innovation would likely have occurred in the but-for world during the conspiracy period.

Weighing these considerations, some would conclude that the model selection procedure associated with forecasting is on balance desirable, especially when damages do not involve dynamic markets. Our analysis suggests, however, that one should consider combining the model selection procedure commonly used in the forecasting approach with the dummy variable approach. To be specific, the model could be chosen based on data prior to the conspiracy period. However, the model could then be estimated using the dummy variable approach and the complete data set that includes the conspiracy period. Propositions 1 through 3 suggest that the dummy variable approach would then work at least as well as the forecasting approach.

VI. Saving Degrees of Freedom in the Dummy Variable Model

An important consideration in the dummy variable model is whether overcharges can be estimated with greater precision by imposing the restriction $\gamma = 0$. While imposing this restriction could in principle either increase or decrease the variability of the overcharges estimate, in most situations imposing the restriction is likely to improve efficiency. To see why it is not necessarily efficient to impose the restriction, observe

⁹A second advantage of the forecasting approach is more technical. If model selection is based purely on the pre-conspiracy period, and one believes that observations before the conspiracy are (approximately) independent of observations during the conspiracy, then the model selection process does not affect the standard errors for the damages estimate.

that the conditional variances of the overcharge estimates including and then excluding $D_t X_t$ are given by

$$V[\widehat{OC}_1] = \left(\sum_t D_t Q_t \right)^2 V[\widehat{\theta}] + \left(\sum_t D_t Q_t X_t \right)' V[\widehat{\gamma}] \left(\sum_t D_t Q_t X_t \right) \quad (10)$$

$$+ \left(\sum_t D_t Q_t X_t \right)' C[\widehat{\gamma}, \widehat{\theta}] \left(\sum_t D_t Q_t \right) + \left(\sum_t D_t Q_t \right) C[\widehat{\theta}, \widehat{\gamma}] \left(\sum_t D_t Q_t X_t \right)$$

$$V[\widehat{OC}_2] = \left(\sum_t D_t Q_t \right)^2 V[\widehat{\theta}_S] \quad (11)$$

where $V[\widehat{\theta}]$ and $V[\widehat{\gamma}]$ represent the variance of the estimator of the dummy variable coefficient and the variance of the estimator of the change in the slope coefficients, respectively. Furthermore, $C[\widehat{\gamma}, \widehat{\theta}]$ is the $k \times 1$ vector of covariances between the slope coefficient estimators and the dummy variable coefficient estimator.

These equations clarify that it is not possible to determine *a priori* whether imposing the restriction that $\gamma = 0$ will improve efficiency. To see why, consider the three terms in Equation (10). The first term, which relates to $V[\widehat{\theta}]$, can be either larger or smaller than $V[\widehat{\theta}_S]$.¹⁰ The second term in Equation (10), summarizing the variability in the estimate of the *change* in the effect of the covariates on price, is strictly positive and typically will be large. The reason is that a precise estimate of the change in the effect of the covariates on price requires sufficient variation in the covariates *both* before and during the alleged conspiracy. Often, covariates that are suspected to have a substantial effect on prices are notably different during the alleged conspiracy period, and there is insufficient variation in the key covariates prior to the alleged conspiracy to obtain a good estimate. The third and final term in Equation (10), pertaining to the covariance between the change in the level of price and the change in the effect of covariates, $C[\widehat{\gamma}, \widehat{\theta}]$, can consist of terms which are all positive, all negative, or a mixture of signs.

With this background, we can now motivate the conclusion that the variance associated with the first approach to overcharges can either be larger or smaller than the variance associated with the second approach. Even if $V[\widehat{\theta}]$ is larger than $V[\widehat{\theta}_S]$, the third and final term in Equation (10), pertaining to covariances, can be negative and large enough in magnitude to offset both the second term in Equation (10) and the difference between $V[\widehat{\theta}]$ and $V[\widehat{\theta}_S]$.¹¹

¹⁰Under some restrictions, such as the Higgins-Johnson restrictions, we have that $V[\widehat{\theta}_S] < V[\widehat{\theta}]$.

¹¹To the best of our knowledge, there is no parametric restriction that guarantees an improvement in precision from imposing the restriction $\gamma = 0$. For example, even in data generating processes where $\gamma = 0$, it can still be more efficient to allow for a change in the effect of the covariates on price. Because of this, we are not aware of any statistical test that would clearly point to whether it was more appropriate to include or exclude the interaction term from the regression, from the point of view of minimizing the variability of the overcharge estimate.

