NBER WORKING PAPER SERIES

THE (MISSING) RELATION BETWEEN ANNOUNCEMENT RETURNS AND VALUE CREATION

Itzhak Ben-David
Utpal Bhattacharya
Stacey E. Jacobsen

Working Paper 27976
http://www.nber.org/papers/w27976

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020, revised April 2024

Previously circulated as "Do Acquirer Announcement Returns Reflect Value Creation?" We are thankful for the helpful comments from Hank Bessembinder, Alon Brav, Nickolay Gantchev, John Graham, Douglas Hanna, Jarrad Harford, Cam Harvey, David Hirshleifer, Gerard Hoberg, Mohammad Irani, Pab Jotikasthira, Adam Kolasinski, Mattia Landoni, Christian Leuz, James Linck, Vojislav Maksimovic, Antonio Macias, Ulrike Malmendier, Darius Miller, Micah Officer, Rik Sen, Wayne Shaw, Jared Stanfield, Mike Stegemoller, René Stulz, Rex Thompson, Kumar Venkataraman, Mike Weisbach, James Weston, Mike Wittry, seminar participants at Baylor University, the Chinese University of Hong Kong, the Hong Kong University of Science and Technology, Southern Methodist University, The Ohio State University, the University of Buffalo, and Texas Tech University. We also thank participants at the American Finance Association Conference, Asian Bureau of Finance and Economic Research (ABFER) Conference, the Midwest Finance Association Conference, and the Financial Research Network (FIRN) seminar series. Bhattacharya acknowledges the generous funding by Hong Kong GRF Grant No. 16500118. Ben-David is with The Ohio State University and the National Bureau of Economic Research (NBER). Bhattacharya is with the Hong Kong University of Science and Technology. Jacobsen is with Southern Methodist University. Ben-David is a co-founder and a partner in an investment advisor that manages investment accounts. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The (Missing) Relation Between Announcement Returns and Value Creation
Itzhak Ben-David, Utpal Bhattacharya, and Stacey E. Jacobsen
NBER Working Paper No. 27976
October 2020, revised April 2024
JEL No. G02,G14,G32,G34

ABSTRACT

Cumulative abnormal returns (CAR) computed during acquisition announcements are widely considered to be market-based assessments of expected value creation. We show that announcement returns do not correlate with commonly used and new measures of ex-post acquisition outcomes. A simple characteristics model using standard information known at announcement can predict outcomes reasonably well, and CAR fails even to capture the prediction from this model. A likely reason is that, because acquisition decisions are endogenous, CAR conveys information about the NPV of the deal as well as the event that triggered the deal announcement. We find that CAR variance is too high to be explained by NPV variance alone, suggesting that other non-NPV information related to this trigger dominates. We conclude that CAR is an unreliable measure of expected value creation.

Itzhak Ben-David
The Ohio State University
Fisher College of Business
606A Fisher Hall
Columbus, OH 43210-1144
and NBER
ben-david.1@osu.edu

Utpal Bhattacharya
and Technology
Clear Water Bay, Kowloon
HONG KONG, S.A.R.
Hong Kong
ubhattac@ust.hk

Stacey E. Jacobsen
Southern Methodist University
Cox School of Business
6212 Bishop Blvd.
Dallas, TX 75275
staceyj@cox.smu.edu

1 Introduction

Cumulative abnormal returns (CAR) computed during acquisition announcements are overwhelmingly favored by financial economists to measure value creation in acquisitions. Over the last five decades, CAR has been used to measure value creation in over 92% of the articles in top finance journals studying value creation in acquisitions.[1] The deep conviction in CAR spills over to business teaching and legal cases (Brealey, Myers, Allen, and Krishnan, 2006; Brav and Heaton, 2015).

