

Using Difference-in-Differences to Estimate Damages in Healthcare Antitrust:
A Case Study of Marshfield Clinic

R. Forrest McCluer*
Greylock McKinnon Associates

Martha A. Starr
American University

March 2013

Abstract

In calculating damages in healthcare antitrust cases, the difference-in-difference (DID) approach provides a potentially valuable means of controlling for lawful factors that influence prices, such as case-mix and quality of care, as distinct from price differentials due to unlawful behavior. After first comparing DID to traditional methods of estimating damages, this paper uses DID to analyze data from a well-known case against Marshfield Clinic, a large multi-specialty group practice that was found to have illegally allocated markets for physician services in Central Wisconsin. Using a specification similar to what was used in the case, we find that illegal behavior accounted for about one-half of the Clinic's extra increase in costs per patient during the damage period. The courts, however, were not persuaded that the analysis adequately controlled for legal factors. We discuss potential pitfalls in using DID to estimate damages suggested by the case, as well as possible ways around them.

Keywords: healthcare antitrust, differences-in-differences, damage estimation
JEL Codes: I11, L4, K21, K41, C20

* Please address correspondence to: R. Forrest McCluer, Greylock McKinnon Associates, One Memorial Drive, Suite 1410, Cambridge, MA 02142, email: fmcccluer@gma-us.com, tel: (703) 237-3010. McCluer worked in support of plaintiff expert witnesses in cases against Marshfield Clinic. We thank Deborah Haas-Wilson, James Langenfeld, Thomas McGuire, William Sage, four anonymous referees, and the editor for valuable comments. The views expressed in this paper are those of the authors and cannot be attributed to Greylock McKinnon Associates or any other employee or affiliate thereof. All errors are our own.

1. Introduction

Estimating damages in antitrust cases depends on being able to estimate what prices of goods or services would have been in the absence of unlawful behaviors, such as price-fixing or illegal market allocations. As such, it requires a method for deriving counterfactual or “but for” prices that is both sound in terms of economic analysis and satisfactory in terms of legal criteria. Estimating damages in healthcare cases can be especially challenging due to the heterogeneous character of health-related ‘products’ and the relatively complex determinants of prices. Notably, in cases related to physician or hospital services, measures of charges, payments or costs of patients’ care are typically used to quantify prices, yet these can vary considerably due to factors other than illegal behavior (McCarthy and Thomas 1995). For example, average payments per patient may be relatively high for providers whose quality of care is relatively high and/or who treat patients with relatively complex or serious conditions, or they may be relatively low for providers who treat patients whose costs are reimbursed at relatively low rates. Thus, regression methods are often used to estimate how illegal behaviors affect prices, controlling for these other types of factors that would be expected to contribute to their variation (Coleman and Langenfeld 2010).

The two traditional sets of methods for estimating damages confront special problems in healthcare antitrust cases. ‘Yardstick’ methods using cross-section data have to deal with problems of unobserved heterogeneities across providers: If some lawful factors affecting prices (like quality of care or complexity of case mix) are difficult to measure and control for in a regression analysis, yet they are thought to vary systematically between providers that were and were not engaged in illegal behaviors, estimated effects of unlawful behaviors on patient costs will be biased. ‘Before and after’ methods using time-series data get around this by using observations from the period before the illegal behavior started to estimate determinants of prices; then out-of-sample forecasts can be used to project how prices would have evolved during the period of the conspiracy, had behavior remained lawful. However, ability to make good out-of-sample forecasts requires adequate time-series data on prices in the ‘before’ period, which is often not available in healthcare cases.

The difference-in-difference (DID) approach is potentially valuable for overcoming these problems in traditional methods for estimating damages. The DID approach estimates effects of illegal behavior by comparing changes in prices for providers who started engaging in illegal behavior between one period and another, to changes in prices of those who did not. As such, changes in prices among providers not engaged in the illegal behavior serve as the counterfactual estimate of how prices would have changed in the absence of conspiracy; then the extra change in prices for those involved in the conspiracy provides the basis for estimating damages. This method abstracts from initial differences in levels of prices between

those who became involved in the conspiracy and those who did not, assuming that whatever differences had prevailed before the conspiracy started reflected other factors contributing to variations in prices apart from illegal behavior.

This paper discusses the use of DID for estimating damages in healthcare antitrust and applies the method to the well-known case of Marshfield Clinic, in which a large multi-specialty group of physicians was found to have illegally divided up markets for physician services with other providers in certain parts of central Wisconsin. In the next section of the paper, we briefly review the difference-in-difference methodology and discuss its advantages over other methods used to estimate damages in healthcare antitrust analysis. The third section of the paper gives an overview of the Marshfield Clinic case, focusing on the remand phase in which calculation of damages figured centrally, and lays out a DID method for estimating damages which parallels the method used in the actual case. The fourth section presents results from the DID analysis showing that, while average payments for physician services of Marshfield clinic patients in the market-allocation area exceeded those of other patients in the state by \$309-326 during the market-allocation period, about one-half of this increase was attributable to illegal behavior. The fifth section discusses the problems that arose in the Marshfield Clinic case in persuading the courts that the DID method correctly isolated the effects of the illegal market allocation. The paper concludes with more general discussion of potential pitfalls in using DID methods to estimate damages in healthcare antitrust cases, as well as possible ways around them.

2. Approaches to estimating damages and the difference-in-difference approach

The basic principle for estimating damages in antitrust cases entails estimating what prices would have been “but-for” the anticompetitive act or set of acts in question.¹ Then total damages to the party or parties injured by the anticompetitive act(s) are estimated by multiplying the difference between actual prices and the “but-for” or counterfactual prices by the relevant measure of quantity. Central to the estimation of counterfactual prices is that price variations due to legitimate and legal factors have to be separated from those associated with unlawful behavior (Rubinfeld 2012). This is a particularly important issue in healthcare, where standard measures of ‘price’-- such as charges or payments per inpatient stay or annual spending for physician services -- are known to reflect hedonic differences in the characteristics of services, such as quality of care, intensity of treatment, or severity of illness. In this case, the method for estimating “but-for” prices has to compellingly separate out differences due to unlawful behavior from differences due to these other factors.

¹ Royall (1997) discusses the need to disaggregate damages when there are multiple illegal acts.

There are two standard sets of methods for estimating “but for” prices. The first, referred to as the “yardstick” or “benchmark” method, relies on cross-section data on prices of the relevant goods or services (McCrary and Rubinfeld 2009, Langenfeld et al. 2010, Hovenkamp 2011, Rubinfeld 2012). Depending on specifics of the case, the approach may either compare prices between geographic areas in which anticompetitive behavior has taken place to prices for comparable items in an otherwise similar area unaffected by the conspiracy, or it may compare prices of products supplied by firms who were party to the conspiracy to equivalent products of other firms who were not.² A common problem here is making sure that all legal sources of differences in prices across areas or firms are well-accounted for. For example, in a regression analysis in which a dummy variable is used to indicate observations in areas where firms’ behavior was anticompetitive, the estimated coefficient on this variable will give a biased estimate of the effect of the conspiracy if included explanatory variables do not fully cover lawful sources of variations in prices, yet these differ systematically across the two types of areas.³

Thus, the second set of methods, known as the “before-and-after” approach, compares prices in the relevant geographical area or of the relevant firms in a period before the unlawful behavior began, to prices in a period when it was in effect.⁴ The model is estimated using data from the “before” period, with an out-of-sample forecast used to estimate what the price would have been in the period when the illegal behavior occurred, had it not occurred (White, Marshall and Kennedy 2006). While this method gets around the problem of unobserved differences in product quality or in characteristics across firms or across markets, estimating a model with sufficient precision to permit out-of-sample forecasting requires a good length of time-series data on prices in the period before the illegal behavior began.⁵

² Various definitions for the approach have been offered. According to Hovenkamp (2011:3-16-2), “The ‘yardstick’ method is similar [to the before/after method] except that instead of comparing a different time period the damages expert compares the plaintiff’s performance in the damages market with its performance in some other market...” Langenfeld et al. (2010:174) write that “...the benchmark approach uses prices in the unaffected market as a control.” On the other hand, as McCrary and Rubinfeld (2010:1) state, “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 “yardstick” and “benchmark” are used interchangeably to refer to the cross-section comparison.

³ See the working paper version of this paper for longer exposition.

⁴ Alternatively, if data are also available for the period after the conspiracy ended, the approach may be “before-during-after.” See Langenfeld et al. (2010: 167-173).