While theory leaves us with a theoretically indeterminate solution, we have found in nearly all the simulations we have examined that the variances associated with the two approaches are quite similar. The most important difference between the two approaches is in their bias, not in their variance. The first approach to estimating overcharges, using the longer specification that allows for an estimate of γ , is more robust in terms of bias because it is consistent even when the effect of covariates on price is different between the pre- and post-conspiracy periods, i.e., when $\gamma \neq 0$. The second approach, using the shorter specification when $\gamma = 0$, is more fragile because it is not consistent with $\gamma \neq 0$. Consequently, our general recommendation is to use the first approach to estimating overcharges, unless one has substantial confidence that the effect of covariates on price is unchanged.

VII. An Example

In this section, we present the results of a simulation study intended to demonstrate the practical relevance of the issues discussed above. We set $T = 100$, with the alleged conspiracy period beginning at $t = 66$. The data generating process is given by equation (1), with ε_t being a sequence of independent and identically distributed (*iid*) normal variables with mean zero and variance one. We also (arbitrarily) choose the following model parameters: $\alpha = 10$, $\beta = 2$ (there is one covariate), $\theta = 4$, and γ is equal to either 0 or 1. With price, P_t , now explicitly the dependent variable, the model is given by:

$$P_t = 10 + 2X_t + 4D_t + \gamma D_t X_t + \varepsilon_t \tag{12}$$

We generate X_t according to

$$X_t = \phi_{1X} + \phi_{2X}t + \phi_{3X}X_{t-1} + e_{tX} \tag{13}$$

where $X_0 = 0$, e_{tX} is distributed *iid* normal with mean 0 and variance 10, and with ε_t and e_{tX} mutually independent. This specification allows for trend and persistence in the covariate, X_t .

The AR(1) specification in Equation (13) does not allow for the possibility that X_t is correlated with the dummy in the conspiracy period. However, because of the trend that is incorporated, in some samples there will be a notable (spurious) correlation. This mimics real world settings in which these methods are used. It will often appear that one or several covariates move differently before the alleged conspiracy period than during, but these apparent differences will potentially be consistent with a complex time series process underlying one or more covariates and with spurious correlation between the covariates and the

outcome variable during the alleged conspiracy period.

We consider two basic specifications of the quantity of sales. The first specification simply holds quantity constant at 100. The second AR(1) specification arises from the model

$$Q_t = \phi_{1Q} + \phi_{2Q}t + \phi_{3Q}Q_{t-1} + \nu P_t + e_{tQ} \tag{14}$$

where ϕ_{1Q} is set to 100, ϕ_{2Q} is set to -0.05, ϕ_{3Q} is set to 0.25, e_{tQ} is distributed *iid* normal with mean 0 and variance 1, and ε_t , e_{tX} , and e_{tQ} are mutually independent. The price elasticity ν is set to either 0 or -1.¹²

The results of this simulation are given in Table 1, which presents estimates of the bias of the overcharges methodologies discussed, using 240,000 replications of data sets of size $T = 100$. We focus on the estimated bias for 4 different models. In all four models, and many others we have examined, differences in variances among the methods are generally quite minor, as compared with differences in bias.

The first model holds quantity constant and generates the covariate X_t according to Equation (13) and price P_t according to Equation (12), with $\gamma = 0$. This corresponds to a setting in which the effect of the covariate on price is the same before and during the alleged conspiracy. The second model is identical to the first, but sets $\gamma = 1$. This implies an increase in the partial correlation between the covariate and price during the conspiracy period, as compared to before. The third model allows quantity to vary according to Equation (13), but there is no partial correlation between quantity and price. The third model continues to use the parameterization $\gamma = 1$. The fourth model is identical to the third, but additionally allows quantity to vary with price by setting the price elasticity, ν , equal to -1.

The first model corresponds to a setting of constant quantity and no change in the effect of covariates on price. Because quantity is constant in this model, Proposition 2 asserts that the first dummy variable approach (\widehat{OC}_1) and the forecasting approach (\widehat{FC}) will lead to identical estimates. This is born out in the simulations. In each of the 240,000 replications, \widehat{OC}_1 and \widehat{FC} are identical. The sample estimates of bias, presented in the first row of the table, are thus also identical. The estimated bias of \$8.30 is statistically insignificant, with a standard error of \$6.70.

