Generalizing Event Studies Using Synthetic Controls:

An Application to the Dollar Tree–Family Dollar Acquisition

Amirhossein Zohrehvand*
Faculty of Science, Science Based Business
Leiden University
Snellius Building, Niels Bohrweg 1, 2333 CA Leiden, Netherlands
a.h.zohrehvand@sbb.leidenuniv.nl

Anil R. Doshi
UCL School of Management
Level 38, One Canada Square
London E14 5AA, United Kingdom

Bart S. Vanneste
UCL School of Management
Level 38, One Canada Square
London E14 5AA, United Kingdom

September 12, 2023

Forthcoming in Longe Range Planning

* Corresponding Author

Acknowledgements

We are grateful for comments offered by Bert De Reyck, Isabel Fernandez-Mateo, Martin Kilduff, Phanish Puranam, Davide Ravasi, Joost Rietveld, and Sruthi Thatchenkery. We also thank seminar attendees at ESMT Berlin and IÉSEG School of Management as well as participants at the Academy of Management Annual Meeting, the Strategic Management Society Conference, the Imperial College Business School PhD Research Day Seminar, and the UCL School of Management Reading Group for their feedback on an earlier version of this paper.

Generalizing Event Studies Using Synthetic Controls:

An Application to the Dollar Tree–Family Dollar Acquisition

ABSTRACT

Event studies, which have significantly advanced mergers and acquisitions (M&A) research, obtain excess returns based on a *theory* linking a firm's shareholder returns to those of the market. For outcomes lacking such a theory, we propose an *empirical* approach using a synthetic control method with machine learning to link outcomes for the acquirer or target to those for a group of comparison firms. We discuss the method's assumptions, its close parallel to event studies, and its difference in weighting comparison firms (based on data versus derived from theory). We provide an illustration of Dollar Tree's acquisition of Family Dollar, by analyzing shareholder returns (to demonstrate consistent results with an event study), realized cost and sales synergies, and customer sentiment (derived from more than 52 million Twitter messages). We highlight this method's potential—for M&A and other areas of strategy research—to open up new lines of inquiry.

Keywords: event studies analysis, mergers and acquisitions, synthetic control method, longitudinal design, regularized regression, elastic net, tweets.

1. Introduction

The event study method has a long history in merger and acquisition (M&A) research (Aalbers et al., 2021; Cording et al., 2010; Datta et al., 1992; King et al., 2004; Mandelker, 1974; for a review, see Harrison and Schijven, 2015). This popular method has spawned a broad literature and enabled a deep understanding of M&As (see Devers et al., 2020; Haleblian et al., 2009). The strength of an event study is that it obtains a credible abnormal outcome by comparing the shareholder returns of the acquirer or target with those of the market. The abnormal outcome relies on a *theory* (i.e., a set of internally consistent statements about the connection between variables (Stinchcombe, 1987))—such as the capital asset pricing model (CAPM) and its associated assumptions about utility functions (Sharpe, 1964)—that links a company's returns to those of the broader market.

Yet for other M&A outcomes, no theory exists to link the outcomes for the acquirer or target to those for a group of comparison firms. Prior literature has identified at least four situations in which studying other outcomes is of interest (Bauer and Matzler, 2014; Zollo and Meier, 2008). First, other outcomes can inform drivers of enhanced shareholder returns (Bauer and Matzler, 2014; Zollo and Meier, 2008). Second, non-equity stakeholders—such as bondholders, employees, or customers (Haleblian et al., 2009)—may fare differently in an M&A than do shareholders (Barney, 2020; Shleifer and Summers, 1988). Third, tradeoffs exist between different outcomes even for a given stakeholder. For example, when it is difficult for public markets to value unique combinations of resources (Benner and Zenger, 2016), an exclusive focus on shareholder returns paradoxically may fail to maximize shareholder value. Fourth, other outcomes can be of interest in their own right; for example, patents (Puranam and Srikanth, 2007). In the absence of a theory, how can we obtain an abnormal outcome?

In order to obtain an abnormal outcome in these situations, we propose an *empirical* approach using Doudchenko and Imbens's (2017) synthetic control (DISC) method. The DISC approach builds on prior synthetic control methods (Abadie and Gardeazabal, 2003; Abadie et al., 2010; for applications, see Conti and Valentini, 2018; Fremeth et al., 2016). It involves comparing the actual outcome of a focal firm (i.e., the firm with the event) with a predicted outcome derived from those of comparison firms (i.e., firms unaffected by the event). In an event study, the predicted outcome is based on a theory of market behavior and the outcome of interest is shareholder returns. In DISC, the predicted outcome for the focal firm is empirically derived using outcomes from a set of comparison firms to create an outcome for what is referred to as a synthetic firm. When obtaining the predicted outcome is viewed as an empirical rather than a theoretical exercise, techniques from the field of machine learning can be applied. DISC uses a machine learning technique called elastic net (Zou and Hastie, 2005), which is a form of regularized regression. As in ordinary least squares (OLS), the model in regularized regression is linear. The coefficients, however, are restrained such that they are shrunk towards zero compared to those of OLS. This adjustment to the model's coefficients results in better predictions. Differences in the outcome between the focal firm and the synthetic firm are analogous to the difference in the focal firm's actual and predicted returns (i.e., abnormal returns) in event studies. Because of the empirical procedure, DISC can be applied to outcomes, other than stock returns.

Whereas the method is useful for non-M&A events, we provide an M&A application to highlight its similarities to an event study. We analyze the 2015 acquisition of Family Dollar by Dollar Tree, two large discount retailers in the general merchandise stores industry. The comparison firms are retailers that are not active in that industry and were not involved in an acquisition. We show how the method is analogous to event study when specifically looking at shareholder returns. Our aim is not to replace event study with DISC for shareholder returns. Rather, similar results would lend credence to DISC as a method that can be used for other events and outcomes.

To illustrate DISC's potential for shedding light on important theoretical and empirical questions, we then apply it to measures of realized synergies—namely, operating costs and sales. There is substantial evidence from event studies pertaining to anticipated synergies from shareholder returns to an M&A announcement, yet much less is known about their realization (Garzella and Fiorentino, 2014). Most stakeholders, including long-term shareholders, ultimately care about the realization of synergies, which is why many M&As are conducted in the first place (Puranam and Vanneste, 2016; Rabier, 2017). To illustrate DISC's versatility, we also analyze a measure of customer sentiment based on data from Twitter. Haleblian et al. (2009) and Bettinazzi and Zollo (2017) note that disruption to customers, who are essential for realizing growth, has been largely ignored. Customers increasingly voice their level of satisfaction online (Gans et al., 2021; Nguyen et al., 2020), which offers an opportunity to study the consequences of an M&A for customers.

This study's key contribution is to generalize event study by using DISC to generate an empirically derived abnormal outcome, which complements situations for which a theoretically derived abnormal outcome is available (as in event study). Thus, we introduce and demonstrate a method, which is novel to M&A research. This method makes it possible to study M&A outcomes besides abnormal returns, thus allowing for the idea of an "event study" to be used more broadly. For instance, we highlight the opportunity to study the realization of cost and sales synergies as distinct from their expectations, and customer sentiment (e.g., as derived from Twitter). One advantage of DISC is that the only data required is the outcome variable for the firms in the sample; no independent variables are needed. Another advantage is that it can be used for analyzing quantitatively the consequences of a discrete, or unique, event. Thus, this study provides a template for applying DISC to the analysis of M&As, and is also applicable to other managerially relevant events.

2. Using Event Study to Understand DISC

Notwithstanding the importance of other outcomes (Bauer and Matzler, 2014; Zollo and Meier, 2008), most M&A studies focus on shareholder returns obtained from event studies (King et al., 2004).[1] In a recent meta-analysis on learning from experience, integration strategy, and M&A performance (Schweizer et al., 2022), 62% of the samples (out of a total of 86) had shareholder returns as outcome. The other samples were about equally split between accounting measures (e.g., ROA) and survey-based measures (e.g., managerial perception of the M&A outcome). We expand on the other outcomes in the discussion section.

A key finding from event studies is that, on average, acquirers do not benefit from M&A in terms of shareholder returns (King et al., 2004; Moeller et al., 2004; Moeller et al., 2005). The target's shareholders, however, gain substantially (Jensen and Ruback, 1983). From a deal perspective, the combined returns are slightly positive. Thus, M&As create value, but often acquirers pay too high a price for the target and are able to capture relatively little value. The

[1] Interpretation of the term "event study" in strategy differs from that in economics (see Freyaldenhoven et al., 2019 for a discussion of the event study in economics). In this study, we illustrate the close analogy between the method that we propose and the strategy interpretation of an event study. As we shall explain, a key characteristic of the event study in strategy is that comparison firms are weighted based on the dependent variable (i.e., shareholder returns). In contrast, event studies in economics typically use regression methods such as difference-in-differences and do not explicitly weigh comparison units (i.e., comparison units are evenly weighted). For example, Card et al. (2013) compare wages of job switchers before and after German reunification. In the Discussion section, we provide a comparison of our proposed method to other empirical methods, including difference-in-differences.

event study is used to quantify the gains or losses in terms of shareholder returns relative to those of the market (Harrison and Schijven, 2015; McWilliams and Siegel, 1997).

An extensive literature has investigated when those shareholder returns are particularly high or low (Haleblian et al., 2009; Schweizer et al., 2022). The typical approach is as follows. First, an event study is used to calculate the returns for each acquisition in the sample (for the shareholders of the acquirer, target, or both). Second, those shareholder returns are regressed on M&A factors to help understand the conditions under which returns are high or low. Popular M&A factors include characteristics of the acquirer, of the target, or of the deal. For example, acquirer characteristics include gaining reputation from praise in outlets such as *Fortune* or *Wall Street Journal* (Haleblian et al., 2017) or being growth oriented (Blagoeva et al., 2020). Target characteristics entail engagement with corporate social responsibility (Tong et al., 2020) or divestiture history (Laamanen et al., 2014). Deal characteristics cover perceptions of the similarity of the acquirer and target (de Groote et al., 2021), the transparency of deal details (Yakis-Douglas, 2017), or the method of payment (Fieberg et al., 2021). Thus, event studies have provided a rich understanding of M&As from the perspective of shareholder returns.

2.1. Event Study

Table 1 summarizes the comparison of the event study and DISC methodologies.

—--

Insert Table 1 about here

—--

An event study is a method for quantifying consequences of an M&A for shareholder returns (MacKinlay, 1997; McWilliams and Siegel, 1997). The goal is to estimate, for an acquirer or target, the abnormal return: the difference between the actual return (with M&A) and a predicted return based on a market index (as an indication of the return absent an M&A). The prediction is made with a theory of market behavior, such as CAPM, that links the focal firm's returns to those of the market. The abnormal return is the part of the focal firm's actual return that is not explained by market movements.

Hence, we can write the abnormal return, *AY,* as the difference between the actual and predicted returns, y and $\hat{y}$, respectively:

$$AY_{e,0t} = y_{e,0t} - \hat{y}_{e,0t}, \tag{1}$$

where the subscript e denotes the event study variables and 0 and t indicate the focal firm and the time episode, respectively. Predicted return in CAPM is:

$$\hat{y}_{e,0t} = a_e + b_e y_{mt}, \tag{2a}$$

where a_e (CAPM's "alpha") is a constant and b_e ("beta") indicates the sensitivity of the focal firm's returns to market returns, y_m. Event studies distinguish between two periods: an estimation window prior to the M&A and an event window that coincides with the M&A. In the estimation window, the market theory from Equation (2a) is estimated with OLS using the returns observed for the focal firm and the market. Then, in the event window, the same market theory is used to calculate abnormal returns.

We set up our analogy between DISC and an event study by showing that the market return reflects a broad market index that consists of a weighted average of several individual firm returns, such as the S&P 500:

$$y_{mt} = \sum_{i=1}^{N_e} w_{e,i} y_{e,it}, \tag{2b}$$

where $w_{e,i}$ is the weight of firm *i*, $y_{e,it}$ is the market return to firm *i* in period *t*, and N_e is the number of firms in the market index. The weight, $w_{e,i}$, of each firm in the index is typically determined by some market factor, such as market capitalization (as is the case with the S&P 500) or share price.