The fact that CAR became the status-quo empirical measure for value creation is surprising, given the disagreement about the underlying theory and the mixed empirical evidence. Campbell, Lo, and MacKinlay (1997) argue that CAR means value creation in a well-functioning market: "...given rationality in the marketplace, the effect of an event will be reflected immediately in asset prices." Early empirical research in finance supported this view.[2] In contrast, Grinblatt and Titman (2002) recognize that CAR could include non-deal information: "The stock returns of the bidder at the time of the announcement of the bid may tell us more about how the market is reassessing the bidder's business than it does about the value of the acquisition." Consistent with this argument, studies found that CAR contains information about the acquirer standalone,[3] CAR is also distorted by leakage, endogeneity, low completion probability, feedback, and price pressure,[4] contains information irrelevant for value creation, and omits information known at the time of the acquisition.[5]

[1] Our review of articles in the *Journal of Finance*, *Journal of Financial Economics*, and *Review of Financial Studies* from 1972 to 2021 reveals that of the 4.8% of articles focusing on topics related to mergers and acquisitions (M&As), 54.8% computed measures of acquisition value creation. Of these, 92.2%—a total of 202 articles—used CAR to measure value creation. We detect no declining trend in its use.

[2] Healy, Palepu, and Ruback (1992) link announcement returns to operating cash flow improvements based on a sample of 50 large acquisitions. Kaplan and Weisbach (1992) show that announcement returns are lower for a sample of 37 transactions that were divested-at-a-loss. Later small-scale European market studies did not find a significant correlation (Papadakis and Thanos, 2010; Schoenberg, 2006).

[3] See Jensen and Ruback (1983), Schipper and Thompson (1983a), Roll (1986), Hietala, Kaplan, and Robinson (2003), Jacobsen (2014), and Pan, Wang, and Weisbach (2016).

[4] See Jensen and Ruback (1983), Malatesta and Thompson (1985), Eckbo, Maksimovic, and Williams (1990b), Fuller, Netter, and Stegemoller (2002), Mitchell, Pulvino, and Stafford (2004), Bhagat, Dong, Hirshleifer, and Noah (2005), Betton, Eckbo, and Thorburn (2008), Viswanathan and Wei (2008), Betton, Eckbo, Thompson, and Thornburn (2014), Edmans, Goldstein, and Jiang (2015), and Wang (2018).

[5] See Mitchell and Stafford (2000), Powell and Stark (2005), Malmendier, Moretti, and Peters (2018), and

Therefore, CAR's usefulness as an empirical measure of deal quality depends on the relevance of the information it contains to value creation. Does CAR primarily reflect information about deals' prospects, or does other non-deal-related information dominate it?

In this study, we systematically assess the validity of CAR as a reliable measure of value creation in the context of acquisitions using a comprehensive sample of over 47,000 acquisition announcements made over almost four decades (1980–2018). In the first part of the paper, we rely on several widely-used measures of ex-post value creation and also devise novel measures and find no meaningful correlation between these measures and announcement returns. We find these measures, instead, are predictable at the time of the announcement using standard deal information known at the announcement. However, CAR also does not correlate with this component, indicating that CAR fails to reflect relevant information at the time of the announcement. In the second part of the paper, we argue that the endogeneity of acquisition announcements is particularly problematic. Using a simple model, we show that even under the most favorable conditions, CAR is theoretically expected to include both information about the net present value (NPV) arising from the transaction *and* non-NPV information related to the event triggering the deal announcement. An empirical analysis of the second moment of CAR reveals that the non-NPV component likely dominates the deal information contained in CAR. We conclude that CAR is an unreliable measure of value creation in acquisitions.

In the first series of empirical tests, we examine whether CAR aligns with observable ex-post transaction- and firm-level outcome measures. Although value creation is unobservable, we construct empirical measures (both new and commonly used) that reflect different aspects of firms' operations and that are derived from different data sources. At the transaction level, we design an indicator to measure acquisition failure. Specifically, we manually collect information on deal-level goodwill impairments, i.e., accounting write-offs, which indicate that the target is no longer worth its original price.[6] At the acquirer level, we employ

Dasgupta, Harford, and Ma (2023).

[6]Unlike other commonly used measures of performance, our goodwill impairment data are linked to

both short- and long-term abnormal return on assets (ROA) measures commonly used in the literature (e.g., Healy et al., 1992; Harford and Li, 2007). Importantly, despite being derived from different sources and capturing both the left-tail and the entire distribution, these ex-post measures are positively and significantly correlated with each other. Following the literature, we also consider whether managers "listen to" CAR and include completion (versus withdrawal) as an additional outcome variable (see Asquith, Bruner, and Mullins, 1983; Jennings and Mazzeo, 1991; Luo, 2005; Kau, Linck, and Rubin, 2008).