Because the effect of covariates on price is constant in this model, the discussion in Section VII indicates that there should be no important difference between the two varieties of the dummy variable approach: the first approach (\widehat{OC}_1) allows the effect of covariates on price to change during the alleged conspiracy and the second approach (\widehat{OC}_2) correctly imposes that the assumption that the effect of covariates on price

¹²All variables are taken to be measured in logarithms.

remain fixed (i.e., that $\gamma = 0$). Consistent with our expectation, the two dummy variable approaches lead to highly similar estimated biases (\$9.70 as compared to \$8.30). The small sample estimate of bias for the second approach is consistent with a hypothesis that the estimator itself is unbiased.

The second model modifies the data generating process to allow the effect of covariates on price to change during the alleged conspiracy period. In this setting, since quantity is constant as in the first model, the first dummy variable approach and the forecasting approach are identical. However, because the effect of the covariates changes during the alleged conspiracy period, the second dummy variable approach (which imposes the constraint that $\gamma = 0$) is inferior to the first approach, which allows γ to be different from zero. Table 1 shows that the bias of the second dummy variable approach is quite large, estimated at roughly \$1,100, or nearly 9 percent of the true parameter.

The third model allows quantity to vary according to Equation (13), as discussed, but does not allow any predictable relationship between quantity and price. This model retains the assumption that the effect of the covariate on prices differs before and during the alleged conspiracy period. Because quantity and price have a zero partial correlation, both the first dummy variable approach and the forecasting approach generate consistent estimates (Proposition 3).¹³ The sample biases are very similar, but not identical. (The estimated sample bias differs in the second decimal place, but the differences are masked by rounding.) However, because the effect of the covariates changes over time, the second dummy variable approach is not consistent. The estimated bias, approximately \$1,450, is just below 9 percent of the true parameter.

The final model allows for quantity and price to be related, with a price elasticity of -1. In this data generating process, this implies a negative partial correlation between the reduced form residuals and quantity. From equation (9), we know that this will lead the forecasting method to exhibit negative bias, and the simulation results bear this out. While the first dummy variable approach continues to perform well in this setting, the forecasting method has a bias nearly as large as the second dummy variable approach. The estimated bias of the forecasting method, -\$341.10, represents roughly 2 percent of the true parameter. In this setting, because the effect of the covariate changes over time, the second dummy variable model performs the least well of the three methods considered, with an estimated bias of \$1,355.40, or over 9 percent of the true parameter.

¹³The standard deviation of the forecasting approach estimate is less than 1 percent larger than that of the the first dummy variable approach (results not shown).

VIII. Conclusion

In this paper, we have reviewed two major approaches to the estimation of overcharges: the dummy variable approach and the forecasting approach. The dummy variable approach is based on a regression model that explains price before and during the alleged conspiracy period. There are two leading variants of the dummy variable approach, corresponding to whether the effects of covariates are allowed to differ before and during the alleged conspiracy period, or are instead imposed to be the same throughout. We consider both of these variants. The forecasting approach formulates a model for price before the alleged conspiracy period and then compares price forecasts with actual prices. For both the dummy variable approach and the forecasting approach, a quantity-weighted difference in prices is used to estimate overcharges.

We show that the first dummy variable approach, in which the effects of covariates are allowed to differ over time, is numerically equivalent to the forecasting approach when quantity is constant. When quantity varies over time, but not in a manner related to unobserved determinants of price, then both the forecasting and the first dummy variable approaches generate consistent estimates of overcharges, but the forecasting approach is less efficient (has more variability) than the first dummy variable approach. This raises a question as to the merits of the forecasting approach.

We also show that there is some justification for the second dummy variable approach. When the effects of covariates on price are indeed constant over time, then the second dummy variable approach can have less variability than the first dummy variable approach. However, this is not guaranteed. Moreover, in simulation results, we do not find important differences in the variability of the two dummy variable approaches. Moreover, there is a substantial problem associated with the second dummy variable approach. If the restriction that the effects of covariates are constant is, in fact, false, then the method is biased, possibly by a large amount.

Overall, our discussion points to an important role for the first dummy variable approach. The primary drawback of the first dummy variable approach is the possibility that analysts will “overfit” the regression model, including a great number of covariates that do not belong in the regression model. This can lead to imprecise overcharges estimates, and perhaps even spurious overcharges estimates if inappropriate covariates are included. To ameliorate these problems, we suggest that further consideration be given to the use of a model selection procedure (such as that currently used in the forecasting approach) in conjunction with the first dummy variable approach.