⁵ Thus, in healthcare antitrust, the before/after method is most often used in cases involving pharmaceuticals, for which time-series data on prices are more commonly available. McCrary and Rubinfeld (2009) show that, in cases where both approaches are possible, the yardstick/benchmark and before/after approaches will yield identical estimates of overcharges due to price-fixing *if* the level of sales of the product does not change during the price-fixing period. Estimates will differ, however, if the level of sales changes in the price-fixing period.

In antitrust cases where repeated cross-section or panel data are available, the difference-in-difference (DID) approach can be used.⁶ To illustrate using an example that parallels the case analyzed in this paper, suppose P_{it} is the price of a given healthcare service obtained by

individual i at time t .⁷ Determinants of price can be specified as follows:

$$P_{it} = \beta_0 + \gamma_0 D_{it} + \gamma_1 M_{it} + \gamma_2 M_{it} D_{it} + X_{it}' \beta + \varepsilon_{it} \quad (1)$$

where M_{it} is a dummy variable equal to one if the individual lived in an area in which firms were engaged in an illegal behavior; D_{it} is a dummy variable equal to one for time periods when the illegal behavior was in effect; and X_{it} is a vector of individual- or product-level characteristics that might affect the price of the service other than the illegal market allocation. Here the parameter γ_1 measures persistent unobserved sources of variation in price between two areas. The parameter γ_0 captures broad-based changes in price between the initial period and the period in which the illegal behavior took place; allowing for broad-based price growth is clearly important in healthcare applications, given widespread robust growth in prices in recent decades.⁸ The parameter γ_2 , reflects the *extra* increase in price in the area and in the period in which firms' behavior was anticompetitive, and as such provides a basis for estimating damages. Note that, for γ_2 to have this interpretation, the set of included explanatory variables, X_{it} , must compellingly capture causes of differential changes in prices between the two areas due to legal factors; otherwise the estimate of γ_2 will reflect both illegal behavior and legal-but-omitted factors that changed differentially between the two areas.⁹

⁶ For general expositions and references, see Wooldridge (2002), Cameron and Trividi (2005), or Angrist and Pischke (2009, 2010).

⁷ The question of how to measure 'prices' of healthcare services is discussed below.

⁸ U.S. Bureau of Labor Statistics (2009),

⁹ See Rubinfield (2010: 74-75) and the working paper version of this paper for discussion of other assumptions required for γ_2 to constitute a clean measure of the conspiracy's effects. Abadie (2005) discusses this 'parallel trend' assumption in DID analysis.

While DID approaches have been widely used in economic research,¹⁰ their use in analysis for antitrust cases has been more limited. In healthcare antitrust specifically, DID has been used mostly in retrospective studies of hospital mergers.¹¹ Vita and Sacher (2001) studied effects of a merger of two California hospitals on prices for acute inpatient care, using a peer group of hospitals of similar size in similar metropolitan areas in the state as a control group. Tenn (2011) examined price effects of a merger between two Oakland/Berkeley-area hospitals, comparing their prices for acute inpatient care to other Bay area hospitals offering similar services. Several studies have used DID to examine whether the quality of care tends to improve after a hospital merger, as parties to mergers often claim. But as the DID approach clarifies, the question is whether quality improves *differentially* in merged hospitals, relative to otherwise similar peers. As discussed in Vogt and Town (2006) and Romano and Balan (2011), there is little evidence thus far that mergers enable firms to improve their quality in this relative sense.¹²

3. The case of Marshfield clinic

Marshfield Clinic is located in Central Wisconsin and is one of the largest private, multi-specialty group practices in the United States. When the antitrust case against the Clinic began in 1994, it was staffed by over 400 physicians working both in the Clinic in Marshfield and a network of two dozen Regional Centers. Physically adjacent to the Clinic was St. Joseph's Hospital, a 524-bed tertiary-care teaching institution. Most of Marshfield Clinics' primary physicians and specialists served on the medical staff at St. Joseph's Hospital. In addition, Marshfield Clinic owned and operated its own health maintenance organization (HMO), Security Health Plan. Coombs (2005) provides a detailed discussion of the establishment and growth of Marshfield Clinic and its affiliated operations.

In 1994, Blue Cross-Blue Shield United of Wisconsin (hereafter 'Blue Cross') and its HMO, CompCare, brought an antitrust suit against Marshfield Clinic and its HMO, alleging that they were monopolizing the market for physician services, fixing prices, and illegally dividing markets with their competitors.¹³ The experts testifying on behalf of Blue Cross analyzed four product markets and an array of geographic markets associated with Marshfield Clinic and its

¹⁰ See Angrist and Pischke (2009, 2010) for general overviews and references. Relevant work in industrial-organization and antitrust economics includes Borenstein (1990), Hastings (2004), Simpson and Schmidt (2008), Coleman and Langenfeld (2008), Simpson and Taylor (2008), Jiménez and Perdiguerro (2009), Langenfeld et al. (2010), and Rubinfeld (2012).

¹¹ See the February 2011 special issue of *International Journal of the Economics of Business*.

¹² See also Hammer and Sage (2002: 564), who remark that, "Not all 'quality' claims are equally meritorious. There is a long history in antitrust law of quality claims being used to rationalize behavior that is patently anticompetitive."

¹³ *Blue Cross and Blue Shield United of Wisconsin v. Marshfield Clinic*, 95-1965 & 95-2140, 883 F. Supp. 1247, 1259 (W.D. Wis. 1995). Formally, this was a Section 2 monopolization charge combined with a Section 1 price-fixing and division-of-markets charge.

branch clinics in which market allocation occurred.¹⁴ Benchmark competitive prices were estimated using methods that controlled for the extra services provided to Marshfield patients and comparisons of prices charged by other large clinics for given physician services.¹⁵ The jury found in favor of the plaintiffs, and Blue Cross was awarded damages of \$16 million (before trebling).¹⁶ However, when the case was appealed to the 7th Circuit Court of Appeals in 1995, Chief Judge Richard Posner threw out the damage estimates on several grounds, but affirmed that Marshfield clinic's illegal division of markets constituted a basis for Blue Cross to be compensated for extra payments it may have made, if any, due to the illegal division.¹⁷ This led to the remand case in 1997, in which the District Court judge issued a summary judgment in favor of Marshfield clinic, saying that Blue Cross had failed to establish that it had suffered any injury as a result of the illegal market allocation. The remand case was in turn unsuccessfully appealed in 1998.¹⁸

The core charges against Marshfield Clinic in the remand phase of the case concerned its practice of making arrangements with other providers in Central Wisconsin intended to reduce competition between them. Other providers with which Marshfield and its HMO were found to have had illegal market-allocation agreements included North Central Health Protection Plan based in Wausau; Rhinelander Medical Center based in Rhinelander; and the Wausau Medical Center in Wausau.¹⁹ While the agreements varied across providers and spread over space and time, they included a variety of practices aiming to restrict competition, including agreements not to open new general practices in each others' core geographic areas, not to compete actively for each others' patients, and not to locate new specialty practices in areas close to

¹⁴ The product markets analyzed in the initial case were primary care, pediatric care, specialty care and HMO services. The courts disagreed that HMO services constituted a separate market. Haas-Wilson (2003) provides in-depth discussion of antitrust implications of managed care, including issues that arose in the first Marshfield Clinic case.

¹⁵ See McCluer and Starr (2013) for further discussion of the litigation against Marshfield Clinic. The first phase included allegations that the higher payments for physician services of Marshfield patients reflected not only prices that were higher than they would have been in the absence of the Clinic's illegal behavior, but also higher levels of service provision.

¹⁶ However, the trial judge reduced the award to \$5.6 million, saying the jury's award was "clearly excessive, ... contrary to the evidence presented at trial, and based on a figure lacking adequate foundation" (Shabaz 1995).

¹⁷ Posner specifically instructed that damage calculations had to omit any extra patient costs that Blue Cross may have paid as a result of Marshfield Clinic's 'legal' market power, i.e. its ability to raise prices because of its dominant position in the main areas it served, as opposed to increases in prices resulting from illegal market allocation.

¹⁸ The original case was *Blue Cross Blue Shield vs Marshfield Clinic* (883 F. Supp. 1247, 1259 (W.D. Wis. 1995)), its appeal (65 F.3d 1406 (7th Cir. 1995)), the remand case (*Blue Cross Blue Shield vs Marshfield Clinic* (980 F. Supp. 1298 (W.D. Wis. 1997))), and its appeal (152 F. 3d 588 (7th Cir. 1998)). A related class action suit was ultimately settled (*Rozema, et al v. Marshfield Clinic, Security Health Plan of Wisconsin, Inc., North Central Health Protection Plan, and Rhinelander Medical Center, S.C.* (977 F. Supp. 1362 (W.D. Wis. 1997)) and *cert. denied*, 525 U.S. 1071, 119 S.Ct. 804 (1999)).

