

Financial Market versus Consumer Market Responses to Emergent Phenomena: An Application to a Social Media-Inspired Boycott

Brian Kelleher Richter^a
Timothy Werner^b

This draft: 11 January 2014

Abstract

We extend behavioral finance theories to explain how and when financial and consumer markets may respond differently to emergent phenomena. The nature of emergent phenomena may be incongruent with the assumptions underlying the commonly employed the financial-market event study, causing it to fail to provide reliable results. Given the problems in applying financial market event studies to emergent phenomena, we propose researchers adopt an additional methodology—synthetic control—that employs fewer and weaker assumptions in their analysis. Synthetic control is flexible enough to be applied to measures of consumer market responses—and therefore more able to provide reliable estimates of the causal effects. We test and find support for our methodological arguments in an application to “United Breaks Guitars,” the first major social media-inspired boycott.

Keywords: Emergent Phenomena, Event Study, Synthetic Control, Social Media, Private Politics, Counterfactuals, Boycotts

^a brian.richter@mcombs.utexas.edu. Office: 512-232-6751. <http://www.briankrichter.com/>. Assistant Professor of Business, Government, and Society, McCombs School of Business, University of Texas at Austin.

^b timothy.werner@mcombs.utexas.edu. Office: 512-232-6844. <http://www.timothywerner.com/>. Assistant Professor of Business, Government, and Society, McCombs School of Business, University of Texas at Austin.

I. INTRODUCTION

Financial markets and consumer markets (e.g., product and service markets) operate in fundamentally different ways. Nevertheless, following the rise to prominence of the efficient markets hypothesis in the finance literature in the early 1970s (Fama, 1970), researchers and analysts seeking to understand the impact of some phenomenon on firm performance assumed financial market performance served as a good proxy for consumer market performance (McWilliams and Siegel, 1997). Beginning in the late-1990s, however, behavioral finance theorists began to challenge the assumptions that financial markets always respond to events in ways consistent with efficient markets (see, Shiller, 2003) or in ways consistent with how consumer markets ultimately respond.

From a managerial perspective, it is important to understand when and how responses in financial markets may not be indicative of responses in consumer markets given long-standing questions about how meaningful any single measure of firm performance is and whether or not we should expect convergence across different types of measures (Venkatraman and Ramanujam, 1986). In this research we extend insights from behavioral finance theory to understand how and when changes in (or a lack of changes in) measures of financial market performance may fail to be indicative of similar effects in consumer markets with regards to the study of some of the very phenomena that set management and strategy scholarship apart from work in economics, psychology, and sociology—i.e., phenomena that are “emergent” (Nag et al., 2007) and likely to have immediate managerial relevance because they represent trends reshaping organizations and their futures.

Emergent phenomena are not narrow in conception, nor are they simply the latest fad. They are appearing with increased frequency in an ever more complex world that is constantly requiring top management teams to respond by re-directing firm resources to maximize their firm’s performance (Drucker, 1995). Examples of such phenomena in the last 50 years reveal

the breadth of the challenges that managers have faced and that management researchers have examined: (i) changing social norms around sex, race, and sexual-orientation have altered employment relationships and have made firms a site of contention in which previously marginalized groups have pressed their rights (e.g., Dobbin, 2009; Briscoe and Safford, 2008); (ii) regulatory challenges and deregulation in the West have increased the need for managers to develop “legal astuteness” (Bagley, 2008); and (iii) political economic reforms in the former socialist bloc and the lowering of barriers on trade in goods and services have not only opened up new markets for the outsourcing of manufacturing and service positions but also increased the need for strategies that mitigate political risk in these markets (e.g., Henisz and Zelner, 2003).

The most prominent emergent phenomenon is technological change, which often provides sufficient stimulus for managers to alter their firms’ routines (Fligstein, 1991). Many innovations in high technology have caused disruptions to traditional managerial models on a massive scale, forcing incumbent firms to alter their practices or risk being overtaken by new entrants (Christensen, 1997; Rosenbloom, 2000). Technological innovations since the mid-1990s include the rise of the Internet, increases in the processor speeds and storage capabilities of personal computers, and the advent of big data. These factors, taken together, have changed the management of all types of organizations ranging from firms to political campaigns. Most recently, the rise of social media has challenged how firms interact not only with consumers, but also with activists who confront the firm.

Despite the importance of studying and understanding the role of emergent phenomena in management and strategy, these phenomena create challenges for researchers, when, as is often the case, they initially affect only a small number of firms (or a single industry) or when they unfold over an ambiguous timeframe. The characteristics of newness and limited initial reach also complicate the use of extant methods to assess a new phenomenon’s effects. One way researchers have dealt with the small sample size problem is to utilize the financial market event

study methods, grounded in counterfactual logic, to examine the impact of specific events on firm value (see, e.g., McWilliams and Siegel, 1997). The assumptions required to measure treatment effects accurately in financial market event studies, however, face fundamental challenges when applied to emergent phenomena. Due to their limited reach, emergent phenomena may not produce information that is quickly distributed to and incorporated by market participants. Due to the newness of emergent phenomena, market participants (similarly to researchers) may not have strong priors (or consensus opinions) on what the exact effect of an emergent phenomenon will be. Hence, as behavioral finance theory suggests, emergent phenomena (even when broadcast widely) are likely to generate disagreement among market participants causing event-specific or firm-specific breakdowns in the efficient markets hypothesis (Hong and Stein, 2007). That is, when market participants are exposed to an emergent factor, they likely hold heterogeneous priors about its potential effect on firm value—opinions which are not likely to be reflected in the share prices of a focal firm immediately but which might instead be reflected in the volatility of trading of that firm's shares.

As a result of the failure of critical assumptions underlying financial market event studies to hold when examining questions about the consequences of emergent phenomena, a new methodological approach is required to assess their causal effects. One solution could be to look to product or consumer markets for effects rather than the financial markets (McWilliams and Siegel, 1997). To do so, a counterfactual approach is needed that allows us to estimate the treatment effect of an event that applies to only a single firm on non-financial outcomes. Synthetic control methodology (Abadie and Gardeazabal, 2003; Abadie et al., 2010) provides one potential approach that is analogous in many ways to financial market event studies. Synthetic control is an unbiased method of constructing counterfactual trends for a single statistical unit, such as a firm, based upon data from a combination of similar units. The approach allows researchers to estimate the effects of an event, such as the emergence of a new

phenomenon, by examining the differences in performance between a synthetic unit (which behaves as if it were unaffected) and the actual, observed unit (which was affected). The approach requires fewer and weaker underlying assumptions than a financial market event study, and thus, is less likely to fail when presented with an event that has the characteristics of an emergent phenomenon.

To build our argument for the value of this new approach to the analysis of emergent phenomena, we proceed by first discussing the role of counterfactual thinking in causal inference—highlighting how financial market event study methods and synthetic control methods are both grounded in this logic. We then compare and contrast differences between the underlying assumptions of the two counterfactual methodologies. Moving on, we discuss why the assumptions underlying financial market event studies are more likely to be violated when applied to an emergent phenomenon than to an established one. In doing so, we also articulate why we might expect to see the effects of a new phenomenon appear in non-financial performance metrics, for which effects can be estimated using synthetic control, despite not finding any share price responses in financial market event studies. This argumentation produces two testable hypotheses regarding what we might expect to see happen in financial markets versus real product and consumer markets around manifestations of emergent phenomena.

To demonstrate the power of synthetic control as a more robust methodology for studying an emergent phenomenon than financial market event studies, we apply our arguments to the first large scale social media-inspired boycott. Specifically, we test our hypotheses in an application to the case of “United Breaks Guitars,” a series of videos that appeared on YouTube in 2009 and 2010 and were promoted heavily through social media such as Twitter. Before delving into the results of this application, we briefly review the related literature on boycotts, introduce our data, and relate our hypotheses to the analyses we will conduct. To preview our findings, we uncover no meaningful stock price effects for the “United Breaks Guitars” videos in

our financial market event study. We are, nevertheless, able to document increased disagreement among financial market participants around the release of the videos which calls into question the validity of this result, and more importantly the validity of the financial market event study approach as applied to emergent phenomena. This disagreement suggests that critical assumptions underlying the financial market event study approach may fail to hold in similar cases. When we employ synthetic control as an alternative methodology, however, we find that United experienced a significant loss of passengers in the wake of the videos, further suggesting that the underlying assumptions in financial market event studies may break down when they are applied to emergent phenomena.

II. COUNTERFACTUAL HISTORIES, POTENTIAL OUTCOMES, AND CAUSAL INFERENCE

Despite the challenges embodied in “learning from samples of one or fewer,” March, Sproull, and Tamuz (1991) note that “in trying to understand unique experiences, organizations make implicit choices between two alternative perspectives on history” (4)—one of which is a “hypothetical” history that can be benchmarked to the history a firm observed. More recently, scholars have echoed similar sentiments, explicitly advocating for the use of counterfactual reasoning in strategic management contexts, particularly when statistical identification of causal effects is important (e.g., Durand and Vaara, 2009).

Although Durand and Vaara’s (2009) and March et al.’s (1991) depiction of counterfactual reasoning suggests it is qualitative in nature, it need not be. The Rubin Causal Model or Rubin Potential Outcomes framework (Rubin, 1974, 2005; Holland, 1986) applies the logic of counterfactual reasoning or hypothetical history to a quantitative model for estimating statistical treatment effects. The framework argues that we can think of causal treatment effects (τ) as being equal to $Y_1 - Y_0$ where Y_1 represents the performance of a treated statistical unit and Y_0 represents the performance of the same statistical unit had it not received treatment.

The intuitive and simplistic nature of counterfactual thinking has a natural appeal—which is why March, Sproull, and Tamuz (1991) suggest its use in the analysis of emergent phenomena—and its representation in the Rubin Causal model helps formalize the logic in quantitative settings. This makes the counterfactual methodological approach particularly relevant with respect to analyzing the effects of emergent phenomena on firm performance.

We review below, in counterfactual terms, the financial market event study which is one of the most widely adopted methodologies for causal inference among management and strategy scholars. We also present a newer counterfactual methodology that can also be applied in cases where only one firm is affected by some phenomena—synthetic control—that is rapidly being adopted across social science disciplines due to its flexibility. Before moving to our theoretical hypotheses about emergent phenomena and our application to a social media inspired boycott, we compare these two methodologies to each other to illustrate their relative advantages and the differences in the assumptions they make. This comparison of assumptions motivates our theoretical hypotheses.

A. Counterfactuals in Event Studies

Event study methodology uses financial market returns around an event to construct expected returns for a focal firm and compares those expected (or counterfactual) returns to the actual returns of the focal firm. The goal of this comparison between expected and actual returns is to uncover what effect, if any, researchers can attribute to the event.

As a counterfactual method, financial market event studies proceed in three steps.¹ First, using data from an estimation period prior to the event of interest, a baseline relationship between the stock market return of the focal firm affected by the event and a market benchmark

¹ We discuss event study methodology in this section as if an event affects only a single firm; obviously, some events affect multiple firms simultaneously, and firms often experience highly similar, if not identical, events at different points in time.

return is estimated using ordinary least squares regression:

$$FF_Ret_t = b_0 + b_1 Mkt_Ret_t \quad (1)$$

where FF_Ret_t represents returns of the focal firm on day t , and Mkt_Ret_t represents either the equal- or value-weighted returns of a portfolio of benchmark firms on day t .² In the second step, the parameter estimates from equation (1) are used to predict an expected return for the focal firm during the event period, i.e. an appropriate window following the event, which is chosen by the researcher to reflect when the event's potential impact should have appeared in the focal firm's share price:

$$FF_Ret_t = \hat{\beta}_0 + \hat{\beta}_1 Mkt_Ret_t \quad (2)$$

Researchers then calculate the treatment effect of the event (τ) on the focal firm by taking the difference between the firm's expected and actual return; i.e., they define the focal firm's abnormal return as:

$$FF_AR_t = FF_Ret_t - FF_Ret_t \quad (3)$$

B. Counterfactuals in Synthetic Control

Synthetic control is a more recently developed counterfactual methodology (Abadie and Gardeazabal, 2003) that Fremeth et al. (2013) have recently shown can be used to examine how a single firm would have performed had it not received some treatment event. The synthetic control approach applies similar counterfactual reasoning as a financial market event study in a quantitative way that is more flexible than financial market event study methodology, as it can be applied to any number of outcome variables, such that its use need not be limited to the study of firms' financial performance. This explains why researchers in other academic disciplines have quickly adopted it to create a diverse set of synthetic statistical units³ so as to answer questions

² These benchmark firms typically consists of the full market or a relevant subset, such as the focal firm's industry. We discuss this point further in an appendix on event study methods.

³ Synthetic control units that have been constructed as the unit of analysis range from geographical units such as countries (e.g., Bilmeier and Nannicini, 2013), states within countries (e.g., Abadie and Gardeazabal, 2003);

about a diverse set of phenomena including terrorism (Abadie and Gardeazabal, 2003; Montalvo, 2011), natural disasters (Coffman and Noy, 2012; Cavallo et al., forthcoming), economic liberalization (Billmeier and Nannicini, 2013), re-unification of countries (Abadie et al., 2012), health policy (Abide et al., 2010), campaign finance laws (Dowling et al., 2012), voting (Montalvo, 2012), affirmative action (Hinrichs, 2012), and performance-based teacher pay (Hudson, 2010).

In essence, synthetic control creates a counterfactual representation of a particular statistical unit, or firm in our case, as a convex combination, or weighted average of non-treated statistical control units. Rather than treating the counterfactual as the expected financial market performance of a firm based on a financial market index, synthetic control constructs a counterfactual based on the actual performance of other firms in the same real world product and consumer markets. In addition to creating a match on outcome variables (Y), synthetic control also attempts to create the closest match possible on a statistical unit's underlying attributes (X), which can be thought of as roughly analogous to predictor variables in a regression model.

