Practitioner's Corner

Justin McCrary* and Daniel L. Rubinfeld

Measuring Benchmark Damages in Antitrust Litigation

Abstract: We compare the two dominant approaches to estimation of benchmark damages in antitrust litigation, the forecasting approach and the dummy variable approach. We give conditions under which the two approaches are equivalent and present the results of a small simulation study.

Keywords: antitrust; damages estimation; law and economics.

***Corresponding author: Justin McCrary,** Professor of Law, U.C. Berkeley; and Faculty Research Associate, NBER, E-mail: jmccrary@law.berkeley.edu
Daniel L. Rubinfeld: Robert L. Bridges Professor of Law and Professor of Economics Emeritus, U.C. Berkeley; Professor of Law, NYU; and Faculty Research Associate, NBER

1 Introduction

The quantitative evaluation of monetary damages from alleged antitrust violations occupies a central place in antitrust litigation. The two most common approaches to evaluating damages 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. We also assume that the appropriate legal remedy involves *overcharges* rather than *lost profits*.[2] Our particular focus is a comparison of the forecasting and dummy variable approaches, which we define in Section 3. Our analysis underscores that these competing approaches to computing benchmark damage estimates often yield similar estimates, despite seemingly different implementation schemes.

We are not the first to consider the advantages and disadvantages of each of these methodologies.[3] However, we believe many of the results comparing the forecasting and dummy variable approaches, while straightforward, are underappreciated. In order to focus on the central methodological issues, we begin in Section 2 by describing the basic regression framework and defining the object of interest. In Section 3, 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 4 presents three 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 5. In Section 6, we return to the dummy variable approach, discussing some important model specification issues. In Section 7, we offer an example that illustrates the differences between the various approaches. Section 8 concludes.

2 The Basic Model

Let Y_t denote the price of the product in question, X_t a vector of exogenous covariates (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.

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

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

3 See, for example, Salkever (1976), Fisher (1980), Rubinfeld and Steiner (1983), Rubinfeld (1985, 2008), and Higgins and Johnson (2003). See especially White, Marshall, and Kennedy (2006); those authors strongly prefer the forecasting approach and are highly critical of the dummy-variable approach.

We assume that there are data both before and during the alleged conspiracy period. Let T_0 denote the last period prior to the beginning of the alleged conspiracy, so that t=1, 2, ..., T_0 corresponds to the pre-conspiracy control period and $t=T_0+1, T_0+2, \ldots, T_0+T_1$ to the conspiracy period. Define the total number of periods as $T=T_0+T_1$. We assume throughout that price is generated according to

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

where ε_t is a mean zero residual that is uncorrelated with X_t, D_t, and $D_t X_t$, i.e., $0=\mathbf{E}[\varepsilon_t]=\mathbf{E}[\varepsilon_t X_t]=\mathbf{E}[\varepsilon_t D_t]=\mathbf{E}[\varepsilon_t D_t X_t]$. This relatively general specification takes into account the possibility that the alleged conspiracy will affect price directly, as given by δD_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 covariates, 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 through market entry.

We assume that the conspiracy does not cause changes to the covariates. 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). These considerations importantly affect the choice of covariates. Note, however, that assuming the conspiracy does not *cause* the covariates to change 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 the pre-conspiracy period than during the conspiracy period.[4]

We focus on what is to be done when the model in (1) is appropriate and there are sufficient data to apply either approach.[5] For simplicity, we assume 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.

The model in Equation (1) can be thought of as a model of counterfactual outcomes, namely

$$Y_t(1)=\alpha+\delta+(\beta+\gamma)' X_t+u_t \tag{2}$$

$$Y_t(0)=\alpha+\beta' X_t+v_t \tag{3}$$

where $Y_t(1)$ is price under conspiracy conditions, $Y_t(0)$ is price under non-conspiracy conditions, and u_t and v_t are mean zero residuals that are uncorrelated with X_t (Rubin 1974; Imbens 2004). We additionally impose the assumptions that u_t and v_t are uncorrelated with D_t, which then implies our earlier orthogonality assumption $\mathbf{E}[\varepsilon_t D_t]=0$.[6]

Under this formulation, observed price is $Y_t=D_t Y_t(1)+(1-D_t)Y_t(0)$ and the price residual from Equation (1) is $\varepsilon_t=D_t u_t+(1-D_t)v_t=v_t+D_t(u_t-v_t)$. The formulation in Equations (2) and (3) is useful for understanding some of the conceptual points we raise below.

A damages award in litigation is typically based on estimated aggregate overcharges, as measured here by the difference in revenues under conspiracy conditions and under non-conspiracy conditions. To simplify, we assume that costs are unaffected by the conspiracy. To define this estimand explicitly, denote quantity as Q_t. Note that quantity will not be included as a covariate, because of controversies over exogeneity – the set of covariates is restricted to be those variables exogenous to the conspiracy. Although quantity is excluded from the regression, it may nonetheless be correlated with some of the covariates and with price. Indeed, this may occur even if the conspiracy did not cause changes to quantity. Average overcharges are aggregate overcharges relative to the number of time periods. Multiple consistent estimators are available for average overcharges, and we focus on the issues associated with the estimation of the relevant parameter or parameters. The population parameter corresponding to average overcharges is

$$OC^*=\mathbf{E}[D_t Q_t\{Y_t(1)-Y_t(0)\}] \tag{4}$$

which can be thought of as the product of the true average overcharge during the conspiracy period, or $\mathbf{E}[(Y_t(1)-Y_t(0))Q_t|D_t=1]$, and the probability that a sampled period is during the conspiracy, or $\mathbf{E}[D_t]$.[7]

4 The 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.

5 There may be too few observations under conspiracy conditions to estimate the parameters $\alpha+\delta$ and $\beta+\gamma$ using the conspiracy period alone.

6 Note that this assumption is justified if the decision to initiate and cease a conspiracy is based largely on factors captured by the covariates, X_t, or if it is based on idiosyncratic factors that are unrelated to the gains from conspiracy. It is *not* justified if the decision to initiate or cease a conspiracy is based on unmeasured factors affecting but-for prices, i.e., v_t, or on the gains to conspiracy, i.e., u_t-v_t.