¹⁹ Allegations concerning agreements with Rice Clinic in Portage County were found to be lacking evidence (Crabb 1997).

already-established practices of competitors.²⁰ Additionally, Marshfield Clinic and the North Central Health Protection Plan made “free flow” arrangements whereby HMO patients of one provider could see the others’ providers without first obtaining a referral, having the effect of reducing competition for enrollees.²¹ Blue Cross argued that these market-allocation agreements reduced competition in the market for physician services in Central Wisconsin, forcing Blue Cross to pay supra-competitive prices for physician services for its enrollees in this part of the state.

Thus, the key question for estimating damages was the extent to which payments for physician services were higher in affected parts of the state compared to what they would have been in the absence of the anti-competitive agreements. As these agreements pertained to the full range of medical specialties practiced by physicians, the relevant product market was taken to be the market for professional services rendered by physicians.²² As the antitrust charges specifically pertained to excess payments made by Blue Cross on behalf of its patients, the basis for estimating damages, which also served as the dependent variable in the regression analysis, was total annual payments per patient made by Blue Cross over the course of the year.

In the analysis that follows, we use individual health-insurance claims data for Blue Cross/Blue Shield subscribers, along with a DID approach similar to what was used in the remand case, to examine whether the market-allocation agreements significantly affected payments for physician services in the areas in which they were in effect. Whereas the analysis in the case used county-level averages in the DID regressions, the analysis presented here uses the individual-level claims data. This enables us to use the rich information on individual patient characteristics to reduce odds that unobserved individual-level heterogeneities may lead to bias in estimated effects of anti-competitive arrangements. Nonetheless, the results differ little qualitatively between the analyses using county-level averages versus individual-level data.²³

4. Difference-in-difference specification and data

Identifying the anti-competitive effect via DID requires delineations between treatment vs. control groups and between before vs. after periods. In terms of the treatment/control

²⁰ In principle, agreements to restrain trade are illegal *per se* although with some distinction made between “naked” restraints on trade and agreements that may facilitate productive activity (see Crabb 1997). In the remand phase of the Marshfield field, it had already been ruled that the agreements in question were illegal, with the focus instead on the extent to which Blue Cross experienced damages resulting from them.

²¹ See Crabb (1997) for details of the facts of the case and Coombs (2005) for extended discussion.

²² Specifically, professional services of doctors of medicine and osteopathy in sole practices or single- or multispecialty group practices.

²³ See the working paper version of this paper for analysis using county-level averages.

delineation, following the terms of the remand order, the analysis identified an eight-county area surrounding Marshfield Clinic as the areas where the illegal market-allocation agreements took place.²⁴ While it would be expected that the agreements increased prices within this area, there was no expectation that they increased prices systematically or uniformly throughout the area; here the allegation is not that the agreements enabled Marshfield Clinic and other parties to monopolize the market for physician services in the area, but rather that Marshfield Clinic in particular was able to charge higher prices for services of its physicians due to these arrangements.

Thus, the primary 'treatment' group in the DID analysis consists of patients of Marshfield Clinic residing in the AOI, as the primary interest is in quantifying the extra amount that Blue Cross paid for their care. In principle, a 'Marshfield Clinic patient' could be defined as a patient who received any care from the Clinic during the year, or attention could be confined to patients for whom at least some minimum percentage of their payments for physician services in a given year went to Marshfield Clinic. In fact, within the AOI, most patients who received some care from the Clinic tended to use it for most of their care; for example, among patients living in the AOI who received any care from Marshfield Clinic, the average share of their payments going to the Clinic was 80.2%. Nonetheless, to ensure that our analysis is robust to alternative definitions, we define Marshfield Clinic patients variously as patients who received over 50%, 75% or 90% of the value of payments for their care from Marshfield Clinic providers.

In selecting the control group for the 'but for' analysis, three possible alternatives could be considered. The first is other 'non-Marshfield' patients residing within the AOI, whose shared geographical location with the 'treatment' group might seem to make them a good candidate for the control group. However, this group cannot serve as the control group because they have also been subject to the 'treatment': that is, because some patients in this group obtained services from the providers with which Marshfield Clinic had its market-allocation agreements, their payments for physician services would also be expected to exceed levels that would have been observed without the anticompetitive behavior. The second possibility is to use all patients residing in the rest of the state of Wisconsin outside of the AOI, as this group will be subject to the same state-wide shifts in economic conditions and state-level economic policies as residents of the AOI, but healthcare providers outside the AOI were not party to the market-allocation agreements with Marshfield Clinic. A potential problem here is that, especially with Marshfield Clinic having a reputation as a high-quality provider, some patients living outside the AOI travel to Marshfield facilities within the AOI for their care, so that their payments for healthcare services also would have been increased by the anticompetitive agreements. In practice, this may not pose a substantial problem for the use

²⁴ The eight counties in the AOI were: Clark, Lincoln, Marathon, Oneida, Portage, Price, Taylor, and Wood.

of all residents of Wisconsin outside the AOI as the control group, because the share of this group acquiring services from Marshfield Clinic is very small (see below). Still, because non-AOI residents who travel to Marshfield facilities may have unusually high annual healthcare spending, it may be preferable to exclude such people from the control group. Thus, a third -- and in our view preferred -- possible control group is Wisconsin residents living outside the AOI who were not patients of Marshfield Clinic in the year in question. It is important to note that the choice of control group only affects which group is treated as the omitted group in the regression analysis; the regression specification incorporates terms intended to capture differences in price levels and price growth across these four groups (i.e. patients inside and outside the AOI, considered to be Marshfield patients or not), so as to correctly isolate the effect of the 'treatment'.

Table 1 presents information on the distribution of patients across these categories for the period covered by the data (1988-95), along with their average annual payments on physician services. Here and throughout the paper, all dollar values are converted to constant 1995 terms using the consumer price index. During this time, patients living in the AOI represented less than 5% of the state population, with 22.9% of patients within the AOI receiving 50% or more of the value of their care from Marshfield Clinic; outside of the AOI, only 0.6% of patients received this amount of care from Marshfield. Shifting the threshold for defining a Marshfield patient to 75% or 90% changes in this picture very little: at the 90% threshold, Marshfield's share of patients within the AOI slips to 18.8% while that outside inches down to 0.4%.

Both inside and outside the AOI, the annual payments of Marshfield patients were considerably higher than those of other patients. Again using the 50% threshold for defining Marshfield patients, average payments were \$1,112 for Marshfield patients inside the AOI compared to \$577 for other patients in the area; in the rest of the state, average payments were \$1,187 for Marshfield patients versus \$648 for others. These differences across patient types are not substantially different if the alternative thresholds for defining Marshfield patients are used. Given the expectation that patients living outside the AOI might have unusually high spending on physician services, it may be surprising to see a relatively small difference in annual payments for Marshfield patients inside and outside the AOI. The explanation may be that 95% of patients from outside the AOI who get at least 50% of their care from Marshfield Clinic reside in counties contiguous to the AOI, so conceivably the mix of conditions for which they seek treatment at Marshfield may not differ substantially from that of Marshfield patients within the AOI.

In terms of defining the 'before' and 'after' periods of the DID analysis, the standard DID approach assumes a clear-cut divide between the period before anticompetitive behavior started and the period after. But as is not uncommon in antitrust applications, the delineation

in the Marshfield Clinic case between a period when prices were uninfluenced by anticompetitive behavior and one in which they were is not so straightforward.²⁵ From the evidence presented in the first phase of the Marshfield litigation, the market-allocation agreements amounted to a series of bilateral meetings and agreements starting as early as 1978 and intensifying in the second half of the 1980s, as the spread of preferred-provider arrangements and managed care prompted healthcare providers to explore ways of reinforcing their competitive positions (Coombs 2005, Greenberg 1998). However, even if data on prices of physician services from the mid-1970s were available, the amount of structural change in the health-care industry between then and the time of concern in the litigation case suggests that prices from the mid-1970s would not provide a compelling basis for deriving the 'but for' prices in the period of interest. At the same time, given the staggered way in which the agreements were negotiated and implemented, it is not necessarily clear *a priori* when one should expect them to have been having their full effects on prices. Valuable here is the date used to define the class period in subsequent class-action legislation related to the Marshfield case, which took July 1992 as the beginning of the class period; this suggests that the higher prices being paid for Marshfield Clinic patients should be expected to have been materially important by this time.