The first step in using the synthetic control methodology is to estimate values for W using a dual optimization computational algorithm to minimize the following equation:

$$\sum_{m=1}^k v_m (X_{1m} - X_{0m}W)^2 \quad (4)$$

\widehat{W} is of critical interest as it represents weights that will be applied to control units to calculate a synthetic counterfactual—and is used much like how the $\hat{\beta}_0$ and $\hat{\beta}_1$ estimates generated in event studies by equation (1) are used to construct expected returns. In the equation above, X_{1m} represents the value of the m -th observable attribute of a focal unit; X_{0m} is a $1 \times j$ vector containing the values of the m -th observable attribute of each of the j potential comparison or

Abadie et al., 2010; Dowling et al., 2012) and even subdivisions within states (Coffman and Noy, 2012) to institutions or organizations such as universities (Hinrichs, 2012), schools (Hudson, 2010), and firms (Fremeth et al., 2013).

control units; W is a vector of weights on control units; and v_m is a vector of weights on attributes of the control units which is optimized to create the closest match possible between the synthetic and focal unit conditional on minimizing W (Abadie et al., 2010).⁴ More technical details on synthetic control for the interested reader appear in an appendix focused on synthetic control issues.

Equation 4 on synthetic control is analogous to Equation 1 on financial market event studies, as each represents the first step in the process—since each is used in its respective methodology to predict how a firm should behave under normal circumstances, i.e., in the pre-treatment or pre-event period. The focal firm’s counterfactual performance in the post-treatment period can be estimated in the synthetic control set-up as:

$$\hat{Y}_o = Y_o \hat{W} \quad (5)$$

where Y_o represents the outcome variable of interest not for the focal unit, or firm, but for the unaffected units or firms. Hence, this formula states that the synthetic represents a weighted average of otherwise similar firms that were not directly affected by some phenomenon of interest in both the pre-treatment/pre-event period and the post-treatment/post-event period.

Following directly from the Rubin Potential Outcomes Framework for quantifying causal effects using counterfactual logic, the treatment effect (τ) of a synthetic can be estimated as:

$$\tau = Y_1 - Y_o \hat{W} \quad (6)$$

where Y_1 represents the outcome variable of interest for a particular focal unit or firm.

C. Assumptions underlying Quantitative Counterfactual Methodologies

We outlined above the steps employed in financial market event studies and synthetic control for constructing a hypothetical representation of a focal firm. Both methodologies are

⁴ Note that while in many ways we can think of the X variables as analogous to predictor variables in a regression model, we cannot interpret variables on the weights applied to them (v_m) in the exact way that we can interpret coefficients in a regression, so caution is needed in trying to interpret values of v_m .

similar in that they find their bases in counterfactual logic, such that each takes a focal firm and estimates how they would have performed had a focal event not happened. Nevertheless, the type of data to which each can be applied, the time horizon over which each generates treatment effects, and the underlying assumptions of each differ in several important ways. Table 1 summarizes these differences.

<Insert Table 1 here>

The most obvious difference to the casual observer between event study methods and synthetic control is that synthetic control can be more flexibly applied to a range of outcome variables, and as such, it does not limit the analyst to the study of financial performance. Other differences that are important, but less obvious at a surface level include the length of the window over which effects can be observed and the basis for the counterfactual. These two features are related since the rational financial markets model basis for the counterfactual implies instantaneous reactions, while the observable similarities of attributes and outcomes between firms tend to be less volatile allowing for longer post-treatment or post-event observation windows in synthetic control.

This brings us to the most important difference between the two counterfactual methodologies: their underlying assumptions, as the results generated by a counterfactual methodology can only be reliable if its underlying assumptions hold in the situation being analyzed. Although both methodologies share common assumptions about the lack of confounding events and the representativeness of the counterfactual generated, they make different assumptions regarding the data generating process for the counterfactual. Synthetic control takes a *lassiez-faire* approach towards what the underlying data generating process for the counterfactual is, assuming simply that firms that are observably similar in certain ways tend to stay observably and unobservably similar in the absence of a shock. Event study methods, on the other hand, root the counterfactual's data generating process in a rational statistical model of

financial markets, assuming that the semi-strong form of the efficient markets hypothesis holds. This assumption requires that (i) “all available information [about a given firm] is costlessly available to all market participants” or at least filters very rapidly into the market, and (ii) that “all [market participants] agree on the implications of current information for the current price [of a given firm]” (Fama, 1970: 387).

The comparison of these two methodologies reveals that the synthetic control method is not only more flexible than the financial market event study method in that it can be applied to a variety of outcome variables but also requires fewer strong underlying assumptions.

III. EFFICIENT MARKETS AND BEHAVIORAL FINANCE MODELS OF DISAGREEMENT

The financial market event study is a powerful tool that requires fairly minimal data collection, even when compared to synthetic control, and thus, has been widely applied throughout the management literature. Examples include applications to a number of frequently observed phenomena including IPOs (Liu et al., 2013), technological acquisitions (Sears and Hoetker, 2013), political connections (Hillman, Zardkoohi, and Bierman, 1999), intercorporate alliances (Stuart, 2000), and strategic decision-making broadly (Woolridge and Snow, 1990).

Nevertheless, the financial market event study method’s appropriateness for generating valid counterfactual trends as a basis for estimating treatment effects has not gone unchallenged.⁵ For example, financial markets may struggle to properly value newly issued stocks in the long-term (Loughran and Ritter, 1995) and, in particular, financial markets had significant difficulties in properly valuing technology stocks in the 1990s. The latter problem was so great that firms could increase their market valuations simply by adding “.com” to their names (Ofek and Richardson, 2003). Given the emergent contexts in which these studies from the finance

⁵ Additionally, event studies have been criticized in the management literature for being poorly implemented; see, e.g., McWilliams and Siegel (1997) or McWilliams, Siegel, and Teoh (1999).

literature are situated, it is clear that the concerns are more pressing when newness is a factor.

Research in behavioral finance provides additional reasons to be skeptical of financial market event studies of emergent phenomenon, as uncertain and unreliable events amplify noise, which, in turn, leads to excess volatility in the trading of affected stocks. As Black (1986) argues, noise represents information that does not relate to a firm's economic fundamentals or future performance. Those who trade on this noise, which emergent phenomena are likely to generate, are often thought of as irrational actors (De Long, Schleifer, Summers, and Waldman, 1990), and as a result, it is difficult to assume that traders are processing all publicly available information about new phenomena accurately and that this processing is reflected in share prices. These problems call into question the most basic assumptions of the semi-strong version of the efficient market hypothesis we discussed earlier—at least with respect to emergent phenomena. Were this critical underlying assumption no longer tenable in a particular situation a researcher wanted to analyze, the financial market event study method would be largely ineffective in modeling counterfactual trends that would enable researchers to identify treatment effects.

In addition to these informational problems, Hong and Stein (2007) argue that some events, such as an emergent phenomenon, are likely to lead the financial markets to express considerable disagreement over the value of a focal firm's stock post-event. If investors hold heterogeneous priors regarding the potential effects of an event on a stock, then the occurrence of a new phenomenon may serve only to increase market participants' disagreement in the post-event period. That is, a "news stimulus may spark increased disagreement among those investors who were already following the stock, and is also likely to grab the attention of those who were not" (Hong and Stein, 2007: 125). This argument suggests that regardless of whether market participants are familiar with a stock or not, the effect of news regarding a new phenomenon likely will not be reflected in a firm's price but rather in its trading volume/share turnover because market participants process new information using different economic models (Kandel

and Pearson, 1995; Harris and Raviv, 1993). Such an increase in turnover and volatility is also consistent with the presence or entrance of irrational “noise” traders into a market (Foucault, Sraer, and Thiesmar, 2011; De Long, Schleifer, Summers, and Waldman, 1990), as such market participants turnover stocks at a rate that is in excess what would be expected from a prototypical rational investor (Shiller, 1981).

Thus, due to the significant issues raised by behavioral finance with regard to the validity of the assumptions underlying financial market event studies in the presence of emergent phenomenon, we hypothesize the following with regard to emergent phenomena, including our test case of a social media-inspired boycott:

H1: Around emergent phenomena financial market participants may have heterogeneous priors, such that the market may exhibit signs of disagreement:

- a. Leading to no significant change in securities prices*
- b. Leading to abnormally high share turnover*

Although financial market event studies may be vulnerable to producing null results when applied to emergent phenomena due to disagreement in markets, we cannot necessarily conclude (i) from a null result that an event has no effect on a firm or (ii) from a statistically significant result that an event has an effect on a firm in product or consumer markets. In particular, financial market event studies may fail to yield reliable inferences about an emergent phenomenon because, as we just outlined, the method’s assumptions do not hold.

Financial market event study results may reflect the statistical weaknesses of the method as applied to certain situations rather than the substantive nature of the phenomenon of interest—particularly if that phenomenon is emergent. As a result, it remains critical to assess whether or not the emergent phenomenon is indeed affecting a firm in product and consumer markets in order to avoid committing a Type I error (i.e., rejecting the null hypothesis incorrectly). We can and should examine non-financial outcomes and should employ methodological approaches,

such as synthetic control, that require fewer and weaker statistical assumptions than the event study. Doing so not only answers the call of McWilliams and Siegel (1997) to examine how events affect firm stakeholders other than shareholders (e.g., consumers) but also allows us to have greater confidence in our findings. This leads to our second hypothesis:

H2: Despite disagreement in financial markets around emergent phenomena, these phenomena may still affect relevant firm performance measures (e.g., in product or customer markets)

IV. APPLICATION TO “UNITED BREAKS GUITARS”

The specific emerging phenomenon to which we apply our argument and its associated methodological suggestions is the social media-inspired boycott. We examine in-depth the “United Breaks Guitars” videos produced by Canadian musical artist Dave Carroll and his band Sons of Maxwell as a case through which to test our hypotheses about divergent financial and product/customer market reactions to emergent phenomena.

Carroll and his band released three videos on YouTube over several months in 2009 and 2010. The videos targeted United Airlines for breaking his guitar at Chicago’s O’Hare International Airport in 2008—and for what he perceived to be the airline’s poor customer service in the months following the original incident. These videos were viewed widely on YouTube, with the first video receiving 190,000 views within two days and over 2.2 million views within five days (Deighton and Kornfeld, 2011). Additionally, the videos were promoted broadly on other social media, including Twitter, and received coverage in traditional media outlets, such as CNN and *Time*. Carroll also became widely sought out as a speaker on customer service issues—and he has even been a speaker at top-ranked business schools. Furthermore, at least two business school teaching cases have been written about “United Breaks Guitars,” including one published by Harvard Business School (Deighton and Kornfeld, 2011) and another published by the University of Toronto (Dunne, 2010). Carroll (2012) himself has even published a popular book on his experience. Nevertheless, no formal analysis of the events’

effects have been published in the scholarly literature to date.

Due to the prominence of the “United Breaks Guitars” videos, they represent the first major example of a social media-inspired boycott and thus present an important case study (Yin, 2013) with which to test whether existing methods can be applied to an emergent phenomenon as in *HI*. The videos additionally present a tough test for our argument that new approaches are needed, as: (i) the emergent phenomena in it is an outgrowth of an existing phenomenon (the traditional boycott) that markets are familiar with; and, (ii) the information produced from the videos was widely available to market participants, satisfying that core assumption of at least the weak version of the efficient markets hypothesis underlying the event study approach. In many ways, we might have expected market participants to react to these videos; for instance, the *Times of London* reported (incorrectly) that United suffered a 10% (or \$180 million) loss in market capitalization due to the first video (Ayers, 2009), in a story that was widely picked up by other media outlets.

Nevertheless, the “United Breaks Guitars” videos also exhibit the hallmarks of an emergent phenomenon that both firms and markets struggle to evaluate. First, the “event” or the video releases occurred over multiple days and at three widely distributed points in time. Second, and stemming from these timing issues, information from the events diffused slowly to market participants despite being publicly available (i.e., views on YouTube are not comparable to the instantaneous flow of information upon which efficient market hypothesis relies). Finally, the videos involved multiple new technologies, making it less obvious to market participants what the extent of the phenomenon would be. Given this set of characteristics, market participants, including equity analysts and traders, could have held heterogeneous priors about the impact of the videos, which, in turn, could have led to disagreement about what the information produced by the events meant and whether or not that information needed to be incorporated into expectations of United’s future performance. Thus, it is unclear whether or not

the market would react to the videos.

In the remainder of this section, we: (i) briefly review the literature on boycotts; (ii) outline the application of our methods to the case; (iii) introduce data sources; and, (iv) discuss our results.

A. Contentious Markets & The Study of Boycotts

Markets are a site of contention between firms and activists, both organizational and individual (King and Pearce, 2010). Activists challenging firms are engaging in what Baron (2005) defines as private politics; that is, “action by interest groups directed at private parties, as in the case of an activist group launching a campaign against a firm. This can be independent of government but generally draws strength from the public, as in the case of a boycott” (340). Firms have been increasingly targeted in private politics (Soule, 2009; Bartley, 2007; Seidman, 2007; Baron and Diermeier, 2007), in part because such direct engagement affords activists opportunities to air grievances against the firm and to encourage management to change or abandon controversial practices (Lenox and Eesley, 2009; Walker, Martin, and McCarthy, 2008).

