Before-and-after analysis: An application of structural break testing to the determination of economic damages

Logan Kelly, Ph.D.[a] & David Sienko, Esq.[b]

[a] *College of Business and Economics, University of Wisconsin, logan.kelly@uwrf.edu*

[b] *Hammarback Law Office, River Falls, WI, dsienko@hammarback-law.com*

January 2018

PRELIMINARY DRAFT

Abstract

Before-and-after analysis is a highly reliable method for determining economic damages and providing support of causation claims in business interruption cases that has been accepted by courts in nearly every jurisdiction, but rudimentary methods, e.g. graphical analysis, simple correlation analysis, etcetera, are not foolproof. We discuss three problems that may arise and how structural break analysis may be used to solve these problems.

Before-and-after analysis: An application of structural break testing to the determination of economic damages

January 2018

PRELIMINARY DRAFT

Abstract

Before-and-after analysis is a highly reliable method for determining economic damages and providing support of causation claims in business interruption cases that has been accepted by courts in nearly every jurisdiction, but rudimentary methods, e.g. graphical analysis, simple correlation analysis, etcetera, are not foolproof. We discuss three problems that may arise and how structural break analysis may be used to solve these problems.

1 Introduction

Before-and-after analysis is a highly reliable method for determining economic damages and providing support of causation claims in business interruption cases (Gaughan 2009) that has been accepted by courts in nearly every jurisdiction. See, e.g., Lloyd (2014) for an extensive list of court decisions regarding the before-and-after method. Broadly speaking this method refers to the comparison of profits during a period without damage, referred to as the "benchmark period," and profits during the period when the business interruption is occurring, referred to as the "damage period."[1]

While the before-and-after method is highly reliable and widely accepted, it is not foolproof. As Tomlin and Merrell (2006) demonstrate, "simple" forms of the before-and-after method may yield "phantom" damages arising from inaccurate or incomplete calculations. This contrasts with the "commonly held perception that simple methods of calculating damages are often more understandable and persuasive to a judge or jury than more complex methods (Tomlin and Merrell 2006, 295)." Tomlin and Merrell go on to conclude,

[1] See Lloyd (2014) and Gaughan (2009), chap. 2 for more a detailed definition.

“Our results indicate that simple methods for calculating lost profits that do not take into account demand and supply factors are capable of being highly inaccurate. In addition, this article has shown that damages methodologies in general are highly manipulable.” Tomlin and Merrell (2006, 321)

Rudimentary methods, e.g. graphical analysis, simple correlation analysis, etcetera, are more commonly employed in forensic analysis than is sophisticated time series analysis, such as structural break models (Gaughan 2009). The simplistic approach, however, may run into three problems that we label: the overt omitted variable problem, the multiple comparisons problem, and the latent omitted variable problem. Fortunately, the court has long allowed more sophisticated methodology, see e.g. Vuyanich v. Republic Nat. Bank of Dallas (1984).

Sophisticated time series models, such as those accounting for structural breaks, do require more data to estimate and more effort to explain, but when appropriately applied they can add powerful evidence to a case. Therefore, in this article, we analyze the problems that may arise from before-and-after analysis, discuss under what conditions those problems may occur, and demonstrate how one possible path to solving these problems may be found in the structural break literature.

2 The problems

The before-and-after method of proving damages can run into three problems that we attempt to solve in this paper: (i) the overt omitted variable problem, (ii) multiple comparisons problem, and (iii) latent omitted variable problem. In this section, we define and provide illustrations of each problem.

2.1 Overt omitted variable problem

The overt omitted variable problem occurs when obviously influential explanatory factors are omitted from the analysis. It is often natural to compare the performance of a business in an undamaged state, i.e. either before the damage occurred or after the damage has been remedied, to the performance of the business during the period of damage (Gaughan 2009, 50). However, it is necessary to account for explanatory factors that may differ between periods. See Lloyd (2014) for a survey of case law regarding the analysis confounding factors.

An example of an expert opinion that failed to account possible different external factors is found in Katskee v. Nevada Bob’s Golf of Nebraska, Inc. (1991). In this case, Nevada Bob's Golf of Nebraska, Inc. counterclaimed that the plaintiff denied its right of first refusal to lease and occupy adjacent space, thereby diminishing its profits. The expert for Nevada Bob's Golf of Nebraska, Inc. estimated lost profit by simply calculating profit per square foot at the alternative location leased and multiplied that by the square footage of the adjacent space. The court rejected this analysis stating,

> *“The expert assumed that the only difference between the two locations was the square footage. No studies or comparisons were made as to differences in the customer base, relative accessibility of the facilities, proximity to recreation areas or other shopping areas, parking, or any other external factors.” Katskee v. Nevada Bob’s Golf of Nebraska, Inc. (1991, 472:379)*

Within the context of econometric modeling, the overt omitted variable problem is simply the omitted variable bias that may occur in regression models when important explanatory variables are omitted, see e.g. Greene (2003). It is import to consider, though, that this problem is not limited to applications of regression. Standard graphical and correlation analysis generally compare only two variables at a time. Examining the co-movement of a pair of variables in isolation from other explanatory factors can lead to erroneous conclusions about the structural relationship between the variables in question. For example, it can be shown that baseball player salary and number of strikeouts in a season are positively correlated. It would clearly be erroneous to then conclude that poor hitters earn higher salaries. More formally, Tomlin and Merrell (2006) demonstrate how easily before-and-after analysis can be used to show loses even in the absence of actionable conduct.

2.2 Multiple comparisons problem

Multiple comparisons problem arises when a statistical analysis encompasses a number of formal comparisons, with the presumption that attention will focus on the strongest differences among all comparisons that are made (see e.g. Miller 1981; Benjamini 2010). Erica P. John Fund, Inc. v. Halliburton Company (2015) is a recent example of the court dealing with multiple comparisons problem.