In the absence of one before-and-after break in the data, but knowing that the materially-important agreements phased in between the late 1980s and 1992, we specify the DID regression in two different ways. The first in effect uses 1988 -- the first year for which electronic files of Blue Cross claims data were available -- as the 'before' year, with year dummies used to capture broad-based growth in healthcare spending over the period.²⁶ More precisely, this approach entails estimating the following regression:

$$Y_{it} = \beta_0 + \gamma_0 D_{it} + \gamma_1 MA_{it} + \gamma_2 MA_{it} D_{it} + \gamma_3 NA_{it} + \gamma_4 NA_{it} D_{it} + \gamma_5 MO_{it} + \gamma_6 MO_{it} D_{it} + X'_{it} \beta + T'_{it} \theta + \varepsilon_{it} \quad (2)$$

where Y_{it} is real average payments for physician services for patient i in year t . D_{it} is a dummy variable equal to 1 if the year of the observation is after 1988. MA_{it} is a dummy variable equal to 1 for Marshfield patients in the AOI; NA_{it} is a dummy variable equal to one for non-

²⁵ Tenn (2011) discusses this issue in the case of staggered expirations of contracts with insurance companies in the analysis of hospital mergers.

²⁶ This parallels the specification used in the analysis in the case.

Marshfield patients in the AOI; and MO_{it} is a dummy variable for Marshfield patients in the rest of the state. In effect, the omitted (control) group is non-Marshfield patients in the rest of the state. X_{it} is a vector of characteristics of the individual patient and the county in which he/she lives that could contribute to lawful sources of variation in payments, and T_{it} is a vector of year dummies.

Here the terms γ_1 , γ_3 and γ_5 capture sources of persistent unobserved heterogeneities in payments for physician services between Marshfield patients in the AOI, non-Marshfield patients in the AOI, and Marshfield patients outside the AOI respectively, relative to non-Marshfield patients outside the AOI. The parameter of primary interest is γ_2 which captures the *extra* increase in average costs for Marshfield Clinic patients in the AOI between 1988 and 1989-95, above and beyond the increase experienced by the control group, where changes in observed explanatory variables in X_{it} aim to control for factors besides the illegal market agreements that could account for differential changes in prices. γ_4 and γ_6 are the analogous terms that quantify any differential change for non-Marshfield patients in the AOI and Marshfield patients outside of the AOI respectively.

Importantly from the point of view of estimating damages, it is straightforward to show that having a 'before' period in which illegal competitive behavior had already started taking shape will tend to result in underestimation of damages due to that behavior; this is because damages are rooted in estimated excess price growth, yet price *growth* will tend to be understated if measured from a point when price *levels* have already increased above a competitive benchmark.²⁷ Thus, while it is not ideal to take 1988 as the 'before' period, this approach is nonetheless valuable for gauging whether prices rose differentially for Marshfield Clinic in the AOI relative to a period when the agreements were more partial and limited in scope.

Alternatively, the DID can also be specified more flexibly, to provide a way of characterizing how differential prices faced by Marshfield Clinic patients in the AOI may have phased in over the 1988-95 period.²⁸ Here we re-estimate the regressions using a full set of interactions between years and patient types as follows:

²⁷ See the working paper version of this paper for full exposition.

²⁸ See also Laporte and Windmeijer (2005).

$$Y_{ijt} = \beta_0 + \sum_{t=1988}^{1995} \sum_{j=1}^4 \alpha_{jt} d_t M_{ijt} + X'_{it} \beta + \varepsilon_{ijt} \quad (3)$$

where d_t is a dummy variable equal to one in year t . Here M_{ijt} is a set of dummy variables indicating the patient t , where $j=1$ is patients of Marshfield Clinic in the AOI, $j=2$ is non-Marshfield patients in the AOI, $j=3$ is Marshfield patients in the ROS, and $j=4$ is non-Marshfield patients in the ROS. The omitted category is non-Marshfield patients outside of the AOI in 1988. The α_{jt} terms and the vector β are parameters to be estimated. Then using the

estimated coefficients α_{jt}^e we can compute how the differential costs of Marshfield patients in the AOI changed relative to those of non-Marshfield patients elsewhere in the state between 1988 and year t :

$$\Delta_{t,1988} \Delta_{1,4} = [\alpha_{1t}^e - \alpha_{1,1988}^e] - [\alpha_{4t}^e - \alpha_{4,1988}^e] \quad (4)$$

where $\alpha_{4,1988}^e = 0$ by assumption. The advantage of this approach over (2) is that it is agnostic as to when precisely during the 1988-1995 period we should think of the illegal agreements as having been fully phased in: the damage estimate implied by (4) expects the effect to be small in 1988 and to have reached full proportion by 1995, but no assumption is made about the timing pattern whereby the effect materialized.

Table 2 provides details on the data used in the regression analysis. The claims data include 40 million professional-services claims filed by all Blue Cross/Blue Shield subscribers in the state of Wisconsin in the period from 1988 to 1995. Aggregating annual payments for individual patients yields 2.26 million individual-year observations for individuals under age 65. In this period, Blue Cross's fee-for-service plan comprised approximately 10-13% of the statewide market for health insurance.²⁹ As mentioned, the dependent variable of the regression is total real payments for physician services in the given year; while the claims data contain both provider charges and payments made by Blue Cross, the latter are of interest for estimating damages experienced by Blue Cross by virtue of overpaying for physician services in the AOI. In addition to the DID terms, the explanatory variables in the regressions also include: key socio-demographic characteristics of individual patients (gender, age, marital status, relationship to the subscriber), the specific type of insurance plan in which

²⁹ Office of the Commissioner of Insurance, State of Wisconsin, 1996.

the patient was enrolled, and a number characteristics of the county in which the patient lived that could correlate with patterns of utilization of physician services and changes therein over time (real per capita income, the unemployment rate, the share of the population with a high school education or more, population density, birth and death rates, and the share of the population receiving medical assistance). Note that the county-level variables are time-varying and so can help explain variations in payments for physician services both cross-sectionally and over time. Standard errors are estimated using the Huber-White method and are heteroskedasticity-robust.

5. Results

Results from the basic specification of the DID regression in (2) are shown in Table 3. The specification in the left-hand panel uses the definition of Marshfield Clinic patients as those having 50% or more of their annual payments for physician services going to Marshfield Clinic patients, while the specification in the right-hand panel uses the analogous 75% threshold.³⁰ Before examining the coefficients on the DID in terms, it is worth noting that many of the estimated coefficients on the included explanatory variables have signs consistent with expectations and are statistically significant (where the latter is not surprising given the large sample size). For example, *ceteris paribus*, annual payments for physician services tend to rise with the patient's age and are considerably higher for patients in high-risk plans; they also tend to be higher in counties with relatively high income per capita and relatively high population density, but relatively low in counties where the share of the population with a high-school education or more is relatively high. However, the r-squareds of the regressions are quite small, consistent with the low explanatory power generally found in analyses of annual healthcare spending or utilization using individual data.

Turning to the estimated coefficients on the DID terms, the estimates of the level differences across patients of different types in 1988 are consistent with findings from the descriptive statistics shown in Table 1: Using the 50% threshold for defining a Marshfield Clinic patient, spending by such a patient who resided in the AOI was \$299 higher than that for a non-Marshfield patient in the rest of the state (ROS) in 1988, *ceteris paribus*, while spending levels for non-Marshfield patients in the AOI and Marshfield patients in the ROS were \$75 below and \$395 above respectively. While the coefficient for Marshfield patients in the rest of the state exceeds that for Marshfield patients in the AOI (consistent with the idea of patients with more complex conditions or more serious illnesses traveling to the Clinic from greater distances), the difference between the two coefficients is not statistically significant. Estimates using the 75% threshold slightly enlarge the differences between the two groups of Marshfield patients relative to non-Marshfield patients outside of the AOI, while narrowing the difference from

³⁰ Results using the 90% threshold are available upon request.

non-Marshfield patients in the AOI; however, the magnitudes of the differences from the specification using the 50% threshold are not large.

In terms of how the *increases* in payments between 1988 and 1989-95 differed across groups, using the 50% threshold definition, this increase was \$158 greater for Marshfield patients in the AOI than it was for non-Marshfield patients elsewhere in the state, *ceteris paribus*, where the coefficient is statistically significant. For Marshfield patients elsewhere in the state, payments rose by an extra \$55, but the coefficient is not statistically significant. Putting the key number for Marshfield Clinic patients in the AOI in perspective, average spending for such patients rose between 1988 and 1989-95 by \$309 more than it did for non-Marshfield patients elsewhere in the state; the regression result suggests that 51.1% ($\$158/\309) of this increase could be traced to the illegal market-allocation.