We can substitute for y_{mt} in Equation (2a) with Equation (2b) so as to show the individual firms that constitute the market return:

$$\hat{y}_{e,0t} = a_e + b_e \sum_{i=1}^{N_e} w_{e,i} y_{e,it}. \tag{2c}$$

Equation (2c) shows how the focal firm's predicted return is based on a weighted average of other firms' actual returns.

2.2. DISC

The logic of DISC is similar to that of an event study; the goal in both is to estimate the "abnormal" outcome for the acquirer or target. For example, abnormal sales is the difference between actual sales (with M&A) and predicted sales based on the sales of a few comparison firms not involved in an M&A (as an indication of the sales without M&A). With DISC, this prediction is derived empirically using a machine learning technique called elastic net. So, analogously to abnormal returns, abnormal sales is that part of the focal firm's actual sales that is not explained by changes in the sales of other firms that have not engaged in an M&A.

Similar to in Equation (1) for the event study, an abnormal outcome, *AY,* in DISC is the difference between the actual and predicted outcomes, y and $\hat{y}$, respectively:

$$AY_{d,0t} = y_{d,0t} - \hat{y}_{d,0t}, \tag{3}$$

where the subscript *d* denotes the DISC variables.

And much as in the event study's Equation (2c), the predicted outcome for the focal firm is based on a linear combination of the outcomes for comparison firms:

$$\hat{y}_{d,0t} = a_d + \sum_{i=1}^{N_d} w_{d,i} y_{d,it}, \quad (4)$$

where N_d is the number of comparison firms, a_d is a constant and, $w_{d,i}$ is the weight on the outcome of comparison firm *i*. The event study selects a beta, b_e in Equation (2c), that is applied to all comparison firms, whereas DISC selects a weight, $w_{d,i}$ in Equation (4), that is specific to each comparison firm. Thus, DISC can be understood as generalizing a CAPM event study.

Comparison firms are selected in two steps. First, the researcher assembles a pool of potential comparison firms. Since the goal is to use the comparison firms to predict the outcome absent an M&A, candidate comparison firms must have neither undergone an M&A themselves nor been affected by the focal firm's M&A. Comparison firms can be private or public. Furthermore, a few comparison firms are often sufficient for estimation. Second, the elastic net algorithm selects the comparison firms from this pool as part of the estimation procedure.

2.2.1. Estimation

The DISC approach, like its event study counterpart, distinguishes between two periods: a pre-period before the M&A and a post-period after the M&A is announced or closed. The prediction model is estimated during the pre-period; then, in the post-period, it is used to calculate the abnormal outcome.

A prediction model is susceptible to overfitting if it is too flexible and therefore "overadjusts" to the pre-period's idiosyncrasies. Since idiosyncrasies differ in the post-period, this tendency means that the model would perform poorly when calculating the abnormal outcome. In event studies, the prediction model is relatively inflexible because the same beta coefficient is used for all comparison firms (see Equation (2a)). In DISC, the prediction model is relatively flexible because it allows each comparison firm to have a different weight (see Equation (4)). To avoid overfitting, DISC uses regularization, which is a standard approach against overfitting.

In particular, DISC uses a form of regularized regression known as elastic net (Zou and Hastie, 2005). This method finds the parameter a (a constant) and the w_i's (the weights) that minimize:

$$\sum_{t=1}^{T} \left(y_{0t} - a - \sum_{i=1}^{N} w_i y_{it} \right)^2 + d \sum_{i=1}^{N} \left(c w_i^2 + (1 - c) \left| w_i \right| \right). \tag{5}$$

The first summation is the sum of squared errors (as in OLS). The second is the regularization, which imposes a penalty on the weights. The parameter d (≥ 0) sets the extent of regularization. If $d = 0$, then we obtain OLS. If $d > 0$, then the weights shrink towards zero relative to those obtained using OLS. A greater value for this parameter leads to a higher penalty, more shrinkage of the weights, and a model that is less likely to overfit. Too high a value, however, results in a model that may underfit (i.e., it would pick up neither noise nor signal).

The parameter c ($0 \leq c \leq 1$) sets the type of regularization. If $c = 0$, then the penalty consists of the absolute values of the weights. As a consequence, weights are not merely shrunk towards zero but towards exactly zero. We obtain a sparse model with some zero weights and some non-zero weights. This type of regularization is called least absolute shrinkage and selection operator, or LASSO (Tibshirani, 1996). If $c = 1$, then the penalty consists of the squared values of the weights. As a result, weights are shrunk towards zero, but may never become exactly zero. This type of regularization is called ridge regression (Hoerl and Kennard, 1970). With values $0 < c < 1$, elastic net seeks to combine the benefits of LASSO (excluding irrelevant comparison firms whose performance is not predictive of the focal firm's performance) and ridge regression (retaining relevant comparison firms whose performances are correlated) (Zou and Hastie, 2005). Given its penalty on the absolute value of weights, LASSO tends to exclude most relevant comparison firms whose performance is correlated with one another (because the retained comparison firms provide similar information). In contrast, ridge regression would keep most of these correlated comparison firms, but also some comparison firms whose performance is not strongly predictive. Elastic net seeks a compromise between these two approaches, and it includes both LASSO and ridge regression as special cases.

The hyperparameters (i.e., parameters determined outside the model) c and d are, in practice, set through cross-validation using only the comparison firms (Doudchenko and Imbens, 2017). For given values of c and d, each comparison firm is selected in turn to act as a "focal" firm.[2] Weights are then calculated using only the pre-period outcomes. The "focal" firm's post-period outcomes are predicted and the mean squared error (MSE) is computed.[3] After cycling through all "focal" firms for a given c and d, the average MSE is recorded as the prediction error for that iteration. This procedure is repeated for different values of c and d. The iteration with the lowest prediction error yields the chosen values of c and d.

As in event study, the core assumption in DISC is that the relationship (as approximated by the linear combination of weights) between the pre-period outcomes of the comparison firms

[2] Because weights are firm-specific, we cannot simultaneously assign multiple firms as being focal. This fact precludes alternative cross-validation techniques, such as K-fold cross-validation.

[3] Doudchenko and Imbens (2017) calculate the prediction error for the post-period's final episode only. To reduce sensitivity to the study's end date, we calculate the prediction error not just for the final episode but over all episodes after the M&A's closing.

and those of the focal firm is the same as that between the post-period outcomes of the comparison firms and those of the focal firm without the M&A (Doudchenko and Imbens, 2017). For example, the relationship between sales of the comparison firms and that of the focal firm in the pre-period would be the same in the post-period if the M&A had not occurred. This assumption is analogous to that in event study, where the sensitivity of the focal firm's returns to those of the market is assumed to be the same in the pre-period and post-period in the absence of the M&A. Unfortunately, we cannot ascertain this assumption's validity because its statement involves an unrealized situation. One way the assumption may be violated is if the focal firm's M&A affects any of the comparison firms. In that event, the comparison firm's outcomes would no longer represent outcomes without the M&A.

Note that DISC (similar to event study) imposes no constraints on the pattern of outcomes for the comparison firms during the pre- and post-period: outcomes may go up or down or both. The requirement is that the relationship between outcomes for the comparison firms and the focal firm (without M&A) remains constant.

2.2.2. Interpretation

The synthetic control literature interprets the abnormal outcome (i.e., the difference between actual and predicted outcome) as an estimation of the effect size of a treatment (Abadie et al., 2010; Doudchenko and Imbens, 2017). This interpretation is similar to that in event studies, whereby an abnormal return (i.e., the difference between an actual and predicted return) provides an estimate of the effect size of an M&A. As applied to synergies, we can interpret the abnormal outcome as an estimate of the realization (or lack thereof) in costs or sales synergies. Just as event study provides an estimate of abnormal returns for a given deal, DISC yields an estimate of abnormal costs or sales specific to a given deal.[4]

A key difference is that event study reflects anticipated synergies and DISC reflects realized synergies. The tradeoff is as follows. An event study allows for a short event window of a few days around the announcement (though many event studies use longer event windows of many months or years [King et al., 2004]). A short event window helps in avoiding confounding events but means we can only anticipate the extent of synergies. In efficient markets, this guess is good in the sense that it is hard to systematically make better guesses but not in the sense that the guess necessarily comes true. DISC estimates realized synergies, which differ from anticipated synergies (Oler et al., 2008). For example, one way to assess whether anticipated costs and sales synergies have been realized is to consider the financial statements (Koller et al., 2020). A longer event window is needed because synergies are not realized instantaneously and thus do not immediately show in the financial statements. However, a longer event window creates the possibility of confounding events. One check for confounding events is to inspect when synergies are realized; for example, after the closing date rather than before.

3. Methods

3.1. Sample

We illustrate DISC with Dollar Tree's acquisition of Family Dollar. Both of these firms are large discount retailers selling a wide range of items that include kitchen supplies, food, beauty products, office materials, and cleaning products. The acquisition was announced on July 28,

[4] If the interest is in understanding differences between industries, then one can (1) obtain specific effects from multiple deals and (2) run a regression with dependent variable abnormal sales (or costs) and industry indicators as independent variables.

2014, and closed on July 6, 2015. Dollar Tree offered a 31% premium over Family Dollar's average share price during the four weeks prior to announcement. At the time of the acquisition, Dollar Tree was generating $8.0 billion in revenue and $1.1 billion in EBITDA from 5,080 stores; Family Dollar was generating $10.4 billion in revenue and $815 million in EBITDA from 8,246 stores.

When selecting this acquisition, we applied the following criteria to ensure that outcome data were available. First, for the shareholder returns, operating costs, and sales measures, we sought an acquisition in which both the acquirer and target were publicly listed. Second, for the customer sentiment measure, we looked for a post-2012 acquisition between US-based firms in retail industries (SIC codes 52 to 59) operating chiefly in a single 2-digit SIC code so that tweets about a company would relate to a single industry. The Dollar Tree–Family Dollar acquisition met these criteria.

We used the same criteria to select comparison firms but with a few modifications to accommodate key DISC assumptions. First, we sought public firms in a primary industry other than that of Dollar Tree and Family Dollar (SIC 53: General Merchandise Stores) to preclude the acquisition from affecting any comparison firms (353 firms). Second, we ensured that comparison firms had more than one million dollars in sales (152 firms) to ensure comparability of accounting measures with those of Dollar Tree and Family Dollar. Third, comparison firms must not have undergone an M&A during the sample period (69 firms). Fourth, to ensure that the sentiment measure for comparison firms is comparable to that of focal firms, we dropped 56 companies: non-American firms, firms with only online operations, and firms with multiple brands or active in multiple industry segments. The search yielded 13 comparison firms: Barnes & Noble, Bed Bath & Beyond, Chipotle Mexican Grill, The Home Depot, Jamba Juice, Lowe's, Nordstrom, Office Depot, Panera Bread, Ross Stores, Ulta Beauty, Whole Foods Market, and Williams-Sonoma.

3.2. Measures

We use four outcome measures. The first is daily shareholder returns (from Compustat) to assess whether DISC yields results similar to those of an event study.

The second is operating costs. A distinction is made between cost and sales synergies (Rabier, 2017; for two other types of synergy, see Bennett and Feldman, 2017, on ongoing synergy and Hernandez and Shaver, 2019, on network synergy). These are improvements in the joint costs and sales that the target and acquirer would not have achieved independently. We use data originating from SEC filings. To analyze operational synergies, we focus on operating costs (Compustat: XOPR), which includes cost of goods sold as well as selling, general, and administrative expenses. The measure is calculated as the joint quarterly operating costs divided by joint sales. We normalize by sales because sales increases may lead to increases in operating costs and thereby mask cost synergies.[5] The third measure is the joint quarterly sales (Compustat: SALEQ) of Dollar Tree and Family Dollar.

The fourth measure is customer sentiment. M&A disruptions are well documented for internal stakeholders (Graebner et al., 2017) but much less so for customers. Sentiment is "overall opinion towards the subject matter" (Pang et al., 2002: 79). We follow previous studies in using Twitter data to measure customer sentiment (Gans et al., 2021; Ma et al., 2015). Using 52 million tweets that mentioned the focal or comparison firms, the measure is the probability that, in a given month, a tweet about a company has positive sentiment. More information on the

[5] An analysis using a measure of operating costs without sales normalization yields conclusions similar to those reported here.

data used for customer sentiment—and on our construction of that measure—is presented in the Online Appendix.