In this case, Erica P. John Fund, Inc. (EPJ Fund) alleged Halliburton made a series of misrepresentations in an attempt to inflate the price of its stock. Halliburton subsequently made a number of corrective disclosures causing its stock price to fall and EPJ Fund. to lose money (see also Halliburton Co. v. Erica P. John Fund, Inc. 2014). Both EPJ Fund and Halliburton provided export reports regarding the impact of the corrective disclosures on Halliburton stock price. Halliburton's expert, Lucy Allen, argued that the multiple comparison issue arose because a large number of price reactions were tested for statistical significance. Chad Coffman argued that Allen's use of a multiple comparison adjustment is novel, improper, and yields erroneous results, because it results in unacceptably high false negatives. The court concluded,

> *"The Court is persuaded that the use of a multiple comparison adjustment is proper in this case because of the substantial number of comparisons, thirty-five comparisons, being tested for statistical significance… Moreover, there is the unverified, but not entirely refuted, specter that Mr. Coffman's predecessor, Ms. Nettesheim, selected her dates by looking for statistically significant dates and then looking for Halliburton-specific news on those dates, from which Mr. Coffman selected the six events in his expert report." Erica P. John Fund, Inc. v. Halliburton Company (2015)*

As can be seen from the court's statement, the danger of appearing to cherry pick breakpoints in before-and-after analysis is real. This is especially true when the economist is called upon to testify about the timing of the damage period. Determining the point when the damage period begins and/or ends requires implicitly, if not explicitly, the testing of every possible breakpoint. Moreover, in the case of an expert who lacks significant training in time series analysis, that expert might not even realize there is a problem. The simple use of graphical analysis to determine a breakpoint is highly susceptible to this criticism.

2.3 Latent omitted variable problem

The latent omitted variable problems occur when there exists an unknown or unobserved factor that is at least partially responsible for changes that occur during the damage period. An example of clear recognition of this problem by the court can be found in the United States District Court decision in re Live Concert Antitrust Litigation (2012), referred to hereafter as Live Concert Antitrust. In this case, Clear Channel Communications, Inc. et al. (Clear Channel hereafter) is alleged to have "engaged in anticompetitive, predatory, and exclusionary practices in an effort to acquire, maintain and extend its monopoly power in a national ticket market for live rock concerts" (Heerwagen v. Clear Channel Communications 2004, 435:223).[2]

In the Live Concert Antitrust decision, the court excluded the testimony of an economist who supported his before-and-after analysis with a simple regression model. The plaintiff's expert, Owen R. Phillips, used ticket prices for all liver rock music concerts in the Denver and Los Angeles markets from 1981 through 1998 to predict the expected average ticket prices from 2000 through 2006. Phillips then compared expected average ticket prices to the actual average ticket prices, and concluded that since actual ticket prices were higher than what Phillips' model predicted, Defendants' market entry caused an increase in ticket prices (In re Live Concert Antitrust Litigation 2012, 863:996–97).

> *"Dr. Phillips' 'Before-and-After' analysis is significantly less robust than the analysis rejected by the Court in In re REMEC Inc. Sec. Litig.[3] There, the expert's analysis accounted for two independent variables (the 'market index' and the 'industry index'), in addition to the corrective disclosures that allegedly caused the declines in stock price. Here, Dr. Phillips' 'Before-and-After' analysis accounts for no independent variables other than time. Moreover, in In re REMEC Inc. Sec. Litig., the expert observed stock declines after several corrective disclosures by the defendants, arguably supporting his hypothesis that the disclosures were, in fact, the cause of the declines. Here, only one 'event' (i.e., the entry of Defendants into the market in 2000) was considered. Finally, as discussed in more detail below, Dr. Phillips improperly excluded data from the year 1999 from his analysis." (In re Live Concert Antitrust Litigation 2012, 863:978)*

The court excluded Philips' testimony stating while a model need not include all measurable variables to be admissible (Bazemore v. Friday 1986), a model may be too incomplete to be admissible (Bickerstaff v. Vassar College 1999).

In his rebuttal report, Phillips conducted a structural break test at the year 2000 using a sample including concerts in both Los Angeles and Denver for the complete sample period (1986-2006). The court rejected this analysis as well stating,

---

[2] Heerwagen v. Clear Channel Communications (2004) sought national class action status. The United States Court of Appeals, Second Circuit upheld the lower court's denial of national class action status, citing, among other reasons, that monopolization claims should be proved with reference to a specific market. Litigation in several districts, including United States District Court, C.D. California, ensued.

[3] In In re REMEC Inc. Securities Litigation (2010) the expert testimony was excluded on similar grounds.

> *“This ‘pooled sample’ analysis is fatally flawed for the same basic reason discussed above; namely, it fails to account for any ‘major factors’ (other than Defendants' entry into the market in 2000) that could have caused and/or contributed to an increase in ticket prices. Moreover, the ‘pooled sample’ analysis fails even to consider whether a so-called ‘structural break’ occurred in any year other than 2000. To the contrary, Dr. Phillips concedes that there may be ‘structural breaks’ in other years; he simply did not test for them.”* Live Concert Antitrust *(2012, 863:982–83)*

The court was clearly concerned about overt omitted variables, but in the court's assessment of Phillips’ analysis, we see two references to the latent omitted variable problem. First, in its assessment of Phillips’ initial report, the court is critical of only one event being considered. Second, in its assessment of Phillips’ rebuttal report, the court is critical of testing for only one structural break date (In re Live Concert Antitrust Litigation 2012). If the actual break occurred in 1999, then a Chow test would likely indicate that a break in 2000 is significant. Thus, testing for a single breakpoint is likely insufficient to establish causation.

3 Structural break testing: solving the problems

3.1 The overt omitted variables problem: Testing for a known structural break

We can deal with the overt omitted variables problem by including all major factors that are measurable in the model. Chow (1960) proposes a method which allows for the comparison of a benchmark period and a damage period and the inclusion of multiple explanatory variables.

Suppose we have a dataset of T observations, N explanatory variables and wish to test if the relationship between the explanatory variables and the dependent variable changes at period t_0, then consider the following simple implementation of a Chow (1960) structural change test. Let y is a $T \times 1$ vector of observations of the dependent variable, X is a $T \times N$ matrix of explanatory variables, Γ is a $T \times 1$ vector indicating the hypothesized structural change, i.e.

$$\Gamma = \begin{cases} 0 & t \in (0, t_0 - 1) \\ 1 & t \in (t_0, T) \end{cases}, \tag{1}$$

and Z be a subset of regressors whose parameter will be allowed to vary between regimes. The Chow test can be conducted first by estimating the following equation

$$y = \begin{bmatrix} X & \Gamma^\circ Z \end{bmatrix} \begin{bmatrix} \beta \\ \gamma \end{bmatrix} + \varepsilon , \tag{2}$$

where $\Gamma^\circ Z$ Hadamard, or element-wise, product, and then testing the hypothesis $\gamma = 0$. Rejecting this hypothesis indicates parameter instability, or a structural break.