In contrast, average payments for non-Marshfield patients in the AOI rose by \$36 more than they did for patients outside of the AOI; this is consistent with the likelihood that the higher prices charged by Marshfield Clinic enabled other providers in the AOI to boost their prices as well, albeit to a much smaller degree. Analogous estimates using the 75% threshold for defining Marshfield Clinic patients differ very little from those using the 50% threshold. They too suggest that the illegal market allocation agreements represented about one-half of the extra increase in prices experienced by Marshfield Clinic patients in the AOI relative to patients elsewhere in the state.³¹ Broadly these findings strongly support the claim that payments for physician services rose more for Marshfield Clinic patients than they did for patients who received services from other providers between 1988 and the 1989-95 period.

Figure 1 shows the key results from the alternative specification in which the four different patient categories are interacted with a full set of year dummies, focusing on the comparison between Marshfield patients in the AOI and non-Marshfield patients in the rest of the state. The solid line shows the estimated differential increases in payments given by (4), and the dotted lines show 95-percent confidence intervals. Results suggest that, while the difference in average annual healthcare costs between Marshfield patients in the AOI and patients in the rest of the state did not change significantly from 1988 to 1991, it widened significantly to \$198 in 1992 and rose further to \$421 in 1995. Averaging across years, the extra increase in average costs for Marshfield clinic patients in the AOI again amounts to \$158, which is again statistically significant. This time profile is consistent with story in which it took time for illegal market-allocation agreements to have broad effects on prices; it is also consistent with the class period of the class-action lawsuit, which computed damages from mid-1992. Broadly these results square well with the simpler specifications that use 1988 as a benchmark. Taken

³¹ Here the average extra increase for Marshfield Clinic patients in the AOI was \$326 between 1988 and 1989-95, while the regression results suggest that \$172 of the increase is attributable to the illegal market-allocation agreements.

together, they show Blue Cross to have paid an extra \$158-172 per Marshfield Clinic patient in the AOI in the 1989-95 period, relative to a benchmark that takes into consideration the other lawful reasons for its higher prices and broad-based increases in spending on physician services in the state during that time.

Conceivably, the extra increases in payments for physician services for Marshfield Clinic patients in the AOI may reflect other types of differential changes in competitive conditions between the AOI and the rest of the state over the period that would enable the Clinic to increase prices at higher rates than providers elsewhere. For example, growth of managed care may have been slower in the AOI than it was elsewhere, so that providers there faced less pressure to moderate growth in prices than providers elsewhere. Alternatively, if Marshfield Clinic's general expansion had tended to increase concentration among healthcare providers in the area, it may have been able to increase its prices more rapidly than providers elsewhere in the state, but by virtue of lawful growth in its market share rather than illegal agreements.

To investigate these possibilities, we extend the basic analysis by including two additional measures of competitive conditions: the Herfindahl index (the sum of squared market shares of firms in a given market) and the HMO enrollment rate, where both are measured at the county level.³² In addition, because it is often mentioned that Marshfield Clinic 'naturally' has relatively high market shares in its core areas because low population density would make it inefficient to have multiple practices in given towns, we re-run the regressions using only observations from those counties having population densities of 200 or less.

Results are shown in Table 4. Adding the Herfindahl index to the regression slightly reduces the estimates of the differential changes in payments for Marshfield Clinic payments in the AOI, while also showing payments per patient to rise as the Herfindahl increases. Including the HMO enrollment rate shifts the DID terms back up while also unexpectedly showing payments for physician services to be higher in counties with relatively high HMO enrollment rates; conceivably, this may reflect a "selection effect" whereby healthy low-cost patients shift into HMOs while less-healthy, higher-cost patients remain in standard insurance plans. Finally, confining the analysis to counties with relatively low population densities causes the magnitudes of the DID term to shift down from around \$155-160 to \$130 in the model with the 50% threshold, and from around \$170 to \$145 in the 75% threshold model; estimated effects remain appreciable and statistically significant. Altogether, then, these modifications leave unchanged the basic finding that payments for physician services rose by significantly more for Marshfield Clinic patients in the AOI compared to non-Marshfield patients elsewhere in the state.

³² See Table 2 for specifics of variable definitions.

6. Discussion: DID in court

A clear problem in using difference-in-difference analyses in antitrust cases is the extent to which judges and juries find it compelling. A recent article asked the question, “Is Antitrust Too Complicated for Generalist Judges?” (Baye and Wright 2011), highlighting possibilities that arguments based on the sorts of relatively complex econometric methodologies used by economists may have trouble withstanding scrutiny of opposing experts, and/or may not be simple and clear enough to be understood and properly interpreted by the court.³³

The use of the DID approach in the Marshfield clinic remand case confronted this issue, with three main problems of interpretation becoming subject to debate. First, it was not clear to the judges who ruled on the case how the DID method controlled for lawful reasons for Marshfield Clinic’s relatively high patient costs. For example, in acting as appeals court judge in the remand case, Judge Posner rendered the opinion that the damage analysis of the Blue Cross expert did not “separate the price effects of collusion from the price effects of the defendants’ lawful market power”³⁴ Posner also upheld the view of the judge in the remand case that Blue Cross had not established that it had been damaged by the Clinic’s unlawful behavior, holding that, “Any non-conspiratorial factors likely to have made the prices charged by the Marshfield Clinic higher than the prices charged by other health-care providers had to be taken into account in order to make a responsible estimate of the prices that Blue Cross would have paid had it not been for the conspiracy.”

Of course, the DID specification does aim to capture persistent, unobserved factors causing payments for physicians services to differ between Marshfield and non-Marshfield patients inside and outside the AOI, for example via the parameters $\{\gamma_1, \gamma_2, \gamma_3\}$ in (2). As these parameters are estimated from price differentials before the illegal behavior began, their causes are by definition assumed to be lawful in nature. With Marshfield thought to be a relatively high-quality provider with a service-intensive treatment style, including these terms is clearly essential for not mis-attributing to illegal behavior that part of the price differential between Marshfield Clinic and other providers that only reflects lawful differences in practice style, quality of care, availability of advanced equipment and facilities, etc. However, the court

³³ As Posner (1999: 96) himself has written, “...econometrics is such a difficult subject that it is unrealistic to expect the average judge or juror to be able to understand all the criticisms of an econometric study, no matter how skillful the econometrician is in explaining the study to a lay audience.” See Sage (1997) on the subject of quantitative analysis in antitrust cases related to managed care.

³⁴ *Blue Cross & Blue Shield v. Marshfield Clinic*, 152 F.3d 588, 593-94 (7th Cir.1998)). Posner also cites himself as precedent in *In RE: Brand Name Prescription Drugs Antitrust Litigation*, 186 F.3d 781 (1999) United States Court of Appeals (7th Cir. 1999).

expected to see specific quantitative measures of such factors on the right-hand side of the regression, and in their absence inferred that account was not being taken of them.

The second problem that arose in the interpretation of the DID results concerned assumptions implied by the approach about *changes* in intensity or quality of care. Marshfield Clinic argued that the DID methodology improperly assumed that Marshfield Clinic's quality remained constant at its 1988 level over the 1989-95 period. Rather Marshfield Clinic argued that its quality continued to improve over time relative to its 1988 *levels*, so that some part of the increased costs for Marshfield Clinic patients in the AOI after 1988 reflected continued improvements in its quality of care rather than illegal exercise of market power. But again, the DID approach does not assume that Marshfield's service intensity and quality of care held constant over time. Rather, it assumes that lawful factors that might be raising Marshfield Clinic's prices were occurring at the same rate as they were elsewhere in the state, as reflected in the term γ_0 in the DID regression. In other words, the 'parallel trend' assumption assumes that the wedge between Marshfield and other healthcare providers due to differences in quality was neither narrowing nor widening over time.³⁵

While this is a maintained hypothesis of the regression specification, one of the experts for Blue Cross presented complementary evidence in support of this assumption. For example, he showed that the share of procedures performed by specialists (as opposed to primary-care physicians) did not increase faster at Marshfield Clinic than elsewhere in the state, as might be expected if the quality of care at Marshfield Clinic was rising more rapidly than it was elsewhere in the state.³⁶ In fact it would have been surprising had Marshfield Clinic's quality increased more rapidly than that of providers around the state after 1988. By then it had clearly established itself as one of the premier tertiary centers in Wisconsin and had an absolute quality advantage over most other healthcare providers in the state; thus, while it is plausible that the Clinic would have continued to invest in improving quality to maintain its premier position, it is difficult to see why it would have invested heavily to further and appreciably enlarge its quality advantages over other providers in the state.³⁷ In any case, this is not what troubled the courts, which took the DID analysis to ignore the possibility that increased payments for Marshfield patients could reflect improvements in the quality of care.