The increasing recognition of reputation as an essential asset (Diermeier, 2011) has lead managers to be more conscious of the impact of activists (see, e.g., Reid and Toffel, 2009) and social ratings on their firms (see, e.g., Chatterji and Toffel, 2010). And, as a result of increased activism, managers have strategically dedicated greater resources to their nonmarket environments in recent decades (Werner, 2012). Although boycotts as a manifestation of private politics have existed since colonial times in the U.S., their recent success rates have been cause for greater managerial attention: for example, one quarter of boycotts that receive national media attention result in a concession of some form by the target firm according to King (2008).

The success rate of these traditional boycotts is as high as it is because of the breadth of their demonstrated effects on firms. Boycotts appear: (i) to enhance firms’ environmental risks

and thus, increase the probability of credit downgrades (Vasi and King, 2012); (ii) to harm firms' images and reputations and thus, to decrease the value of these intangible assets (McDonnell and King, 2013; Basdeo, Smith, Grimm, Rindova, and Derfus, 2006); (iii) to reduce firms' income by disrupting their routine practices (Luders, 2006) and supply chains (Seidman, 2007); and, (iv) to affect consumer decisions that in turn affect firms' revenues (Friedman, 1999) and marketing expenses (Garrett, 1987).

In addition to these effects, traditional boycotts may negatively affect firm market value. Financial market event studies have been widely used by boycott researchers, who have largely found that traditional boycotts depress market value (see, e.g., Bartley and Child, 2011; King 2011; Pruitt, Wei, and White, 1988; Pruitt and Friedman, 1986; c.f. Teoh, Welch, and Wazzan, 1999; Koku, Akhigbe, and Springer, 1997). More broadly, King and Soule (2007) examine social movement protests, including boycotts, against firms using an event study approach and find that such protests are associated with declines in firm value if they target issues related to unions or consumers, generate greater media coverage, and last longer. Together, the findings of existing research suggest that the United Breaks Guitars videos, as inspiration for consumers to boycott United, ought to have negatively affected the firm's market value.

Given that they are an emerging phenomenon, no existing financial market event studies of boycotts or protests examine social media-inspired forms of these movements.⁶ However, there is a growing body of empirical literature on how social media may affect the political and economic opportunity structures of activists that target corporations. Both Schurman (2004) and

⁶ There is a small literature in computer science that has demonstrated a correlation between social media sentiment (measured using language processing of Tweets) and share prices/trading volumes (see, e.g., Evangelopoulos, Magro, and Sidorova, 2012). However, none of these studies have demonstrated a causal effect of social media on share prices nor have any of them examined the interaction of social media with other forms of anti-firm contention, such as boycotts or protests.

Walker, Martin, and McCarthy (2008) argue that social media increase activists' ability to monitor firms. King and McDonnell argue that activists can use social media to mobilize against firms when negative practices are exposed (2011) and the Internet more generally to increase significantly the speed and amount of negative media coverage of an attacked firm (2013). These findings also suggest that social media-inspired boycotts are likely to have a tangible impact on a target firm but do not identify how that effect might be measured.

B. Methodological Approaches

Our methodological approaches to analyzing the effects of the “United Breaks Guitars” videos as representative of emergent phenomena are grounded in counterfactual logic; however, they differ depending upon whether the outcome we are interested in studying is a financial market outcome, as in *H1a* and *H1b*, or a product or customer market outcome, as in *H2*. Critically, for the hypotheses in question, the methodological approaches also differ in the assumptions underlying them.

The first method that we employ to analyze the financial market reactions to the “United Breaks Guitars” YouTube series, is a financial market event study. The premise is simple: we can construct a counterfactual for a given firms' financial market returns by generating what the firms' estimated returns should have been as a function of the firm's relation to a given financial market index or portfolio in the pre-treatment period. We can then assess the effects of the financial market event study by comparing the actual returns to the expected returns we calculated, assuming the efficient markets hypothesis holds. Also, as we noted earlier, financial market event studies have already seen wide use in the study of traditional boycotts—which allows us to apply them as a benchmark for existing methods in our application to a social-media inspired boycott, further allowing us to demonstrate the power of the synthetic control method.

In addition to using the event study method to analyze the impact of “United Breaks

Guitars” on financial market returns, we also look for signs of disagreement in the markets. Visually we look for increased noise/volatility in daily returns along with evidence of abnormal share turnover. We supplement these approaches with qualitative evidence on actual market participants’ opinions, or lack thereof, about the effects of the social-media inspired boycott.

The final method we employ to analyzing the effect of the “United Breaks Guitars” video series is better suited to analyzing customer market outcomes, particularly ones that evolve more slowly than financial market event studies. The synthetic control technique creates a counterfactual United that performs as if the YouTube videos targeting United had never been posted or gone viral. The synthetic control method does so by creating a weighted average of other airlines using data prior to the event and projects that data forward keeping the weights on the other airlines fixed. This approach is similar to financial market event studies; however, the underlying assumptions embedded in the methodological techniques differ.

C. Data

For the analyses of financial market reactions to the “United Breaks Guitars” event, we used data from the Center for Research in Security Prices (CRSP) for the 16 firms that appear in either of the two SIC 4-digit codes (4510 and 4512) that capture airlines which provided scheduled, domestic passenger service. We collected daily pricing and turnover data. In our analyses we employed data from 100 market days before the first YouTube video was posted (July 6, 2009) to 15 market days after the third and final YouTube video was posted.

For analyses of the customer market reaction to the events, we collected data from the Department of Transportation (DOT). The DOT maintains a rich set of data on airlines monthly operations that includes passenger, freight, and airport information for each route each major airline flies in its TransStats Dataset. The data include operational information on all 14 airlines that captured at least 1 percent of all airline passenger volume (which is a cutoff level for some

DOT data collection efforts and explains the two firm difference across methods).⁷ Our outcome variable of interest is monthly passengers by air carrier, as this what the “United Breaks Guitars” YouTube videos were targeting most directly in their boycott efforts. Our predictor or airline attribute variables include monthly data on characteristics of firm size and scope: passengers carried, pounds of freight carried, number of routes flown, number of international routes flown, and number of short routes (less than 500 miles) flown. Other predictor attributes included variables measured in June 2009 on the airlines’ market position, including: the number of routes for which a given airline had a monopoly, the number of airports an airline served, and the number of airports where an airline was a dominant carrier (defined as controlling greater than 50 percent of routes).

D. Results

To preview our findings, in the emergent phenomenon case of “United Breaks Guitars,” our event study and turnover financial market-based results support *H1a* and *H1b*, and our synthetic control customer market-based results support *H2*.

1. Financial Market Results

To conduct our event study, we calculated our abnormal returns using the three-step procedure described earlier. For our market benchmark, we employed a value-weighted index of the 16-firm industry portfolio described in the data section, as this portfolio allowed us to build a counterfactual performance for United using data from the publicly traded firms that are most similar to it.⁸ Our estimation window consisted of the 100 trading days prior to the release of the

⁷ Given the frequent M&A that has occurred in the airline industry over the last several decades, we had to find a way to deal with this reality in the data. For the purposes of consistency, we chose to treat airlines that merged at any point as the entity that remained in July 2009 since that was the date of the focal event we are analyzing. With regard to the two-firm difference across methods’ samples, our financial market event study results are robust to dropping the two firms that were publicly traded but not included in the DOT data.

⁸ In Appendix A, we present the results of a financial market event study that uses market-wide performance, but we note that the statistical and substantive findings presented are insensitive to this choice of benchmark.

first video on July 6, 2009. We thus regressed the industry benchmark (*Ind_Ret*) on United's share price (*UAL_Ret*) and produced the following equation (standard errors in parentheses):

$$UAL_Ret_t = -0.006 + 1.4976 Ind_Ret_t \quad (7)$$

(0.004) (0.081)

Equation 7's R^2 was 0.78 indicating that the industrial portfolio explained a substantial degree of the variation in United's share price in the estimation window. Having established this goodness-of-fit, we used this regression to predict our counterfactual, the expected returns for United absent the impact of the "United Breaks Guitars" videos in the relevant event window. We then computed abnormal returns for United by calculating the difference between the airline's actual returns and its expected (counterfactual) returns.

<Insert Figures 1 and 2 Here>

Figures 1 and 2 illustrate the last two steps of building our financial market counterfactual. Figure 1 plots the actual and predicted performance of United across time, and Figure 2 plots the difference between these two trends, the abnormal return, for United. In both figures, the dashed vertical lines demarcate the release dates of the three videos.

As is standard, we calculated cumulative abnormal returns as the simple sum of abnormal returns over the relevant event window. We examined the release of each of the three videos as separate events, as well as the cumulative impact of all three events. To ensure our results were not sensitive to our choice of event window, we estimated and reported four different event windows, ranging from 2 to 16 days (two calendar weeks) in length. Because the videos were posted by an actor who was not a significant market participant prior to their production and who posted them without prior publicity, we start the event windows on day 0; that is, on the date the video was released. Additionally, to assess the statistical significance of these abnormal returns, we utilized three different tests: the standard asymptotic approach based upon the standard

deviation in the estimation window, the bootstrapping approach advocated for by McWilliams and Siegel (1997), which is related to randomization-based forms of inference, and the SQ test developed by Gelbach, Helland, and Klick (2013), which is specifically designed for events that only effect one firm.⁹

Table 2 reports the results of our event studies. For the first and second “United Breaks Guitars” videos, there are no significant effects on United’s market value, across all four events windows and using any of the three tests for statistical significance. For the third video, there are significant effects for the two longest event windows (10 and 15 days), but only when we assess significance using the SQ test. Although this last point is interesting, it is not representative of the overall event study results.¹⁰ Moreover, it is further weakened by the null results in the cumulative event analysis—which along with the first video might be where we expect the largest (negative) effect. Overall, we find no significant and meaningful impact of the three “United Breaks Guitars” videos on United’s share price, regardless of event window length or statistical test.

<Insert Table 2 Here>

The results of the event study suggest that “United Breaks Guitars,” although a popular sensation, had few strategic implications for the management of the airline. However, this statement only holds if the assumptions behind an event study hold and if the event study is the appropriate tool with which to assess an emergent phenomenon. Figure 2 provides some evidence that the markets disagreed about and struggled to value the effects of United Breaks

⁹ We describe each of the latter two approaches in greater detail in Appendix A but note that the latter two are robustness checks of the traditional event study approach.

¹⁰ Caution should be used before accepting the alternative hypothesis based on this single result, however, as: (i) we do not have comparable findings when we use a traditional *t*-test nor when we use the bootstrap method; (ii) the results from the SQ test are only significant at the 0.10 level; and, (iii) the signs on the coefficient are opposite of expectation (i.e., these results suggest that United experienced positive cumulative abnormal returns in the wake of the third video).

Guitars, as after the release of the first video, in particular, the range in the abnormal returns increases, with the abnormal returns being significantly above the value predicted by the industry portfolio on some days and significantly below it on others, suggestive of additional noise.

We investigated a possible cause of this wider range in the second part of our financial market analysis. Following Hong and Stein (2007), we analyzed whether the financial markets disagreed about the potential effects of “United Breaks Guitars,” by examining United’s trading volume following the release of the first video.¹¹ We plotted abnormal turnover data against event time in market days to illustrate whether the turnover ratio increases around the event, as such an increase would be consistent with market participants having heterogeneous priors and thus considerable market disagreement over the implications of the new phenomenon of the social media-inspired boycott. In addition, we also plotted the cumulative number of YouTube views to demonstrate the speed with which the event diffused.

<Insert Figure 3 Here>

Figure 3 plots the industry-adjusted abnormal turnover ratio in United’s stock in event time around the first “United Breaks Guitars” video. The results of our turnover analysis suggest that relying on the financial markets to assess strategic implications of emergent phenomena may be inappropriate, which is consistent with both parts of *HI*. The spike in abnormal turnover and YouTube views, taken together, suggest that social media-inspired boycotts, as a manifestation of emergent phenomena, do lead to increased disagreement in markets due to participants holding heterogeneous priors about the potential effects of the new phenomena. The figure

¹¹ To do so, we first calculated the daily turnover for United by dividing the number of shares traded (volume) by the total number of shares outstanding. Second, to benchmark United’s turnover, we calculated an abnormal turnover ratio by (a) subtracting from United’s turnover the average turnover over the prior year in its industry (defined as the same 16 firms we employed in the event study, with the prior year ending 6 days pre-event) and (b) dividing this difference by the industry average turnover used in (a). We calculated this ratio for a window extending from five days prior to the event to 10 days after it in order to assess pre- and post-event turnover trends.

shows that as the first video accumulated additional views on YouTube, United's turnover ratio became more abnormal in comparison to its industry (at least through the first seven market days after the event). Further, the figure reveals that the financial market not only disagreed about the potential impact of the video but that the market's reaction to the event lagged the actual event by several days, which provides further evidence that the assumptions of event studies may break down when analyzing emergent phenomena.

Qualitative evidence further corroborates *H1b* by showing that there was disagreement among financial market participants. Although media reports suggested the social-media inspired boycott should have had a negative impact on United (Ayers, 2009), stock analysts did not appear to share this view of the event's impact: none of the sell-side analysts at investment banks or independent equity research firms who covered United (and whose reports appear in the Thomson ONE database) mentioned the videos in any of the reports they issued in the six months following the release of the first video. This absence suggests that, in contrast to more traditional risks for the airline (e.g., fuel prices, labor disputes, debt financing, and terrorism), analysts viewed the videos as relatively inconsequential.

The null results from the event study, together with the descriptive evidence from the abnormal turnover analysis, raise serious questions regarding whether the event study is an appropriate tool for studying emergent phenomena. In the particular case of the social media-inspired boycott, the financial markets expressed considerable disagreement over the meaning of the event and further, were behind in sending signals of any nature as to what the event meant for United's future performance. Due to these factors, it would be unwise for researchers or managers to interpolate from null results in a financial market event study of a new phenomenon that it has no strategic implications for a firm—as such null result may have been driven entirely by an underlying assumption failing to hold in a particular situation, namely the efficient markets hypothesis breaking down. At the very least, such null results should be analyzed along with

information from other assessments of the event’s potential impact to avoid a Type I error.