7 In some applications, price will be modeled in logs, in which case the object of interest may be redefined as $\mathbf{E}[D_t Q_t\{\ln Y_t(1)-\ln Y_t(0)\}Y_t(0)]$ or $\mathbf{E}[D_t Q_t\{\ln Y_t(1)-\ln Y_t(0)\}Y_t(1)]$, for example.

3 The Dummy Variable and Forecasting Approaches

One standard approach to the evaluation of overcharges 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?

When the time period or periods in which the alleged antitrust behavior affected prices is sufficiently long and the necessary data are available, a second standard approach to the evaluation of overcharges is a two-step procedure. 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 impact period.[8] This approach is conventionally referred to as a forecasting (or "before-after") approach.

To apply the dummy variable approach, we estimate Equation (1) for the entire time period. The estimation may or may not use quantity weights.[9] When the estimation does use quantity weights, we assume that the model is correctly specified in the sense that the earlier orthogonality conditions are modified to be $0=\mathbf{E}[\varepsilon_t Q_t]=\mathbf{E}[\varepsilon_t X_t Q_t]=\mathbf{E}[\varepsilon_t D_t Q_t]=\mathbf{E}[\varepsilon_t D_t X_t Q_t]$.

Continuing to assume that the impact of the covariates on price is unaffected by the conspiracy, δ measures the temporally constant 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. Estimates of average overcharges are a quantity-weighted average of the difference in prices with and without the conspiracy, or

$$\widehat{OC}_1=\frac{1}{T}\sum_{t=1}^{T}D_t Q_t\{\hat{Y}_t(1)-\hat{Y}_t(0)\} \tag{5}$$

where $\hat{Y}_t(1)=\hat{\alpha}+(\hat{\beta}'+\hat{\gamma}')X_t+\hat{\delta}$ and $\hat{Y}_t(0)=\hat{\alpha}+\hat{\beta}'X_t$ are the regression fitted values during the conspiracy and non-conspiracy periods, respectively.

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

$$\widehat{OC}_2=\frac{1}{T}\sum_{t=1}^{T}D_t Q_t\{\tilde{Y}_t(1)-\tilde{Y}_t(0)\} \tag{6}$$

where $\tilde{Y}_t(1)=\tilde{\alpha}+\tilde{\delta}+\tilde{\beta}'X_t$ and $\tilde{Y}_t(0)=\tilde{\alpha}+\tilde{\beta}'X_t$ are the fitted values corresponding to the conspiracy and non-conspiracy periods, respectively, where all coefficients are estimated subject to the restriction that the interaction between covariates and the conspiracy period dummy is properly excluded from the regression, i.e., that $\gamma=0$. Our main focus is on $\widehat{OC}_1$, but we discuss $\widehat{OC}_2$ in Section 6. Note that both of these estimators can be rewritten in terms of sample means and estimated regression coefficients, i.e., $\widehat{OC}_1=\hat{\pi}_1\hat{\delta}+\hat{\pi}_X'\hat{\gamma}$ and $\widehat{OC}_2=\hat{\pi}_1\tilde{\delta}$, where $\hat{\pi}_1=\frac{1}{T}\sum_{t=1}^{T}D_tQ_t$ and $\hat{\pi}_X=\frac{1}{T}\sum_{t=1}^{T}D_tQ_tX_t$.[10]

Because the regression model in Equation (1) interacts the covariates with the dummy for conspiracy, the coefficients $\hat{\alpha}$ and $\hat{\beta}$ can be obtained equivalently by running a regression of price on covariates during the non-conspiracy period alone. The fitted values $\hat{Y}_t(0)=\hat{\alpha}+\hat{\beta}'X_t$ are then the in-sample predictions for periods t with $D_t=0$ and the out-of-sample forecasts for periods t with $D_t=1$. The forecasting approach to estimating average overcharges takes the quantity-weighted difference between actual and forecasted prices, or

$$\widehat{FC}=\frac{1}{T}\sum_{t=1}^{T}D_t Q_t\{Y_t-\hat{Y}_t(0)\} \tag{7}$$

As with the dummy variable approach, the forecasting approach estimate can be rewritten in terms of sample means and the estimated regression coefficients, i.e., $\widehat{FC}=\hat{\pi}_Y-\hat{\pi}_1\hat{\alpha}-\hat{\pi}_X'\hat{\beta}$, where $\hat{\pi}_Y=\frac{1}{T}\sum_{t=1}^{T}D_tQ_tY_t$. In the next section, we discuss the issues involved in choosing between the two approaches.

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

9 A variety of considerations are involved in the decision of whether to use quantity weights, including data quality, heteroskedasticity, efficiency, strong trends in quantity (particularly for narrowly defined products), and robustness, among others. We focus on the case where weights are not used, but note the implications of using weights where relevant.

10 As noted by Wooldridge (2002, section 18.3.1), covariates can be de-meaned prior to estimation without changing the estimated regression coefficients except for the constant and with essentially negligible effects on the standard errors. This means that we can ensure that $\hat{\pi}_X$ is by construction zero, which is computationally convenient. In that case, the coefficient on the dummy variable needs only to be scaled up by $\hat{\pi}_1$ in order to obtain $\widehat{OC}_1$.

4 When Do These Two 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

$$Y_t = \hat{\alpha} + \hat{\beta}' X_t + \hat{\delta} D_t + \hat{\gamma}' D_t X_t + \hat{\varepsilon}_t \tag{8}$$

where $\hat{\varepsilon}_t$ is a fitted price residual which in the sample has zero correlation with the covariates by construction. We can use this decomposition to connect the forecasting and the dummy variable approaches.

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

$$\widehat{FC} = \widehat{OC}_1 + \frac{1}{T}\sum_{t=1}^{T} D_t Q_t \hat{\varepsilon}_t$$

The Lemma 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, one presumption is that the forecasting estimate of overcharges will be negatively biased relative to the dummy variable estimate of overcharges. However, note that Equation (1) is not typically interpreted as an inverse demand equation, but rather a reduced form model for price. Consequently, there may be no economic basis for the assumption that quantity and unexplained price deviations are negatively related.