3.3. Pre-period and Post-period

For shareholder returns, the estimation window (or pre-period) is 250 trading days and ends 60 trading days prior to announcement: [−310, −61]. The event window (or post-period) begins 10 trading days prior to announcement and ends 10 trading days thereafter: [−10, 10] (Cuypers et al., 2017). Data for sales and operating costs are available quarterly, so the pre-period runs from 2010Q1 to 2014Q2 (i.e., the quarter preceding announcement) and the post-period runs from 2014Q3 (the announcement quarter) to 2017Q1. We exclude the closing quarter (2015Q3) because of a shift in Family Dollar's reporting cycle. Finally, for customer sentiment—where data are aggregated by month—the pre-period is from January 2010 to June 2014 (the month prior to announcement) and the post-period is from July 2014 (the announcement month) to March 2017 (to ensure that data are available for all measures and to avoid any acquisitions by comparison firms). The R code for the analysis of operating costs is given in the Online Appendix. The same code is used when analyzing the other measures.[6]

3.4. Placebo Test

Given the small sample size, traditional hypothesis tests of statistical significance are not feasible in DISC. Instead, DISC, like prior synthetic control methods (Abadie et al., 2010), uses a placebo test that yields a p-value analogue. The idea is to compare the effect size obtained for the focal firm to the placebo effect size that arises if, instead, a comparison firm is viewed as the focal firm. The placebo test calculates the placebo effect size for each comparison firm in turn,

[6] For researchers wanting to apply DISC but who do not use R, we also provide Stata code in the Online Appendix. For generality, the Stata code considers only pre- and post-periods and not intermediate periods (e.g., between announcement and completion). Furthermore, note that R and Stata use different optimization methods, so that results of these applications may differ slightly.

while using only the data for the comparison firms. The statistic used is the ratio of the root mean squared predicted error (RMSPE) in the post-period to that in the pre-period (Abadie et al., 2015). If the magnitude of the RMSPE ratio is higher for the focal firm than for the comparison firms, then the result is less likely due to chance. A p-value indicates how unusual the data are without an effect. Likewise, the proportion of placebo effects whose magnitudes exceed that of the focal firm (i.e., the p-value analogue) captures the rarity of the focal firm's observed effect. If this proportion is high, then the observed effect for the focal firm (with an M&A) is typically equaled or exceeded for firms without an M&A, which suggests that magnitude of the outcome for the focal firm is not unusual (which may call into question the presence of the M&A effect).

4. Results

Figure 1 compares DISC with an event study using the first outcome measure: shareholder returns. For the event study, we used CAPM (see Equation (2a)).[7] As is common, we summed the returns over time to yield cumulative abnormal returns (CAR). The estimated CAR are typical: close to zero (1%) for the acquirer (Dollar Tree, the figure's left panel) and both large and positive (26%) for the target (Family Dollar, right panel). More importantly, the DISC results are similar (−1% and 22%). The consistency of these results suggests that DISC is a viable method for studying additional outcomes.

—--

Insert Figure 1 about here

—--

[7] Carhart's (1997) four-factor market model yields similar results.

Figure 2 presents DISC results and additional analyses for the second outcome measure: the combined (normalized) operating costs of Dollar Tree and Family Dollar for the period 2010–2017. The upper left panel plots actual and predicted operating costs. In the pre-period (before announcement), actual operating costs are cyclical (as allowed by DISC) and closely matched by predicted operating costs; this indicates that the DISC prediction model works well for the pre-period. In the post-period (after announcement), actual operating costs are lower than predicted costs, consistent with the realization of cost synergies. The divergence did not begin until after the acquisition had closed and it increased over time to more than two percentage points in 2017. The figure's upper right panel shows that 9 of the 13 comparison firms had nonzero weights and that the weights for The Home Depot and Office Depot are negative.

The figure's middle row presents the placebo test. The left panel shows that the fit of Family Dollar and Dollar Tree is better than the fit of the placebos in the pre-period and is modest in the post-period. Together, these two facts yield a ratio of post-RMSPE to pre-RMSPE that is the third-most negative—and fifth-most extreme overall—of the placebo comparison results, as shown in the middle right panel. That panel suggests a value for the observed result of $p = 0.23$ (one-sided; calculated as 3/13) or $p = 0.38$ (two-sided; calculated as 5/13). Thus, changes in operating costs are substantial but not unusual.

—---

Insert Figure 2 about here

—---

Finally, the figure's lower left panel plots results from the leave-one-out procedure, which investigates the sensitivity of results to the selected comparison firms. The leave-one-out procedure assesses the extent to which the results change when using a smaller set of comparison firms. The procedure iteratively drops one comparison firm from the pool at a time, then re-runs the estimation. The dropped comparison firm is the one that received the highest weight in the previous iteration.

Results after dropping the first comparison firm give consistent, albeit more pronounced, improvements in operating costs. Dropping additional comparison firms yields no estimation (i.e., all remaining comparison firms get zero weight), as shown by the horizontal dashed line. This suggests that, in this case, after the second comparison firm is dropped there are too few remaining to collectively approximate operating costs in the pre-period. Thus, the results are insensitive when dropping one comparison firm and cannot be established with fewer comparison firms.

—---

Insert Figure 3 about here

—---

Figure 3 plots the DISC results for the third outcome: combined sales for Dollar Tree and Family Dollar. Actual sales are lower than predicted sales, consistent with sales dis-synergies. The divergence began only after the acquisition had closed and amounted to per-quarter sales that were several hundred million dollars lower, or about 6% on average. Compared with the placebos, the sales decline for Dollar Tree and Family Dollar is the second-most extreme (p = 0.15, or 2/13). In the leave-one-out analysis, dropping one comparison firm yields a consistent decline in sales but dropping more suggests a sales increase. However, the sensitivity analysis for the post-period result is not valid, because the model fits (between actual sales and predicted sales based on subsequently smaller pools of comparison firms) in the pre-period are poor. In each iteration of the leave-one-out analysis, predicted sales deviates from actual sales well before the end of the pre-period. Without a better fit in the pre-period, the post-period has no clear interpretation.

—---

Insert Figure 4 about here

—---

—---

Insert Figure 5 about here

—---

Figures 4 and 5 present the results of the DISC analysis for the final outcome measure: customer sentiment. Because data on Family Dollar's customer sentiment are available throughout the period of study, we can assess this outcome separately for target and acquirer even after the acquisition closes. For both Dollar Tree and Family Dollar, actual customer sentiment after the announcement increases relative to predicted customer sentiment.[8] These changes are more extreme than for any of the placebos for both firms (p = 0.08, or 1/13). The leave-one-out analyses indicate that these observed results are not sensitive to the choice of comparison firms.

5. Discussion

The DISC approach generalizes event studies by allowing for the study of outcomes other than shareholder returns. In this paper, we use DISC to study several such outcomes of the Dollar Tree–Family Dollar acquisition. First, the cumulative abnormal returns are close to zero for Dollar Tree and more than 20% for Family Dollar. Given that Dollar Tree and Family Dollar had

[8] In the Dollar Tree analysis, weights for the comparison firms are small because the intercept approximates reasonably well the flat pre-period trend. We also conducted an analysis with the intercept term excluded, which is equivalent to setting $a = 0$ in Equation (5). The results are similar to those reported here.

roughly similar market values before the announcement, this finding indicates that the market anticipate substantial synergies in this deal. Second, we find some evidence of actual operating cost synergies but no evidence of actual sales synergies (in fact, the evidence suggests sales dissynergies). Hence, additional study of other M&As is warranted because managers frequently offer sales synergies as a rationale for acquisitions (Bauer et al., 2021; Cerrato et al., 2016; Rabier, 2017). Third, post-acquisition customer sentiment of both target and acquirer increased relative to that of the comparison firms. So in this case, we do not find that customers experienced spillovers from the internal disruption due to post-merger integration (Graebner et al., 2017). An opportunity exists for future research to explore the conditions under which such spillovers might occur.

We provide several design considerations when implementing DISC in practice. With respect to comparison firms, three filters will assist with their selection. First, comparison firms ought not have executed an M&A over the study period. Second, they ought not have experienced any spillovers from the focal firm's M&A transaction. Third, to avoid confounding events, comparison firms ought not have experienced any idiosyncratic shocks during the study period. One check on the selection of comparison firms is whether there is a good fit between the focal and synthetic firms in the pre-period. A poor fit suggests that the model is not suitable. Thus, a tradeoff exists for the number of comparison firms. When ensuring that undesirable comparison firms are excluded from the pool, one may be left with too few firms that yield unstable results. A potential instability can be assessed by performing a leave-one-out analysis, in which comparison firms are dropped one by one.

With respect to the number of periods in the study (Abadie, 2021; Abadie et al., 2015), it is preferable to include a greater number of pre-periods. With few pre-periods, a close match between the synthetic and the focal firms might arise due to chance. In that case, any relationship established between the synthetic and focal firm in the pre-period is unlikely to carry over into the post-period. Using many pre-periods, a close match is unlikely to arise due to chance, suggesting that the relationship between the two is likely to carry over into the post-period had the M&A not occurred. Including relatively few post-periods precludes the possibility that other events might confound results. Thus, a tradeoff exists if DISC is used to understand the impact on an M&A outcome over longer periods. With many post-periods, more insight can be gained on the M&A outcome, but with greater risk of confounding events.

However, we note that avoiding confounding factors is not always the goal in event studies. A recent stream of M&A research on impression management (e.g., Busenbark et al., 2017; Gamache et al., 2019; Graffin et al., 2016), focuses on acquisitions with confounding announcements. Firms intentionally attempt to alter market reactions to acquisition announcements by releasing information unrelated to the acquisition. As a consequence, abnormal returns reflect the effects of both the M&A and the confounding event. Impression management can also affect the interpretation of DISC. With impression management, the abnormal outcome for the acquirer or target represents not only the acquisition but also the release of unrelated information. In such instances, DISC can help to isolate the consequences of the acquisition by investigating the effects on different stakeholders (e.g., on customers and their sentiment).

5.1. Applications of DISC and Other Methods

A number of empirical methods are used in strategy research, including the original implementation of synthetic controls, coarsened exact matching, and difference-in-differences. In the Appendix, we provide a comparison of these methods, including assumptions, derivation of weights, data requirements, advantages, and disadvantages. Because each method has different strengths and weaknesses, we discuss here the choice of which one to use based on the research question and context.

DISC is useful for evaluating the firm-specific effect over time of a strategic event or policy, even when the focal firm is unique and when only outcome data are available. The absence of weight constraints offers two benefits. First, a focal firm or its strategy can be unique because dissimilar comparison firms can be inverted or scaled (as described previously) to create a good pre-period match. Second, direct competitors—which are likely most similar to the focal firm—can be excluded to mitigate the spillover risk of the focal firm's strategy affecting the comparison firms. One concern is that outside events may confound the effect if the post-period window is long. So depending on the nature of the strategic event, we may have more confidence in earlier than in later results in the post-period.

Event studies are frequently used to study the effects of M&As. When used to study a single firm, an event study estimates an M&A's firm-specific effect on the firm's cumulative abnormal returns. However, an event study is usually applied on a large sample to assess the average effect of some M&A characteristic (e.g., target and acquirer relatedness) on cumulative abnormal returns. Because returns are forward-looking, the post-period (i.e., event window) can be short, which minimizes the possibility of other events confounding the effect. Since the underlying theoretical model focuses only on shareholder returns, it follows that the firms studied must be publicly traded.

Prior implementations of synthetic controls for policy evaluations (Abadie and Gardeazabal, 2003; Abadie et al., 2010; Abadie et al., 2015) have been used in the management literature (Conti and Valentini, 2018; Fremeth et al., 2016). This version was designed to bridge small-sample qualitative analyses and large-sample quantitative analyses. It yields an intuitive synthetic control for studying firm-specific effects over time when the number of focal and comparison firms is small. The availability of independent variables allows for an alternative pre-period model, in the event that a dependent variable-only model results in a weak fit. However, the constraints on weights render this approach less applicable when the focal firm is unique. The method is also susceptible to spillovers arising from the need to choose similar comparison firms and to confounding if the post-period window is too long. These characteristics suggest that this early implementation of synthetic controls is especially suited for policy evaluation.