2. *Consumer Market Results*

To conduct our synthetic control analysis, we used the STATA “synth” package.¹² We created the synthetic representation of United using data prior to July 2009, since we did not want the relative weightings to reflect information from after the first YouTube video was posted. We used 50 months of data over which to optimize and hence, created the synthetic using data from September 2005 through June 2009.¹³

The primary weights used to create the synthetic United appear in Table 3. Although these were generated objectively through the dual optimization process synthetic control uses, they are easy to reconcile with intuition about what other firms are similar to United. Three firms received non-zero weights.¹⁴ American Airlines has the largest weight (0.495) of any other firm in the synthetic, which perhaps is not surprising as those familiar with the industry would recognize it as the other legacy carrier with whom United competes most directly, as they both are of similar size, operate on hub models, and even share some hubs (including Chicago O’Hare). Frontier receives the next largest weight (0.453); while the similarities to United may not be as obvious as for American at first, Frontier actually shares a number of routes and focus airports, such as Denver International, with United that may be subject to similar seasonal fluctuations. Finally, US Air receives a smaller, but still positive weight (0.052), as it and United are similar in size; moreover, both are Star Alliance members who often codeshare.

<Insert Table 3 here>

¹² We plan on making data and code available on the authors’ websites so others can learn from our example. We also provide additional robustness checks in an appendix but omit them here to economize on space.

¹³ We used data back until September 2005 because this is the date that United exited a Chapter 11 bankruptcy reorganization. We did not want to create a counterfactual with data during the bankruptcy reorganization period, as we did not want to use data from a period that might not be representative of the period when the event occurred.

¹⁴ On theoretical grounds, we exclude Continental as a potential control firm given their eventual merger with United in January 2012. Nevertheless, even we they are not excluded, they receive a zero weight.

Table 4 shows that applying these weights to generate a synthetic control unit provides a reasonable proxy of United’s underlying characteristics. The attributes of United and the synthetic United are very similar when averaging over the pre-treatment period. There are only a few marginal differences. Those are that the synthetic flies more short domestic routes, while the actual United flies more long, international routes. Nevertheless, on the most important attribute, monthly passengers, both airlines have approximately the same level of performance.

The primary results from our synthetic analysis appear in Figure 4, which graphs the actual firm’s performance against the synthetic firm’s performance—just like we did in Figure 1 for the event study approach. Again, the three vertical dashed lines in the figure represent the release date for each of the videos in the “United Breaks Guitars” series. Unlike in Figure 1, for the event study approach; however, following the release of the first “United Breaks Guitars” video, we see that the synthetic United and the actual United begin to diverge, implying that the event had a causal impact on the actual firms’ performance relative to how the firm would have performed in the absence of the social-media inspired boycott.

<Insert Figure 4 here>

Figure 5 clarifies this treatment effect, as it graphs the difference between the actual and counterfactual United’s passenger volumes—similar to our approach in Figure 2 in which we graph United’s abnormal stock returns. Figure 5 reveals that within a month or so of the release of the first YouTube video, United saw fewer passengers taking its flights. This implies that unlike in the event study results—which rely on strong and potentially untenable assumptions—in the synthetic control results, there is an adverse effect on firm performance since fewer passengers are flying on United than the counterfactual suggests should have been.

<Insert Figure 5 here>

Although the divergence is interesting in and of itself and suggests that we have found support for *H2*, it is important to determine whether or not this divergence is larger than we

would have expected otherwise. That is, is the effect statistically significant? One way to assess this visually is to compare the performance of United’s synthetic to synthetics constructed around the same event for airlines that were not targeted by the social-media inspired boycott. This comparison is called a placebo-across-units test. We conducted this test for synthetics that we could create good pre-period matches for and plotted the results in Figure 6.¹⁵

<Insert Figure 6 here>

The figure reveals that the effect on United was of a uniquely large magnitude, as the solid blue line representing United diverges further in a negative direction than any of the dashed grey-lines representing the synthetics for the control firms that were not targeted by the boycott.

The data presented in Figure 6 is not scale-independent, however, so it may not provide a completely accurate depiction. Hence, in Figure 7, we present data showing the distribution of how firms perform relative to their synthetics using a scale independent measure that is also signed.¹⁶ Examining it, we see that United occupies the most negative position on this graph, corroborating the interpretation that the social-media inspired boycott had a negative and real effect on United’s ability to attract passengers. We can also use the data in Figure 7 to construct a pseudo *p*-value of $1/14 \approx 0.07$. Although this pseudo *p*-value is based upon randomization inference rather than frequentist inference, it can be interpreted similarly to mean that the chance of obtaining an effect as negative and large as the one we observed for United is 0.07 if we were to select an airline from the pool of airlines at random.

<Insert Figure 7 here>

¹⁵ Note that one of the limitations of synthetic control is that it cannot feasibly create matches for all possible statistical units. This is particularly true for the largest and smallest statistical units by outcome, given that no weighted combination of other units that sums to one can possibly generate a value as large as the largest unit or as small as the smallest unit. In order to make the graph cleaner, we exclude these poor matchers here with RMSPE (root mean square predicted error) values in the pre-period greater than 2x that for United.

¹⁶ Fremeth, Holburn, and Richter (2013) innovated this signed approach as a modification to Abide et al.’s (2010) development of a scale-independent measure. The scale independent measure developed by Abide et al. (2010) reflects the ratio of the post-treatment root mean squared prediction error (RMPSE) over the pre-treatment RMPSE—since different units of different sizes may have prediction errors with relatively different magnitudes.

We save for an appendix additional falsification tests that bolster the strength of the synthetic control analysis and its application. These include: (i) a placebo-in-time test which demonstrates the validity of the method, (ii) a leave-one-out test that checks the sensitivity of the result to the inclusion of specific control unit that may be biasing the results, and (iii) an out-of-sample test that shows that at least over some representative sample periods, we can treat the weightings in synthetic control as if they are constant over some future window.

V. DISCUSSION/CONCLUSION

The analysis of emergent phenomena is an important aspect of management research and practice. To date, however, tools for engaging in such analyses have been built upon fundamental assumptions that could become untenable, particularly around the emergent events of the greatest interests to both scholars and practitioners. To build more reliable research designs, we have advocated for the use of synthetic control methods in place of or along side event study methods. Using insights from behavioral finance, we have shown in an application to the “United Breaks Guitars” social-media inspired boycott, how and why event study methods can break down, yet how and why synthetic control methods can still yield significant, and more reliable, results about firm performance.

The divergent results between the event study method and the more reliable synthetic control method (in this case) have implications for management practice. They suggest that managers cannot look to the financial markets alone for information about how emergent phenomena may be affecting them but rather, that they should be monitoring other real time metrics of performance. This may be particularly true at organizations for which reputation is their “most valuable asset” and with respect to threats that can originate in the social media realm but then be amplified in the traditional media (Diermeier, 2011)—given the results of our application which also have implications for management practice. Firms need to be proactive in

dealing with threats in the non-market environment, as the bargaining power of customers rises when they have credible means of broadcasting their complaints (Baron, 1995), as is the case in a world increasingly connected through social media. The proactive nature of responding to social media must go beyond having a social media engagement team. In many cases, avoiding incidents that prompt consumers such as Dave Carroll to broadcast their complaints can be achieved simply by listening and responding to angry customers. Nevertheless, such a level of responsiveness may be difficult to achieve in larger organizations, given the advantages to them of maintaining strict operating routines (Nelson and Winter, 1982). Giving lower level employees the discretion required to do what makes sense in a given situation might be increasingly valuable, however, if the small costs of such action can help avoid the larger costs individual customers can now impose upon a firm through social media broadcasts of complaints (Phillips et al., 2010; Hambrick and Finkelstein, 1987).

Our analysis of the “United Breaks Guitars” incident also has implications for researcher. We did not choose to analyze this event merely because social media boycotts are interesting phenomena, but rather because it helped illustrate generic hypotheses about how to analyze the effects of emergent phenomena, and it allowed us to check whether the assumptions underlying existing event study methods break down. We have shown that it is difficult for researchers who do not find a significant effect for an emergent event using financial market event study techniques to say with confidence that they have not merely uncovered a false negative result. To ensure that they have not made a Type I error, they should go one step further to see if there are any indications of disagreement in markets, such as higher volatility or share turnover. If these indications are found, we implore future researchers to investigate the same emergent phenomena by looking at performance metrics in customer or product markets using the synthetic control method. To summarize, at a minimum, in the presence of a null event study finding, we advocate that researchers examine non-financial outcomes as a robustness check on

event study assumptions, and more boldly, we advocate that such examinations should become a core part of the analysis of emergent phenomena.

Although we apply our arguments to the specific manifestation of the social media-inspired boycott, this paper examines the methods scholars can employ to analyze the potential strategic implications of new managerial phenomena broadly. The outcomes, events, and unit of analysis that can be studied as emergent phenomena are not limited. In the boycotts space alone, synthetic control could be used to evaluate the outcomes of boycotts related to pro-social claims (McDonnell and King, 2013) or to corporate sponsored boycotts (McDonnell, 2012). The synthetic control method could also be used to analyze many other types of events including whether accidents, natural disasters, or litigation affect firm performance. Moreover, the method could be used to study the effects of events on other outcomes, such as whether downgrades in social ratings have a meaningful impact on consumer markets rather than financial markets, how the adoptions of codes of ethics or other social responsibility practices are perceived by customers, or how bringing in a new manager affects employee morale. Although synthetic control may improve upon the reliability of analyses of emergent and other phenomena in management, it is not a panacea, as there must be a focal event and a focal target of that event which may not always be the case with phenomena reshaping business such as climate change.

Having demonstrated paired use of synthetic control and event studies to study emergent phenomenon, we hope to see others adopt this methodological approach, as it will improve the reliability of answers to questions of concern to managers and strategy researchers alike. To this end we are sharing synthetic control data and code on public websites as well as our event study code including that for calculating standard errors under different approaches. Although Durand and Vaara (2009) may be correct to claim that causation is often poorly understood in strategy research, the only way to improve the quality and reliability of causal claims is for new methods, such as synthetic control, to be diffused as widely and as possible.

REFERENCES

- Abadie A, Gardeazabal J. 2003. The economic costs of conflict: A case study of the Basque country. *American Economic Review* **93**(1): 113–132.
- Abadie A, Diamond A, Hainmueller J. 2010. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American Statistical Association* **105**(490): 493–505.
- Abadie A, Diamond A, Hainmueller J. 2012. Comparative politics and the synthetic control method. Working paper available at <http://ssrn.com/abstract=1950298>.
- Ayers C. 2009. Revenge is best served cold—on YouTube. *The Times of London*. July 22, 2009.
- Bagley CE. 2008. Winning legally: The value of legal astuteness. *Academy of Management Review* **33**(2): 378–390.
- Baron, DP. 1995. Integrated strategy: Market and non-market components. *California Management Review* **37**(2): 47–65.
- Baron DP. 2005. Competing for the public through the new media. *Journal of Economics and Management Strategy* **14**(2): 339–376.
- Baron DP, Diermeier D. 2007. Strategic activism and nonmarket strategy. *Journal of Economics and Management Strategy* **16**(3): 599–634.
- Bartley T, Child C. 2011. Movements, markets, and fields: The effects of anti-sweatshop campaigns on U.S. firms, 1993–2000. *Social Forces* **90**(2): 425–451.
- Bartley T. 2007. Institutional emergence in an era of globalization: The rise of the transnational private regulation of labor and environmental conditions. *American Journal of Sociology* **113**(2): 297–351.
- Basdeo DK, Smith KG, Grimm CM, Rindova VP, and Derfus PJ. 2006. The impact of market actions on firm reputation. *Strategic Management Journal* **27**(12): 1205–1219.
- Bilmeier A, Nannicini T. 2013. Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics* **95**(3): 983–1001.
- Black F. 1986. Noise. *Journal of Finance* **41**(3): 529–543.
- Brealey RA, Myers SC, Allen F. 2010. *Principles of Corporate Finance*. 9th ed. McGraw-Hill Irwin: Boston, MA.
- Briscoe F, Safford S. 2008. The Nixon in China effect: Activism, imitation, and institutionalization of contentious practices. *Administrative Science Quarterly* **53**(3): 460–491.
- Carroll D. 2012. *United Breaks Guitars: The Power of One Voice in the Age of Social Media*. Hay House: Carlsbad, CA.
- Cavallo E, Galiani S, Noy I, Pantano J. Forthcoming. Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*.
- Chatterji AK, Toffel MW. 2010. How firms respond to being rated. *Strategic Management Journal* **31**(9): 917–945.
- Christensen CM. 1997. *The Innovator’s Dilemma: When New Technologies Cause Great Firms to Fail*. Harvard Business School Press: Boston, MA.
- Coffman M, Noy I. 2012. Hurricane Iniki: Measuring the long- term economic impact of a natural disaster using synthetic control. *Environment and Development Economics* **17**(2): 187–205.
- De Long JB, Shleifer A, Summers LH, Waldmann RJ. 1990. Noise trader risk in financial markets. *Journal of Political Economy* **98**(4): 703–738.