Proposition 1: The forecasting and dummy variable approaches yield numerically identical overcharge estimates if either (a) quantity is constant over the conspiracy period, or (b) the regression in Equation (1) is quantity weighted.

Proposition 1 follows directly from the Lemma. If quantity is constant, then the difference between the two approaches is proportional to $\frac{1}{T}\sum_{t=1}^{T} D_t \hat{\varepsilon}_t$, which is zero by the sample orthogonality conditions of regression. If the regression is quantity weighted, then $\frac{1}{T}\sum_{t=1}^{T} D_t Q_t \hat{\varepsilon}_t$ is zero since that is then precisely the sample orthonality condition for the weighted regression. Proposition 1 means that, despite often being a major point of disagreement between opposing expert witnesses, there is no distinction between the dummy variable and forecasting method when the two methods use the same covariates and quantity weights. Proposition 1 does not, however, indicate the relationship between the approaches when the regression models are not weighted by quantity, as they often will not be, and it does not indicate whether either approach measures the parameter of interest. These considerations are covered by the next proposition, which gives a variety of sufficient conditions for the dummy variable and forecasting approaches to be consistent for average overcharges. Three of these sufficient conditions are more detailed than the others, and we discuss them briefly before stating the proposition. These three assumptions, which pertain to the covariance during the conspiracy period between quantity and unmeasured factors influencing actual or but-for prices, are given by:

Assumption 1: $\mathbf{C}[u_t, Q_t \mid D_t = 1] = \mathbf{C}[v_t, Q_t \mid D_t = 1] = 0.$

Assumption 2: $\mathbf{C}[u_t, Q_t \mid D_t = 1] = \mathbf{C}[v_t, Q_t \mid D_t = 1].$

Assumption 1′: $\mathbf{C}[v_t, Q_t \mid D_t = 1] = 0.$

Since u_t corresponds to $Y_t(1)$ and v_t corresponds to $Y_t(0)$, Assumption 1 asserts zero covariance between quantity and unmeasured factors affecting but-for and actual prices during the conspiracy. Assumption 2 asserts that the covariance during the conspiracy between quantity and unmeasured factors affecting but-for price is equal to the covariance during the conspiracy between quantity and unmeasured factors affecting actual price. That is, the covariance does not have to be zero, but it must be the same for actual and but-for prices.[11] Assumption 1′ implies that there is zero covariance between quantity and the unmeasured factors affecting but-for prices during the conspiracy. This is a weaker version of Assumption 1, in the mathematical sense of being implied by it. Assumption 2 is not implied by Assumption 1′, but is implied by Assumption 1. After stating our main proposition regarding consistency, we discuss whether there is a sense in which Assumption 1′ is stronger than Assumption 2.

We now state our main proposition regarding consistency.

11 As noted by the editor, Assumption 2 is also implied the somewhat stronger restriction that $\mathbf{C}[Y_t(1) - Y_t(0), Q_t \mid X_t, D_t = 1] = 0$. In words, this restriction is that the covariance between quantity and price is the same during the damages period as long as the covariates are the same, regardless of whether it is with or without the conspiracy.

Proposition 2: If $(Y_t, X_t, D_t, Q_t)'$ is a vector ergodic stationary process with existence of sufficient moments, then both the forecasting and dummy variable approaches are consistent for OC^* if either (a) quantity is constant over the conspiracy period or (b) the regression in Equation (1) is quantity weighted. In addition, the forecasting approach is consistent if Assumption 1 or Assumption 1′ is met, and the dummy variable approach is consistent if Assumption 1 or Assumption 2 is met.

The proof of Proposition 2 is given in the appendix.[12] The primary conclusion of the Proposition is that both the forecasting and the dummy variable approaches can be consistent for average overcharges, but under slightly different conditions in general.[13] Whether Assumption 2 or Assumption 1′ is more plausible is a matter of judgment, as neither implies the other. Assumption 1 asserts that quantity is uncorrelated with unmeasured factors affecting both actual and but-for prices during the conspiracy. This is plainly a strong assumption, in the sense that it implies both Assumptions 2 and 1′.

Assumption 2 is notably weaker. This assumption allows for quantity to be correlated with unmeasured factors affecting actual and but-for price during the conspiracy; it is justified if adjustments to quantity are due to observable factors controlled for in the regression and idiosyncratic variation. Stated differently, Assumption 2 is justified if adjustments to quantity ignore the unobservable price improvements available from conspiracy, i.e., u_t-v_t.

Turning to Assumption 1′, we find it hard to justify the assumption that there is zero covariance during the conspiracy period between quantity and the unmeasured factors affecting the but-for price, without also being willing to assume that there is zero covariance between quantity and actual price. That is, it seems to us to be hard to justify Assumption 1′ without appealing to Assumption 1. Consequently, in our judgment, there is a sense in which Assumption 1′ is stronger than Assumption 2, despite the fact that neither assumption implies the other, strictly speaking. Note that there is a natural restriction which implies that unmeasured factors affecting price would be the same under conspiracy and non-conspiracy conditions: $v_t=u_t=\varepsilon_t$. In words, this would imply that treatment effects would depend on covariates at best, but not residuals. In this case, Assumption 2 is satisfied automatically, but Assumption 1′ may not be. Proposition 2 suggests, therefore, that the dummy variable approach is likely to be more robust than the forecasting approach.

5 Should One Forecast?

In this section, we discuss the potential advantages and disadvantages 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".[14] 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").

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.

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

12 Informally, an ergodic stationary process is a process that will not change its properties over time and whose properties can be deduced from a sufficiently long sample of the process.

13 While it is not our focus in this paper, we note that if one found Assumption 1 to be justified, then there are two consistent estimators for average overcharges, in which case a more efficient estimator can be obtained by combining the two estimators. For example, the linear combination $(\omega_1\widehat{OC}_1+\omega_2\widehat{FC})/(\omega_1+\omega_2)$ is also consistent for the average overcharge and has asymptotic variance of $1/(\omega_1+\omega_2)$ where $\omega_1\equiv 1/(V_{OC_1}-c)$ with V_{OC_1} the asymptotic variance of the dummy variable estimate and c the asymptotic covariance between it and the forecasting estimate, and $\omega_2\equiv 1/(V_{FC}-c)$ with V_{FC} the asymptotic variance of the forecasting estimate. On the other hand, obtaining good estimates of V_{FC} and V_{OC_1} is challenging and this may limit the practicality of this approach.