Coarsened exact matching (CEM) is useful when model dependence is a concern. Because CEM can be used with many estimation methods, it can be broadly applied across research questions pertaining to the average effect of a strategy on an outcome of interest. Because firms are dropped from the sample in order to achieve a better match among remaining focal and comparison firms, the method is more effective with larger samples. CEM addresses the endogeneity that results when the functional form of observables is misspecified. Yet CEM cannot, by itself, address the endogeneity that arises from unobservables.

Finally, difference-in-differences (DiD) is used in large-sample studies when a policy change affects the focal firms but not the comparison firms (which are otherwise similar to them). In these instances, DiD is used to estimate the policy's average effect on the outcome of interest for focal firms. This method exploits experiment-like conditions and the policy change is typically referred to as a shock. However, identifying a clean shock is a nontrivial task because the shock should neither be anticipated by the focal firms nor affect the comparison firms. DiD relies on the parallel trends assumption but, unlike the other methods, it cannot render focal and comparison firms more similar. So, in practice, DiD is often combined with CEM. Like other methods, DiD also incurs the risk of outside events affecting the focal or comparison firms if the post-period window is too long.

5.2. Implications for Research

We see potential for applications of DISC to research on M&A, strategy, and organizations, which we discuss in turn. A rich literature exists on measuring M&A performance (Bauer and Matzler, 2014; Zollo and Meier, 2008). The dominant approach is to use abnormal shareholder returns derived from an event study (Schweizer et al., 2022). An event study's appeal is that it gives a near immediate measure of M&A performance. This instantaneity is both its strength and its weakness. It allows for isolating the impact of an M&A from other internal or external changes that may occur during the lifetime of the M&A. Yet, it is based on a subjective up-front assessment of a complex organizational process that plays out over a long period (Schijven and Hitt, 2012). It is difficult for managers to predict whether an M&A will be successful, let alone for investors with less information. Indeed, shareholders' long-term returns to M&As deviate from their short-term returns (DeLong and DeYoung, 2007; Oler et al., 2008).

Outcomes in M&A research other than shareholder returns include accounting measures and survey-based measures (Schweizer et al., 2022). An example of the former is return on assets (Castellaneta et al., 2018) and of the latter is managerial perceptions of integration performance (Heimeriks et al., 2012). These outcomes have typically been studied using regressions. For survey-based measures for which repeated data collection over time is rare, DISC is typically not applicable because it requires time series data. In this case, DISC's data requirements are more stringent. For accounting measures, DISC works well, as we have illustrated when analyzing cost and sales synergies. Compared to regressions, one benefit of DISC is that it requires no independent variables. Thus, one may have access to time series outcome data but not to time-series independent variables (or at least not on the same interval). In this case, DISC's data requirements are less stringent.

DISC can help advance our understanding of M&A performance. First, the literature shows that anticipated and realized M&A performance may differ (Garzella and Fiorentino, 2014; Schijven and Hitt, 2012). DISC can help direct attention to realized M&A performance and its drivers. For example, a general notion is that cost synergies are more easily achieved than sales synergies, though the academic evidence is lacking. Second, M&A performance is multi-faceted and the literature has argued for the consideration of non-shareholder measures (Cording et al., 2010; Zollo and Meier, 2008). We showed the versatility of DISC by analyzing customer sentiment from tweets. We see an interesting opportunity to use DISC to study post-merger integration (Schweizer et al., 2022); for example, by considering (the number of) job postings on digital platforms or (the number of) product launches. Third, the literature shows that public acquirers capture more value from private than from public targets (Capron and Shen, 2007). However, M&A outcomes are less well understood when the acquirer is private. Because DISC does not rely on public shareholder returns, it provides an interesting opportunity to study private firms, which outnumber public firms.

We see potential for applications of DISC to broader strategy research. For instance, strategy research suggests that firms adopt unique strategies, yet analyzing them is challenging because of their uniqueness. The literature features three main approaches. First, some studies focus on the non-unique aspects of a strategy, such as generic strategies (Campbell-Hunt, 2000; Porter, 1980; Shinkle et al., 2013). Second, some scholars categorize the firm-specific consequences of a non-unique strategy. For instance, it is possible to classify firms as either diversified or non-diversified and then to estimate a firm-specific outcome of diversification (Mackey et al., 2017) or to classify firms based on their unique corporate social responsibility strategy and study returns to that unique strategy (Nardi et al., 2022). Third, researchers have conducted qualitative analyses of a firm's strategy; one example is Siggelkow's (2002) investigation into Vanguard's strategy evolution. We propose that DISC offers a useful fourth approach—a quantitative approach for studying a strategy's consequences, even if the strategy is adopted by just a single firm and that firm differs from others. Hence, DISC can yield insights into the consequences of unique strategies and events.

Beyond strategy, DISC can be applied in organizational research more broadly. For example, the impact of individual employees' personal events (Chen et al., 2021) on creativity or productivity can be assessed. Also, events that affect team cohesion or other behaviors (Laulié and Morgeson, 2021) could be analyzed. In general, the application of DISC merely requires one or a few units of observation to experience the event, some comparison units that do not experience it, and realizations of the outcome of interest over time. Moreover, the visualizations of the results are easy to interpret and provide an indication of the reliability of the result. DISC expands the principle of an event study and allows organizational researchers to quantify how a variety of events affect individuals, groups, or organizations.

5.3 Implications for Practice

In addition to its usefulness for researchers, we posit that several of DISC's characteristics make it valuable for managers looking to understand the impact of their strategic decisions and policies. First, DISC's light data requirement—requiring only one treated unit and a limited number of comparison units—facilitates apt analysis for managers. Managers are typically not interested in understanding average effects within a large sample. Rather, a manager is often focused on the individual effect on a single focal observation, such as her own firm. Another use case is as a tool for competitive intelligence, whereby specific competitors can be analyzed. For example, in our analysis, we relied strictly on publicly available information. Second, the results can be presented in a graphical format that is easy to interpret. The output from DISC graphically shows the difference between the outcome for the focal unit of observation and the group of comparison firms. Because DISC uses time series, it provides insight into changes over time and can be applied in real time (if data are available). DISC could serve managers looking to engage in a rigorous, data-driven approach to assess their strategies. Third, although we focused on an M&A application, DISC is not limited to such applications. Other business use cases include a new product launch, a change in the compensation scheme, or an organization redesign.

6. Conclusion

This study demonstrates DISC's potential to broaden the scope of event studies in M&A research. DISC effectively generalizes the event study framework and facilitates analysis that extends beyond shareholder returns. The method offers a versatile and empirically derived approach to obtain abnormal outcomes, making it applicable to a diverse set of outcomes, including realized synergies and outcomes for other stakeholders. DISC opens opportunities to study questions pertaining to outlier or unique M&As, events, and strategies more broadly.

References

Aalbers, R.H., McCarthy, K.J., Heimeriks, K.H., 2021. Market reactions to acquisition announcements: the importance of signaling 'why' and 'where.' Long Range Planning, 54(6), 102–105.

Abadie, A., 2021. Using synthetic controls: feasibility, data requirements, and methodological aspects. Journal of Economic Literature, 59(2), 391–425.

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., 2015. Comparative politics and the synthetic control method. American Journal of Political Science, 59(2), 495–510.

Abadie, A., Gardeazabal, J., 2003. The economic costs of conflict: a case study of the Basque country. American Economic Review, 93(1), 113–132.

Barney, J.B., 2020. Measuring firm performance in a way that is consistent with strategic management theory. Academy of Management Discoveries, 6(1), 5–7.

Bauer, F., King, D.R., Schriber, S., Kruckenhauser, C., 2021. Navigating challenging contexts: costs and benefits of codified acquisition experience. Long Range Planning, 54(6), 102088.

Bauer, F., Matzler, K., 2014. Antecedents of M&A success: the role of strategic complementarity, cultural fit, and degree and speed of integration. Strategic Management Journal, 35(2), 269–291.

Benner, M.J., Zenger, T., 2016. The lemons problem in markets for strategy. Strategy Science, 1(2), 71–89.

Bennett, V.M., Feldman, E.R., 2017. Make room! Make room! A note on sequential spinoffs and acquisitions. Strategy Science, 2(2), 100–110.

Bettinazzi, E.L.M., Zollo, M., 2017. Stakeholder orientation and acquisition performance. Strategic Management Journal, 38(12), 2465–2485.

Blagoeva, R.R., Kavusan, K., Jansen, J.J.P., 2020. Who violates expectations when? How firms' growth and dividend reputations affect investors' reactions to acquisitions. Strategic Management Journal, 41(9), 1712–1742.

Busenbark, J.R., Lange, D., Certo, S.T., 2017. Foreshadowing as impression management: illuminating the path for security analysts. Strategic Management Journal, 38(12), 2486–2507.

Campbell‐Hunt, C., 2000. What have we learned about generic competitive strategy? A meta-analysis. Strategic Management Journal, 21(2), 127–154.

Capron, L., Shen, J.C., 2007. Acquisitions of private vs. public firms: private information, target selection, and acquirer returns. Strategic Management Journal, 28(9), 891–911.

Card, D., Heining, J., Kline, P., 2013. Workplace heterogeneity and the rise of West German wage inequality. Quarterly Journal of Economics, 128(3), 967–1015.

Card, D., Krueger, A. B., 1994. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. American Economic Review, 84(4), 772–793.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance, 52(1), 57–82.

Castellaneta, F., Valentini, G., Zollo, M. 2018. Learning or inertia? The impact of experience and knowledge codification on post-acquisition integration. Industrial and Corporate Change, 27(3), 577–593.

Cerrato, D., Alessandri, T., Depperu, D., 2016. Economic crisis, acquisitions and firm performance. Long Range Planning 49(2), 171–185.

Chen, Y., Liu, D., Tang, G., Hogan, T. M., 2021. Workplace events and employee creativity: a multistudy field investigation. Personnel Psychology, 74(2), 211–236.

Connelly, B.L., Li, Q., Shi, W., Lee, K., 2020. CEO dismissal: consequences for the strategic risk taking of competitor CEOs. Strategic Management Journal, 41(11), 2092–2125.

Conti, R., Valentini, G., 2018. Super partes? Assessing the effect of judicial independence on entry. Management Science, 64(8), 3517–3535.

Cording, M., Christmann, P., Weigelt, C., 2010. Measuring theoretically complex constructs: the case of acquisition performance. Strategic Organization, 8(1), 11–41.

Cuypers, I.R.P., Cuypers, Y., Martin, X., 2017. When the target may know better: effects of experience and information asymmetries on value from mergers and acquisitions. Strategic Management Journal, 38(3), 609–625.

Datta, D.K., Pinches, G.E., Narayanan, V.K., 1992. Factors influencing wealth creation from mergers and acquisitions: a meta-analysis. Strategic Management Journal, 13(1), 67–84.

de Groote, J.K., Kleindienst, I., Hoegl, M., Schweizer, D., Laamanen, T., 2021. Similarity perceptions in investor reactions to acquisition announcements. Long Range Planning, 54(1), 101946.

DeLong, G., DeYoung, R., 2007. Learning by observing: information spillovers in the execution and valuation of commercial bank M&As. Journal of Finance, 62(1), 181–216.

Devers, C.E., Wuorinen, S., McNamara, G., Haleblian, J., Gee, I. H., Kim, J., 2020. An integrative review of the emerging behavioral acquisition literature: charting the next decade of research. Academy of Management Annals, 14(2), 869–907.

Doudchenko, N., Imbens, G.W., 2017. Balancing, regression, difference-in-differences and synthetic control methods: a synthesis. NBER Working Paper 22791, National Bureau of Economic Research, Cambridge, MA.

Fieberg, C., Lopatta, K., Tammen, T., Tideman, S.A., 2021. Political affinity and investors' response to the acquisition premium in cross-border M&A transactions—a moderation analysis. Strategic Management Journal, 42(13), 2477–2492.

Fremeth, A.R., Holburn, G.L F., Richter, B. K., 2016. Bridging qualitative and quantitative methods in organizational research: applications of synthetic control methodology in the U.S. automobile industry. Organization Science, 27(2), 462–482.

Freyaldenhoven, S., Hansen, C., Shapiro, J. M., 2019. Pre-event trends in the panel event-study design. American Economic Review, 109(9), 3307–3338.

Gamache, D.L., McNamara, G., Graffin, S.D., Kiley, J., Haleblian, J., Devers, C.E., 2019. Impression offsetting as an early warning signal of low CEO confidence in acquisitions. Academy of Management Journal, 62(5), 1307–1332.