It is worth noting that this is not the typical textbook version of the Chow test, see e.g. Green (2003, 130). However, this version yields two advantages for forensic analysis. First, it produces individual t-statistics for each parameter that is allowed to vary. Thus, an analysis of the stability of each parameter in the model is possible. Second, equation (2) may be used to determine normal values of y but for the damage that occurred.

Given that multiple explanatory factors may be included in (2), the Chow method avoids the overt omitted variable problem of simple before-and-after analysis that may lead to the exclusion of testimony as in Katskee v. Nevada Bob's Golf of Nebraska, Inc. (1991). This method does not, however, solve the multiple comparisons problem in that the validity the Chow test is dependent on knowing, a priori, the date of the structural change. Therefore, if the break date is chosen via graphical or other data based analysis, then the test statistic resulting from the Chow test has a non-standard distribution. Moreover, unless the existence of an unknown or unobserved factor that can explain any structural breakpoints can be eliminated, testing a single breakpoint can provide only weak evidence in an argument for causation.

3.2 The multiple comparisons problem: Testing for an unknown breakpoint

In order to solve the multiple comparisons problem, we must be able to account for the fact that when testing for the existence of a breakpoint, we may (either implicitly or explicitly) be selecting the one breakpoint out of many possible breakpoints that is most likely to yield a statistically significant result. Thus to solve this problem we need to be able to test for a breakpoint that is not known in advance.

Quandt (1960) proposed generalizing the Chow method to test for an unknown breakpoint by first calculating Wald statistic for each of the possible breakpoints within a range of dates within the sample. Then finding the supremum of the individual Wald statistics (sup-Wald statistic hereafter). However, this approach clearly suffers from the multiple comparisons problem; thus, the sup-Wald statistic calculated has a non-standard distribution. Andrews (1993) derived the distribution of this sup-Wald statistics,[4] solving the multiple comparisons problem, and Hansen (1997) determined the approximate p-values for the sup-Wald statistic. Thus, the Quandt-Andrews sup-Wald test both avoids the overt omitted variable problem and accounts for multiple comparisons problem,[5] but this method does not account for the latent omitted variable problem as Quandt-Andrews tests for a single breakpoint.

3.3 The latent omitted variable problem: Testing for multiple unknown structural breaks

In the Live Concert Antitrust decision, the court was concerned about testing for a single breakpoint. Implicitly, the court was concerned that the change in ticket price was due to a factor for which the expert for the plaintiff did not account. Therefore, to account for an unobserved, or possibly unobservable, factors we must allow for multiple unknown breakpoints. Moreover, we must allow for more breakpoints than what the facts of the case suggest as a prior.

Bai (1997) and Bai and Perron (1998) further extend the Quandt-Andrews test to allow for multiple unknown breakpoints. Consider a standard multiple linear regression with T observations and m potential breakpoints and $m+1$ potential regimes. For regimes $j = 0,\ldots,m$, define $T_j \in \{T\}_m$ to be the first date of each regime. Then for the j^{th} regime, i.e. the subsample $T_j,\ldots,T_{j+1}-1$ we have

$$y_t = X_t{}'\beta + Z_t{}'\gamma_j + \varepsilon_t, \tag{3}$$

where X is a matrix of regressors whose parameters are regime invariant and Z is a matrix of regressors whose parameters are allowed to vary between regimes.

Global Maximizer Tests. Bai and Perron (1998) develop a procedure for testing the hypothesis of m breaks versus 0 breaks. They begin by obtaining the least squares estimate

[4] Andrews (1993) also showed that a sup-LM test to be asymptotically equivalent to the sup-Wald test; however, Vogelsang (1999) showed that in finite samples for the special case of a change in mean the sup-LM test can lose power rapidly as the change in mean increases. Hence our focus on the sup-Wald test.

[5] An alternative solution to the multiple comparisons problem is the CUSUMS test, but Andrews (1993) and others have noted that the CUSUMS test suffers serious power problems.

(β, γ) for all possible sets of breakpoints, and then select the optimum set of breakpoints by choosing the set of breakpoints, $\{\hat{T}\}_m$, that minimizes the sum of squared residuals across all possible sets of breakpoints,[6] i.e. they choose $\{\hat{T}\}_m$ to minimize

$$S\left(\hat{\beta}, \hat{\gamma} \middle| \{\hat{T}\}_m\right) = \sum_{j=0}^{m} \left[\sum_{t=\hat{T}_j}^{\hat{T}_{j+1}-1} y_t - \left(X_t{}'\hat{\beta} + Z_t{}'\hat{\gamma}_j \right) \right], \tag{4}$$

where $(\hat{\beta}, \hat{\gamma})$ are the least squares estimate of (β, γ) for a given set of breakpoints. Next Bai and Perron formulate a test of m breaks versus 0 breaks by calculating an F-statistic to evaluate the null hypothesis that $\hat{\gamma}_0 = \hat{\gamma}_1 = \ldots = \hat{\gamma}_m$.

Double maximum testing. If the number of breaks is unknown, then Bai and Perron (1998) show it is possible to test the null of no structural break versus an unknown number of breakpoints up to some upper bound by extending the above procedure to include various values of m. In other words, the global maximize F-statistic is calculated for $l = 1, \ldots, m$ breaks. Then these test statistics are aggregated either by selecting the maximum value, i.e. UDMax test statistic (see e.g. Andrews, Lee, and Ploberger 1996), or by using a weighting scheme, i.e. WDMax test statistic (see Bai and Perron 1998). This type of testing, referred to as double maximum testing, results in a test statistic with a non-standard distribution for which Bai and Perron (2003b) provide critical values.

Sequential tests. Finally, Bai and Perron (1998) develop a test of l versus $l+1$ breaks, which can be used as the basis of a sequential testing procedure to estimate the number of breakpoints. For a model with l breakpoints, each of the $l+1$ regimes is tested for an additional breakpoint. One can then sequentially test $l = 0$, $l = 1, \ldots$ until a non-rejection occurs. There also exist information criteria methods (see e.g. Liu, Wu, and Zidek 1997; Yao 1988); however

[6] For large T and/or m, this process is computationally intensive. Bai and Perron (2003a) develop practical algorithms for computing global optimizers for multiple breakpoint models.

simulation results of Bai and Perron (2006) show that sequential testing of l versus $l+1$ breakpoints preforms much better than the information criteria methods.