14 This is a benefit emphasized by White, Marshall, and Kennedy (2006).

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 2 years, and also that the rate of innovation is growing over time. Suppose also that the conspiracy period is 4 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. A point which is perhaps underappreciated, however, is that one could of course choose the model based only on data prior to conspiracy conditions, as with forecasting, and then having chosen a model, estimate the parameters of the model using the dummy variable approach.

6 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 $\gamma=0$. Imposing this restriction could in principle either increase or decrease the variability of the overcharges estimate, as we now explain.

Conditional on $\hat{\pi}_1$ and $\hat{\pi}_X$, the variances of the the two dummy variable approaches are

$$\mathbf{V}[\widehat{OC}_1|\hat{\pi}_1,\hat{\pi}_X]=\hat{\pi}_1^2\mathbf{V}[\hat{\delta}|\hat{\pi}_1,\hat{\pi}_X]+\hat{\pi}_X'\mathbf{V}[\hat{\gamma}|\hat{\pi}_1,\hat{\pi}_X]\hat{\pi}_X+2\hat{\pi}_X'C[\hat{\gamma},\hat{\delta}|\hat{\pi}_1,\hat{\pi}_X]\hat{\pi}_1 \tag{9}$$

$$\mathbf{V}[\widehat{OC}_2|\hat{\pi}_1,\hat{\pi}_X]=\hat{\pi}_1^2\mathbf{V}[\tilde{\delta}|\hat{\pi}_1,\hat{\pi}_X] \tag{10}$$

Recall that $\hat{\delta}$ is the dummy variable coefficient from the regression including interactions between the dummy variable and the covariates, that $\hat{\gamma}$ is the vector of coefficients on those interactions, and that $\tilde{\delta}$ is the dummy variable coefficient from the regression that excludes the interactions. Using the fact that $\hat{\delta}$ and $\tilde{\delta}$ are connected as $\tilde{\delta}=\hat{\delta}+\hat{\gamma}'\hat{\eta}$, where the kth element of $\hat{\eta}$ is the coefficient on D_t in a regression of the kth element of D_tX_t on a constant, D_t, and X_t, where X_t has K elements, we can write $\widehat{OC}_2=\hat{\pi}_1\hat{\delta}+\hat{\pi}_1\hat{\eta}'\hat{\gamma}$. Consequently, these two approaches differ in the implicit adjustment to $\hat{\gamma}$; $\widehat{OC}_1$ uses $\hat{\pi}_X$ and $\widehat{OC}_2$ uses $\hat{\pi}_1\hat{\eta}$.

Comparing equations (9) and (10), one can see 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 (9). The first term is proportional to $\mathbf{V}[\hat{\delta}|\hat{\pi}_1,\hat{\pi}_X]$. However, this can be either larger or smaller than $\mathbf{V}[\tilde{\delta}|\hat{\pi}_1,\hat{\pi}_X]$.[15] The second term in Equation (9), 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 final term in Equation (9), pertaining to the covariance between the change in the level of price and the change in the effect of covariates, 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 $\mathbf{V}[\hat{\delta}|\hat{\pi}_1,\hat{\pi}_X]$ is larger than $\mathbf{V}[\tilde{\delta}|\hat{\pi}_1,\hat{\pi}_X]$ (as for example when $\gamma=0$), the third and final term in Equation (9) can be negative and large in magnitude.[16] This leads to indeterminacy in the relative magnitudes of the conditional variances of the two approaches, and this indeterminacy carries over to the case of unconditional variances.

On the other hand, it is often possible to estimate the model using both the first and second dummary variable approaches. Assuming the economist is willing to impose the additional assumptions needed for inference (e.g., existence and finiteness of fourth moments or the

15 Higgins and Johnson (2003) consider some restrictions that guarantee $\mathbf{V}[\tilde{\delta}|\hat{\pi}_1,\hat{\pi}_X]<\mathbf{V}[\hat{\delta}|\hat{\pi}_1,\hat{\pi}_X]$, chief among these being $\gamma=0$. This is an old result; see for example Kmenta (2000, section 11–2).

16 To the best of our knowledge, there is no parametric restriction that guarantees an improvement in precision of estimated average overcharges 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.

variance structure of the error term), it may be possible to get a sense in the sample of which estimator is more variable. However, probably the strongest reasons to consider both the first and second dummy variable approaches are prior information, specific data settings, and robustness. The economist might be have a strong prior view that one or more covariates have the same partial effect on price before and during the conspiracy (i.e., an element of γ is zero) and suspect that imposing the restriction will improve efficiency; the economist may not have enough observations to estimate the effect of variables believed to be important in both the before and during periods; or the economist might have enough data to do so, but be worried about the robustness of a model that is deemed close to overfitting.

7 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 roughly two-thirds of the way through the sample, i.e., $D_t=1(t>t^*)$, where for each sample t^* is a single draw from the binomial distribution with parameters T and 2/3. This implies that $\mathbf{E}[D_t]=2/3$. The covariate X_t is generated according to

$$X_t=1-0.015t+0.25X_{t-1}+e_t \tag{11}$$

where we initialize X_t as $X_0=0$ and e_t is distributed independently and identically (*iid*) standard normal. This specification allows for trend and persistence in the covariate. The AR(1) with trend model can exhibit notable (spurious) correlations with the dummy for the conspiracy period. That is, the conspiracy does not cause changes to X_t but may be associated with it. 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 simulate prices according to Equations (2) and (3), with u_t and v_t independent heteroskedastic error terms generated as $u_t=Z_t\tilde{u}_t$ and $v_t=Z_t\tilde{v}_t$, where $\tilde{u}_t$ and $\tilde{v}_t$ are distributed *iid* bivariate normal with means of 0, variances 10, and correlation of zero, and Z_t is distributed *iid* standard normal and independent of $\tilde{u}_t$, $\tilde{v}_t$, and e_t. We also (arbitrarily) choose the following model parameters: α=10, β=2 (there is one covariate), δ=4, and γ is equal to either 0 or 1. In summary, the model for price is given by:

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

where $\varepsilon_t=D_t u_t+(1-D_t)v_t$. In a typical simulated sample from this data generating process, a regression of Y_t on X_t, D_t, and D_tX_t yields an R^2 of about one-third, which is typical of this context.