³⁵ Abadie (2005) discusses this parallel trend assumption.

³⁶ For example, a higher share of procedures performed by specialists may correlate with better outcomes because practitioners who perform relatively high volumes of procedures tend to have higher success rates. It may also reflect better local availability of well-credentialed and highly-trained staff.

³⁷ In this sense, to the degree that Marshfield Clinic's level of quality was initially higher than that of other providers, but may grown somewhat more slowly over the subsequent period, the DID damage estimates were conservative.

The third problem concerned the definition of the 'before' period. The appeals court judge for the remand case (again Judge Posner) attempted to summarize the methodologies used by Blue Cross as follows:

The usual way to measure damages in [a market division] case would be to compare the prices that the Marshfield Clinic charged Blue Cross before and during the conspiracy, or inside and outside the region covered by the conspiracy, or during the conspiracy and after it ended ... correcting by various statistical techniques for any non-conspiratorial factors that might have caused the prices that are being compared to be different from each other. ... *This method or congeries of methods was unavailable to Blue Cross, however, because the division of markets embraced the entire period and region in which the necessary data are obtainable; at least Blue Cross made no effort to show otherwise.* Instead it compared the Marshfield Clinic's prices for medical services between 1988 and 1995 with the prices charged by other providers of medical care for the same services during the same period elsewhere in Wisconsin, on the theory that those other prices, properly adjusted, are what Blue Cross would have had to pay the Clinic and the Clinic's competitors had it not been for the conspiracy [references excluded, emphasis added].³⁸

While it is true that market allocation agreements were in effect in some places within the AOI during the entire period for which there was data, it is simply not correct to say that, because some illegal market allocation agreements were already in effect in 1988, the DID analysis boiled down to a 'yardstick' analysis that had no basis for distinguishing between lawful and unlawful causes of the higher payments for physician services made by patients of the Marshfield Clinic in the AOI. On the contrary, it was to solve the problem of bias in the yardstick effect that the DID approach was used. In effect, Posner was disinclined to put any weight on results that could not starkly distinguish between the time and space in which illegal behaviors were held to have taken place, and other times and spaces where they did not. Posner's negative evaluations have colored opinions and assessments of every salient issue considered in this case. In the end, Blue Cross's application for an appeal to the Supreme Court was denied, which effectively brought the essential elements of this case to an end.

7. Conclusions and implications for the use of DID in antitrust

More than 15 years after the *Marshfield Clinic* cases concluded, the healthcare spending and practice patterns that were questioned in the cases remain visible in the data. As of 2006, the Marshfield, Wisconsin, area ranked as the highest-spending commercial healthcare hospital-referral region in the United States (Chernew et al. 2010). As long as healthcare spending continues to receive antitrust attention, it will remain important to devise methods and approaches that properly control for issues of case mix and quality, while also being both

³⁸ *Blue Cross Blue Shield vs Marshfield Clinic* (152 F.3d 588 (1998)).

robust enough to withstand scrutiny by opposing experts and simple and clear enough to be understood and properly interpreted by the courts.

As the difference-in-difference approach has become more widely used, some of the misunderstandings of it that came up in the Marshfield case might not come up today. As discussed in Coleman and Langenfeld (2008), the concept of 'natural experiments' has been increasingly accepted in antitrust analysis. The DID method has been added into the Federal Judicial Training Center curriculum and reference documents. Criteria for the sufficiency of data and facts under *Daubert* requirements have been spelled out (Coleman and Langenfeld 2008). A variety of important papers have explored various issues that arise in implementing DID approaches in antitrust contexts and interpreting their results (e.g. Rubinfeld 2010: 71-75, 2012; Langenfeld et al. 2010).

At the same time, the Marshfield Clinic case points to some important potential pitfalls of relying on the DID method which remain a problem for its successful use as a basis for estimating damages. First is the issue of implicit versus explicit controls for hard-to-measure but lawful factors that contribute to persistent differences in patient costs between the 'treatment' and 'control' groups. As mentioned, the DID method's tactic of controlling implicitly for lawful factors contributing to persistent differences (e.g. via the parameter α_1 in (2)) can

seem out of line with standards of good-quality regression analysis, which in general holds that no important explanatory variable should be omitted from the regression.³⁹ Thus, unlike the relatively parsimonious DID specification used in the Marshfield Clinic case, recent uses of DID in healthcare analyses have tended to include more extensive controls for lawful causes of variations in prices.⁴⁰ Of course, ability to find compelling explicit measures of factors like 'quality' varies across fields of healthcare, for example, tending to be more readily available for hospitals than for physicians.⁴¹ To the extent that inability to control explicitly for quality represents a handicap in estimating damages, it works against plaintiffs and in favor of the defense.

The second potential pitfall concerns the use of changes in payments for physician services in the control group to approximate what would have happened in the treatment group in the

³⁹ Coleman and Langenfeld's (2008: 753-4) discussion of criteria that analyses based on natural experiments must satisfy to be considered valid evidence under *Daubert* standards includes the question of whether the analysis controls for other significant factors.

⁴⁰ For example, Haas-Wilson and Garmon's (2011) retrospective analysis of two hospital mergers in the Chicago area uses individual-level claims data and alternative controls for the resource-intensity of the patient's condition and the severity of his/her case, along with a vector of hospital characteristics and dummy variables indicating the patient's health plan.

⁴¹ See Gaynor and Town (2011: 568-596) for an overview of the literature on 'quality' and hospital competition, where the measure of quality is often mortality or readmission following treatment for a heart attack or other serious condition.

absence of illegal behavior. The 'parallel trend' assumption is anyway one of the more contested aspects of DID analysis (Abadie 2005), and all the more so in antitrust applications where it constitutes the basis of the 'but for' change in prices. In general, to establish the validity of using price changes in the control group to indicate how prices would have changed in the treatment group, one needs to select the control group so as to maximize the odds that supply and/or demand shocks would be similar between the two groups, with effects of given shocks on prices also expected to be similar across the groups (Simpson and Schmidt 2008). As there is no compelling way to test whether a given control group decisively meets these criteria, there can be substantial debate about whether the control group has been correctly chosen.⁴² A potentially valuable strategy here is to run the analysis using alternative definitions of the control groups. For example, Haas-Wilson and Garmon (2011) use four different categories of Chicago-area hospitals as control groups, with and without patients not covered by private health insurance; estimated cost increases varied somewhat according to the control group, but they all fell within an 11-17 percent range and were uniformly statistically significant.⁴³

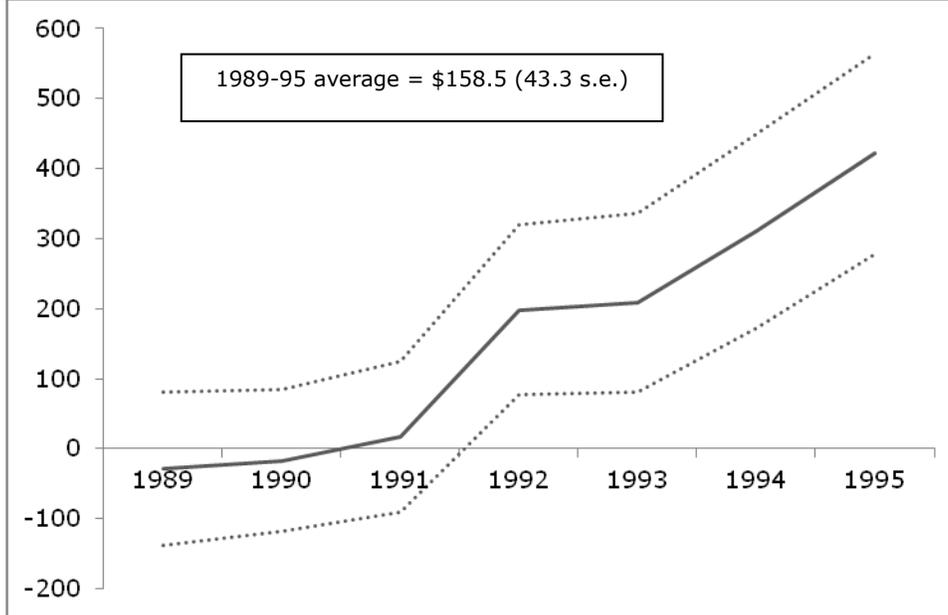
The final issue to mention concerns the role of changes in quality, case-mix or service intensity in contributing to unusual increases in spending on healthcare services. As discussed by Hammer and Sage (2002: 564), defendants in healthcare-antitrust cases commonly invoke claims about rising quality to explain why their patient costs rose unusually after a merger or during a period when they were engaged in unlawful behavior. Such claims can be difficult to support or refute, both because it is difficult to define and measure 'quality' and because different dimensions of quality may change in complex ways in periods when providers are re-organizing their services.⁴⁴ Nonetheless, given the predictability with which this issue is offered as an objection to DID estimates of the 'but for' price change, compelling estimates of damages from illegal behavior have to address the possible role of quality in accounting for differential changes in costs between the treatment and control groups. Moreover, as the Marshfield Clinic case illustrates, it is just as important to ensure that findings with respect to quality of care and related lawful factors are explained clearly and intuitively, so that their implications for estimated damages are clear to the courts.