General procedure. Though the sequential test outperforms other methods, Perron (2006) warns that it is possible that these tests will select fewer than the true number of breaks, and thus should not be used mechanically. Perron recommends first using a double maximum test to determine if any breaks are present, and he argues that it is important to begin with the double maximum test for three reasons:

(i) Some types of structural breaks are difficult to detect with single breakpoint models. This is especially relevant to before-and-after analysis as it is particularly difficult to detect change when the first and third regimes are identical. This may well be the case if the sample contains an undamaged period both before and after the damage period.

(ii) Tests for a fixed number of breakpoints may have power problems if the actual number of breakpoints is greater than the number tested. Moreover, to avoid the latent omitted variable problem, we should allow for more structural breaks than our prior belief in order to allow for an unknown change agent.

(iii) The simulation results of Bai and Perron (2006) show that the power of double maximum tests is nearly as high as when using a test that accounts for the correct number of breakpoints.

Once the existence of breakpoints is established, a sequential test starting at some value of l greater than zero is used to determine the number of breakpoints. The starting value of l should be consistent with the facts of the case, i.e. if the sample period is suspected to have before, during and after periods, it is reasonable to set the starting value of l at least one and test to a maximum number of break of at least four.

4 Using structural break analysis as evidence of causation

It is not possible to use inferential statistics, absent a statistically designed experiment, to establish a causal relationship between variables with absolute certainty. Take, for example, the Granger causality test, which establishes a structural relationship but cannot determine causation in the legal sense (Granger 1969). However, as Gaughan states, "[E]ven in the liability phase [of a trial], there can also be important economic and financial issues for which the expert [economist] may provide evidence" Gaughan (2009, 54). While econometric techniques may not be able to establish causation with absolute certainty, it is possible to either lend significant support to the case for causation or raise significant doubt of a causal relationship. Moreover, structural break analysis can provide support for, or help to refute, the period in which damages were incurred.

While structural break analysis can provide strong support for an argument, for or against, causation, there are at least three objections that can be raised with this type of analysis. One, because timing of a breakpoint is a key part of the evidence there is risk of committing the post hoc ergo propter hoc fallacy. Two, more than one causal event might occur in a single time period. And three, there may be a lag between cause and effect.

Post hoc ergo propter hoc. An obvious challenge to the use of structural break analysis as evidence of causation is that we are relining on timing evidence, i.e. condition 'A' ensues, then effect 'B' is observed, thus 'A' causes 'B.' This line of reasoning may well suffer either the post hoc ergo propter hoc or the cum hoc ergo propter hoc fallacies, but it should be noted that any before-and-after analysis that relies on timing evidence alone is subject to this criticism (see e.g. Young v. Hickory Business Furniture 2000). Therefore, before-and-after analysis, regardless of method, should never rely on time as its only explanatory variable. Further, structural break analysis allows the expert to control for both major explanatory factors that may differ between periods (see e.g. Katskee v. Nevada Bob's Golf of Nebraska, Inc. 1991; Bickerstaff v. Vassar College 1999) and possibility of unknown and/or unobserved causal factors (see e.g. In re Live Concert Antitrust Litigation 2012).

Data frequency. At issue is whether more than one casual factor could have occurred within the same time period, or within a small neighborhood of time periods. This challenge is most apparent when working with annual data. For example, consider Live Concert Antitrust. In this case, the expert for the plaintiff conducted a structural break test at the year 2000 because that was the year the defendant entered the market. The expert then concluded that because the year 2000 was a significant breakpoint, the defendant's entry into the market was the cause of increased ticket prices. Clearly, there are several problems with this analysis (several of which are discussed above), but if we assume away the issues already discussed, we are left with at least one additional problem. How do we know something else did not happen during the year 2000 that could increase ticket prices? This challenge can be met with higher frequency data. It is much less likely that an unknown casual event will happen in the same month as our hypothesized cause than it would be to occur in the same year.

Causal lag. A third challenge to using structural break analysis as evidence of causation is that ‘A’ may cause ‘B,’ but does so with a lag. In this situation, we would not expect to see a breakpoint occur in the same time period as the hypothesized causal event. In fact, for sufficiently high-frequency data, i.e. the data’s periodicity is shorter than the causal lag, the breakpoint occurring at the same time as the hypothesized causal event would tend to refute, rather than support, an argument for causation. This problem is less severe in lower frequency, but the better solution is to avoid conducting statistical analysis in a vacuum. Good analysis will be done in reliance on other expert witness, underlying scientific theory and industry norms to determine reasonable causal lag.

5 An empirical example: Poppler v. Wright Hennepin Co-Op. Electric (2014)

In this section, we present an empirical example involving damages to a dairy farm due to a phenomena known as “stray voltage.” Stray voltage is the presence of electrical current on items that are not part of the electrical system that injures dairy livestock by suppressing milk production, reproduction, and longevity (see e.g. Hultgren 1990; James v. Beauregard Elec. Co-op., Inc. 1999). Key to determining economic loss in these cases is the determination of lost milk production but for the presence of stray voltage. Thus, we present a model of milk production and apply structural break analysis to support a case for causation and to determine the damage period.[7]

5.1 Background

The plaintiffs, Poppler, et. al., own and operate a dairy of approximately 200 registered Holstein dairy cows located a farm near Waverly, Minnesota, which they purchased in 2003. After several years of operation, the herd began exhibiting increased herd-health issues, decreased pregnancy rates, higher mortality, and reduced milk production. After eliminating other possible

[7] Note that is beyond the scope of this paper to provide a complete model of economic damages in dairy litigation. Here we provide only part of such analysis in order to demonstrate the usefulness of structural break analysis. For further discussion of dairy litigation see e.g. Kelly and Sienko (2016).

causes, the plaintiffs hired an independent electrical consultant who determined that Wright-Hennepin's electrical distribution system caused stray voltage on the Poppler dairy in levels that are problematic to dairy cows. In the ensuing litigation, the court ruled in favor of the plaintiff and awarded damages. For further details, see Kelly and Sienko (2016) and Poppler v. Wright Hennepin Co-Op. Electric (2014). A partial remedy, installation of an isolation transformer, was implemented July 2009, and the definitive remedy, installation of a new 3-phase electrical supply line, was implemented June 2012.