As emphasized by Proposition 2, the relationship between quantity of sales and unmeasured factors affecting actual and but-for prices during the conspiracy period relates in highly specific ways to the consistency of the forecasting and dummy variable approaches. Consequently, we consider several different specifications for quantity. Our baseline specification holds quantity constant at 150, where quantity is measured in thousands of units sold. We also consider more complicated specifications based on an AR(1) model with an error term that depends on unmeasured factors affecting actual and but-for prices:

$$Q_t=75+0.5Q_{t-1}+v_u u_t+v_v v_t+\epsilon_t \tag{13}$$

with Q_t initialized to $Q_0=0$. We let ϵ_t be distributed *iid* standard normal, and ε_t, e_t, and ϵ_t be mutually independent.

The coefficients v_u and v_v in Equation (13) control whether Assumptions 1, 2, or 1′ are met, or whether none of them are met. The coefficient γ controls whether interaction terms are needed in the regression model. Table 1 shows the models for Q_t that we consider, the configurations of the v_u, v_v, and γ parameters, and the implications of these choices for the validity of Assumptions 1, 2, and 1′ and for estimator consistency.

Model 1 generates price Y_t according to Equation (12), with γ=0. This corresponds to a setting in which the effect of the covariate on price is the same before and during the alleged conspiracy. Quantity is constant at 150,000 units. Model 2 is identical to the first, but sets γ=1. This implies an increase in the partial correlation between the covariate and price during the conspiracy period, as compared to before. Models 3 through 6 allow quantity to vary according to Equation (13), but the parameters v_u and v_v vary from being equal and zero (Model 3), to being equal and non-zero (Model 4), to being different from one another (Models 5 and 6). In each of the six models, the covariate X_t is simulated according to Equation (11).

Note that Models 2 through 6 involve changes to the data generating process for Q_t, but not for Y_t. Thinking of the structure of the two approaches, we recognize that changes to the data generating process for Q_t affect the forecasting approach in a somewhat more direct way than they do the dummy variable approach. That is, the dummy variable estimate is a function of $\hat{\alpha}$, $\hat{\beta}$, $\hat{\delta}$, $\hat{\gamma}$, $\hat{\pi}_1$, and $\hat{\pi}_X$. As emphasized by the Lemma, however, the forecasting approach is additionally affected by $T^{-1}\sum_{t=1}^{T} D_t Q_t \hat{\varepsilon}_t$, and this means that

Table 1 Overview of Simulation Experiments.

Model	Model for Q_t	ν_u	ν_v	Valid assumptions	γ	Consistent estimators
1	Constant	–	–	1 and 2	0	$\widehat{FC}$, $\widehat{OC}_1$, $\widehat{OC}_2$
2	Constant	–	–	1 and 2	1	$\widehat{FC}$, $\widehat{OC}_1$
3	AR(1)	0	0	1 and 2	1	$\widehat{FC}$, $\widehat{OC}_1$
4	AR(1)	3	3	2	1	$\widehat{OC}_1$
5	AR(1)	3	0	1′	1	$\widehat{FC}$
6	AR(1)	0	3	None of the above	1	None of the above

Note: The table describes the simulation experiments we conduct. Parameters ν_u and ν_v correspond to the model for Q_t, and the parameter γ corresponds to the model for Y_t.

Table 2 Simulation Estimates of Mean and Standard Deviation of Estimators for Average Overcharge.

Model	$\widehat{OC}_1$	$\widehat{OC}_2$	$\widehat{FC}$	OC^*
1	199.89	199.94	199.89	200.00
	(47.88)	(46.58)	(47.88)	
2	183.36	198.97	183.36	183.49
	(49.27)	(50.44)	(49.27)	
3	183.47	198.96	183.47	183.49
	(49.53)	(50.67)	(49.54)	
4	184.37	199.54	193.45	183.49
	(51.95)	(53.24)	(53.06)	
5	184.39	199.55	193.46	193.51
	(51.83)	(53.11)	(52.84)	
6	183.45	198.95	183.46	173.50
	(49.66)	(50.80)	(49.76)	

Note: The table presents simulation estimates of mean and standard deviation (parentheses) of sampling distribution for three estimators of average overcharges. All figures are in thousands of dollars.

the probability limit of the forecasting approach depends on the data generating process for Q_t in a more direct way than does that of the dummy variable approach.

The results of these simulation experiments are given in Table 2, which presents estimates of the mean and standard deviation of the estimators for average overcharge discussed, using 240,000 replications of data sets of size T=100. For reference, we also display the estimand, OC^*, for each simulation experiment.[17] As quantity is measured in thousands of units, the dollar figures in the table are in thousands of dollars. Since T=100 here, aggregate overcharges are 100,000 times as large as the quantities in Table 2, or roughly \$20 million in each scenario. This is a typical damages award for a small to moderate case of this nature; in recent years, it has become common to see damages estimates of \$1 billion or more.[18]

While the table contains the figures for the standard deviation of these estimators, we focus on the simulation estimates of bias. This is because in all six models, and many others we have examined, differences in standard deviation among the methods are generally minor, as compared with

17 For Models 1 through 4 we have $\mathbb{E}[(u_t\text{-}v_t)D_tQ_t]=0$ and the estimand reduces to $\pi_1\delta+\pi'_X\gamma$, but for Models 5 and 6 the estimand is more complicated to calculate. In all instances, we approximate the estimand by taking 7.2 million samples of size 100 and averaging the sample means $\frac{1}{100}\sum_{t=1}^{100}D_tQ_t\{Y_t(1)\text{-}Y_t(0)\}$. The margin of error for the simulation estimate of the estimand is ±0.03 for Models 1, 5, and 6, and ±0.02 for Models 2, 3, and 4, where we take advantage of the fact that the estimand is the same and average the three resulting simulation estimates.