We document that announcement returns are largely uncorrelated with non-impairment and short- and long-term abnormal ROA. We find no meaningful correlation in univariate and multivariate settings, either in- or out-of-sample, across multiple announcement return windows and estimation techniques. Similar to Luo (2005), announcement returns are positively and significantly related to deal completion. However, the economic magnitudes of these effects are very small. Across all outcome variables, CAR explains at best 0.2% of the variation in the probability of impairment, 0.04% of the variation in abnormal ROA, and 0.03% of the variation in completion.

We find that CAR continues to fail to correlate with outcomes in simple subsamples based on different time periods and based on an extensive number of acquirer characteristics (e.g., serial vs. first-time bidders), target characteristics (e.g., public vs. private), and transaction characteristics (e.g., cash vs. stock), and even in subsamples where we expect CAR to perform better. We conduct a brute force data-mining effort, searching among complex subsample formations to find the "golden subset" in which CAR consistently correlates with outcomes. Unfortunately, we cannot identify a group of transactions for which CAR is a reliable predictor of outcomes. We therefore conclude that the lack of reliability of CAR is

specific transactions rather than at the overall acquirer level. In Internet Appendix C, we validate that goodwill impairment is a robust signal of value destruction by relating it to several indirect symptoms of failure: poor stock and operating performance, distressed delisting, and management turnover. Our measure is similar in spirit to that of Mitchell and Lehn (1990), Kaplan and Weisbach (1992), and Berger and Ofek (1996), who construct a transaction failure measure based on transactions for which targets are divested at a loss. However, their sample is small because it is conditioned on future disposal and they lack data on divestiture prices.

not limited to specific subsets of data or specific time periods. In other words, CAR issues are systematic.

Given that CAR does not correlate with "realized" outcomes, we next explore the nature of the negligible correlation between announcement returns and acquisition outcomes by relating CAR to another ex-ante measure. We construct a simple benchmark using standard deal and acquirer characteristics, also known at the time of the announcement. We find that while outcomes can be predicted reasonably well using deal and acquirer characteristics (both in-sample and out-of-sample), CAR performs poorly relative to this simple benchmark. In out-of-sample tests, we assess the relation between CAR and the "predictable" component of acquisition outcomes and find that CAR does not even correlate with outcomes predicted by characteristics known at the time of the announcement. These results indicate that announcement returns do not reflect all relevant information at the time of the announcement.

We corroborate our inference of this wide disparity between the predictive ability of CAR and a characteristics model by forming trading strategies that take long positions in the acquirers predicted by the CAR model to perform the best (i.e., CAR predicts more favorable acquisition outcomes) and take short positions in the acquirers predicted to do the worst by the CAR model. We repeat these tests using the characteristics model, and then compare the results. The performance spread in the five-year DGTW-adjusted buy-and-hold returns between the top and bottom three deciles defined by CAR ranges between -0.8% and 2.8%. In contrast, the return spread between the top and bottom three deciles, as determined by characteristics, is large and ranges from 7.9% to 10.7%. These results are consistent with the recent work of Campbell, Elfrink, Huang, and Lu (2024) showing that post-acquisition long-run returns are predictable based on characteristics.

The weak relation between completion outcomes and announcement returns is consistent with feedback effects, i.e., managers take action (e.g., cancel the deal, work harder) in response to negative or positive CAR. We therefore devise a test to shed light on the magnitude of this feedback effect. Since completion outcomes can be predicted reasonably well out-of-sample using deal and acquirer characteristics, we first document that the lack of correlation between CAR and outcomes (non-impairment, ROA) persists even for a sample of deals that have a high likelihood of completion. This result suggests that the feedback effect cannot explain the lack of correlation between CAR and acquisition outcomes. Second, we consider the benefits of "listening" to CAR. We find that withdrawing (versus completing) negative-CAR deals and completing (versus withdrawing) positive-CAR deals generates a long-term return loss of 5%. In contrast, "listening" to the benchmark characteristics model generates a long-term return spread of more than 20%. We conclude that feedback effects, while present, are unlikely to be the main driver of the lack of correlation between CAR and outcomes.