5.2 The data

In order to model milk production, we use data obtained from two sources: (i) the National Dairy Herd Improvement Association (DHIA), which is a service that provides herd records and testing to the dairy industry (see http://www.dhia.org/). The data relevant to our analysis is reported the monthly herd summary (either report number DHI-202 or DHI-302). And (ii) the monthly milk sale summary, commonly referred to as the milk check.

5.3 Modeling milk production

Annualized monthly average of pounds of milk produced per cow (MHA), was modeled using production data from January 2008 through May 2014 ($T = 77$). Equation (5) is the basic model of MHA that will be tested for structural breaks:

$$\widehat{MHA}_t = \alpha_1 + \alpha_2 TREND + \beta_1 HS_t + \beta_2 SCC_t + \beta_3 DIM_t + \beta_4 L3_t + \varepsilon_t \tag{5}$$

where $TREND$ is a deterministic time trend, HS is the heard size, SCC is the somatic cell count, DIM is the average number of days lactating, referred to as "days in milk," and $L1$ is the percentage of cows in their first lactation cycle. The constant, α_1, and the trend coefficient, α_2 are allowed to break while the parameters $\beta_1, \ldots, \beta_4$, coefficients on control variables, are not allowed to break.

Figure 1 plots each of the control variables. Research in dairy science regarding stray voltage (see e.g. Gustafson and Albertson 1982; A. Lefcourt 1982; Appleman and Gustafson 1985; A. M. Lefcourt 1991) and predicting milk production (see e.g. Ray, Halbach, and Armstrong 1992) indicates that the indicators of animal health most likely to impact milk production are somatic cell count, days in milk, and stage of lactation.

[Figure 1 about here]

Visual inspection of figure 1 indicates the possibility of changes in somatic cell count, and days in milk within the range of July 2009 and June 2012. Before-and-after analysis is complicated in this case because:

(i) Both a partial fix and a definitive occurred, thus there is a before, transition, and after periods. In somatic cell count, and days in milk three periods can be seen with breakpoints roughly in line with the remedies.

(ii) The Poppler's, due to financial stress resulting from lower milk production, were forced to sell 2nd lactation cows. Leading a herd with a higher percentage of cows in their 1st lactation. See figure 3. Cows produce more milk with each lactation cycle they go through (Ray, Halbach, and Armstrong 1992).

(iii) Cows that have been exposed to stray voltage for an extended period do not immediately recover upon secession of exposure.

[Figure 3 about here]

5.4 Before-and-after analysis

We check parameter stability of (5) using the double maximum structural break testing procedure as suggested by Perron (2006) with trimming set to 20% and the maximum number of breaks set to three. The WDMax test statistic calculated to be 40.20 (10.98 critical value at 95% confidence level).[8] Table 1 reports the results of the double maximums structural break analysis. Both UDMax and WDMax statistics indicate one significant, $\alpha = 0.5$, breakpoint at October 2009. The breakpoint is consistent with a recovery beginning in 2009.

[Table 1 about here]

[8] Note that test statistics employ HAC covariances (Quadratic-Spectral kernel, Andrews bandwidth).

Figure 4 plots the breakpoint model of MHA, equation (5), and table 2 parameter estimates. We see that the trend coefficient is negative in the first regime (2008M01 – 2009M09) and the trend coefficient is positive but insignificant during the second regime (2009M10 – 2014M05). This result is consistent with the conjecture that the installation of an isolation transformer in July 2009 began the recovery process for the dairy.

[Figure 4 about here]

[Table 2 about here]

6 Conclusion

Before-and-after analysis is a highly reliable method for determining economic damages and providing support of causation claims in business interruption cases, but though widely accepted, three problems may arise: the overt omitted variable problem, the multiple comparisons problem, and the latent omitted variable problem. Structural break models can provide solutions to these problems. Though structural breaks do require more data to estimate and more effort to explain.

References

Andrews, Donald W. K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61 (4): 821–56.

Andrews, Donald W. K., Inpyo Lee, and Werner Ploberger. 1996. "Optimal Changepoint Tests for Normal Linear Regression." *Journal of Econometrics* 70 (1): 9–38. doi:10.1016/0304-4076(94)01682-8.

Appleman, RD, and RJ Gustafson. 1985. "Source of Stray Voltage and Effect on Cow Health and Performance." *Journal of Dairy Science* 68 (6): 1554–1567.

Bai, Jushan. 1997. "Estimation of a Change Point in Multiple Regression Models." *The Review of Economics and Statistics* 79 (4): 551–63.

Bai, Jushan, and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66 (1): 47–78. doi:10.2307/2998540.

———. 2003a. "Computation and Analysis of Multiple Structural Change Models." *Journal of Applied Econometrics* 18 (1): 1–22.

———. 2003b. "Critical Values for Multiple Structural Change Tests." *Econometrics Journal* 6 (1): 72–78. doi:10.1111/1368-423X.00102.

———. 2006. "Multiple Structural Change Models: A Simulation Analysis." *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, 212–237.

Bazemore v. Friday. 1986, 478 US 385. Supreme Court.

Benjamini, Yoav. 2010. "Simultaneous and Selective Inference: Current Successes and Future Challenges." *Biometrical Journal* 52 (6): 708–21. doi:10.1002/bimj.200900299.

Bickerstaff v. Vassar College. 1999, 196 F. 3d 435. Court of Appeals, 2nd Circuit.

Chow, Gregory C. 1960. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28 (3): 591–605.

Erica P. John Fund, Inc. v. Halliburton Company. 2015. Dist. Court.

Gaughan, Patrick A. 2009. *Measuring Business Interruption Losses and Other Commercial Damages*. 2nd ed. Hoboken, N.J: John Wiley & Sons.

Granger, C. W. J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37 (3): 424–38. doi:10.2307/1912791.

Greene, William H. 2003. *Econometric Analysis*. 5th ed. Upper Saddle River, N.J: Prentice Hall.

Gustafson, RJ, and VD Albertson. 1982. "Neutral-to-Earth Voltage and Ground Current Effects in Livestock Facilities." *Power Apparatus and Systems, IEEE Transactions on*, no. 7: 2090–2095.