Gans, J.S., Goldfarb, A., Lederman, M., 2021. Exit, tweets, and loyalty. American Economic Journal: Microeconomics, 13(2), 68–112.

Garzella, S., Fiorentino, R., 2014. A synergy measurement model to support the pre-deal decision making in mergers and acquisitions. Management Decision, 52(6), 1194–1216.

Gharehgozli, O., 2017. An estimation of the economic cost of recent sanctions on Iran using the synthetic control method. Economics Letters, 157, 141–144.

Graebner, M.E., Heimeriks, K.H., Huy, Q.N., Vaara, E., 2017. The process of postmerger integration: a review and agenda for future research. Academy of Management Annals, 11(1), 1–32.

Graffin, S.D., Haleblian, J., Kiley, J.T. 2016. Ready, AIM, acquire: impression offsetting and acquisitions. Academy of Management Journal, 59(1), 232–252.

Haleblian, J., Devers, C.E., McNamara, G., Carpenter, M.A., Davison, R.B., 2009. Taking stock of what we know about mergers and acquisitions: a review and research agenda. Journal of Management, 35(3), 465–502.

Haleblian, J.J., Pfarrer, M.D., Kiley, J.T., 2017. High-reputation firms and their differential acquisition behaviors. Strategic Management Journal, 38(11), 2237–2254.

Harrison, J.S., Schijven, M., 2015. Event-study methodologies in the context of M&As: a reorientation., Iin: A. Risberg, D.R. King, O. Meglio (Eds.), The Routledge Companion to Mergers and Acquisitions. Routledge, London, pp. 221-241.

Heimeriks, K.H., Schijven, M., Gates, S., 2012. Manifestations of higher-order routines: the underlying mechanisms of deliberate learning in the context of postacquisition integration. Academy of Management Journal, 55(3), 703–726.

Hernandez, E., Shaver, J.M., 2019. Network synergy. Administrative Science Quarterly, 64(1), 171–202.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: biased estimation for nonorthogonal problems. Technometrics, 12(1), 55–67.

Iacus, S.M., King, G., Porro, G., 2011. Multivariate matching methods that are monotonic imbalance bounding. Journal of the American Statistical Association, 106(493), 345–361.

Jensen, M.C., Ruback, R.S., 1983. The market for corporate control: the scientific evidence. Journal of Financial Economics, 11(1), 5–50.

King, D.R., Dalton, D.R., Daily, C.M., Covin, J.G., 2004. Meta-analyses of post-acquisition performance: indications of unidentified moderators. Strategic Management Journal, 25(2), 187–200.

Koller, T., Goedhart, M., Wessels, D., 2020. Valuation: measuring and managing the value of companies, 7th ed. Wiley, Hoboken, NJ.

Laamanen, T., Brauer, M., Junna, O., 2014. Performance of acquirers of divested assets: evidence from the U.S. software industry. Strategic Management Journal, 35(6), 914–925.

Laulié, L., Morgeson, F.P., 2021. The end is just the beginning: turnover events and their impact on those who remain. Personnel Psychology, 74(3), 387–409.

Ma, L., Sun, B., Kekre, S., 2015. The squeaky wheel gets the grease—an empirical analysis of customer voice and firm intervention on Twitter. Marketing Science, 34(5), 627–645.

Mackey, T.B., Barney, J.B., Dotson, J. P., 2017. Corporate diversification and the value of individual firms: a Bayesian approach. Strategic Management Journal, 38(2), 322–341.

MacKinlay, A.C., 1997. Event studies in economics and finance. Journal of Economic Literature, 35(1), 13–39.

Mandelker, G., 1974. Risk and return: the case of merging firms. Journal of Financial Economics, 1(4), 303–335.

McWilliams, A., Siegel, D., 1997. Event studies in management research: theoretical and empirical issues. Academy of Management Journal, 40(3), 626–657.

Miric, M., Jeppesen, L.B., 2020. Does piracy lead to product abandonment or stimulate new product development? Evidence from mobile platform-based developer firms. Strategic Management Journal, 41(12), 2155–2184.

Moeller, S.B., Schlingemann, F.P., Stulz, R.M., 2004. Firm size and the gains from acquisitions. Journal of Financial Economics 73(2), 201–228.

Moeller, S.B., Schlingemann, F.P., Stulz, R.M., 2005. Wealth destruction on a massive scale? A study of acquiring-firm returns in the recent merger wave. Journal of Finance, 60(2), 757–782.

Nardi, L., Zenger, T., Lazzarini, S. G., Cabral, S., 2022. Doing well by doing good, uniquely: materiality and the market value of unique CSR strategies. Strategy Science, 7(1), 10–26.

Nguyen, H., Calantone, R., Krishnan, R., 2020. Influence of social media emotional word of mouth on institutional investors' decisions and firm value. Management Science, 66(2), 887–910.

Oler, D.K., Harrison, J.S., Allen, M.R., 2008. The danger of misinterpreting short-window event study findings in strategic management research: an empirical illustration using horizontal acquisitions. Strategic Organization, 6(2), 151–184.

Pang, B., Lee, L., Vaithyanathan, S., 2002. Thumbs up? Sentiment classification using machine learning techniques. Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing–volume 10. Association for Computational Linguistics, Stroudsburg, PA, pp. 79–86.

Peri, G., Yasenov, V., 2019. The labor market effects of a refugee wave: synthetic control method meets the Mariel boatlift. Journal of Human Resources, 54(2), 267–309.

Porter, M.E., 1980. Competitive strategy: techniques for analyzing industries and competitors. Free Press, New York.

Puranam, P., Srikanth, K., 2007. What they know vs. what they do: how acquirers leverage technology acquisitions. Strategic Management Journal, 28(8), 805–825.

Puranam, P., Vanneste, B., 2016. Corporate strategy: tools for analysis and decision-making. Cambridge University Press, Cambridge, UK.

Rabier, M.R., 2017. Acquisition motives and the distribution of acquisition performance. Strategic Management Journal, 38(13), 2666–2681.

Rietveld, J., Seamans, R., Meggiorin, K., 2020. Market orchestrators: the effects of certification on platforms and their complementors. Strategy Science, 6(3), 244–264.

Roesel, F., 2017. Do mergers of large local governments reduce expenditures? Evidence from Germany using the synthetic control method. European Journal of Political Economy, 50, 22–36.

Schijven, M., Hitt, M.A., 2012. The vicarious wisdom of crowds: toward a behavioral perspective on investor reactions to acquisition announcements. Strategic Management Journal, 33(11), 1247–1268.

Schweizer, L., Wang, L., Koscher, E., Michaelis, B., 2022. Experiential learning, M&A performance, and post-acquisition integration strategy: a meta-analysis. Long Range Planning, 55(6), 102212.

Sharpe, W.F., 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. Journal of Finance, 19(3), 425–442.

Shinkle, G.A., Kriauciunas, A. P., Hundley, G., 2013. Why pure strategies may be wrong for transition economy firms. Strategic Management Journal, 34(10), 1244–1254.

Shleifer, A., Summers, L.H., 1988. Breach of trust in hostile takeovers. In: Auerbach, A.J. (Ed.), Corporate takeovers: causes and consequences. University of Chicago Press, Chicago, IL, pp. 33–68.

Siggelkow, N., 2002. Evolution toward fit. Administrative Science Quarterly, 47(1), 125–159.

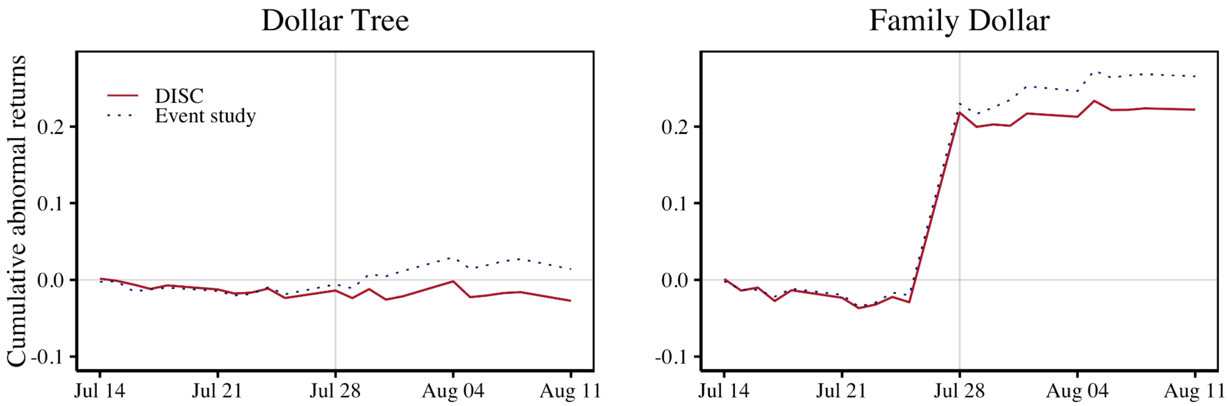
Stinchcombe, A. L., 1987. Constructing social theories. University of Chicago Press, Chicago, IL.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 58(1), 267–288.
Tong, L., Wang, H., Xia, J., 2020. Stakeholder preservation or appropriation? The influence of target CSR on market reactions to acquisition announcements. Academy of Management Journal, 63(5), 1535–1560.
Yakis-Douglas, B., Angwin, D., Ahn, K., Meadows, M., 2017. Opening M&A strategy to investors: predictors and outcomes of transparency during organisational transition. Long Range Planning, 50(3), 411–422.
Zollo, M., Meier, D., 2008. What is M&A performance? Academy of Management Perspectives, 22(3), 55–77.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2), 301–320.

Table 1. Comparison of Event Study and DISC Methodologies

	Event study	DISC
Outcome variable	Shareholder returns	Any time-series outcome (including shareholder returns)
Comparison firms	Market index (many public firms)	Few firms (private and public)
Abnormal outcome	= actual outcome − predicted outcome $AY_{e,0t} = y_{e,0t} - \hat{y}_{e,0t}$	= actual outcome − predicted outcome $AY_{d,0t} = y_{d,0t} - \hat{y}_{d,0t}$
Predicted outcome	Based on a theoretical model. In CAPM: Predicted outcome = constant + linear combination of comparison firms $\hat{y}_{e,0t} = a_e + b_e \sum_{i=1}^{N_e} w_{e,i} y_{e,it}$	Based on an empirical model. Predicted outcome = constant + linear combination of comparison firms $\hat{y}_{d,0t} = a_d + \sum_{i=1}^{N_d} w_{d,i} y_{d,it}$
Coefficients for comparison firms	Weights (w_i) are from a market index; beta (b) is estimated from data.	Weights (w_i) are estimated from data.
Estimation method	Ordinary least squares	Elastic net
Statistical significance	Frequentist inference	Placebo test

Note: The y_{ot} term is the actual outcome for the focal firm (subscripted *0*) in episode *t*, and y_{it} is the actual outcome for comparison firm *i* in episode *t* (event study and DISC are subscripted by *e* and *d*, respectively.