Following the large existing literature that examines the "types" of transactions that create or destroy value, we consider how inferences are altered due to the lack of association between announcement returns and ex-post outcomes. Our four ex-post outcomes are associated with similar deal and acquirer characteristics. Strikingly, we find no association (in terms of sign and relative importance) between the characteristics for which CAR predicts failure or success and those associated with failure or success ex-post. Using the four most common characteristics used in the literature (form of payment, the target's public status, acquirer size, and relative transaction size), we form 16 "clusters" of transactions and find minimal overlap in the performance of clusters based on CAR versus our ex-post outcomes. For example, the cluster considered to create the most value according to CAR has the poorest ex-post outcomes among the 16 clusters, and the cluster considered to destroy the most value according to CAR is associated with ex-post outcomes in the top half of the quality distribution. These results indicate that inferences generated from CAR regarding deal quality are unreliable.

In the final part of the study, we delve deeper into the informational content of CAR. Using a simple model, we explore the theoretical underpinnings of CAR as a measure of

NPV. Acquisitions are not random events; internal and external triggers prompt them. The model shows that even under classical assumptions (efficient market, no agency costs, no anticipation, no feedback), CAR always contains at least two value-relevant components: the deal's NPV and value-relevant information about the triggering event that led to the deal announcement. Our simple model implies that CAR's informativeness about NPV depends on the relative importance of these two components.

Next, we use a novel empirical approach to assess the magnitude of these components that does not rely on ex-post outcomes but instead measures the information content of CAR via its variance. We find that CAR is far too volatile to convey only NPV information. For example, in 27% of deals with negative CAR, dollar losses exceed the amount invested; likewise, in 16% of deals with positive CAR, the value to the acquirer is more than twice what was invested. Further, the sensitivity of CAR variability to acquirer size is about eight times higher than its sensitivity to deal size. As NPV is related primarily to deal characteristics, these results imply that other information related to the acquirer, such as the information that triggered the announcement, likely dominates NPV in influencing CAR. While there may be other noise in CAR that affects the accuracy of its estimation of NPV, our results suggest that endogeneity may be a major issue.

To conclude, across multiple methodologies and samples, we find that CAR is an unreliable measure of NPV. It appears to be swamped with information unrelated to the value created by the deal itself. Researchers should, therefore, reconsider economic inferences based on CAR.

2 Sample and Outcome Measures

In this section, we describe the construction of our acquisition sample and the measures we use to capture the performance of the acquisition. We employ ex-ante and ex-post measures that capture transaction- and acquirer-level outcomes.

2.1 Acquisition Sample Construction

Our sample of mergers and acquisitions comes from the Thomson Reuters Securities Data Company (SDC) Domestic Merger and Acquisition database. Our sample begins in 1980 and ends in 2018, which allows us to track acquisition outcomes over the five years following the transaction. We include transactions that satisfy the following criteria: (a) The merger or acquisition was announced on or after January 1, 1980, and was effective by December 31, 2018; (b) the acquirer is a U.S. company; (c) the acquirer is a publicly-traded firm; (d) the deal is not classified as a leveraged buyout, spinoff, repurchase, self-tender, recapitalization, privatization, stake purchase, or acquisition of partial or remaining interest; (e) the percentage of shares acquired (or sought for not completed deals) is at least 50%; (f) the percentage of shares held by the acquirer six months before the announcement is less than 50%; (g) Compustat has accounting data on the bidder, and the Center for Research in Security Prices (CRSP) database has stock data for the month of the deal announcement; and (h) the deal value is non-missing in the SDC database. These requirements result in an initial sample of 47,543 deals, of which 42,354 are completed, 2,227 are withdrawn (the deal outcome is known in these cases), and 2,962 are not completed but not withdrawn (e.g., the transaction may be pending or the outcome is unknown; these we exclude from the main analysis in Section 3 but include in robustness tests and retain in Section 5).

2.2 Acquisition Performance Measures

For each transaction, we compute acquirer announcement returns. We estimate daily abnormal returns using a market model and a value-weighted index. The market model parameters, α_i and β_i, are estimated from 361 to 61 trading days before the deal announcement day, and r_{mt} is the CRSP value-weighted index. CARs are then computed by summing the daily abnormal returns over various event horizons. Following the existing literature, (e.g., Betton et al., 2008), we estimate CARs over three days $[-1, 1]$ and 11 days $[-5, 5]$ surrounding each acquisition announcement. CAR may understate absolute value expectations if the probability of deal completion is uncertain. Thus, we also estimate "Deal CARs" over the entire acquisition process beginning two days before the announcement and ending two days following the deal completion $[\text{Announcement} - 2, \text{Close} + 2]$. The advantage of this longer window is that uncertainty regarding deal completion is resolved. However, the disadvantage is that returns are measured over a long window and may include other acquirer-specific information. Therefore, we focus primarily on the short-term CAR measures.