18 Freed (2012) notes that "[e]stimates of the potential cost of a settlement of the [Visa] antitrust case vary dramatically – from a few billion dollars into the hundreds of billions." Visa eventually settled for \$4 billion (Touryalai 2012). Other settlement amounts are cited in Marshall (2008), Schoenberger (2009), and Clark (2011). The only review ever conducted along these lines looked at 40 cases (Lande and Davis 2008). The average recovery for plaintiffs among those 40 was \$450 million under one set of assumptions and \$491 million under another set of assumptions, but the cases studied were those known to prominent antitrust attorneys and hence more likely to be ones involving large dollar amounts.

differences in bias. The simulation estimates of the standard deviation of each estimator are nonetheless of interest for computing the margin of error of the estimated means of the sampling distributions. These are given by $\pm 1.96\hat{\sigma}/\sqrt{R}$, where $\hat{\sigma}$ is the estimated standard deviation of the sampling distribution and R is the number of replications, here 240,000.[19] For each of the three estimators and for each model, the margin of error for the mean is roughly ±0.2.

Model 1 corresponds to a setting of constant quantity and no interaction term in the population between covariates and the dummy variable. For this model, Proposition 1 asserts that the first dummy variable approach ($\widehat{OC}_1$) and the forecasting approach ($\widehat{FC}$) are numerically identical. This is borne out in the simulations. In each of the 240,000 replications, the first dummy variable approach and the forecasting approach are identical. The simulation estimated means, presented in the first row of the table, are thus also identical. Because the effect of covariates on price is constant in Model 1, the discussion in Section 6 indicates that there should be no important difference between the two varieties of the dummy variable approach: the first approach allows the effect of covariates on price to change during the alleged conspiracy, and the second approach ($\widehat{OC}_2$) correctly imposes the assumption that the effect of covariates on price is the same over time (i.e., that γ=0). Consistent with our expectation, the two dummy variable approaches perform quite similarly in terms of bias. It is interesting to note that the second dummy variable approach is not particularly precise relative to the first. This will not be true in every setting, as imposing true restrictions can often improve efficiency. Overall, for all three estimators, the simulation estimates of the means are extremely close to the target parameter of $200,000.

Model 2 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 γ=0) is inferior to the first approach. Table 2 shows that the second dummy variable approach has a bias of approximately $15,000, or just over 8% of the true parameter of $183,000.

Model 3 allows quantity to vary according to Equation (13) and retains the assumption that the effect of the covariate on prices differs before and during the alleged conspiracy period. However, this model does not allow any predictable relationship between quantity and price, either under conspiracy conditions or under non-conspiracy conditions. That is, this model conforms to Assumption 1, which implies Assumption 2 holds as well. Proposition 2 implies that both the first dummy variable approach and the forecasting approach should be consistent for average overcharges in this setting. The simulation experiments corroborate this. While it is no longer true that the first dummy variable approach and the forecasting approach are numerically identical, their sampling distributions are nearly identical. In particular, the mean of the two sampling distributions differs in the third decimal place and the standard deviation differs in the second decimal place. Since both distributions are essentially normal, and since the first two moments are essentially identical, it is hard to prefer one estimator over the other in this context on statistical grounds. However, because the effect of the covariates changes over time, the second dummy variable approach is not consistent, with a bias that is again roughly 8% of the true parameter.

Model 4 is the same as Model 3, except for a change to the parameters in Equation (13), which governs quantity. In particular, this model now allows for quantity to be related to unmeasured factors affecting price, which violates Assumption 1. However, by restricting the correlation to be equal under conspiracy and non-conspiracy conditions, Assumption 2 is met. As noted, an easy way to understand this setting is that quantity may be related to but-for prices, but not to the gains from conspiracy. As indicated by the schematic in Table 1, the only consistent estimator in this setting is the first dummy variable approach. The simulation experiments bear this prediction out, with both the second dummy variable approach and the forecasting approach being badly biased, by roughly 8 and 5% of the true parameter, respectively. In contrast, the first dummy variable approach has a sampling distribution mean that is less than one-half of 1% above the true parameter.[20]

19 This margin of error may be justified either by appealing to the central limit theorem applied to the estimators, or to normality of the sampling distribution of each estimator. A detailed examination of the sampling distribution confirms that for these simulation experiments, the sampling distribution is approximately normal. For example, each estimator in each model exhibits skewness of roughly 0.2 and kurtosis of roughly 3.2.

20 While the effect is small, we were somewhat surprised that the first dummy variable approach was not as close to the target in this model as it was in Models 1, 2, and 3. We note that the conclusion of Proposition 2 is not that the estimator is unbiased, but rather that it is consistent. On the other hand, we conducted a similar experiment with a slightly larger sample size of T=200 and encountered similar results – a simulation estimate of the mean that is roughly one-half of 1% above the true parameter, where the true parameter is not inside the confidence region for the simulation estimate.

Model 5 highlights the performance of these estimators when Assumption 1′ is met but Assumption 2 is violated. In this new context, the estimand is no longer \$183,000, but is instead \$194,000. The simulation experiments confirm the prediction of Proposition 2; now only the forecasting approach performs well in terms of proximity to the true parameter. The first and second dummy variable approaches have biases of approximately –5 and 3%, respectively. On the other hand, as discussed, we find it difficult to imagine a real-world justification for Assumption 1′ that would not also imply the validity of Assumption 1, and so this model may be viewed as somewhat artificial.

Finally, Model 6 emphasizes that there is no guarantee that one of these these approaches will estimate average overcharges successfully. Now, the parameters of Equation (13) are such that neither Assumption 1, nor Assumption 2, nor Assumption 1′ is met. In this context, the true parameter is \$174,000 and the first and second dummy variable approaches have approximate biases of 6 and 15%, respectively, and the forecasting approach has an approximate bias of 6%.