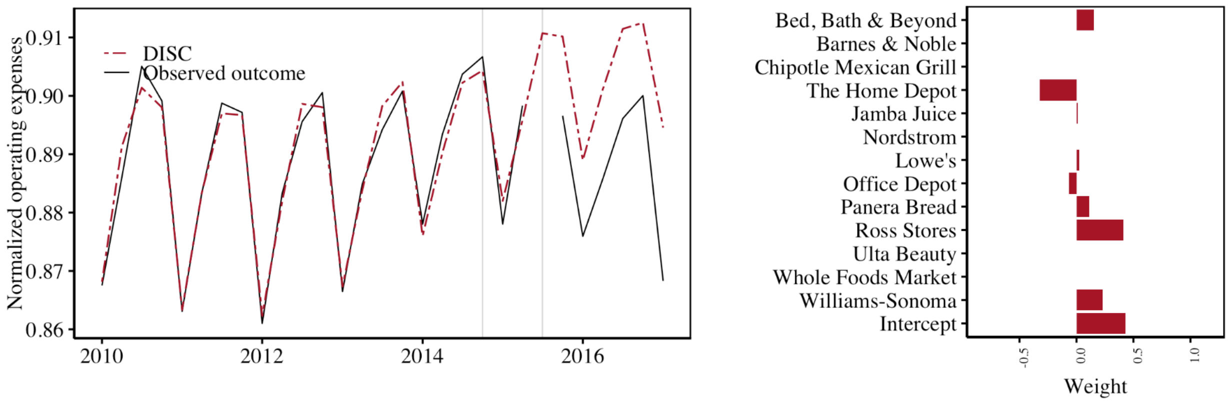
Figure 1. DISC and Event Study Compared: Cumulative Abnormal Returns Estimates



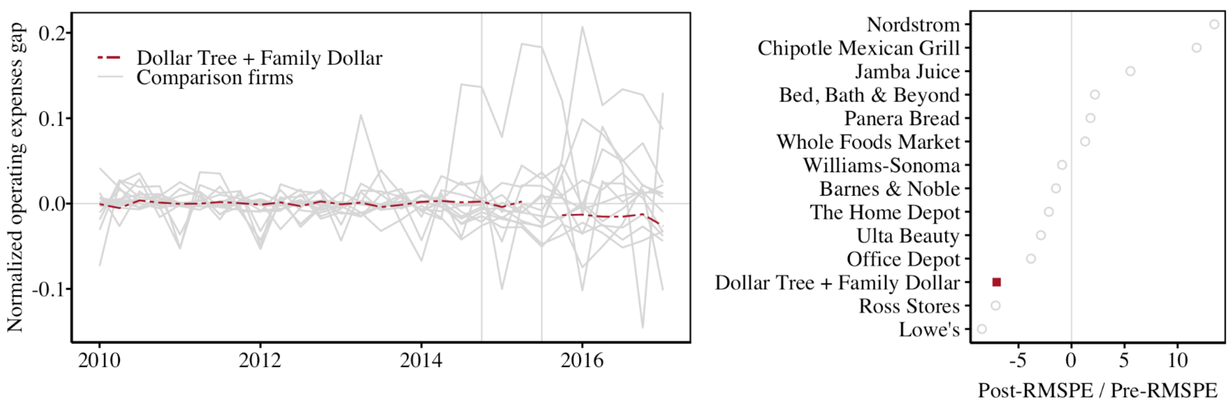
Note: The vertical line in each panel marks the acquisition's announcement date (July 28, 2014).

Figure 2. Analysis of Combined Operating Costs for Dollar Tree and Family Dollar

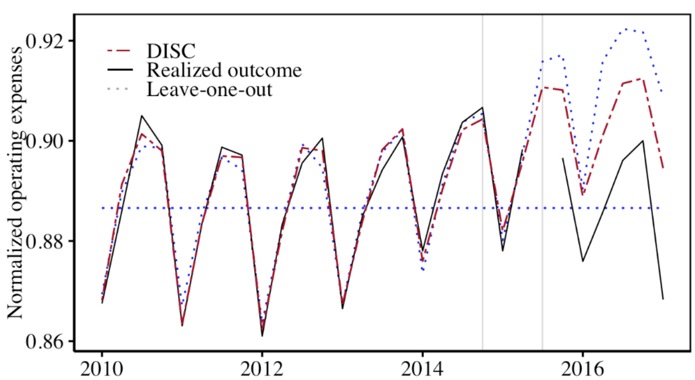
DISC analysis



Placebo test

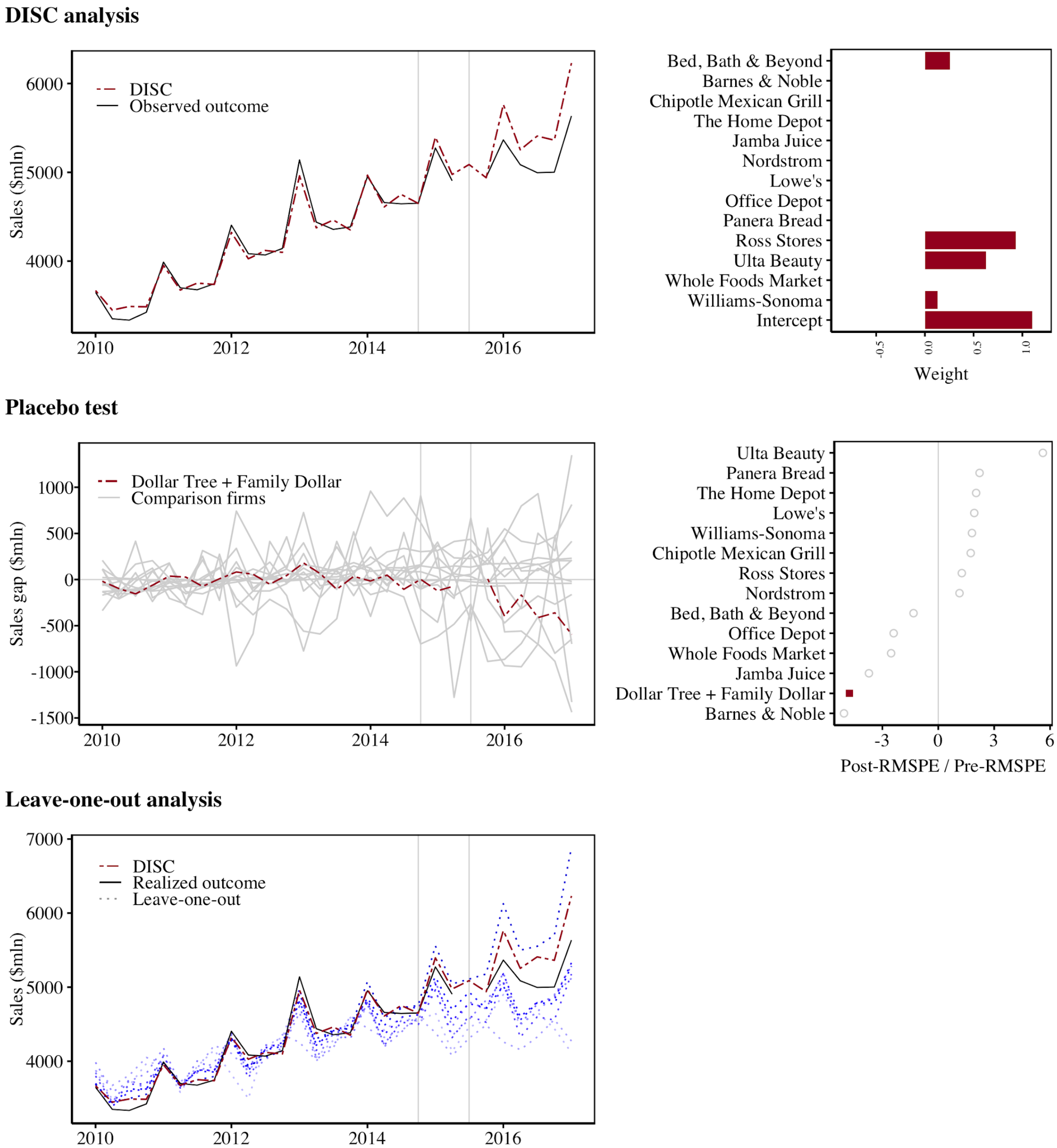


Leave-one-out analysis



Note: The vertical lines in the left column mark the announcement date (July 28, 2014) and closing date (July 6, 2015).

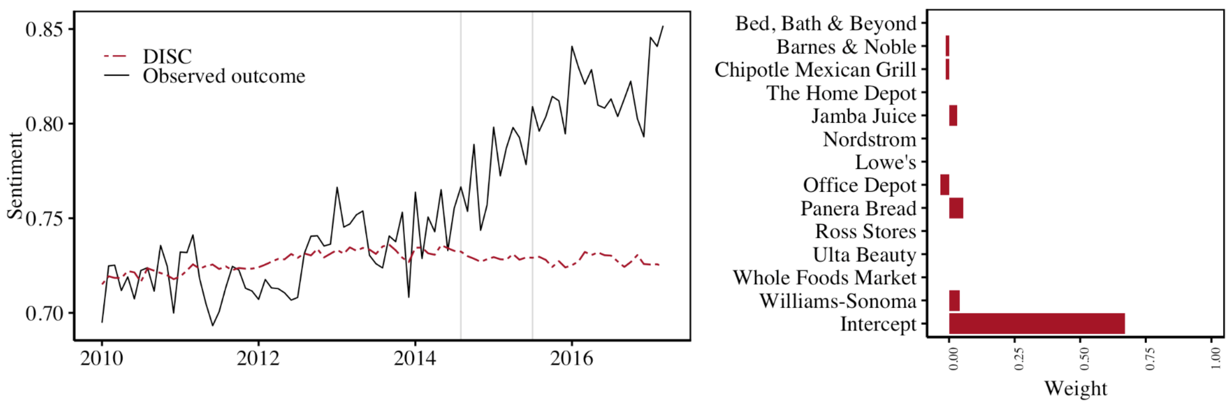
Figure 3. Results of Combined Sales for Dollar Tree and Family Dollar



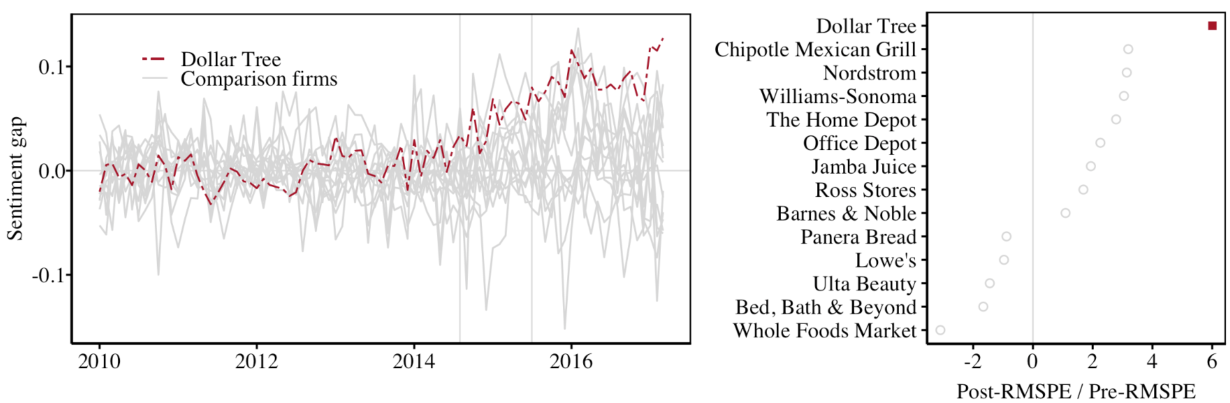
Note: The vertical lines mark the announcement date (July 28, 2014) and closing date (July 6, 2015). The weight of the intercept (858.32) is beyond the range in the upper right panel.