We construct both transaction-level and firm-level proxies for acquisition outcomes to assess the core relation between announcement returns and value creation. Due to differences in data availability across outcome measures, the sample sizes vary for each measure. We provide further details on sample filters and the number of observations for the various outcome variables in Internet Appendix A.

2.2.1 Transaction-level Ex-post Measure: Goodwill Impairment

Measuring the extent to which specific acquisitions create or destroy value for the acquiring firm is challenging. Because the target is typically merged into the acquiring entity, we cannot directly observe the ex-post performance of the target or the synergies generated from the combined firms. To overcome this issue, we rely on increased transparency in accounting rules for goodwill impairment to construct a new transaction-level measure of acquisition failure. We construct an indicator of whether the goodwill associated with the transaction was materially impaired within five years of the deal's completion date.

We manually collect a sample of transactions with goodwill write-downs identified at the transaction level. These data offer a direct and quantifiable representation of ex-post value destruction in the acquiring firm for at least three reasons. First, goodwill, defined as the portion of the purchase price over the fair value of the target's identifiable net assets, reflects the going concern value of the target, the value of expected synergies, and overpayment. Therefore, the write-down of goodwill reflects value destruction caused by the following factors: overvaluation of existing assets, overestimated synergies, or the inability to realize synergies due to firm, industry, or economy-wide shocks. Second, the quality of goodwill impairment data has improved in recent years. The Statement of Financial Accounting Standards 142, passed in 2001, was implemented with the intent that unsuccessful acquisitions would be reflected more precisely and more quickly in firms' financial statements. After the completion of an acquisition, firms must conduct impairment tests following "material" events, and for many years in our sample, firms were required to conduct routine annual impairment tests to check for reductions in the value of goodwill.[7] The new standard also requires increased transparency for goodwill and impairment reporting at the reporting unit rather than at the firm level, making linking impairment to a specific triggering transaction easier. Third, prior research has documented that goodwill impairment events are value-relevant.[8]

We provide evidence that goodwill impairment is a signal of value destruction by relating our impairment measure to several indirect symptoms of acquisition failure. First, acquirers that impair goodwill are more likely to experience distressed delisting and poor operating and stock performance in the years following the acquisition relative to acquirers without impairment (Internet Appendices C.2 and C.3). Second, the market reaction to earnings announcements that contain goodwill impairment news is negative and large in magnitude, −2.8% on average (Internet Appendix C.1).[9] Third, CEOs are more likely to be fired in the period surrounding goodwill impairments than following negative CARs surrounding the original acquisition announcements (Internet Appendix C.4), indicating that the labor market regards impairment as an important signal for managerial discipline.[10]

One drawback of goodwill impairment as a measure of acquisition failure is the poten-

[7]In September 2011, the Financial Accounting Standards Board (FASB) modified SFAS 142 such that formal valuations to produce comparisons of fair value and carrying value of a reporting unit are only required when certain qualitative indicators of impairment exist.

[8]See Henning and Stock (1997), Chen, Kohlbeck, and Warfield (2004), Bens, Heltzer, and Segal (2011), Gu and Lev (2011), and Li, Shroff, Venkataraman, and Zhang (2011).

[9]Note that impairment news is a strictly negative piece of news about an event that has already happened. The fact that the market reaction is negative, given this stale negative news, does not conflict with the main finding of the study that CAR is not associated with future changes in cash flows.

[10]Of course, there are settings where impairment may not imply value destruction. For example, a target may be shut down once a target technology is exploited or a competitor is eliminated (e.g., Cunningham, Ederer, and Ma, 2021). Our results indicate that goodwill impairment is, in the vast majority of settings, associated with value destruction.

tial for subjectivity. Researchers have documented managerial discretion in the write-down decision, mainly about the amount and timing of the impairment.[11] In this paper, we focus on substantial goodwill impairments, a setting in which strategic manipulation is less viable because extreme losses must be revealed at some point.[12] Further, we focus on an indicator for impairment; thus, our results are less sensitive to the amount and timing of impairment.