Halliburton Co. v. Erica P. John Fund, Inc. 2014, 134 S. Ct. 2398. Supreme Court.

Hansen, Bruce E. 1997. "Approximate Asymptotic P Values for Structural-Change Tests." *Journal of Business & Economic Statistics* 15 (1): 60–67. doi:10.2307/1392074.

Heerwagen v. Clear Channel Communications. 2004, 435 F. 3d 219. Court of Appeals, 2nd Circuit.

Hultgren, J. 1990. "Small Electric Currents Affecting Farm Animals and Man: A Review with Special Reference to Stray Voltage." *Veterinary Research Communications* 14 (4): 287–298.

In re Live Concert Antitrust Litigation. 2012, 863 F. Supp. 2d 966. Dist. Court.

In Re REMEC Inc. Securities Litigation. 2010, 702 F. Supp. 2d 1202. Dist. Court.

James v. Beauregard Elec. Co-op., Inc. 1999, 736 So. 2d 353. Court of Appeals, 3rd Circuit.

Katskee v. Nevada Bob's Golf of Nebraska, Inc. 1991, 472 NW 2d 372. Supreme Court.

Kelly, Logan, and David Sienko. 2016. "The Determination of Economic Damages in Dairy Litigation."

Lefcourt, Alan. 1982. "Behavioral Responses of Dairy Cows Subjected to Controlled Voltages." *Journal of Dairy Science* 65 (4): 672–674.

Lefcourt, Alan M. 1991. "Effects of Electrical Voltage/current on Farm Animals. How to Detect and Remedy Problems." *Agriculture Handbook (Washington)*, no. 696.

Liu, Jian, Shiying Wu, and James V. Zidek. 1997. "On Segmented Multivariate Regression." *Statistica Sinica*, 497–525.

Lloyd, Robert. 2014. "Proving Damages for Lost Profits: The Before-and-After Method." *College of Law Faculty Scholarship*, May. http://trace.tennessee.edu/utk_lawpubl/71.

Miller, Rupert G. 1981. *Simultaneous Statistical Inference*. 2d ed. Springer Series in Statistics. New York: Springer-Verlag.

Perron, Pierre. 2006. "Dealing with Structural Breaks." In *Palgrave Handbook of Econometrics*, 1:278–352. https://www.researchgate.net/profile/Pierre_Perron/publication/4998524_Dealing_with_Structural_Breaks/links/0deec528ab21d602ce000000.pdf.

Poppler v. Wright Hennepin Co-Op. Electric. 2014, 845 NW 2d 168. Supreme Court.

Quandt, Richard E. 1960. “Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes.” *Journal of the American Statistical Association* 55 (290): 324–30. doi:10.2307/2281745.

Ray, DE, TJ Halbach, and DV Armstrong. 1992. “Season and Lactation Number Effects on Milk Production and Reproduction of Dairy Cattle in Arizona.” *Journal of Dairy Science* 75 (11): 2976–2983.

Tomlin, Jonathan T., and David R. Merrell. 2006. “Accuracy and Manipulability of Lost Profits Damages Calculations: Should the Trier of Fact Be Reasonably Certain, The.” *Transactions: The Tennessee Journal of Business Law* 7: 295.

Vogelsang, Timothy J. 1999. “Sources of Nonmonotonic Power When Testing for a Shift in Mean of a Dynamic Time Series.” *Journal of Econometrics* 88 (2): 283–99. doi:10.1016/S0304-4076(98)00034-7.

Vuyanich v. Republic Nat. Bank of Dallas. 1984, 723 F. 2d 1195. Court of Appeals, 5th Circuit.

Yao, Yi-Ching. 1988. “Estimating the Number of Change-Points via Schwarz Criterion.” *Statistics & Probability Letters* 6 (3): 181–189.

Young v. Hickory Business Furniture. 2000, 538 SE 2d 912. Supreme Court.

Table 1: Double maximum structural break analysis of somatic cell count

Number of Breaks	*F-statistic*	*Scaled F-statistic*	*Weighted F-statistic*	*Critical Value*
1 *	**20.10**	**40.20**	**40.20**	**10.98**
2 *	12.48	24.96	30.52	8.98
3 *	10.10	20.19	31.10	7.13
UDMax statistic*	40.20	critical value**		11.16
WDMax statistic*	40.20	critical value**		12.15

** Significant at the 0.05 level.*

*** Bai-Perron (Econometric Journal, 2003) critical values.*

Estimated break dates:
1: 2009M10
2: 2009M10, 2012M07
3: 2009M10, 2011M05, 2012M12

Table 2: Breakpoint Model of MHA

Variable	*Coefficient*	*Std. Error*	*t-Statistic*	*Prob.*
2008M01 - 2009M09 -- 21 obs				
Constant	30811.85	4331.852	7.112859	0.0000
Trend	-232.1087	36.19975	-6.411888	0.0000
2009M10 - 2014M05 -- 56 obs				
Constant	29066.22	4148.954	7.005675	0.0000
Trend	11.60879	15.79990	0.734738	0.4650
Non-Breaking Variables				
HS	-36.29178	15.70141	-2.311371	0.0238
SCC	-2284.438	650.7707	-3.510358	0.0008
DIM	53.44965	14.99331	3.564900	0.0007
L3	30056.98	5229.953	5.747085	0.0000
R-squared	0.573894	Mean dependent var		29704.69
S.E. of regression	1325.250	S.D. dependent var		1934.447
F-statistic	13.27595	Prob(F-statistic)		0.000000