Figure 4. Results of Customer Sentiment for Dollar Tree

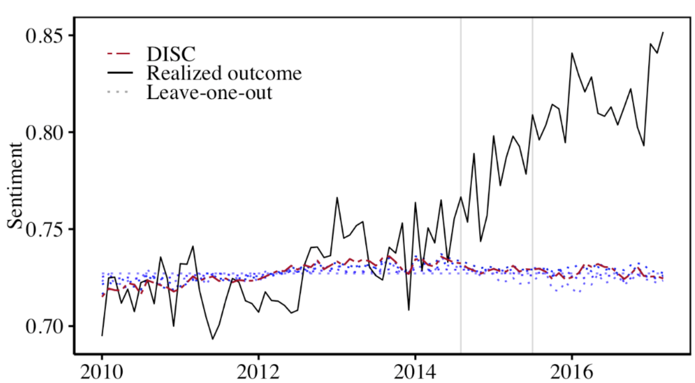
DISC analysis



Placebo test



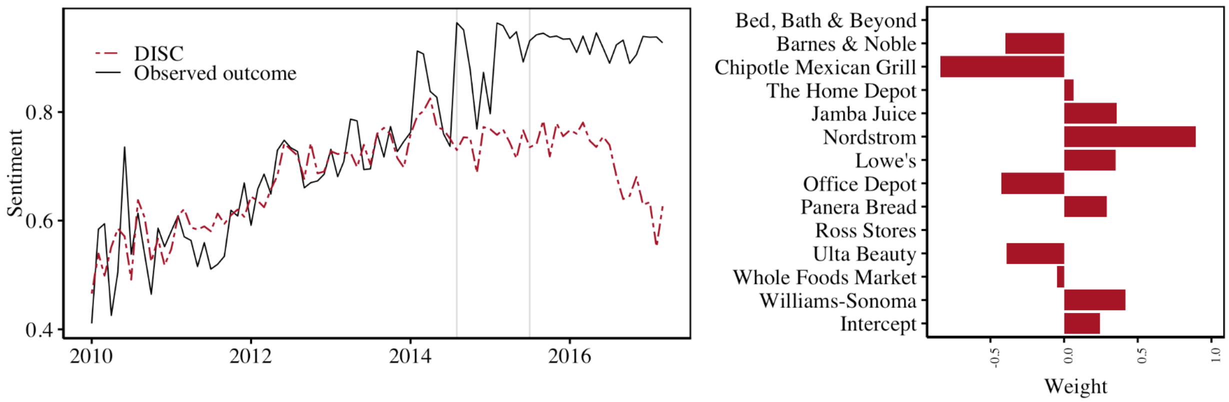
Leave-one-out analysis



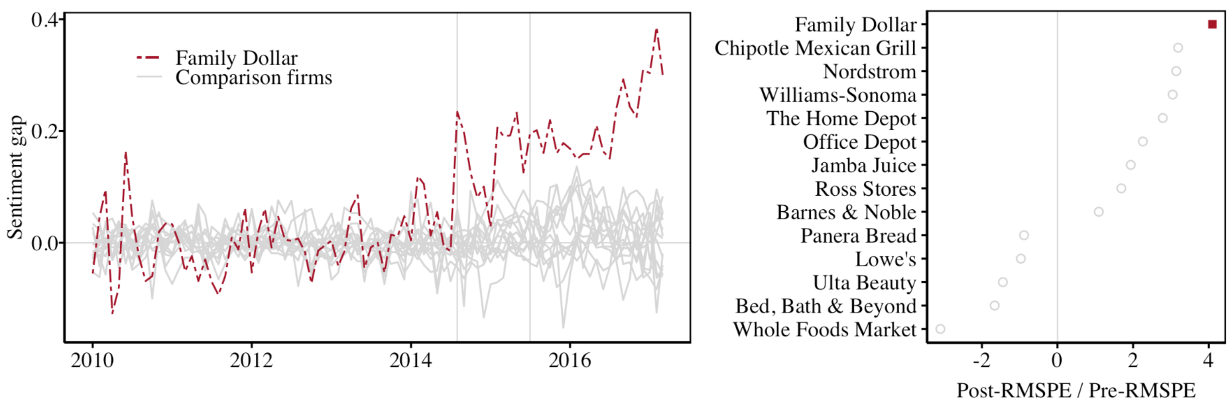
Note: The vertical lines mark the announcement date (July 28, 2014) and closing date (July 6, 2015).

Figure 5. Results of Customer Sentiment for Family Dollar

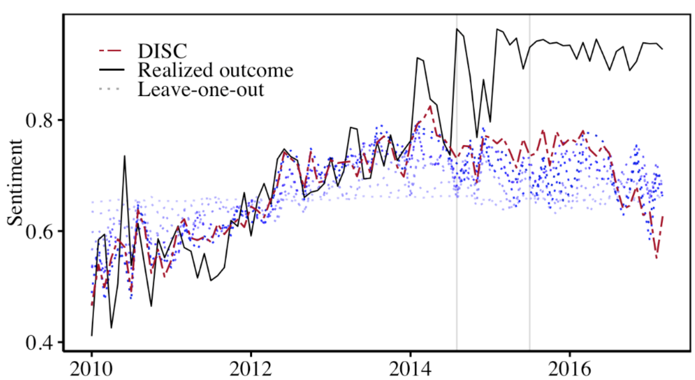
DISC analysis



Placebo test



Leave-one-out analysis



Note: The vertical lines mark the announcement date (July 28, 2014) and closing date (July 6, 2015).

APPENDIX

A. Other Methods Using Comparison Firms

Given that DISC is concerned with selecting comparison firms, we compare this approach with other methods that make the comparison group explicit. In addition to event study, popular methods in strategy research that do so are the original implementation of synthetic controls, coarsened exact matching, and difference-in-differences (see Abadie et al., 2015; Iacus et al., 2011; and Card and Krueger, 1994 for a detailed description of each respective method). We highlight how each method selects and weights comparison firms. Because we are interested in strategy applications, we adopt a firm-level unit of analysis in our discussion (but note that these methods are not exclusive to these applications or levels). Table A.1 provides an overview of the comparisons. Columns 1 and 2 summarize event study and DISC, which are covered in the paper. We elaborate below on early versions of synthetic controls, coarsened exact matching, and difference-in-differences.

—--

Insert Table A.1 about here

—--

A.1. Abadie Synthetic Controls

DISC is based on the synthetic control approach devised by Abadie and co-authors (Abadie and Gardeazabal 2003, Abadie et al. 2010, Abadie et al. 2015), which we denote “ASC.” Their insight is to compare a single focal firm with a synthetic firm (a “synthetic control”) constructed from a weighted (rather than a simple) average of comparison firms. The synthetic firm is created based on the pre-period values of the independent and dependent variables. The effect size is the difference in dependent variable between the focal and synthetic firm in the post-period.

ASC has been used in the management literature (e.g., Fremeth et al. (2016) on the impact of a government intervention and a product recall on automobile sales, and Conti and Valentini (2018) on the effect of judicial independence on entrepreneurial entry). The method has enjoyed particular success in economics for studying policy implications (e.g., Gharehgozli (2017) investigates the effect of sanctions on Iran's GDP, Peri and Yasenov (2019) study the impact of a refugee wave on local wages in Miami, and Roesel (2017) analyzes how mergers of local municipality governments affected municipality spending and voter behavior). Although most synthetic control applications have one or a few focal units, recently this approach has been applied also in settings with multiple focal firms (e.g., Connelly et al. 2020, Rietveld et al. 2020).

ASC differs from DISC in two noteworthy ways. First, DISC uses only the dependent variable to compute the weights whereas implementations of ASC use both independent and dependent variables. In fact, only a few (rather than all) pre-period episodes of the dependent variable are used in order to simplify computations (Abadie et al. 2010). If independent variables are required, then more stringent data requirements are imposed. Second, ASC requires that weights be nonnegative and to sum to one. Although the resulting synthetic control has an intuitive interpretation, these constraints imply that the focal firm must be similar to comparison firms. Absent these constraints, it is unnecessary for a focal firm to resemble a comparison firm because DISC can apply a negative weight to a comparison firm to yield its "opposite" or apply a weight greater than 1 to yield a larger version of a comparison firm in terms of the outcome variable (for example, a comparison firm with \$8 million in sales could mimic a focal firm with \$10 million in sales by applying a weight of 1.25). As a result, using ASC to study a unique or outlier focal firm—which is of particular interest in strategy—is less feasible. Moreover, similar firms are most likely found in the same industry, which may raise a spillover concern: when the focal firm's strategic event also affects the comparison firms. The option of choosing dissimilar comparison firms (e.g., from different industries) can alleviate this concern.

A.2. Coarsened Exact Matching

Synthetic controls are typically used when there is one (or a few) focal firm(s), but matching methods, such as coarsened exact matching (CEM), are used when there are many focal firms. The goal of CEM is to reduce bias that arises from model dependence. CEM retains only similar focal and comparison firms. Consequently, the choice of the model to account for differences between focal and comparison firms is less important.

CEM drops focal firms for which there is no similar comparison firm as well as comparison firms for which there is no similar focal firm. Similarity is based on grouping each firm into bins that are created after coarsening each independent variable into ranges of values. Firms in bins that contain only focal or comparison firms receive a weight of zero, which excludes them from the sample. Weights for matched comparison firms balance the number of focal and comparison firms within a bin and adjust for the total number of matched firms. CEM accounts for differences in observables, not in unobservables. Because CEM is a data preprocessing approach, it is often combined with an estimation method (e.g., regression, difference-in-differences).

A.3. Difference-in-Differences

Whereas CEM uses observables to reduce model dependence based on observables, difference-in-differences (DiD) is a large-sample approach that seeks to account for time-invariant unobservables. A policy change or shock (e.g., intellectual property theft (Miric and Jeppesen 2020)) affects the focal firms but not the comparison firms. The central idea is to compare the difference in the outcome of interest for the focal firms (pre- and post-period) with the difference in outcome for the comparison firms. By assessing the difference between focal and comparison firms relative to their pre-period baselines, DiD accounts for the possibilities that, regardless of shock, firms are different beforehand and that outcomes change over time.

DiD is implemented using regression, so comparison firms are evenly weighted. A key assumption is that of parallel trends: the outcomes for the focal and comparison firms would have followed the same trajectory in the post-period were it not for the shock. This counterfactual assumption is untestable, so empirical researchers instead test for whether pre-period time trends are similar for focal and comparison firms. If so, then it is more plausible that post-period time trends would also have been similar. Justification for the validity of a chosen comparison group ultimately depends on the institutional context.

Table A.1. Comparison between DISC and Other Empirical Methods

	Event study	Doudchenko and Imbens synthetic control (DISC)	Abadie synthetic control (ASC)	Coarsened exact matching (CEM)	Difference-in-differences (DiD)
Method					
Justification for comparison group validity	Theoretical	Empirical	Empirical	Resulting similarity to focal firm(s) in terms of independent variables	Context-driven
Assessment of comparison group validity	Pre-period fit (e.g., R^2)	Pre-period fit	Pre-period fit	Statistical test of balance	Parallel trends
Justification for post-period model	Pre-period model holds in the post-period (in the absence of a strategy event)	Pre-period model holds in the post-period (in the absence of a strategy event)	Pre-period model holds in the post-period (in the absence of a strategy event)	Distributions of independent variables are similar between treatment and controls (no explicit distinction between pre- and post-period)	Pre-period model holds in the post-period (in the absence of a strategy event)
Endogeneity assumptions	For each M&A, no confounding events; for regressions using CAR as dependent variable, no unobserved factors	For each strategic event, no confounding events	Allows for constant unobserved factors with time-varying effects if pre-period is sufficiently long	No unobserved factors	No time-varying unobserved factors (firm fixed effects account for constant unobserved factors)
Statistical testing	Frequentist inference	Placebo test	Placebo test	Frequentist inference	Frequentist inference
Weights					
Variables used to compute weights	Dependent variable	Dependent variable	Dependent and independent variables	Independent variables	None (evenly weighted)
Firm-varying weights?	Yes (because market weight, w_i, varies by firm)	Yes	Yes	No (for matched firms within each stratum)	No

	Event study	Doudchenko and Imbens synthetic control (DISC)	Abadie synthetic control (ASC)	Coarsened exact matching (CEM)	Difference-in-differences (DiD)
Negative weights possible?	Yes (if $b < 0$, then $b \times w_i$ is negative for all firms)	Yes	No	No	No
Sum of weights equals one?	No	No	Yes	No	Yes
Illustrative weights of three comparison firms and one focal firm	$b \times 0.04$	5/4	3/4	5/4	1/3
	$b \times 0.02$	0	1/4	5/4	1/3
	$b \times 0.01$	–1/3	0	0	1/3
Data					
Type of data	Longitudinal	Longitudinal	Longitudinal	Cross-sectional or longitudinal	Longitudinal
Minimum number of focal firms needed for method	One	One	One	Many	Many
Minimum number of comparison firms needed for method	Many (index)	Few	Few	Many	Many
Application					
Representative strategy research question	What is the firm-specific effect of an M&A on a firm's cumulative abnormal returns? What is the average effect of an M&A characteristic on the CAR for a group of firms?	What is the firm-specific effect of a strategy or policy change on a strategic outcome for a firm that may be unique or an outlier?	What is the firm-specific effect of a strategy or policy change or strategy on a strategic outcome for a firm?	What is the average effect of a strategy on an outcome for a group of firms?	What is the average effect of a policy change on a strategic outcome for a group of firms?