Linking goodwill impairment to specific transactions is not straightforward because goodwill and impairment data reported on financial statements are based on *aggregate* firm-level data. As a result, we manually collect goodwill and impairment (if any) data. We begin by identifying all sample firms with firm-level goodwill impairments indicated in Compustat. For these "potentially" impaired transactions, we use the Notes to Consolidated Financial Statements in both the acquisition and impairment years to determine whether and how much of the impairment is due to the specific transaction in our sample. We focus on impairment within five years of the deal's effective date.[13]

For our analyses involving goodwill impairment, we impose additional filters on the 42,354 completed deals described in Section 2.1. First, our sample starts in 2003, when we can begin tracking goodwill impairment at the transaction level due to the implementation of SFAS 142 in 2002. Second, we require the transaction value to exceed \$10 million and to be at least 5% of the acquirer's market capitalization at the end of the fiscal year before the deal is announced. This filter allows a more precise measure of impairment: for very small deals (both in dollar and relative terms), it is difficult to determine the source of the impairment and, in many instances, the amount of goodwill originally produced from the transaction. These filters result in 8,367 deals. We exclude deals that have missing or zero

[11]See Elliott and Shaw (1988), Francis, Hanna, and Vincent (1996), Beatty and Weber (2006), Ramanna and Watts (2012), and Li and Sloan (2017).

[12]Our initial sample of potentially impaired deals requires firm-level impairment of at least 5% of acquirers' assets.

[13]To our knowledge, we are the first to construct a comprehensive dataset that includes transaction-specific goodwill balances and transaction-specific impairment outcomes in the post-SFAS 142 period. Hayn and Hughes (2006) also trace initial goodwill balances and subsequent impairments at the transaction level, but they exclude 55% of transactions due to insufficient information. Overall, they focus largely on the pre-SFAS 142 period, a time when the disclosure of initial goodwill and the source of the impairment were generally less comprehensive.

Compustat goodwill balances in both the year of and the year after the completion date, which yields 6,767 deals. Of these, we can reliably classify acquisition outcomes (transaction-level impairment or not) for 6,437 deals, of which 6,128 have the required announcement return and control variables.

Internet Appendix IA.A3 provides further details about the data collection procedure and shows that we successfully linked impairment events to specific transactions. As reported in Appendix Table IA.A4, goodwill impairments are relatively common: 14.8% of transactions in our sample experience an impairment event. These events are substantial: the average impairment constitutes 83% of total transaction-level goodwill, 57% of the total purchase price, and 11% of acquirer assets.

2.2.2 Firm-level Ex-post Measure: Abnormal Return-on-assets

We follow existing studies that approximate the contribution of acquisitions to the acquirers' cash flows by calculating their abnormal ROA (e.g., Healy et al., 1992; Chen, Harford, and Li, 2007; Fu, Lin, and Officer, 2013). The motivation is that the change in the acquirer's cash flows can be detected relative to an industry counterfactual.

We follow the procedure in Chen et al. (2007) and compute abnormal ROA over the years following the acquisition. To measure abnormal ROA, we regress the post-acquisition industry-adjusted n-year average ROA $(t+1, t+2, \ldots, t+n)$ on the pre-acquisition corresponding measure $(t-n, \ldots, t-2, t-1)$ and a constant:

$$\frac{1}{n}\sum_{t=1}^{n}[\mathrm{ROA}_{i,t} - \mathrm{ROA}_{Industry,t}] = \alpha + \beta\frac{1}{n}\sum_{t=-n}^{-1}[\mathrm{ROA}_{i,t} - \mathrm{ROA}_{Industry,t}] + \varepsilon_i, \tag{1}$$

where the residual ε_i measures the abnormal ROA.

For our "short-term abnormal ROA" measure, we define the post-acquisition (pre-acquisition) period as the three years after (before) the deal's effective date. We use three years as a plausible horizon because the median acquirer with goodwill impairment writes down by the third year following the acquisition. We also allow six years to capture a longer horizon for synergy realization—captured by "long-term abnormal ROA", in which we change the post-acquisition period to years four, five, and six after the deal's effective date.