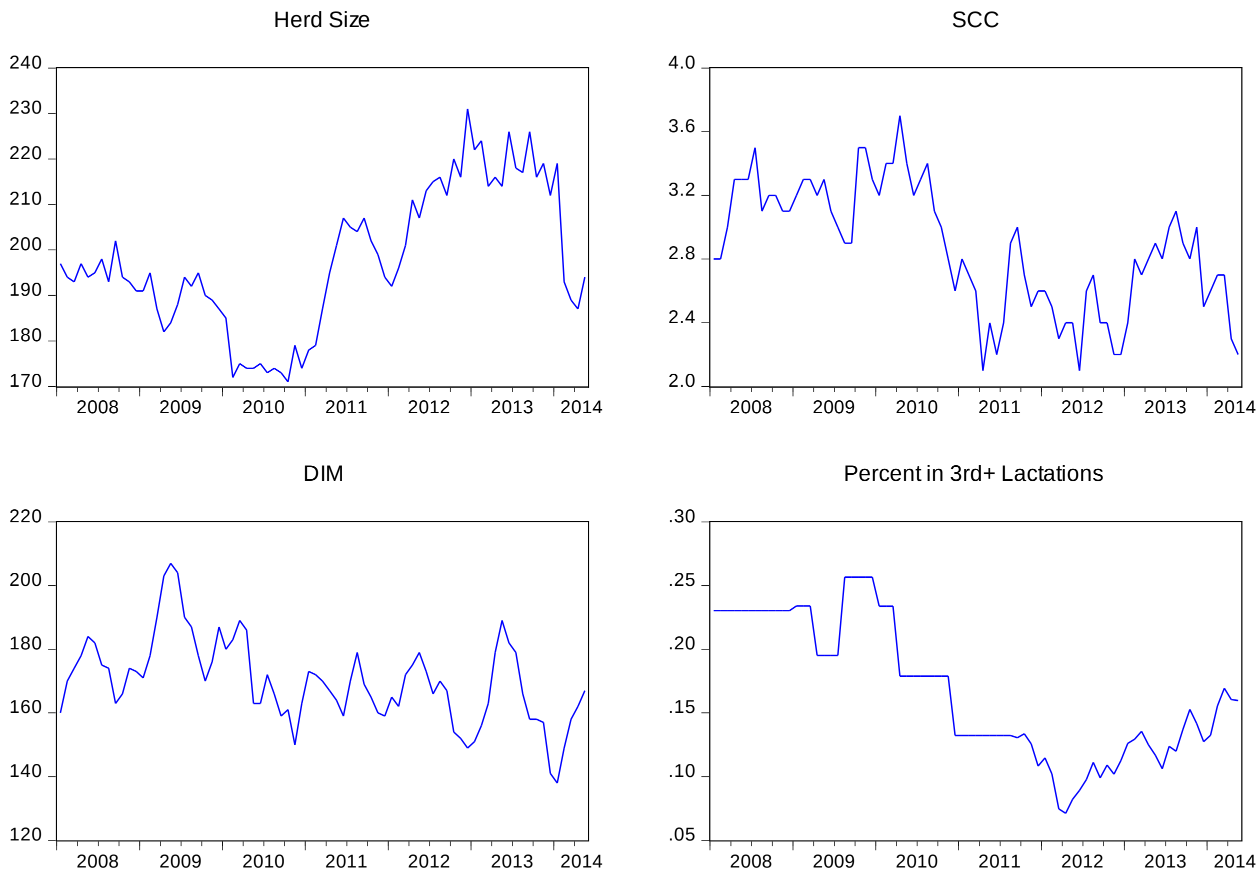

Figure 1: Milk model explanatory variables

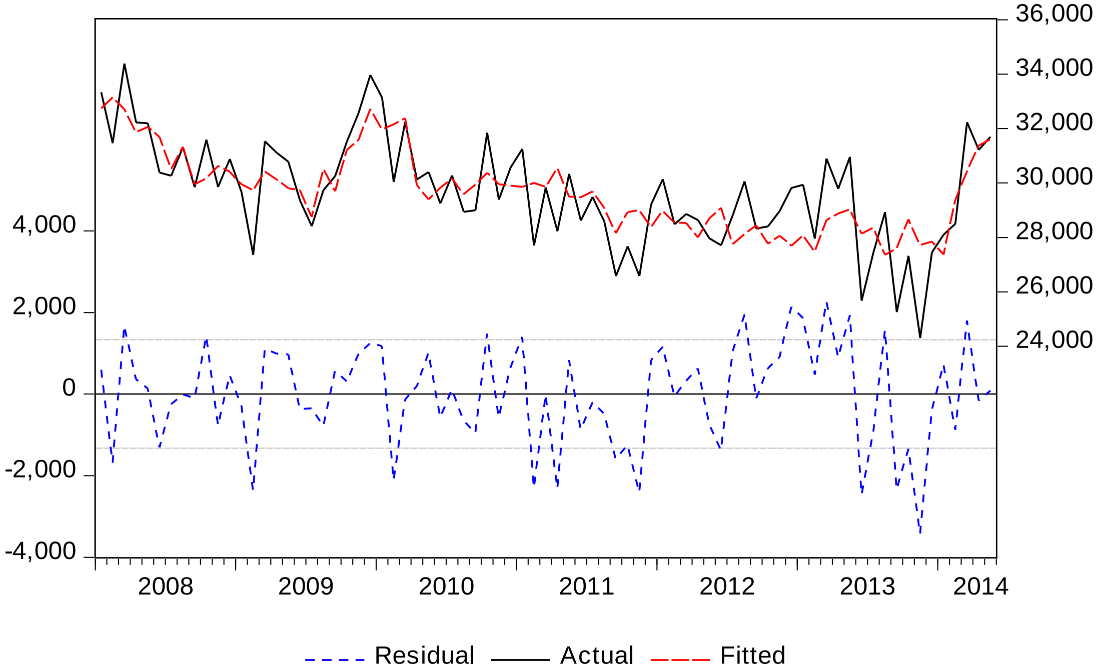

Figure 2: Actual vs. fitted values for MHA model

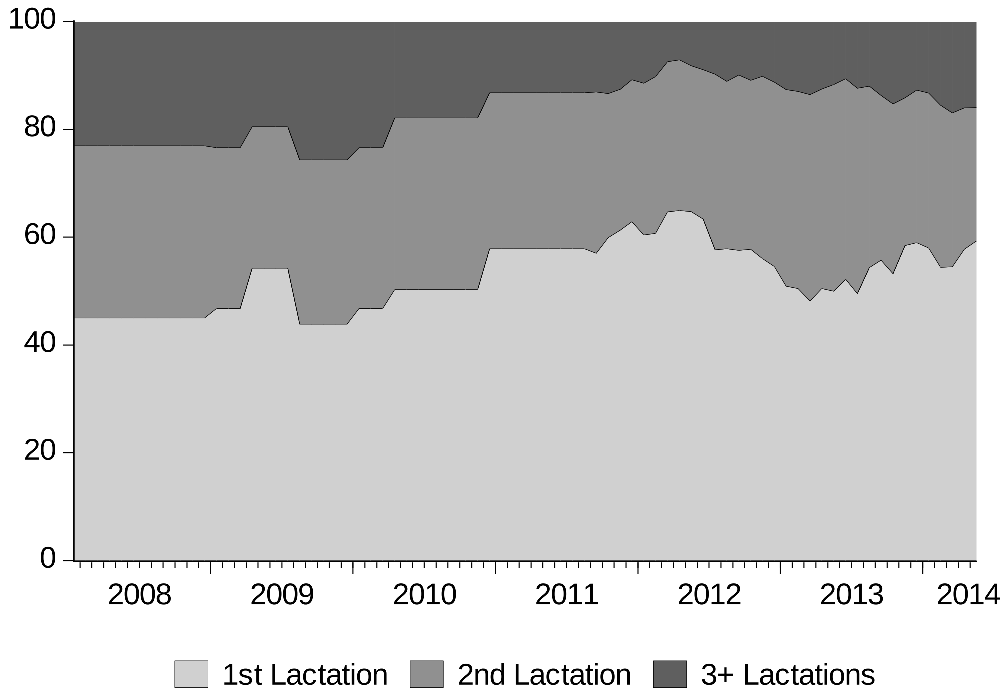