- Deighton J, Kornfeld L. 2011. *United Breaks Guitars*. Harvard Business School: Boston, MA.
- Diermeier D. 2011. *Reputation Rules: Strategies for Building Your Company's Most Valuable Asset*. McGraw-Hill: New York, NY.
- Dobbin F. 2009. *Inventing Equal Opportunity*. Princeton University Press: Princeton, NJ.
- Dowling CM, Enos RD, Fowler A, Panagopoulos C. 2012. Does public funding chill political speech? Exploiting a court injunction as a natural experiment. *Election Law Journal* **11**(3): 302–315.
- Drucker P. 1995. *Managing in a Time of Great Change*. Dutton: New York.
- Dunne D. 2010. *United breaks guitars*. University of Toronto Rotman School of Management: Toronto, ON.
- Durand R, Vaara E. 2009. Causation, counterfactuals, and competitive advantage. *Strategic Management Journal* **30**(12): 1245–1264.
- Evangelopoulos N, Magro M, Sidorova A. 2012. The dual micro/macro informing role of social media sites: Can Twitted macro messages help predict stock prices? *Informing Science* **15**: 247–268.
- Fama E. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* **25**(2): 383–417.
- Fligstein N. 1991. The structural transformation of American industry: An institutional account of the causes of diversification in the largest firms. In *The New Institutionalism in Organizational Analysis*, Powell WW, DiMaggio PJ (eds). University of Chicago Press: Chicago, IL; 311–336.
- Foucault T, Sraer D, Thiesmar DJ. 2011. Individual investors and volatility. *Journal of Finance* **66**(4): 1369–1406.
- Fremeth A, Holburn GLF, Richter BK. 2013. Making causal inferences in small samples using synthetic control methodology: Did Chrysler benefit from government assistance? Working paper available at <http://ssrn.com/abstract=2135294>.
- Friedman M. 1999. *Consumer Boycotts: Effecting Change Through the Marketplace and the Media*. Routledge: New York.
- Garrett DE. 1987. The effectiveness of marketing policy boycotts: environmental opposition to marketing. *Journal of Marketing* **51**(2): 46–57.
- Gelbach JB, Helland E, Klick J. 2013. Valid inference in single-firm, single-event studies. *American Law and Economics Review*, Forthcoming.
- Hambrick DC, Finkelstein S. 1987. Managerial discretion: A bridge between polar views of organizational fates. In *Research in Organizational Behavior*, Staw B, Cummings LL (eds). JAI Press: Greenwich, CT; 369–406.
- Harris M, Raviv A. 1993. Differences of opinion make a horse race. *Review of Financial Studies* **6**(3): 473–506.
- Henisz W, Zelner BA. 2003. The strategic organization of political risks and opportunities. *Strategic Organization* **1**(4): 451–460.
- Hillman A, Zardkoohi A, Bierman L. 1999. Corporate political strategies and firm performance: Indications of firm-specific benefits from personal service in the U.S. government. *Strategic Management Journal* **20**(1): 67–81.
- Hinrichs P. 2012. The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *Review of Economics and Statistics* **94**(3) 712–722.
- Hudson S. 2010. The effects of performance-based teacher pay on student achievement. Stanford

- Institute for Economic Policy Research Discussion Paper No. 09-023.
- Holland PW. 1986. Statistics and causal inference. *Journal of the American Statistical Association* **81**(396): 945-960.
- Hong H, Stein JC. 2007. Disagreement and the stock market. *Journal of Economic Perspectives* **21**(2): 109–128.
- Kandel E, Pearson ND. 1995. Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy* **103**(4): 831–872.
- King, BG. 2008. A political mediation model of corporate response to social movement activism. *Administrative Science Quarterly* **53**(3): 395–421.
- King BG. 2011. The tactical disruptiveness of movements: Sources of market and mediated disruption in corporate boycotts. *Social Problems* **58**(4): 491–517.
- King BG, McDonnell M-H. 2011. Case study – Delta Airlines boycott. *Financial Times* November 21, 2011.
- King BG, Pearce NA. 2010. The contentiousness of markets: Politics, social movements, and institutional change in markets. *Annual Review of Sociology* **36**: 249–267.
- King, BG, Soule, S. 2007. Social movements as extra-institutional entrepreneurs: The effect of protests on stock price returns. *Administrative Science Quarterly* **52**: 413–442.
- Koku PS, Akhigbe A, Springer TM. 1997. The financial impact of boycotts and threats of boycotts. *Journal of Business Research* **40**(1): 15–20.
- Lenox M, Eesley C. 2009. Private environmental activism and the selection and response of firm targets. *Journal of Economics and Management Strategy* **18**(1): 45–73.
- Liu K, Arthurs JD, Nam D, Mousa F-T. 2013. Information diffusion and value redistribution among transaction partners of the IPO firm. *Strategic Management Journal*, Forthcoming.
- Loughran T, Ritter JR. 1995. The new issues puzzle. *Journal of Finance* **50**(1): 23–51.
- Luders, JE. 2006. The economics of success: Business responses to civil rights mobilization. *American Journal of Sociology* **111**(4): 963–998.
- March JG, Spourll LS, Tamuz M. 1991. Learning from samples of one or fewer. *Organization Science* **2**(1): 1–13.
- McDonnell M-H. 2012. If you can't beat them, join them: Corporate sponsorship of social movement boycotts. Working paper available at <http://ssrn.com/abstract=2137371>.
- McDonnell M-H, King BG. 2013. Keeping up appearances: Reputation threat and prosocial responses to social movement boycotts. *Administrative Science Quarterly* **58**(3): 387–419.
- McWilliams A, Siegel D. 1997. Event studies in management research: Theoretical and empirical issues. *Academy of Management Journal* **40**(3): 626–657.
- McWilliams A, Siegel D, Teoh SH. 1999. Issue in the use of the event study methodology: A critical analysis of corporate social responsibility studies. *Organizational Research Methods* **2**(4): 340–365.
- Montalvo JG. 2011. Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections. *Review of Economics and Statistics* **93**(4): 1146–1154.
- Nag R, Hambrick DC, Chen M-J. 2007. What is strategic management, really? Inductive derivation of a consensus definition of the field. *Strategic Management Journal* **28**(9): 935–955.
- Nelson RR, Winter SG. 1982. *An Evolutionary Theory of Economic Change*. Belknap Press: Cambridge, MA.
- Ofek E, Richardson M. 2003. DotCom mania: The rise and fall of internet stock prices. *Journal of Finance* **58**(3): 1113–1137.

- Phillips RA, Berman SL, Elms H, Johnson-Cramer ME. 2010. Strategy, stakeholders and managerial discretion. *Strategic Organization* **8**(2): 176–183.
- Pruitt SW, Friedman M. 1986. Determining the effectiveness of consumer boycotts: A stock price analysis of their impact on corporate targets. *Journal of Consumer Policy* **9**(4): 375–387.
- Pruitt SW, Wei KCJ, White RE. 1988. The impact of union-sponsored boycotts on stock prices of target firms. *Journal of Labor Research* **9**(3): 285–289.
- Reid E, Toffel M. 2009. Responding to public and private politics: Corporate disclosure of climate change strategies. *Strategic Management Journal* **30**(11): 1157–1178.
- Rosenbloom RS. 2000. Leadership, capabilities, and technological change: The transformation of NCR in the electronic era. *Strategic Management Journal* **21**(10-11): 1083–1103.
- Rubin D. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66**(5): 688–701.
- Rubin, D. 2005. Causal inference using potential outcomes. *Journal of the American Statistical Association* **100**(469): 322–331.
- Schurman R. 2004. Fighting “Frankenfoods”: Industry opportunity structures and the efficacy of the anti-biotech movement in Western Europe. *Social Problems* **51**(2): 243–268.
- Sears J, Hoetker G. 2013. Technological overlap, technological capabilities, and resource recombination in technological acquisitions. *Strategic Management Journal*, Forthcoming.
- Seidman G. 2007. *Beyond the Boycott: Labor Rights, Human Rights, and Transnational Activism*. Russell Sage Foundation: New York, NY.
- Shiller RJ. 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* **71**(3): 421–436.
- Shiller RJ. 2003. From efficient markets theory to behavioral finance. *Journal of Economic Perspectives*. **17**(1): 83–104.
- Soule S. 2009. *Contention and Corporate Social Responsibility*. Cambridge University Press: New York, NY.
- Stuart TE. 2000. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic Management Journal* **21**(8): 791–811.
- Teoh SH, Welch I, Wazzan CP. 1999. The effect of socially activist investment policies on the financial markets: Evidence from the South African boycott. *Journal of Business* **72**(1): 35–89.
- Vasi IB, King B. 2012. Social movements, risk perceptions, and economic outcomes: The effect of primary and secondary stakeholder activism on firms’ perceived environmental risk and financial performance. *American Sociological Review* **77**(4): 573–596.
- Venkatraman N, Ramanujam V. 1986. Measurement of Business Performance in Strategy Research: A Comparison of Approaches. *Academy of Management Review*. **11**(4): 801–814.
- Walker ET, Martin AW, McCarthy, JD. 2008. Confronting the state, the corporation, and the academy: The influence of institutional targets on social movement repertoires. *American Journal of Sociology* **114**(1): 35–76.
- Werner T. 2012. *Public Forces and Private Politics in American Big Business*. Cambridge University Press: Cambridge, UK.
- Woolridge JR, Snow CC. 1990. Stock market reaction to strategic investment decisions. *Strategic Management Journal* **11**(5): 353–363.
- Yin R. 2013. *Case Study Research: Design and Methods*. 5th ed. Sage Publications: Los Angeles,

CA.

FIGURES AND TABLES

FIGURE 1 – Actual vs. Counterfactual in Event Study

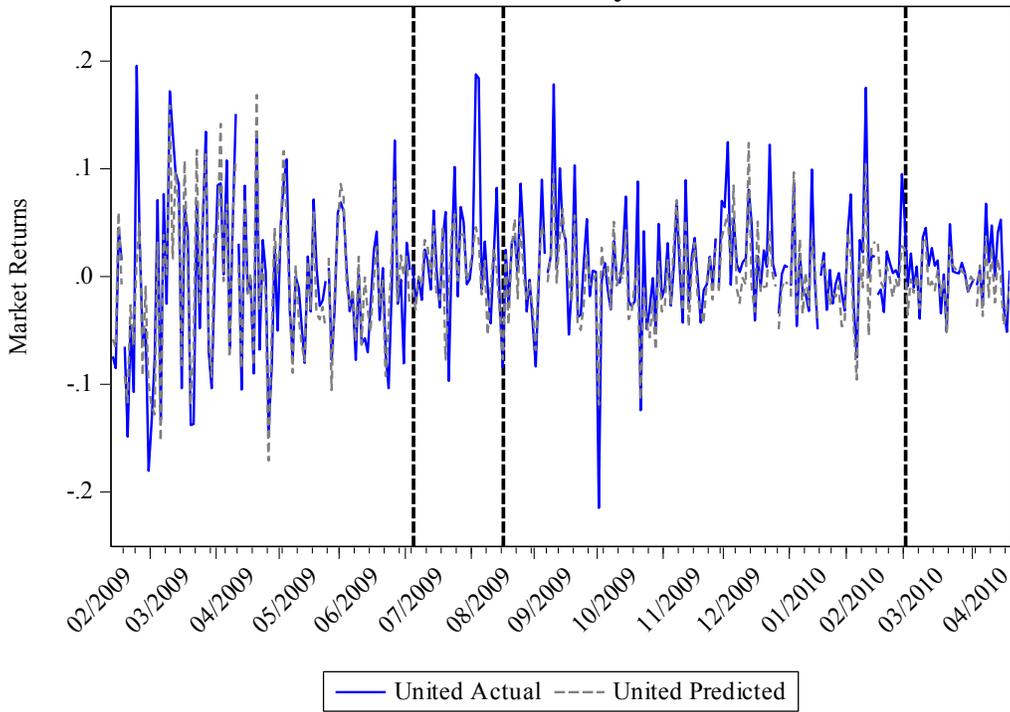


FIGURE 2 – Abnormal Returns in Event Study (i.e. the Treatment Effect or Lack Thereof)

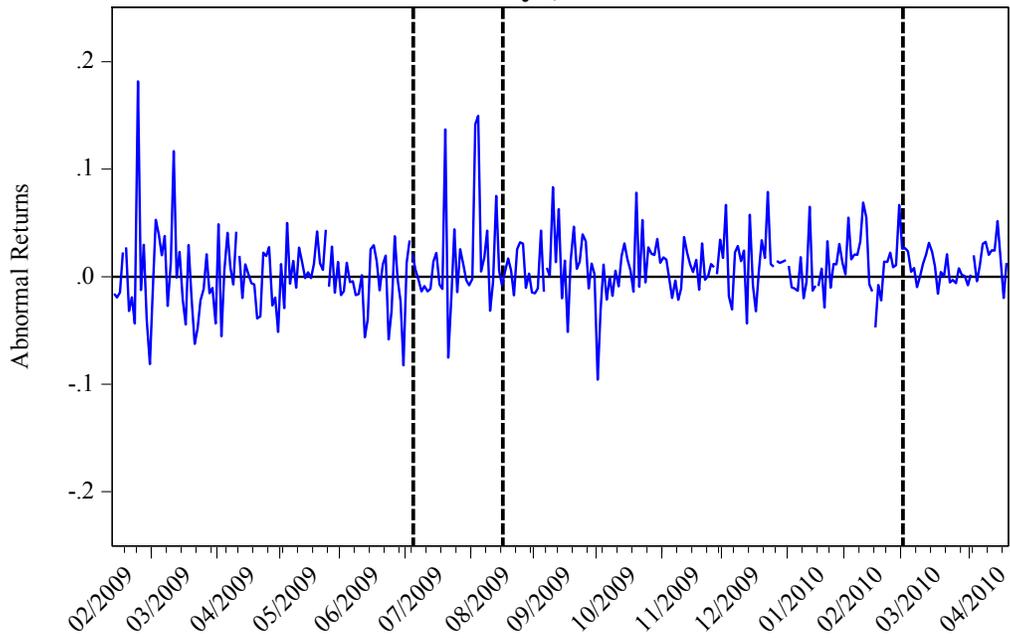


FIGURE 3 – Abnormal Turnover in United Stock around the Release of the First “United Breaks Guitars” Video, July 6, 2009