8 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 between prices under conspiracy and non-conspiracy conditions 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 or when the regressions themselves are quantity-weighted. 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. However, when quantity is related to unobserved determinants of price, one sufficient condition leads to consistency of the forecasting approach and another sufficient condition leads to consistency of the first dummy variable approach. Neither of these sufficient conditions is implied by the other. However, we have argued that the sufficient condition for consistency of the forecasting approach in this case is somewhat artificial, suggesting slightly greater robustness of the first dummy variable 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. On the other hand, if the restriction that the effects of covariates are constant is, in fact, false, then the second dummy variable approach can be biased.

Overall, our discussion points to a particularly important role for the first dummy variable approach, particularly when there are sufficient data to estimate the model reliably. 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.

Previously published online October 26, 2013

Appendix

Before proving the lemma and propositions, we replicate the key equations from the text and give a synopsis of the maintained assumptions underlying them. Recall the definitions

$$OC^* = \mathbf{E}[\{Y_t(1) - Y_t(0)\}D_tQ_t] \tag{A.1}$$

$$Y_t(1) = \alpha + \delta + (\beta + \gamma)'X_t + u_t \tag{A.2}$$

$$Y_t(0) = \alpha + \beta'X_t + v_t \tag{A.3}$$

$$\hat{Y}_t(1) = \hat{\alpha} + \hat{\delta} + (\hat{\beta} + \hat{\gamma})'X_t \tag{A.4}$$

$$\hat{Y}_t(0) = \hat{\alpha} + \hat{\beta}'X_t \tag{A.5}$$

$$\widehat{FC} = \frac{1}{T}\sum_{t=1}^{T}\{Y_t - \hat{Y}_t(0)\}D_tQ_t = \hat{\pi}_Y - \hat{\pi}_1\hat{\alpha} - \hat{\pi}'_X\hat{\beta} \tag{A.6}$$

$$\widehat{OC}_1 = \frac{1}{T}\sum_{t=1}^{T}\{\hat{Y}_t(1) - \hat{Y}_t(0)\}D_tQ_t = \hat{\pi}_1\hat{\delta} + \hat{\pi}'_X\hat{\gamma} \tag{A.7}$$

and the decomposition into fitted values and fitted residuals:

$$Y_t = \hat{\alpha} + \hat{\beta}'X_t + \hat{\delta}D_t + \hat{\gamma}'D_tX_t + \hat{\varepsilon}_t \tag{A.8}$$

In a context where the regression models are unweighted, we assume

$$0 = \mathbf{E}[u_t] = \mathbf{E}[u_tD_t] = \mathbf{E}[u_tX_t] = \mathbf{E}[u_tD_tX_t] \tag{A.9}$$

$$0 = \mathbf{E}[v_t] = \mathbf{E}[v_tD_t] = \mathbf{E}[v_tX_t] = \mathbf{E}[v_tD_tX_t] \tag{A.10}$$

These assumptions imply that the regression residuals $\varepsilon_t = u_tD_t + v_t(1-D_t)$ satisfy $0 = \mathbf{E}[\varepsilon_t] = \mathbf{E}[\varepsilon_tX_t] = \mathbf{E}[\varepsilon_tD_t] = \mathbf{E}[\varepsilon_tD_tX_t]$, i.e., that $\alpha + \beta'X_t + \delta D_t + \gamma'D_tX_t$ is the best linear predictor of $Y_t = D_tY_t(1) + (1-D_t)Y_t(0)$ given X_t, D_t, and D_tX_t. When the regression models are weighted, we modify the assumptions above to

$$0 = \mathbf{E}[u_tQ_t] = \mathbf{E}[u_tD_tQ_t] = \mathbf{E}[u_tX_tQ_t] = \mathbf{E}[u_tD_tX_tQ_t] \tag{A.11}$$

$$0 = \mathbf{E}[v_tQ_t] = \mathbf{E}[v_tD_tQ_t] = \mathbf{E}[v_tX_tQ_t] = \mathbf{E}[v_tD_tX_tQ_t] \tag{A.12}$$

In other words, we assume that, when the economist chooses to weight, weighting is in fact appropriate. We now prove the lemma and propositions from the main text.

Proof of Lemma: Using Equations (A.5), (A.6), and (A.8), we have

$$\widehat{FC} = \frac{1}{T}\sum_{t=1}^{T}D_tQ_t(\hat{\delta}D_t + \hat{\gamma}'D_tX_t + \hat{\varepsilon}_t) = \hat{\pi}_1\hat{\delta} + \hat{\pi}'_X\hat{\gamma} + \frac{1}{T}\sum_{t=1}^{T}D_tQ_t\hat{\varepsilon}_t$$
$$= \widehat{OC}_1 + \frac{1}{T}\sum_{t=1}^{T}D_tQ_t\hat{\varepsilon}_t. \qquad \square$$

Proof of Proposition 1: Applying the Lemma, note that when quantity is constant, say $Q_t = \bar{Q}$, we have $\frac{1}{T}\sum_{t=1}^{T}D_tQ_t\hat{\varepsilon}_t = \bar{Q}\frac{1}{T}\sum_{t=1}^{T}D_t\hat{\varepsilon}_t = 0$ by the orthogonality of fitted residuals and covariates. The same holds for a weighted regression, but the orthogonality condition is then precisely that $\frac{1}{T}\sum_{t=1}^{T}D_tQ_t\hat{\varepsilon}_t = 0$. $\square$

Proof of Proposition 2: The assumption that (Y_t, X_t, D_t, Q_t) is a vector ergodic stationary process with existence of sufficient moments implies that moments such as $\mathbf{E}[D_tQ_tY_t]$ exist, are finite, and are time invariant; and also that the corresponding sample mean converges in probability to that expectation. Together with Equations (A.9) and (A.10), this implies that the regression coefficients $\hat{\alpha}, \hat{\beta}, \hat{\delta}$, and $\hat{\gamma}$ are consistent for α, β, δ, and γ, respectively, and that the averages $\hat{\pi}_1$, $\hat{\pi}_X$, and $\hat{\pi}_Y$ are consistent for $\pi_1 \equiv \mathbf{E}[D_tQ_t]$, $\pi_X \equiv \mathbf{E}[D_tQ_tX_t]$, and $\pi_Y \equiv \mathbf{E}[D_tQ_tY_t]$, respectively.