⁴² See, for example, Adams and Noether's (2011) comment on Haas-Wilson and Garmon's (2011) analysis, as well as Garmon and Haas-Wilson's (2011) reply.

⁴³ As they write, "If the empirical results are robust across control groups, then it is less likely the estimated treatment effect is driven by unobserved characteristics in the control group" (Garmon and Haas-Wilson 2011: 43). This approach is consistent with Coleman and Langenfeld's (2008: 759-772) elaboration of characteristics that analyses based on natural experiments need to have to meet *Daubert* criteria.

⁴⁴ Mutter, Romano and Wong's (2011) careful analysis of changes in various dimensions of quality following hospital mergers finds very little in the way of systematic patterns, leading them to conclude that antitrust authorities need to evaluate implications of hospital mergers for quality of care on a case-by-case basis.

Figure 1. Current year's difference between average annual costs of Marshfield clinic patients in the AOI vs. non-Marshfield patients in the rest of the state, minus the difference in 1988 (in constant 1995 dollars)



Note: Dotted lines represent 95% confidence intervals. Standard errors are computed robustly.

Table 1. Average annual payments for physician services, 1988-95, constant 1995 dollars; 50%, 75% and 90% thresholds used to define Marshfield patients						
	% of total patient-year records			Average annual payments ('95 \$)		
	Threshold share:			Threshold share		
	50%	75%	90%	50%	75%	90%
<i>Patients residing in:</i>						
Area of influence	4.8	4.8	4.8	705	705	705
Marshfield patients	1.1	1.0	0.9	1,112	1,106	1,080
Other patients	3.7	3.8	3.9	577	596	616
Rest of state	95.2	95.2	95.2	651	651	651
Marshfield patients	0.6	0.5	0.4	1,187	1,158	1,067
Other patients	94.6	94.7	94.8	648	649	650
Threshold share refers to the share of total annual payments for physician services that went to Marshfield clinic providers that was used to define Marshfield patients.						

Table 2. Variable definitions	
Variable	Definition
Average payments per patient	Average annual payments for physician services, constant 1995 dollars
<i>DID terms</i>	
D1989-95	=1 if year is 1989-95; 0 otherwise
Marshfield in AOI	=1 for Marshfield patients in the AOI; 0 otherwise
Non-Marshfield in AOI	=1 for Non-Marshfield patients in the AOI; 0 otherwise
Marshfield in ROS	=1 Marshfield patients in the ROS; 0 otherwise
Non-Marshfield in ROS	=1 Non-Marshfield patients in the ROS; 0 otherwise (omitted)
<i>Individual characteristics</i>	
Female	=1 if individual is female; 0 otherwise
Age	Age of individual
Married	=1 if individual is married & over 18; 0 otherwise
Spouse	=1 if individual is spouse of subscriber; 0 otherwise
Dependent	=1 if individual is other dependent of subscriber; 0 otherwise
<i>Plan types:</i>	
FFS/PPO	Standard fee-for-service/PPO insurance plan (omitted)
Administrative services only	BCBS receives fixed rate to administer claims for another insurer
Cost-Plus Contracts	Plans administered by BCBS that reimburse providers for costs plus a specified margin
Federal Employee Program	BCBS plan offered to Federal Employees
National Programs	Other traditional BCBS insurance plans based elsewhere
High Insurance Risk Sharing	State-organized plan offering insurance to people unable to obtain coverage due to preexisting conditions
Other BCBS plans	BCBS plans other than those listed above
<i>County characteristics*</i>	
Per capita income	Per capita personal money income ('000 of constant 1995 \$)
Unemployment rate	Unemployment rate (%)
% of pop. w/HS education or more	% of county population with high school education or more
Log of population density	Log of population per square mile (in '00)
Birth rate	Births per 1,000 population
Death rate	Deaths per 1,000 population
Medical assistance per capita	Per capita medical-assistance payments ('000 of constant 1995 dollars)
Log Herfindahl	Log of the sum of squared market shares of all providers used by county residents, measured in terms of billing events
HMO enrollment	Share of county's population enrolled in an HMO
Time dummies	=1 for given year; zero otherwise
*Data from U.S. Bureau of Economic Analysis (per capita income); U.S. Department of Health and Human Services' Area Resource File (unemployment, education), Wisconsin Interactive Health Statistics (population density, birth and death rates), the State of Wisconsin's <i>Blue Book</i> (medical assistance per capita), and the BCBS claims data base (Herfindahl)	

Table 3. Determinants of annual payments for physician services in real 1995 dollars: 50% and 75% thresholds for identifying Marshfield patients				
	50% threshold		75% threshold	
	Coeff.	s.e.	Coeff.	s.e.
Constant	542.3*	(52.1)	522.1*	(52.0)
<i>Geography, patient type, & period :</i>				
D1989-95	25.2*	(3.8)	25.4*	(3.8)
Marshfield patients in AOI	299.1*	(39.3)	280.4*	(42.3)
Marshfield patients in AOI x D1989-95	158.1*	(43.3)	171.9*	(46.5)
Non-Marshfield patients in AOI	-75.5*	(10.0)	-63.8*	(10.1)
Non-Marshfield patients in AOI x D1989-95	35.6*	(11.4)	38.7*	(11.6)
Marshfield patients in ROS	394.8*	(58.6)	398.6*	(66.4)
Marshfield patients in ROS x D1989-95	54.7	(63.7)	26.0	(72.2)
<i>Controls for individual characteristics:</i>				
Male 5-17	-131.5*	(4.4)	-131.4*	(4.4)
Male 18-24	-37.0*	(6.2)	-37.1*	(6.2)
Male 25-34	9.5	(7.1)	9.3	(7.1)
Male 35-44	129.3*	(7.2)	129.1*	(7.2)
Male 44-54	428.1*	(8.7)	428.4*	(8.7)
Male 55+	939.3*	(11.1)	940.0*	(11.1)
Female <5	-76.2*	(5.5)	-76.3*	(5.5)
Female 5-17	-138.3*	(4.4)	-138.4*	(4.4)
Female 18-24	65.7*	(5.3)	65.7*	(5.3)
Female 25-34	290.2*	(6.8)	290.2*	(6.8)
Female 35-44	280.1*	(7.0)	280.1*	(7.0)
Female 45-54	416.6*	(7.7)	417.0*	(7.7)
Female 55+	636.6*	(8.8)	637.3*	(8.8)
Married	44.4*	(8.8)	44.4*	(8.8)
Spouse	-12.5*	(9.0)	-12.5	(9.0)
Dependent	25.8*	(7.2)	25.6*	(7.2)
<i>Insurance plan type:</i>				
Administrative Services Only	-1.8	(3.5)	-1.5	(3.5)
Cost-Plus Contracts	8.7*	(4.7)	9.0*	(4.7)
Federal Employee Program	299.2*	(14.6)	299.7*	(14.6)
National Programs	-87.2*	(5.7)	-87.8*	(5.7)
High Insurance Risk Sharing	1,057.9*	(34.9)	1,060.9*	(34.9)
Other BCBS plans	-27.4*	(3.0)	-27.2*	(3.0)
<i>Controls for county characteristics:</i>				
Per capita income	1.9*	(0.9)	1.9*	(0.9)
Unemployment rate	3.8*	(1.8)	4.1*	(1.8)
% of pop. w/HS education or more	-3.0*	(0.8)	-2.8*	(0.8)
Log of population density	20.6*	(2.1)	19.7*	(2.1)
Birth rate	-7.2*	(1.6)	-6.9*	(1.6)
Death rate	-5.1*	(1.2)	-4.5*	(1.2)
Medical assistance per capita	2.5	(6.0)	3.8	(6.0)
Year dummies	Yes		Yes	
R-squared	0.0401		0.0401	
* = statistically significant at 5% level. + = statistically significant at 10% level. Robust standard errors in parentheses. N=2,264,308				

Table 4. Estimated effects of competition-related variables: Alternative specifications using 50% and 75% thresholds				
	Whole state			Counties with population density <200/sq mi.
	(a)	(b)	(c)	(d)
50% threshold				
Differential increase for Marshfield patients in the AOI	158.1* (43.3)	156.0* (43.3)	159.3* (43.3)	131.7* (43.6)
Log Herfindahl		4.2+ (2.3)	4.9* (2.3)	-7.6+ (4.6)
HMO enrollment			51.1* (13.0)	-29.8 (31.6)
75% threshold				
Differential increase for Marshfield patients in the AOI	171.9* (46.5)	169.3* (46.5)	172.9* (46.5)	144.5* (46.8)
Log Herfindahl		5.2* (2.3)	6.0* (2.3)	-3.3 (4.6)
HMO enrollment			56.0* (13.0)	-21.5 (31.6)

Notes: *= significant at a 5% level or better. += significant at 10% level or better. All regressions include individual and county-level characteristics as explanatory variables. The number of observations is 2,264,308 in specifications (a)-(c) and 612,242 in (d). Heteroskedasticity-robust standard errors in parentheses.