Industry definitions are based on the Fama-French 48 industries (Fama and French, 1997). As discussed in Chen et al. (2007), this model considers the possibility that pre-acquisition operating performance could predict post-acquisition operating performance. Because of data availability throughout the entire 1980 to 2018 sample period, of the 42,354 completed acquisitions, we can compute short-term (long-term) abnormal ROA and required control variables for 28,710 (22,577) transactions.

The acquirer-level ROA performance measure has advantages and disadvantages relative to our transaction-level goodwill impairment indicator. The transaction-level deal failure indicator is binary and captures extreme value loss. In contrast, these acquirer-level measures, like CAR, are continuous, can take either positive or negative values, and may potentially capture nuanced outcomes. However, firm or market factors unrelated to the transaction may also impact these measures.

2.2.3 Transaction-level Ex-post Measure: Deal Withdrawal

If announcement returns reflect expected value creation from the transaction, managers should utilize this signal to continue or cancel the acquisition. Following the literature, we construct an indicator for whether the deal was completed or withdrawn (e.g., Asquith et al., 1983; Jennings and Mazzeo, 1991; Luo, 2005; Kau et al., 2008).[14] Unlike impairment and ROA, which are measured proxies of value creation, deal withdrawal is a realized outcome for which we observe both the occurrence and timing of withdrawal.

Earlier studies found mixed evidence about the correlation between CAR and withdrawal propensity. Jennings and Mazzeo (1991) find no such correlation, while Luo (2005) and Kau et al. (2008) find that deal withdrawal is more likely following negative CAR, particularly in

[14] We include only completed and withdrawn deals in this analysis as the outcome is often uncertain for deals that are not completed but not formally withdrawn.

settings where managerial "learning" is likely more important. We expand these studies by including a large sample of 39,585 transactions (of completed and withdrawn deals with non-missing control variables) in the period 1980 to 2018, and unlike previous studies (e.g., Luo, 2005), we include both public and private targets. We study the relation between CAR and withdrawal in- and out-of-sample and examine whether CAR fails to capture all information regarding withdrawal probability at announcement.[15]

The use of withdrawn/completed deals brings an additional benefit. We can now examine the dampening effects of "feedback" effects–managers withdraw the bid if CAR is too negative–on the core relation between CAR and outcomes. We discuss feedback in more detail in Section 3.

2.3 Descriptive Statistics of Outcome Measures

We report our sample summary statistics and correlations in Table 1. We winsorize all continuous variables at the 1% level to reduce the effect of outliers. Panel A shows summary statistics for our acquisition performance measures. On average, 85% of transactions do not experience firm-level impairment; of the subset of failed acquisitions, Appendix Table IA.A4 shows that 79% of impairments occur within the third year following the deal effective date, and the remaining 21% of impairments happen in the fourth and fifth year. The average acquirer has a small negative short-term abnormal ROA of -0.22% and long-term abnormal ROA of -0.21%. In our sample, 94% of transactions are completed and not withdrawn.[16] On average, announcement returns immediately surrounding the announcement are positive: the three-day and 11-day CARs range from 0.83% to 0.88%, whereas the return over the

[15] Of course, some deal cancellation decisions are outside the control of the acquiring firm's management (e.g., transactions canceled due to regulatory pressure or a failed vote by target shareholders). Jacobsen (2014) shows that of withdrawn deals, 14% are canceled due to regulatory or judicial obstacles, and 13% are canceled due to target shareholders blocking or voting against the deal.

[16] This completion rate is slightly higher than the rate reported in Luo (2005), Kau et al. (2008), and Ellahie, Hshieh, and Zhang (2021), who focus either on only public targets or on an earlier period. For example, in our sample, only 82% of public target deals are completed, whereas 97% of private target deals are completed. We find similar completion rates if we look at the sample period used in Luo (2005) and Kau et al. (2008), which is mostly in the 1990s.

announcement to deal-closing period (Deal CAR) is −1.12%[17].

We also measure characteristics-adjusted cumulative buy-and-hold monthly returns (DGTW-adjusted BHAR used by Daniel, Grinblatt, Titman, and Wermers, 1997), in line with the literature studying the long-term performance of acquirers (e.g., Mitchell and Stafford, 2000; Dong, Hirshleifer, Richardson, and Teoh, 2006; Ben-David, Drake, and Roulstone, 2015).[18] We report buy-and-hold returns over the 60 months beginning the month before the deal's announcement date. The average 60-month DGTW-adjusted buy-and-hold returns are −9.6%.