Figure 3: Percent of herd in various lactation stages

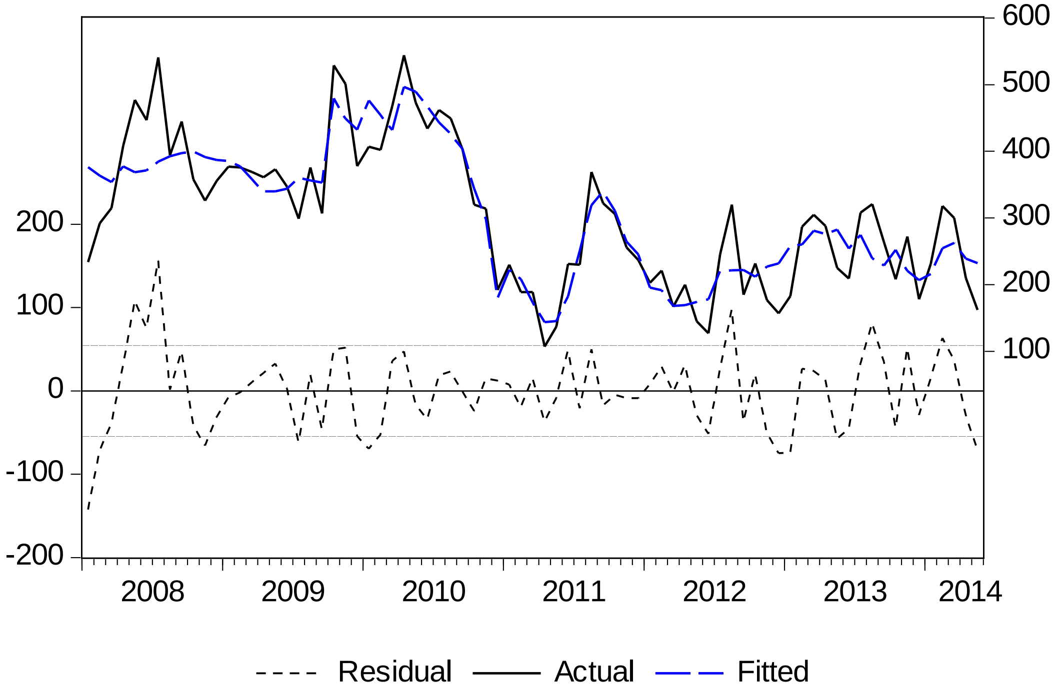

Figure 4: Three breakpoint model of somatic cell count

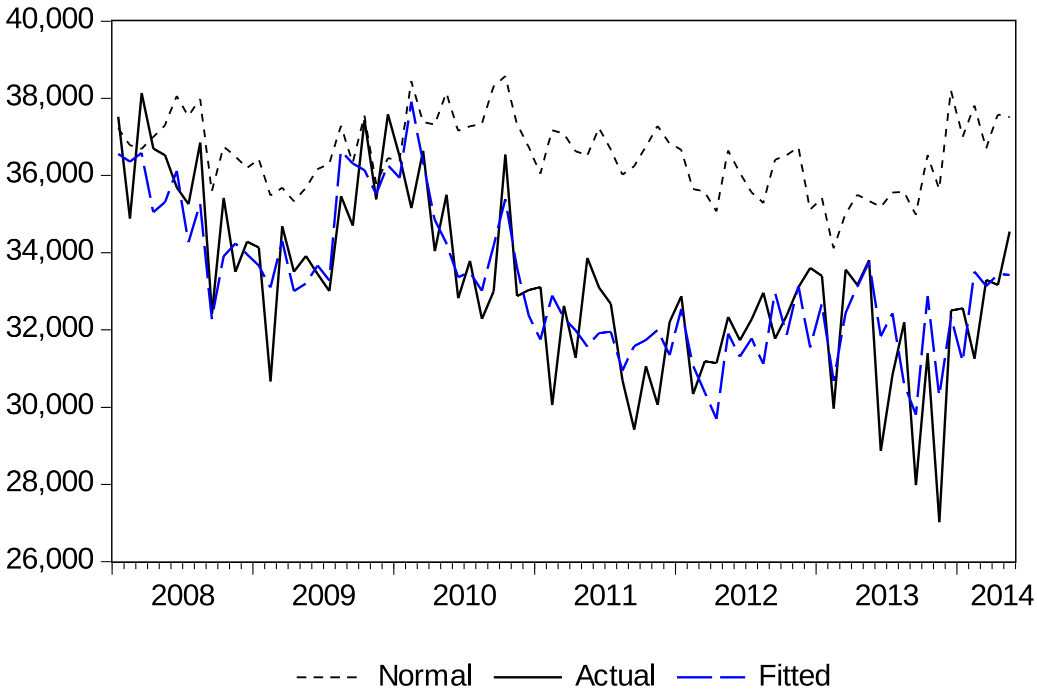

Figure 5: Normal, actual, and fitted average milk production per lactating cow