Appendix

Proof of Proposition 1: We begin by noting that $\widehat{Y}_t^f = \widehat{\alpha} + \widehat{\beta}X_t$, where $\widehat{\alpha}$ and $\widehat{\beta}$ are identical to the dummy variable regression estimates. Then, using Equation (8), we have

$$\widehat{FC} = \sum_t D_t Q_t \left(\widehat{\theta} D_t + \widehat{\gamma}' D_t X_t + \widehat{\varepsilon}_t \right) = \widehat{\theta} \sum_t D_t Q_t + \widehat{\gamma}' \sum_t D_t X_t Q_t + \sum_t D_t Q_t \widehat{\varepsilon}_t \quad (\text{A.1})$$

$$= \widehat{OC}_1 + \sum_t D_t Q_t \widehat{\varepsilon}_t \quad (\text{A.2})$$

□

Proof of Proposition 2: Applying Proposition 1, note that when quantity is constant, say $Q_t = \overline{Q}$, we have $\sum_t D_t Q_t \widehat{\varepsilon}_t = \overline{Q} \sum_t D_t \widehat{\varepsilon}_t = 0$ by the orthogonality of fitted residuals and covariates. □

Proof of Proposition 3: Applying Proposition 1, the dummy variable approach and the forecasting approach share a probability limit whenever

$$\sum_t D_t Q_t \widehat{\varepsilon}_t = \sum_t D_t Q_t \varepsilon_t - (\widehat{\alpha} - \alpha) - (\widehat{\beta} - \beta)' X_t - (\widehat{\theta} - \theta) - (\widehat{\gamma} - \gamma)' X_t \quad (\text{A.3})$$

converges in probability to zero. Since each of $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\theta}$, and $\widehat{\gamma}$ is consistent, this condition holds whenever $\sum_t D_t Q_t \varepsilon_t$ converges in probability to zero. Note then that a sufficient condition is $0 = E[D_t Q_t \varepsilon_t] = E[Q_t \varepsilon_t | D_t = 1] / P(D_t = 1)$, which holds whenever Q_t and ε_t are uncorrelated in the population during the alleged conspiracy (since ε_t is mean zero). Finally, note that the variance of $\sum_t D_t Q_t \widehat{\varepsilon}_t$ will be strictly positive unless Q_t fails to predict prices in each sample. □

References

- Fisher, Franklin, “Multiple Regression in Legal Proceedings,” *Columbia Law Review*, 1980, pp. 702–736.
- Higgins, Richard and Paul Johnson, “The Mean Effect of Structural Change in the Dependent Variable is Accurately Measured by the Intercept Change Alone,” *Economics Letters*, 2003, 80, 255–259.
- Hovenkamp, Herbert, *Federal Antitrust Policy: The Law of Competition and Its Practice*, St. Paul, MN: Thomson-West, 2005.
- Rubinfeld, Daniel L., “Econometrics in the Courtroom,” *Columbia Law Review*, June 1985, pp. 1048–1097.
- , “Quantitative Methods in Antitrust,” in Dale Collins, ed., *Issues in Competition Policy*, ABA Antitrust Section, 2008, chapter 30.
- and Peter O. Steiner, “Quantitative Methods in Antitrust Litigation,” *Law and Contemporary Problems*, 1983, 26, 69–141.
- Salkever, David, “The Use of Dummy Variables to Compute Predictions, Prediction Errors, and Confidence Intervals,” *Journal of Econometrics*, 1976, 4 (4), 393–397.
- White, Halbert, Robert Marshall, and Pauline Kennedy, “The Measurement of Economic Damages in Antitrust Civil Litigation,” *ABA Antitrust Section, Economic Committee Newsletter*, Spring 2006, 6 (1).

Table 1. Bias of Overcharge Methodologies

| Description of Data Generating Process | Estimation Approach | | |
|--|---------------------|------------------|-----------------|
| | OC ₁ | OC ₂ | FC |
| <u>Model 1</u> Q _t constant; $\gamma=0$ | 8.3 (6.7) | 9.7 (6.4) | 8.3 (6.7) |
| <u>Model 2</u> Q _t constant; $\gamma=1$ | 8.3 (6.7) | 1,134.7 (6.5) | 8.3 (6.7) |
| <u>Model 3</u> Q _t an AR(1) unrelated to prices (demand elasticity is 0); $\gamma=1$ | 10.6 (8.5) | 1,445.7 (8.3) | 10.6 (8.5) |
| <u>Model 4</u> Q _t an AR(1) related to prices (demand elasticity is -1); $\gamma=1$ | -14.7 (7.4) | 1,355.4 (7.2) | -341.1 (7.4) |

Notes: Table presents simulation estimates of bias and standard deviation of three approaches to estimating overcharges. Estimates based on 160,000 simulated data sets, each with 100 observations. See text for details.