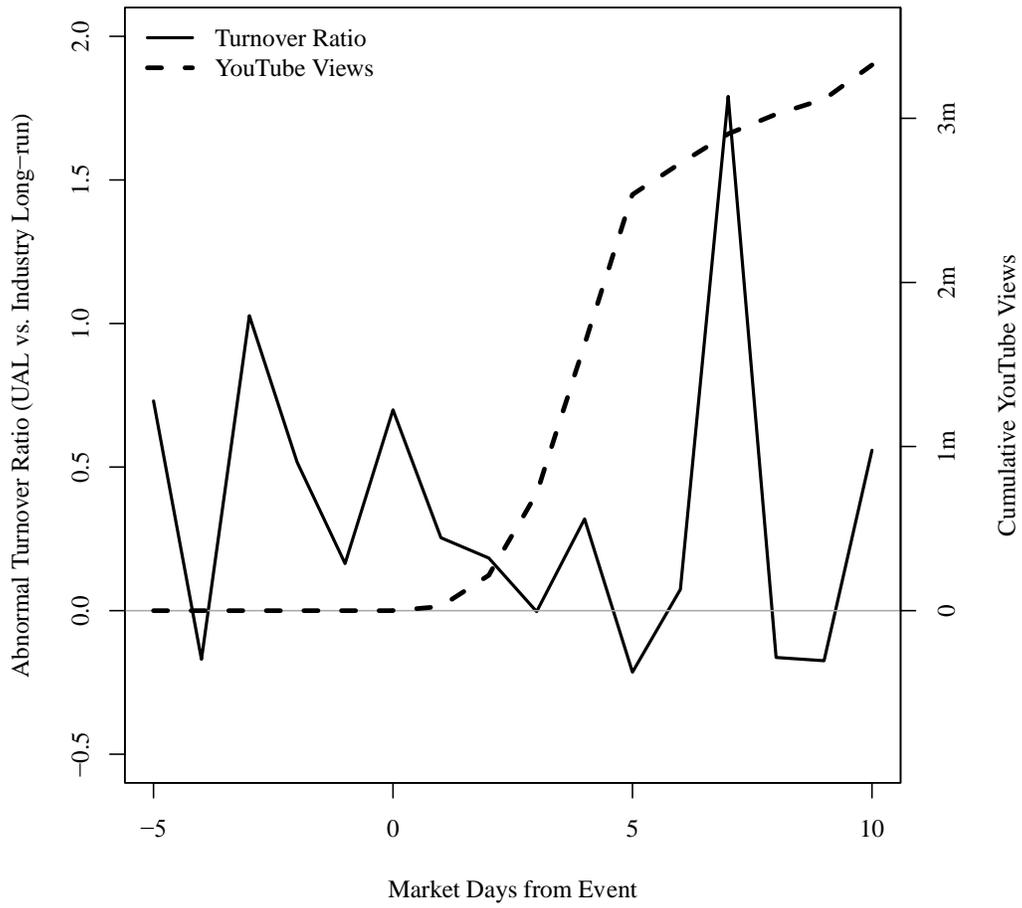


FIGURE 4 – Baseline Synthetic Control Results: Actual vs. Counterfactual United

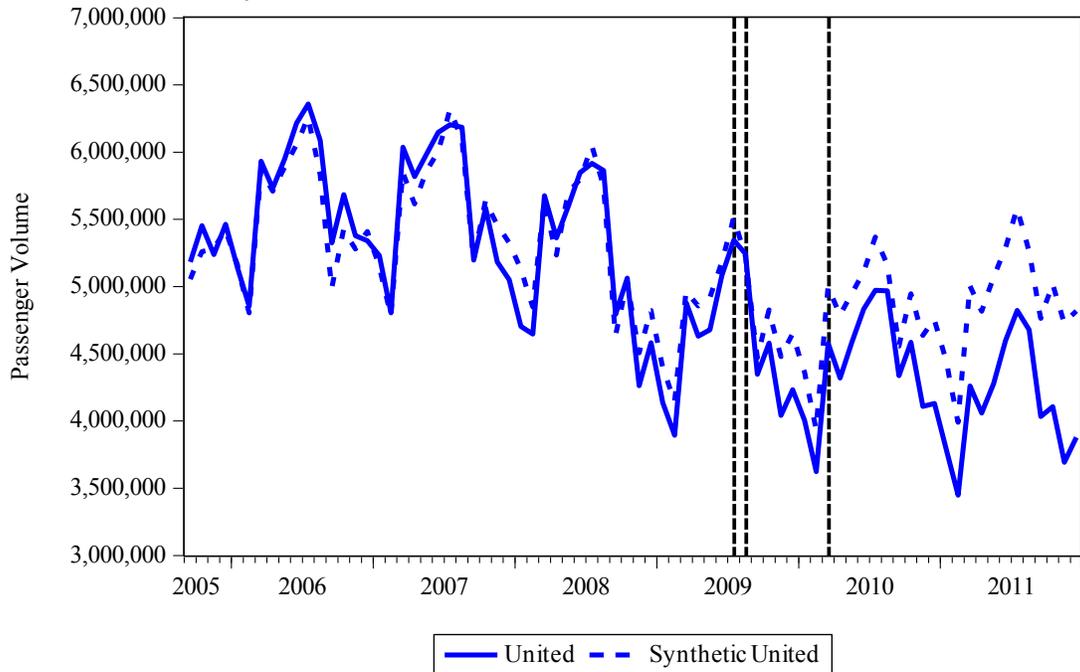


FIGURE 5 – Synthetic Control Difference Results: Actual minus Counterfactual United

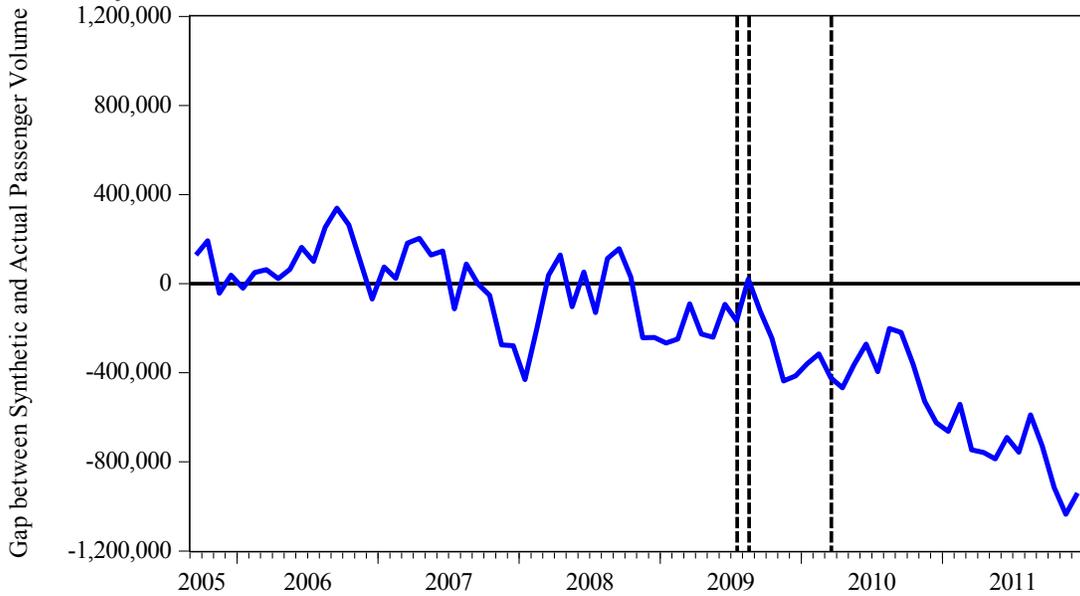


FIGURE 6 - Placebo Test Among Units (Excluding those with RMPSE 2x United)

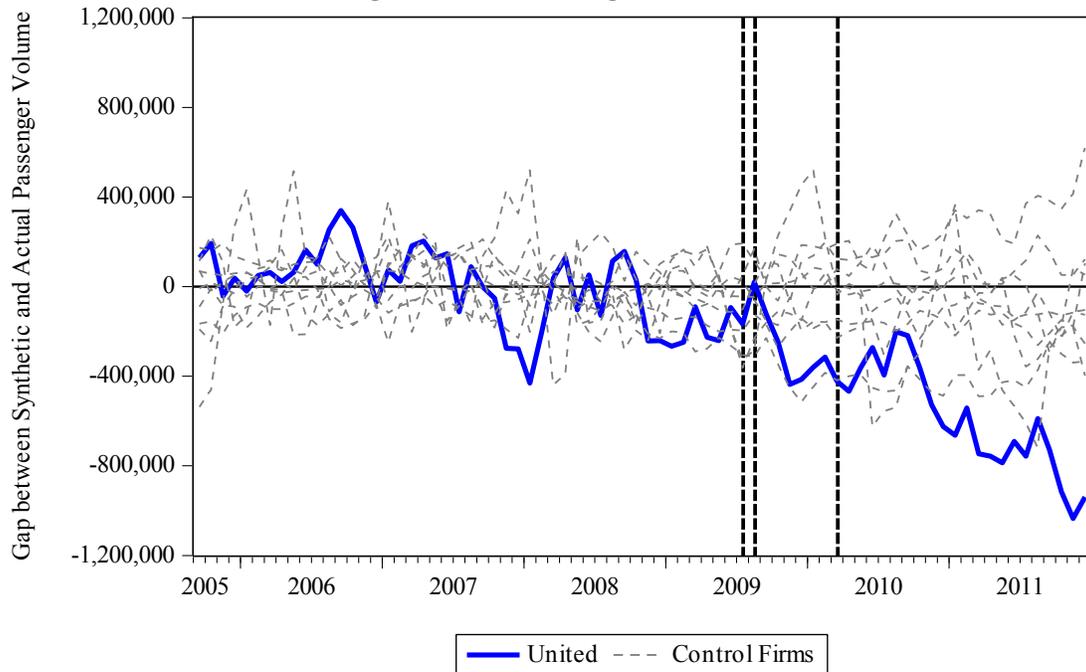


FIGURE 7 – Treatment Extremity Distribution

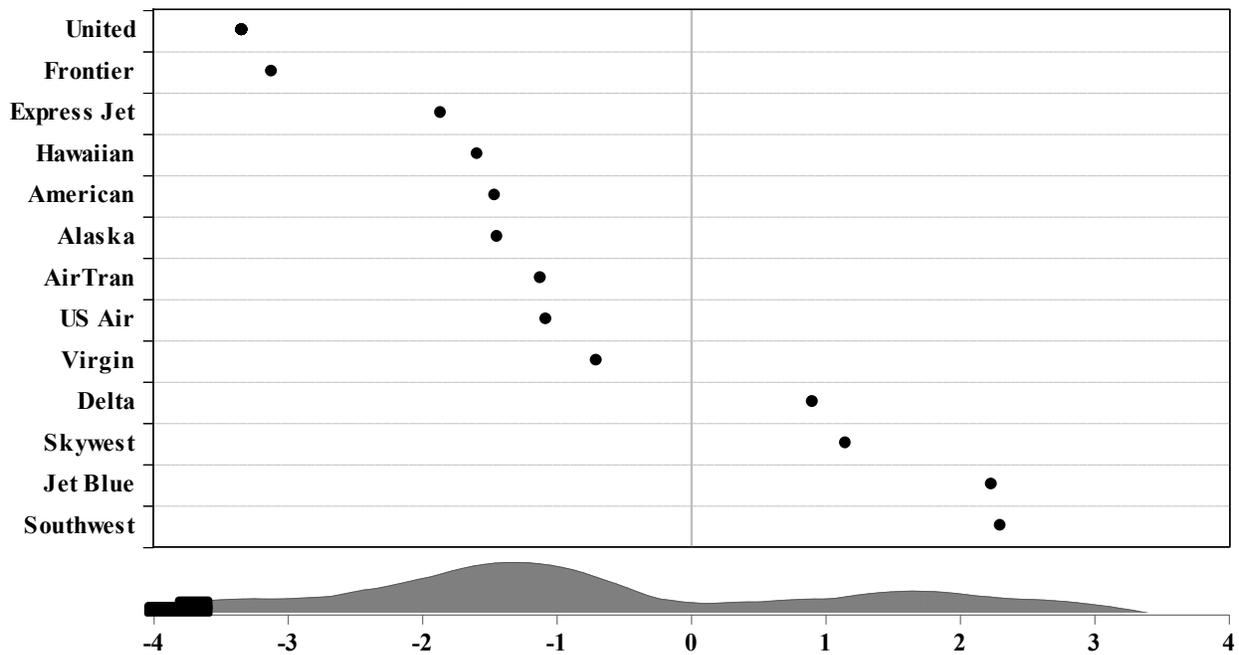


TABLE 1 – Comparison of Single Firm Counterfactual Methodologies

	EVENT STUDIES	SYNTHETIC CONTROL
Data	<ul style="list-style-type: none"> Financial market data (i.e., returns) 	<ul style="list-style-type: none"> Product or consumer market data (e.g., sales)
Window for Observing Effects	<ul style="list-style-type: none"> Immediate/short 	<ul style="list-style-type: none"> Unlimited/unconstrained
Basis for Counterfactual	<ul style="list-style-type: none"> Rational model of financial market returns 	<ul style="list-style-type: none"> Observable similarities between firms
Major Assumptions Underlying Statistical Tests of Treatment Effect	<ul style="list-style-type: none"> No major confounding events Representativeness of counterfactual Semi-strong form of the efficient markets hypothesis holds: <ul style="list-style-type: none"> Market participants have homogenous expectations Quick information dissemination/ reactions (no leaks of information) 	<ul style="list-style-type: none"> No major confounding events Representativeness of counterfactual Firms that are observably similar tend to stay observably similar

TABLE 2 - Event Study Results

Event Window p-value	Video 1 July 6, 2009	Video 2 August 17, 2009	Video 3 March 1, 2010	Cumulative All Video Dates
[0,1]	0.028	-0.003	0.052	0.077
t-test	0.596	0.949	0.321	0.400
Bootstrapped	0.502	0.876	0.218	0.476
SQ-test	0.260	0.430	0.110	0.190
[0,5]	-0.011	0.028	0.079	0.096
	0.901	0.759	0.656	0.544
	0.999	0.790	0.450	0.464
	0.550	0.400	0.190	0.230
[0,10]	-0.005	0.068	0.167	0.230
	0.967	0.580	0.181	0.285
	0.999	0.474	0.322	0.262
	0.510	0.280	0.060*	0.230
[0,15]	0.068	0.079	0.186	0.333
	0.646	0.595	0.216	0.201
	0.654	0.632	0.144	0.186
	0.320	0.300	0.070*	0.200

* Significant at 0.10

Cumulative abnormal returns were calculated using a value-weighted industry-specific market model. The three sets of *p-values* were calculated using asymptotic *t*-tests employing the estimation window's standard deviation, the bootstrap method outlined in McWilliams and Siegel (1997), and the SQ test developed by Gelbach, Helland, and Klick (2013), respectively.