	Event study	Doudchenko and Imbens synthetic control (DISC)	Abadie synthetic control (ASC)	Coarsened exact matching (CEM)	Difference-in-differences (DiD)
Pros (questions, data, context)	Shareholder returns are forward-looking, which allows for short post-periods and does not require independent variables M&A-related questions Large sample (typically) Context is M&As	Does not require independent variables; comparison firms can be dissimilar; shows how effect changes over time Questions pertaining to individual effects Small sample Diverse contexts in which comparison firms are unaffected	Intuitive and explicit synthetic control, shows how effect changes over time Questions pertaining to individual effects Small sample Diverse contexts in which comparison firms are similar to and unaffected by focal firm	Reduces model dependence; can be used in conjunction with other methods Many questions Large sample Broadly applicable across many contexts	Approximates experimental conditions using observational data Policy-related questions Large sample Contexts involving an exogenous shock
Cons	Public firms only, underlying theoretical model focuses on shareholder returns only	Strategy or policy change cannot affect comparison group, long post-periods may be confounded by other events	Comparison and focal firms must be similar, else difficult to obtain good pre-period fit; long post-periods may be confounded by other events	Cannot account for unobserved factors	Difficulty in finding an exogenous shock; long post-periods may be confounded by other events

ONLINE APPENDIX

A. Constructing The Customer Sentiment Measure

To compute the customer sentiment measure, we first wrote a Python script for the web scraping of tweets. Using Twitter's advanced search options, the following instructions were provided. First, each company name was used as a search term. For names with more than one word, both the separated and concatenated strings (e.g., "dollar tree" and "dollartree") were considered. The company name could be used anywhere in the tweet, including as a hashtag (#, or topic) or as a handle (@, i.e., username, for addressing a tweet). Second, we excluded tweets sent by any of the companies in our sample. Third, tweets could only be written in English. Fourth, tweets had to be sent between January 1, 2012 and September 26, 2017, when Twitter initiated a change in the maximum tweet length from 140 to 280 characters (Rosen and Ihara, 2017). The analyses are based on 52,486,229 tweets collected through March 2017 (see Table A.1).

To investigate the possibility that tweets are not from customers but from others (e.g., journalists), we read thousands of tweets and found that nearly all of them relate to customer interactions. Even though we cannot rule out that others impersonate customers, we found no evidence that a small percentage of users were responsible for a large percentage of tweets (for both Family Dollar and Dollar Tree, 50% of users account for 80% of tweets).[9] Some customers are more vocal than others (Gans et al., 2021; Ma et al., 2015). By design, then, this measure is interpreted as the average sentiment not of all customers but rather of the most vocal customers.

The sentiment measure is the average probability that a tweet about a company has positive sentiment in a given month. We estimate this probability in two steps. First, we train a machine learning model (or "classifier") using a separate dataset of 1.6 million tweets that were

[9] This split compares favorably to the split across all of Twitter, where the top 10% of users account for 80% of tweets (Wojcik and Hughes, 2019).

previously labeled as having either positive or negative sentiment (Go et al., 2009). We use a Bernoulli naïve Bayes classifier, which "naïvely" assumes that the features are mutually independent, conditional on the tweet's sentiment, which enables fast training of the classifier even when there are millions of features. Despite this strong assumption, this algorithm performs well in general and can outperform more complicated methods (Das and Chen, 2007). Because of its good performance and short training time, naïve Bayes is often used for natural language processing tasks, and for sentiment analysis in particular (Wang and Manning, 2012).

We randomly select 1.4 million tweets for the training set and use the remaining 200,000 tweets for the test set. For each tweet in the training set, we extract each word, or unigram, and every pair of consecutive words, or bigram (Pak and Paroubek, 2010). This approach yields a total of more than 4.5 million unigrams and bigrams. The classifier then learns which of these unigrams and bigrams predicts positive sentiment. The predicted value yields a probability of positive sentiment that ranges from 0 to 1. The classifier achieves a correct classification rate (or accuracy) of 80.4% in the test set (using a cut-off of 0.5 for positive sentiment).

Figure A.1 shows the ROC curve for the test set. This curve plots the performance of the classifier when lowering the cut-off from 1 at (0,0) to 0 at (1,1). A higher true positive rate goes hand in hand with a higher false positive rate. Better classifiers deviate more from the baseline (i.e., the diagonal); in other words, their true positive rate is higher (for a given false positive rate) or their false positive rate is lower (for a given true positive rate). A measure that captures difference from the baseline is the area under the curve (AUC), which is 0.88 for this classifier. An ROC curve can be used to select a cut-off for a desired level of true and false positive rates. A cut-off is not required in our case because we work directly with the probability outcomes.

Second, we use the classifier just described to predict the probability of positive sentiment for each tweet in our data. Table A.2 gives some examples of tweets from the data along with their predicted probability of having positive sentiment. To arrive at the measure, we average these probabilities by month and company. The monthly measure for the focal firms and each of the comparison firms is plotted in Figure A.2. Of note, Family Dollar's customer sentiment is an outlier, relative to our comparison firms, in terms of both slope and intercept. We therefore cannot use the original implementation of synthetic controls with weight constraints.

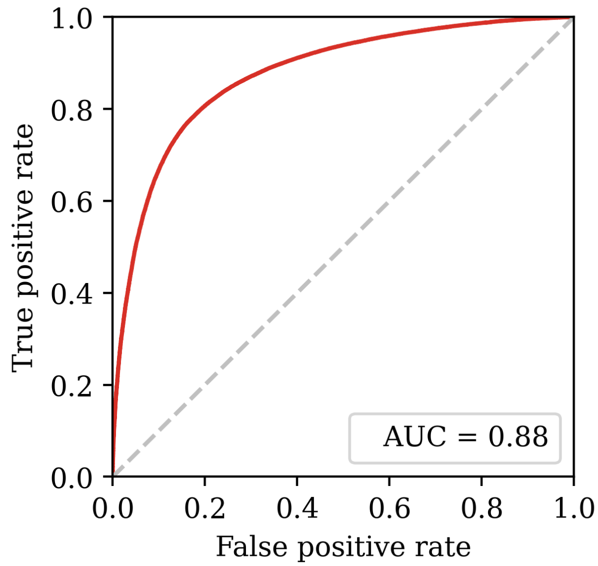
Table A.1. Number of Tweets by Firm

Firm	**Tweets**
Barnes & Noble	2,438,692
Bed Bath & Beyond	572,777
Chipotle Mexican Grill	27,518,415
Dollar Tree	*1,226,460*
Family Dollar	*209,329*
The Home Depot	4,915,138
Jamba Juice	304,482
Lowe's	3,689,943
Nordstrom	3,241,814
Office Depot	1,169,012
Panera Bread	710,414
Ross Stores	71,568
Ulta Beauty	441,418
Whole Foods Market	5,813,397
Williams-Sonoma	163,370
Total	*52,486,229*

Table A.2. Dollar Tree Tweet Examples and Estimated Positive Customer Sentiment Probabilities

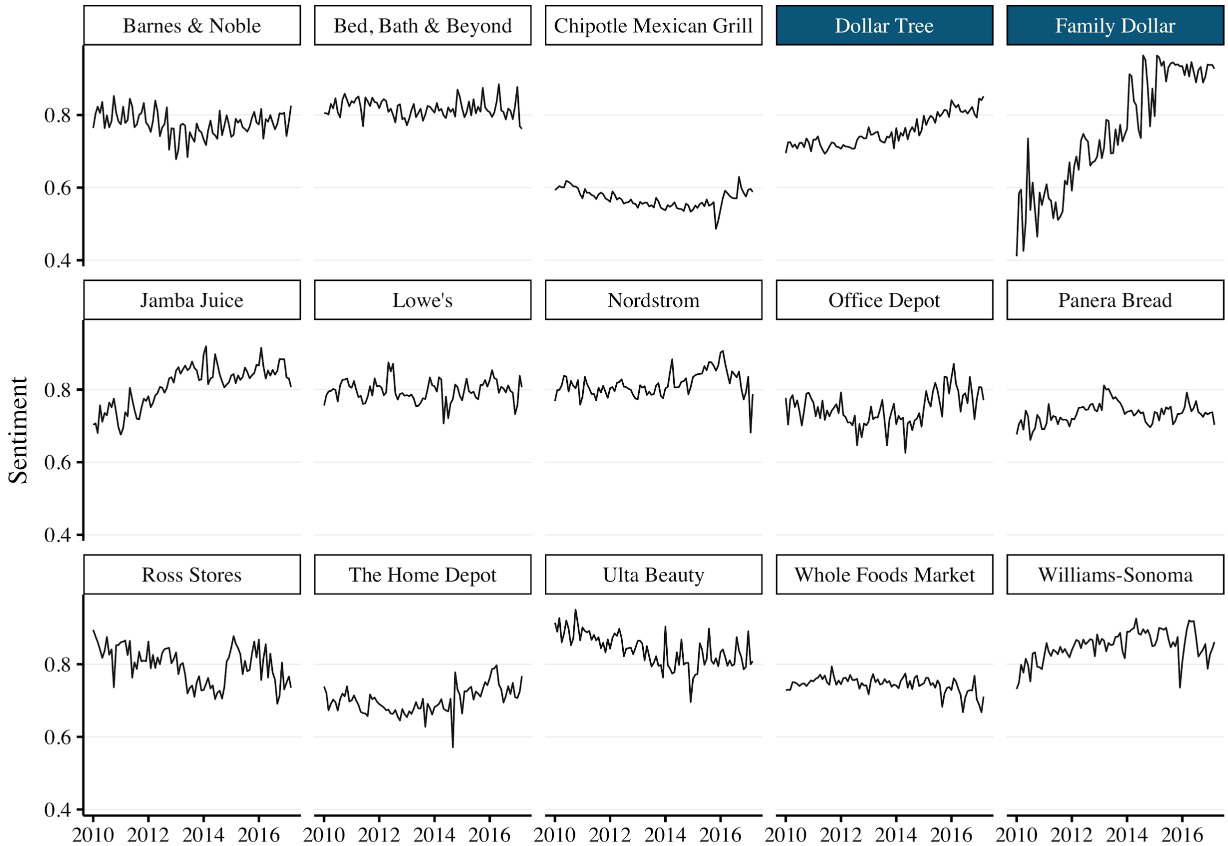
Tweet	**Probability**
God bless dollar tree	0.99
I find the best things. @ DOLLAR TREE	0.98
Dollar Tree Has Everything!!!! @ DOLLAR TREE	0.76
I feel rich at the dollar tree	0.73
Dollar tree white cheddar popcorn is the shit	0.68
I'm at @DOLLARTREE (Jacksonville, FL)	0.54
Picking up a few things I forgot yesterday (at @DOLLARTREE)	0.47
I don't like dollar tree candy	0.32
That's what I get for buying shades from Dollar Tree. #BROKED	0.23
This place sucks cheap stuff but cashiers and lines awful!! (@ DOLLAR TREE)	0.01

Figure A.1. ROC Curve of the Bernoulli Naïve Bayes Classifier for the Test Set



Note: AUC is the area under the curve.

Figure A.2. Customer Sentiment by Firm, 2010–2017



References

Das, S. R., Chen, M. Y., 2007. Yahoo! for Amazon: sentiment extraction from small talk on the web. Management Science, 53(9), 1375–1388.

Gans, J. S., Goldfarb, A., Lederman, M., 2021. Exit, tweets, and loyalty. American Economic Journal: Microeconomics, 13(2), 68–112.

Go, A., Bhayani, R., Huang, L., 2009. Twitter sentiment classification using distant supervision. CS224N Project Report, Stanford, 1(12), 2009.

Ma, L., Sun, B., Kekre, S., 2015. The squeaky wheel gets the grease—An empirical analysis of customer voice and firm intervention on Twitter. Marketing Science, 34(5), 627–645.

Pak, A., Paroubek, P., 2010. Twitter as a corpus for sentiment analysis and opinion mining. In: Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., Tapias, D. (Eds.), Proceedings of the seventh international conference on language resources and evaluation: LREC 2010. European Language Resources Association, Valletta, Malta, pp. 1320–1326.

Rosen, A., Ihara, I., 2017, September 26. Giving you more characters to express yourself. Retrieved from https://blog.twitter.com/official/en_us/topics/product/2017/Giving-you-more-characters-to-express-yourself.html

Wang, S., Manning, C. D., 2012. Baselines and bigrams: Simple, good sentiment and topic classification. In: Strube, M. (Ed.), Proceedings of the 50th annual meeting of the Association for Computational Linguistics: Short papers - volume 2. Association for Computational Linguistics, Stroudsburg, pp. 90–94.

Wojcik, S., Hughes, A., 2019, April 24. How Twitter users compare to the general public. Pew Research Center: Internet, Science & Tech. Retrieved from https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/