To discuss consistency, it is helpful to characterize the estimand under our assumptions. We utilize two such characterizations, one for the forecasting approach and the other for the dummy variable approach. That for the forecasting approach is given by

$$OC^* = \mathbf{E}[Y_t(1)D_tQ_t] - \mathbf{E}[Y_t(0)D_tQ_t] = \mathbf{E}[Y_tD_tQ_t]$$
$$= -\mathbf{E}[(\alpha + \beta'X_t + v_t)D_tQ_t] = \pi_Y - \pi_1\alpha - \pi'_X\beta - \mathbf{E}[v_tD_tQ_t]$$

To see how this characterization is related to consistency of the forecasting approach, note that $\widehat{FC} = \hat{\pi}_Y - \hat{\pi}_1\hat{\alpha} - \hat{\pi}'_X\hat{\beta}$ converges in probability to $\pi_Y - \pi_1\alpha - \pi'_X\beta$ by continuity of probability limits. So consistency of $\widehat{FC}$ for OC^* follows if

$$\mathbf{E}[v_tD_tQ_t] = 0$$

Sufficient conditions for this conclusion include: quantity is constant at $\bar{Q}$ as then $\mathbf{E}[v_tD_tQ_t] = \bar{Q}\mathbf{E}[v_tD_t]$ [cf., Equation (A.10)]; the regression is quantity-weighted [cf., Equation (A.12)]; or the covariance during the conspiracy between quantity and unmeasured influences on the but-for price is zero, i.e., $0 = \mathbf{C}[v_t, Q_t | D_t = 1]$, since by the law of total probability we have $\mathbf{E}[v_tD_tQ_t] = \mathbf{E}[D_t]\mathbf{E}[v_tQ_t | D_t = 1] = \mathbf{E}[D_t]\mathbf{C}[v_t, Q_t | D_t = 1]$ (cf., Assumptions 1 and 1′).

Turning to the dummy variable approach, we give our second characterization of the estimand. We have

$$OC^* = \mathbf{E}[\{\delta + \gamma'X_t + u_t - v_t\}D_tQ_t] = \delta\pi_1 + \gamma'\pi_X + \mathbf{E}[(u_t - v_t)D_tQ_t] \tag{A.13}$$

To see how this characterization is related to consistency of the dummy variable approach, note that

$\widehat{OC}_1 = \hat{\delta}\hat{\pi}_1 + \hat{\gamma}'\hat{\pi}_X$ converges in probability to $\delta\pi_1 + \gamma'\pi_X$ by continuity of probability limits. So consistency of the dummy variable approach follows if

$$\mathbf{E}[(u_t - v_t) D_t Q_t] = 0$$

Sufficient conditions for this conclusion include: quantity is constant at $\bar{Q}$ [cf., Equations (A.9) and (A.10)]; the regression is quantity-weighted [cf., Equations (A.11) and (A.12)]; the covariance between quantity and unmeasured influences on price during the conspiracy is zero, both for actual price and for but-for prices (cf., Assumption 1); or the covariance during the conspiracy between quantity and unmeasured influences on price would have been the same for actual and but-for price, since by the logic above $\mathbf{E}[(u_t - v_t) D_t Q_t] = \mathbf{E}[D_t]\{\mathbf{C}[u_t, Q_t | D_t = 1] - \mathbf{C}[v_t, Q_t | D_t = 1]\}$ (cf., Assumption 2). □

References

Clark, Don. 2011. "Rambus Loses Antitrust Case." *Wall Street Journal*, November 17.

Fisher, Franklin. 1980. "Multiple Regression in Legal Proceedings." *Columbia Law Review* 702–736.

Freed, Dan. 2012. "Bank of America, Citigroup Face Billions In Losses in Antitrust Case (Update 1)." *The Street*, January 12.

Higgins, Richard, and Paul Johnson. 2003. "The Mean Effect of Structural Change in the Dependent Variable is Accurately Measured by the Intercept Change Alone." *Economics Letters* 80: 255–259.

Hovenkamp, Herbert. 2005. *Federal Antitrust Policy: The Law of Competition and Its Practice*. St. Paul, MN: Thomson-West.

Imbens, Guido W. 2004. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *Review of Economics and Statistics* 86 (1): 4–29.

Kmenta, Jan. 2000. *Elements of Econometrics*. 2nd ed. Ann Arbor, MI: The University of Michigan Press.

Lande, Robert H., and Joshua P. Davis. 2008. "Benefits From Private Antitrust Enforcement: An Analysis of Forty Cases." *University of San Francisco Law Review* 42: 817–918.

Marshall, Kevin S. 2008. *The Economics of Antitrust Injury and Firm-Specific Damages*. Tucson: Lawyers & Judges.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688–701.

Rubinfeld, Daniel L. 1985. "Econometrics in the Courtroom." *Columbia Law Review* 1048–1097.

Rubinfeld, Daniel L. 2008. "Quantitative Methods in Antitrust." In *Issues in Competition Policy, ABA Antitrust Section*, edited by Dale Collins, American Bar Association, chapter 30.

Rubinfeld, Daniel L., and Peter O. Steiner. 1983. "Quantitative Methods in Antitrust Litigation." *Law and Contemporary Problems* 26: 69–141.

Salkever, David. 1976. "The Use of Dummy Variables to Compute Predictions, Prediction Errors, and Confidence Intervals." *Journal of Econometrics* 4 (4): 393–397.

Schoenberger, Robert. 2009. "Eaton Might Owe Damages in Billions in Antitrust Case." *The Plain Dealer*, October 20.

Touryalai, Halah. 2012. "Senator Dick Durbin Hates The $7 Billion Visa, Mastercard Settlement," *Forbes*, August 7.

White, Halbert, Robert Marshall, and Pauline Kennedy. 2006. "The Measurement of Economic Damages in Antitrust Civil Litigation." *ABA Antritrust Section, Economic Committee Newsletter, Spring* 6 (1): 17–22.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and* Panel Data. Cambridge: MIT Press, 2002.