TABLE 3 – Synthetic Control Weights

<i>Company</i>	<i>Weight</i>	<i>Company</i>	<i>Weight</i>
AirTran	0	Hawaiian	0
Alaska	0	Jet Blue	0
American	0.495	Skywest	0
Continental	<i>n/a</i>	Southwest	0
Delta	0	United	<i>focal</i>
Express Jet	0	US Air	0.052
Frontier	0.453	Virgin	0

Notes:

Continental is excluded as a potential match on theoretical grounds, given its later merger with United. Nevertheless, even when Continental is allowed to be a match, it receives 0 weight.

TABLE 4 – Synthetic Control Attributes

Comparison between United and Synthetic United (Prior to July 2009)		
<i>Attribute</i>	<i>United</i>	<i>Synthetic United</i>
Passengers Carried, Monthly	5,356,163	5,361,038
Pounds of Freight Carried, Monthly	15,200,000	11,000,000
Number of Routes, Monthly	1042	972
Number of International Routes, Monthly	318	241
Number of Short Routes, Monthly	134	177
Number of Monopoly Routes	476	419
Number of Airports Served	133	145
Number of Airports where Dominant Carrier	8	15

APPENDIXES FOR
**“Financial Market versus Consumer Market Responses to Emergent Phenomena:
An Application to a Social Media-Inspired Boycott”**

APPENDIX A - EVENT STUDY APPENDIX

In this appendix, we present (i) a robustness check of our event study analysis that uses a market-wide performance benchmark in place of the industry-based index presented in the paper and (ii) discuss the technical details behind the calculation of the three different sets of standard errors that we report in the event studies here and in the paper.

A. Market-wide Benchmark

In the text, we present results from an event study in which we constructed our counterfactual using a market model based upon an industry portfolio composed of the 16 publicly traded firms in the same industry as United at the time of the United Breaks Guitars videos. An alternative approach would have been to employ a market model based upon market-wide value-weighted returns. We choose to emphasize the former, as in constructing a counterfactual using observational data, we aimed to maximize the comparability of our focal firm and the components of its counterfactual; that is, we choose “most similar” firms in its industry in order to strengthen the assumption of unit homogeneity, which provides us a much stronger basis from which to estimate treatment effects (Rosenbaum, 2005).

<Insert Table A1 Here>

As a robustness check, we report the results of our market-wide analysis in Table A1. The results of this analysis are similar to those using the industry-wide portfolio in that there is little evidence that would lead us to reject the null hypothesis that the effect of United Breaks Guitars on United’s financial performance was insignificant. Only when analyzing the three videos cumulatively and only when using the SQ test did we find a statistically significant effect and only for two of the four event windows. Given the lack of results for the other two statistical tests and the other two event windows, as well as a positive sign on the coefficients (opposite of expectation), these results are not particularly robust and provide fairly weak evidence in favor

of a positive effect for the United Breaks Guitars videos. Thus, these results, along with the superiority of the industry portfolio market model, do not alter the conclusions we draw in the text.

B. Event Study Standard Errors

As mentioned in the main text, we calculated three sets of standard errors in our event studies. First, we conducted standard asymptotic t -tests based upon the standard deviation of the prediction errors in the 100-day estimation window. We used the prediction errors from the estimation window as the variability of the prediction errors during the event window is affected by the event itself, and serial correlation can occur when errors are calculated from consecutive event days. Second, we also conducted significance tests using the bootstrapping method employed outlined by McWilliams and Siegel (1997). To do so, for each event, we first calculated cumulative abnormal returns using data from the estimation window for the same number of consecutive days as used in the given event window (sample A); second, we drew a sample (sample B) of 1,000 returns with replacement from sample A; and third, we calculated two-tailed p -values, where $p = \frac{\text{number of observations in sample B with values greater (lower) than the cumulative abnormal return}}{1,000} \times 2$, if the cumulative abnormal return was ≥ 0 (if < 0).

Third, because our study is a single-firm event study, we also calculated the SQ test developed by Gelbach, Helland, and Klick (2013). As Gelbach and his co-authors note, empirical inferences derived from standard tests for statistical significance in single-firm event studies (our first test of statistical significance) have a high likelihood of producing an invalid result. Using return data over eight years to test the traditional approach, Gelbach, Helland, and Klick found that it is prone to both Type I and Type II errors. As a result, they argue that analysts should instead find where their event's cumulative abnormal return falls in a distribution of cumulative abnormal returns calculated in a comparable manner using pre-event data. If the

event's cumulative abnormal return is in the top or bottom 0.05 percentile of this distribution, then we would be able to reject the null hypothesis. They refer to this approach as the SQ test, and we implemented it in our analysis by comparing our events' cumulative abnormal returns to a distribution constructed of 100 comparable cumulative abnormal returns calculated from estimation window data.

Additional Reference

Rosenbaum PR. 2005. Heterogeneity and causality: Unit heterogeneity and design sensitivity in observational studies. *The American Statistician* **59**(2): 147–152.

TABLE A1 - Event Study Results

Event Window p-value	Video 1 July 6, 2009	Video 2 August 17, 2009	Video 3 March 1, 2010	Cumulative All Video Dates
[0,1]	0.048	-0.013	0.025	0.060
t-test	0.295	0.444	0.388	0.346
Bootstrapped	0.562	0.890	0.688	0.718
SQ-test	0.280	0.450	0.360	0.210
[0,5]	0.053	-0.006	-0.001	0.046
	0.366	0.515	0.498	0.440
	0.686	0.972	0.998	0.804
	0.340	0.470	0.500	0.210
[0,10]	0.028	0.088	0.149	0.266
	0.554	0.335	0.236	0.230
	0.570	0.440	0.344	0.432
	0.290	0.220	0.160	0.090*
[0,15]	0.068	0.133	0.150	0.351
	0.393	0.397	0.275	0.209
	0.344	0.298	0.280	0.416
	0.190	0.150	0.140	0.010*

* Significant at 0.10

Cumulative abnormal returns were calculated using a value-weighted market-wide market model. The three sets of *p-values* were calculated using asymptotic *t*-tests employing the estimation window's standard deviation, the bootstrap method outlined in McWilliams and Siegel (1997), and the SQ test developed by Gelbach, Helland, and Klick (2013), respectively.

APPENDIX B - SYNTHETIC CONTROL APPENDIX

In this appendix, we present (i) additional resources for strategy researchers interested in applying synthetic control to their research questions and (ii) additional analyses and robustness checks of our application to the “United Breaks Guitars” case. In the case of the latter, we present these results in the belief that all researchers using synthetic control should conduct comparable checks of the robustness of their results.

A. Synthetic Control Resources

One broader goal of this paper is to help interested readers find the resources necessary to learn synthetic control and to understand its limitations, such that they can apply the method in their own research.

1. Data and Code

We firmly believe that the best way to learn how to apply synthetic control is experientially. That is, interested researchers ought to download data and code from an existing, published example. To that end, we are sharing our data and code.

In order to run synthetic control, the necessary software must be installed for one of three statistical software packages: STATA, R, or MATLAB. Although strategic management researchers less commonly use the latter two, STATA is widely adopted package throughout the social sciences. All three of these pieces of software are downloadable at MIT political scientist Jens Hainmuller’s website,¹⁷ or alternatively, the synthetic control software can be directly installed from within STATA by inputting “ssc install synth, replace all” into the command prompt. One advantage of downloading the software for STATA is that it includes the data set from Abadie, et al.’s (2010) study of the effect of California’s Tobacco Control Program. Further, the help commands for synthetic control within STATA illustrate how to technically

¹⁷ <http://www.mit.edu/~jhainm/synthpage.html>

implement the approach using commands that can be run on this dataset. Working through this example and the code for our own example of “United Breaks Guitars” ought to aid researchers wanting to use the approach with their own data.

2. Primers to Read

Although the first published paper using synthetic control is Abadie and Gardeazabl’s (2003) study of the costs of terrorism in the Basque Country, it is not the best initial paper to read if one is interested in developing deeper technical knowledge of synthetic control methods, as the approach was in its infancy and robustness checks and means of assessing “confidence” were not fully developed by 2003. Instead, we suggest that readers start with Abadie et al.’s (2010) California smoking paper, which provides a full technical introduction but does so succinctly. Readers interested in even more detail on the method itself could then turn to Abadie et al.’s (2012) study on the reunification of Germany, which although more detailed particularly with respect to robustness checks, remains highly accessible. Finally, for additional information on the nuances of synthetic control in strategic management contexts, researchers can read Fremeth, et al. (2013).

B. Additional Synthetic Control Analysis and Robustness for “United Breaks Guitars”

In order to avoid distracting readers who were wholly unfamiliar with synthetic control, we minimized the scope of the analysis presented in the paper, specifically by limiting the number of robustness checks we presented. Thus, in the remainder of this appendix, we present the results of three additional robustness checks of our synthetic control results.

1. In-Time Placebo Test

An in-time placebo test, which examines whether or not synthetic control methods produce treatment effects when a pseudo-treatment is assigned to the affected unit at a time other than the actual event, is the first robustness check we conducted. This test does not appear often

in the literature applying the method because it only serves as a *prima facie* check on whether or not the method works. Technically, it does little to bolster the strength of inferences we might draw with regard to the actual event, especially in contrast to the across-unit placebo tests that we presented in the paper.

In this particular application, we assigned an in-time placebo to occur ten years prior to the actual “United Breaks Guitars” videos. That is, we modeled a series of three false interventions, with the first occurring in July 1998, and the subsequent two false interventions appearing at the same spacing in time as the actual second and third videos. Otherwise, the steps we employed in our analysis remained the same.

<Insert Figures B1 and B2 Here>

As Figures B1 and B2 plot the result of our in-time placebo. Figure B1 plots the actual performance of United between mid-1994 and the end of 2000, as well as the performance of the synthetic United during this time period. Figure B2 plots the difference between these two. Together, these plots reveal what we expected: our false interventions had no effect on United’s passenger volume. The one significant (positive) difference between the actual and synthetic United that occurs in the post-intervention period in Figure B2 appears to be completely unrelated to the timing of our intervention (and could even reflect an underlying data issue not caught by the Department of Transportation).

2. *Leave-One-Out Tests*

The second robustness check we employed was a series of leave-one-out tests. The goal of these tests was to ensure that the results of our analysis were not overly sensitive to presence of a single control firm. To conduct, these tests, we repeated the same steps we used to conduct to our main analysis three times and each time we rotated out one of the control airlines with a positive weight in our baseline analysis (i.e., the analysis presented in the body of the paper).

<Insert Figures B3 and B4 Here>

Figure B3 and B4 present the results of these tests, and as with Figures B1 and B2, they respectively present the result of United and its synthetics and the differences between these. Just as with our in-time placebo checks, our results hold up when we rotate out our control units one-by-one. The overlap between the lines representing our main analysis and these alternative estimations that leave out control units one-by-one is very strong visually, indicating that our results are particularly insensitive to the other airlines that are in the pool from which we construct our baseline synthetic United.

3. *Out of Sample Tests*

Our final robustness check is an out-of-sample test. The goal of out-of-sample tests is to use periods other than the immediate run-up to the event to assess whether or not we can treat the weights on the control firms that we use to construct our synthetic as if they were approximately constant (or at least relatively stable) across time. This robustness check was particularly challenging to apply in the airline industry given events such as mergers and acquisitions, numerous bankruptcies, and the September 11, 2001 terrorist attacks—the first two of which affected different firms at different points in time. To implement this test, we gathered out of sample data from 50 periods from the late 1990s and early 2000s; we ended our sampling of periods in August 2001 to avoid issues related to September 11, as the terrorist attacks, which involved two United aircraft, would make data from that period unrepresentative.¹⁸

<Insert Figures B5 and B6 Here>

Figures B5 and B6 parallel Figures B1/B3 and B2/B4 in terms of information conveyed. B5 in particular reveals that the performance of the synthetic control method is not as strong when we used data from the out of sample period to construct our synthetic. However, the two

¹⁸ In theory, the attacks could provide an interesting additional event to study using synthetic control.

panels still demonstrate that the “United Breaks Guitars” events had a negative effect on United’s passenger volume. That is, although Figure B5 reveals that the synthetic and actual United do not fit particularly well in the pre-treatment period (again, because the construction of synthetic United is based on data drawn from 7 years prior), the gap or difference between the two grows larger in the wake of the “United Breaks Guitars” events.