References

- Abadie, Alberto. 2005. "Semiparametric difference-in-difference estimators," *Review of Economic Studies*, 72(1): 1-19.
- Adams, Gregory and Monica Noether. 2011. "Comment on 'Hospital Mergers and Competitive Effects: Two Retrospective Analyses'," *International Journal of the Economics of Business*, 18(1): 33-40.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, 24(2): 3-30.
- Baye, Michael R., and Joshua D. Wright. 2011. "Is Antitrust Too Complicated for Generalist Judges? The Impact of Economic Complexity & Judicial Training on Appeals," *Journal of Law and Economics*, 54(1): 1-24.
- Bornstein, Severin. 1990. "Airline Mergers, Airport Dominance, and Market Power," *American Economic Review*, 80(2): 400-404.
- Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. New York: Cambridge University Press.
- Chernew, Michael D., Lindsay M. Sabik, Amitabh Chandra, Teresa B. Gibson, and Joseph P. Newhouse. 2010. "Geographic Correlation between Large-Firm Commercial Spending and Medicare Spending," *American Journal of Managed Care*, 16(2): 131-137.
- Coleman, Mary and James Langenfeld. 2008. "Natural Experiments," *Issues in Competition Law and Policy*, ABA Section of Antitrust Law, Vol. 1: 743-772.
- Coombs, Jan G. 2005. *Rise and Fall of HMOs: An American Health Care Revolution*. Madison, WI: University of Wisconsin Press.
- Crabb, Judge Barbara. 1997. "Order and Opinion, Rozema v. Marshfield Clinic." 977 F.Supp 1362 (W.D. Wisc 1997), Oct. 2.
- Garmon, Christopher and Deborah Haas-Wilson. 2011. "The Use of Multiple Control Groups and Data Sources as Validation in Retrospective Studies of Hospital Mergers," *International Journal of the Economics of Business*, 18(1):41-44.
- Gaynor, Martin and Robert Town. 2011. "Competition in Health Care Markets." In Mark Pauly, Thomas McGuire, and Pedro Barros, eds., *Handbook of Health Economics*. Oxford: North Holland/Elsevier, pp. 499-637.

- Greenberg, Warren. 1998. "Marshfield Clinic, Physician Networks, and the Exercise of Market Power," *Health Services Research* 33:1461-1476.
- Hammer, Peter J., and William M. Sage. 2002. "Antitrust, Health Care Quality, and the Courts," *Columbia Law Review*, 103(3): 545-649.
- Haas-Wilson, Deborah. 2003. *Managed Care and Monopoly Power*. Cambridge, MA: Harvard University Press.
- _____ and Christopher Garmon. 2011. "Hospital Mergers and Competitive Effects: Two Retrospective Analyses," *International Journal of the Economics of Business*, 18(1): 17-32.
- Hastings, Justine. 2004. "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California," *American Economic Review*, 94(1): 317-328.
- Imbens, Guido W. and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47(1): 5-86.
- Jiménez, Juan Luis and Jordi Perdiguer. 2009. "Mergers, Differences-in-Differences and Concentrated Markets: Why Firms Do Not Increase Prices?" Accessed electronically at: <http://www.ub.es/graap/merger.pdf>
- Langenfeld, James, Wenqing Li, Greg Leonard, and John Morris. 2010. "Econometrics and Regression Analysis." In *Proving Antitrust Damages: Legal and Economic Issues*, 2nd edition. Chicago, IL: American Bar Association, pp. 125-192.
- Laporte, A., and F. Windmeijer. 2005, "Estimation of Panel Data Models with Binary Indicators when Treatment Effects are not Constant over Time," *Economics Letters*, 88(3): 389-396.
- McCarthy, Thomas and Scott Thomas. 1995. "Analyzing damages in health care antitrust cases." NERA unpublished working paper, <http://www.nera.com/extImage/3938.pdf>
- McCluer, R. Forrest and Martha Starr. 2013. "Disaggregating prices and quantities in health care antitrust damages." American University working paper.**
- McCrary, Justin and Daniel L. Rubinfeld. 2010. "Measuring Benchmark Damages in Antitrust Litigation," Paper prepared for a workshop on quantification of antitrust damages organized by the European Union, Directorate-General of Competition (Jan.), http://ec.europa.eu/competition/antitrust/actionsdamages/rubinfeld_mccrary.pdf
- Mutter, Ryan, Patrick Romano, and Herbert Wong. 2011. "The Effects of U.S. Hospital Consolidations on Hospital Quality," *International Journal of the Economics of Business*, 18(1): 109-126.

- Posner, Richard A. 1999. "The Law and Economics of the Economic Expert Witness," *Journal of Economic Perspectives*, 13(2): 91-99.
- Romano, Patrick and David Balan. 2011. "A Retrospective Analysis of the Clinical Quality Effects of the Acquisition of Highland Park Hospital by Evanston Northwestern Healthcare," *International Journal of the Economics of Business*, 18(1): 45-64.
- Royall, M. Sean. 1997. "Disaggregation of Antitrust Damages." *Antitrust Law Journal*, 65: 331-352.
- Rubinfeld, Daniel L. 2010. "Econometric issues in antitrust analysis," *Journal of Institutional and Theoretical Economics*, 166(1): 62-77.
- _____. 2012. "Antitrust Damages." In Einer Elhauge, ed., *Research Handbook on the Economics of Antitrust Law*. Northampton, MA: Edward Elgar, pp. 378-394.
- Sage, William M. 1997. "Judge Posner's RFP: Antitrust Law and Managed Care," *Health Affairs*, 16(6): 44-61.
- Shabaz, John C. 1995. "Order, Blue Cross & Blue Shield United of Wisconsin and Compcare Health Services v. Marshfield Clinic and Security Health Plan of Wisconsin," U.S. District Court, Western District of Wisconsin (March 22).
- Simpson, John and David Schmidt. 2008. "Difference in Differences Analysis in Antitrust: A cautionary note," *Antitrust Law Journal*, 75(2): 623-636.
- _____ and Christopher Taylor. 2008. "Do Gasoline Mergers Affect Consumer Prices? The Marathon Ashland Petroleum and Ultramar Diamond Shamrock Transaction," *Journal of Law and Economics* 51(1): 135-152.
- Tenn, Steven. 2011. "The Price Effects of Hospital Mergers: A Case Study of the Sutter-Summit Transaction," *International Journal of the Economics of Business*, 18(1): 65-82.
- U.S. Bureau of Labor Statistics, 2009. "BLS Spotlight on Statistics: Healthcare." (Nov.) Accessed electronically at http://www.bls.gov/spotlight/2009/health_care/pdf/health_care_bls_spotlight.pdf.
- Vita, Michael G., and Seth Sacher. 2001. "The competitive effects of not-for-profit hospital mergers: A case study," *Journal of Industrial Economics*, 49(1): 63-84.
- Vogt, William and Robert Town. 2006. "How has hospital consolidation affected price and quality of hospital care?" Robert Wood Johnson Foundation, Synthesis Project Issue No. 9 (Feb.). <http://www.rwjf.org/pr/product.jsp?id=15231>
- White, Halbert, Robert Marshall, and Pauline Kennedy. 2006. "The Measurement of Economic Damages in Antitrust Civil Litigation," American Bar Association, Antitrust Section, Economic Committee Newsletter, 6: 17-22.

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