Figure B6 illustrates this effect more clearly and provides us with additional evidence in favor of a causal interpretation of the videos’ effects. In contrast to the prior difference figures (B2 and B4), we adjusted the scale on B6 such that the midpoint on the vertical axis is the mean of the displayed pre-intervention period rather than 0. This adjustment allows us to more easily assess the impact of the “United Breaks Guitars” videos, and the range of the values presented on the axis is still consistent with that presented in Figures B2 and B4. The takeaway lesson from this robustness check is that despite the tumult in the airline industry throughout the 1990s and 2000s, it is valid, in this context, to treat the weights we estimated in W as if they were approximately stable over ranges of time that mirror those that we analyze. Substantively, the results of this out-of-sample test bolster our claim that the “United Breaks Guitars” videos had an adverse consumer market effect on United, in the form of decreased passenger volume.

Given the results of these robustness checks, as well as the across-unit placebo tests in the paper, we are confident in the negative performance effect on United Airlines that we have uncovered for the first social media-based boycott of “United Breaks Guitars.”

FIGURE B1- In-Time Placebo Test (Primary) – False Video Releases start in July 1998

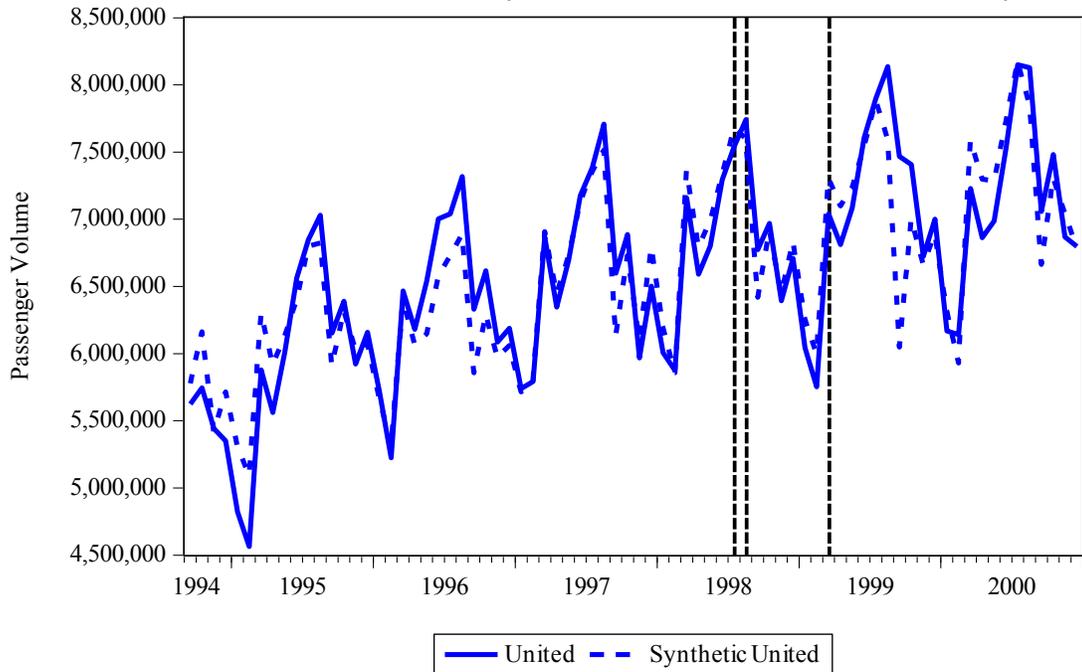


FIGURE B2- In-Time Placebo Test (Difference) – False Video Releases start in July 1998

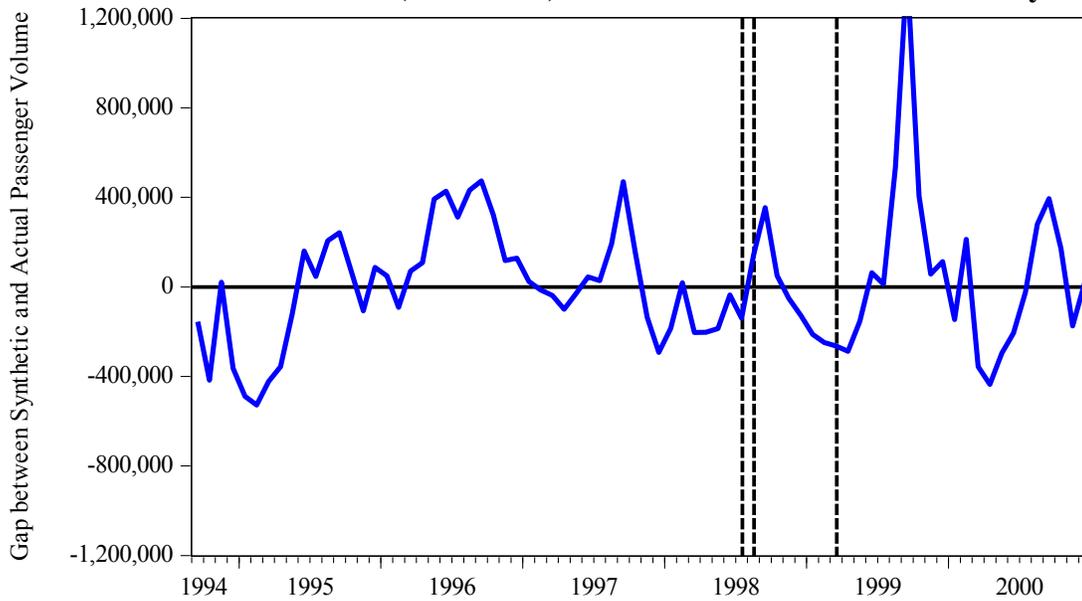


FIGURE B3- Leave-one-Out Test (Primary)

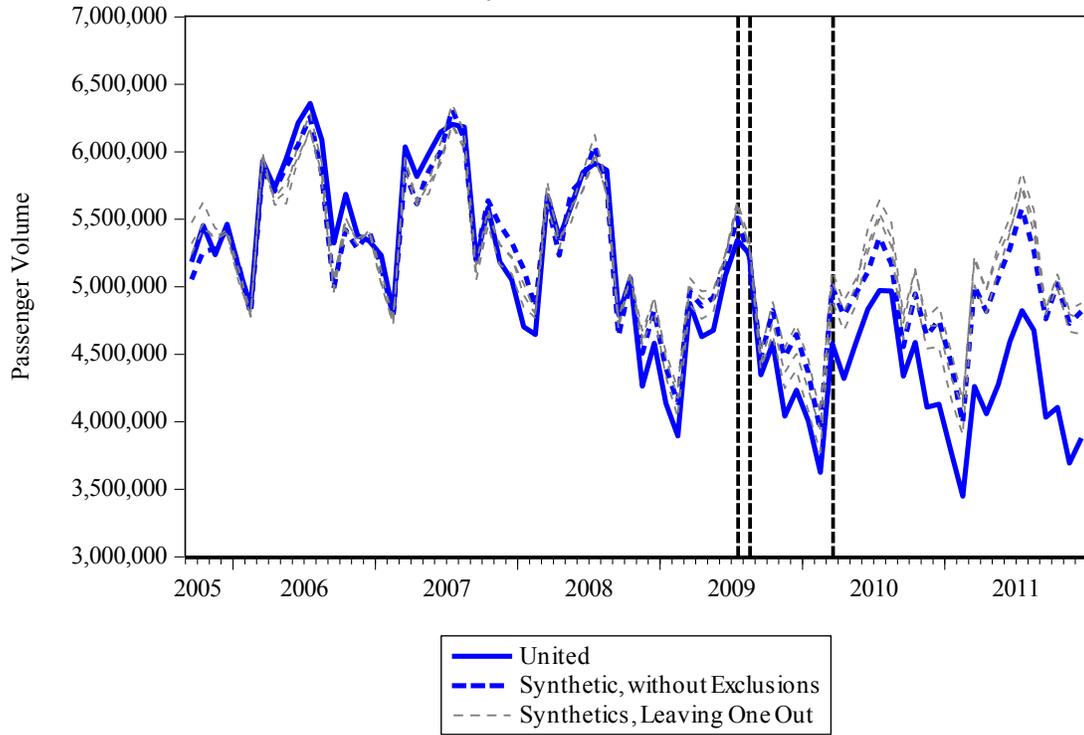


FIGURE B4- Leave-one-Out Test (Difference)

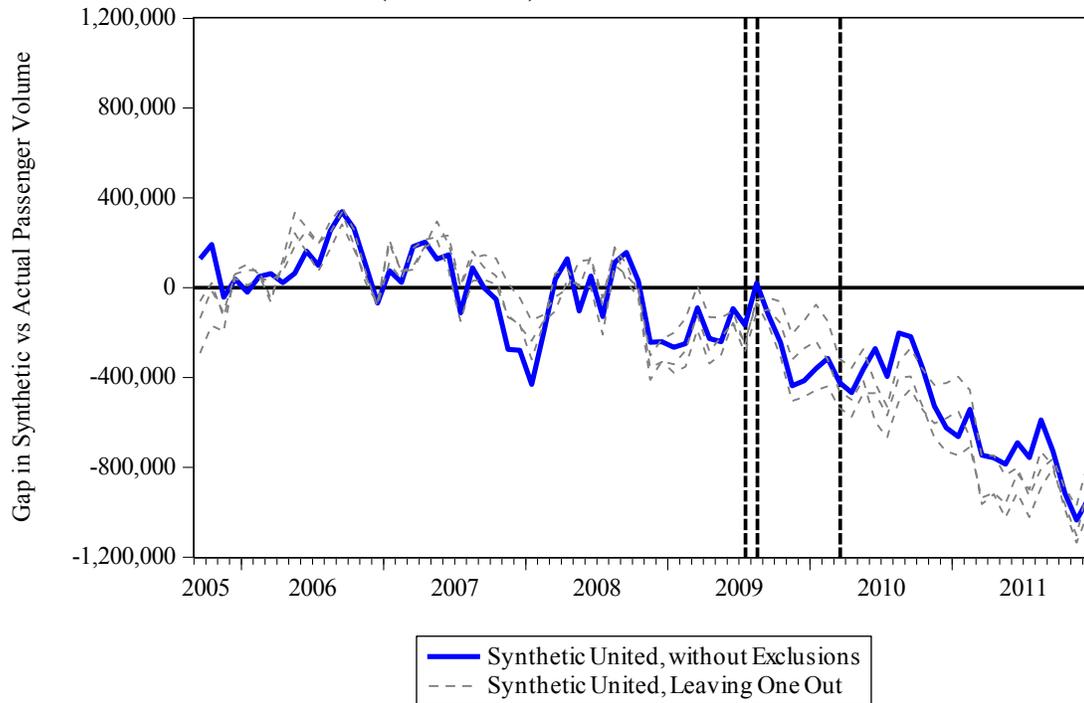


FIGURE B5- Out-of-Sample Test (Primary) (using 50 periods, pre-September 2001)

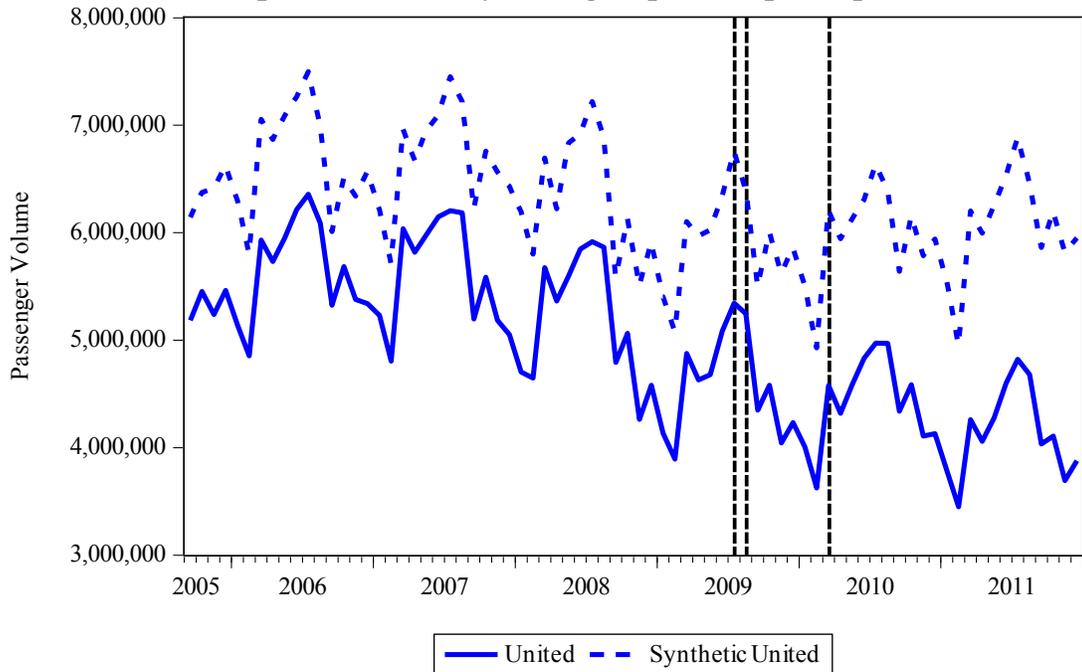


FIGURE B6- Out-of-Sample Test (Difference) (using 50 periods, pre-September 2001)