Table 1: **Descriptive Statistics of the Measures**

This table reports descriptive statistics of the ex-ante and ex-post measures of acquisition quality. Panel A presents summary statistics, and Panel B shows correlations. All continuous variables are winsorized at the 1% level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Summary Statistics

	N	Mean	Std Dev	P10	P50	P90
Non-impairment	6,128	0.852	0.355	0	1	1
Short-term abnormal ROA (ST abROA)	28,710	−0.002	0.082	−0.078	−0.006	0.091
Long-term abnormal ROA (LT abROA)	22,577	−0.002	0.091	−0.086	−0.007	0.097
Completion	39,585	0.944	0.230	1	1	1
CAR$[-1,1]$	28,710	0.009	0.066	−0.056	0.003	0.080
CAR$[-5,5]$	28,710	0.008	0.100	−0.099	0.002	0.122
DealCAR$[\mathrm{A}-2,\mathrm{C}+2]$	28,710	−0.011	0.199	−0.218	−0.001	0.183
DGTW-adjusted BHAR (Adj BHAR)	27,355	−0.096	1.225	−1.289	−0.293	1.214

Panel B: Correlations

	Non-impair	ST abROA	LT abROA	Comp-letion	CAR $[-1,1]$	CAR $[-5,5]$	DealCAR $[A-2,C+2]$	Adj BHAR
Non-impairment	1							
ST abROA	0.133***	1						
LT abROA	0.127***	0.671***	1					
Completion	—	—	—	1				
CAR$[-1,1]$	−0.004	0.003	−0.010	0.015***	1			
CAR$[-5,5]$	0.002	−0.002	−0.005	0.015***	0.628***	1		
DealCAR$[\mathrm{A}-2,\mathrm{C}+2]$	0.047***	0.020***	0.009	—	0.354***	0.398***	1	
Adj BHAR	0.252***	0.221***	0.257***	0.018***	0.040***	0.070***	0.121***	1

Since expected value creation/destruction from a particular deal is not observable, we

[17]Our 3-day CAR estimate is similar to Betton et al. (2008)who document a mean 3-day CAR of 0.73%.

[18]The DGTW adjustment procedure involves adjusting returns by the returns of benchmark portfolios based on characteristics. We form $5 \times 5 \times 5$ portfolios each month based on size, the book-to-market ratio, and 12-month past returns.

utilize the above multiple outcome measures derived from different sources (each with its strengths and weaknesses), and importantly, we find they are correlated. Table 1, Panel B shows that the correlation coefficients across the four ex-post outcome variables range between 0.127 and 0.671. These correlations dramatically exceed correlations with CAR for each outcome variable; correlation coefficients with 3-day and 11-day CAR range from −0.010 to 0.015. Correlations between outcomes and Deal CAR, measured over a longer window, range from 0.009 to 0.047. We also find that our outcome variables are correlated with long-term returns: correlations between DGTW-adjusted BHAR and non-impairment, and the two ROA measures range from 0.221 and 0.257. In contrast, the correlation between DGTW-adjusted BHAR and 3-day and 11-day CAR ranges from 0.040 to 0.070, and the correlation between DGTW-adjusted BHAR and Deal CAR is 0.121. Indeed, correlations between the three CAR definitions are also not strong in a relative sense: the correlation between three-day CAR and CAR from announcement to close (Deal CAR) is only 0.354, whereas the correlation between short- and long-term abnormal ROA is 0.671.

To summarize, across four ex-post acquisition outcome measures—transaction-level impairment, short-term and long-term abnormal ROA, and deal completion probability—we observe significant correlations but only very weak correlations between CAR and these acquisition outcome measures. We now turn to formal tests of these correlations.

3 Predicting Acquisition Outcomes

In this section, we test the ability of announcement returns to capture acquisition value creation by relating CAR to the observable ex-post measures described in Section 2. We follow a multipronged approach to test our null hypothesis that CAR measures NPV, i.e., the net present value of cash flows arising from the acquisition.

We first test the correlation between CAR and realized acquisition outcomes (non-impairment, short- and long-term abnormal ROA, and completion). Because the measures capture realized rather than “expected” outcomes and because there is no clear guide for the level of correlation that deems CAR an adequate measure of NPV, we construct a simple benchmark measure using data also available at the time of the acquisition announcement, based on the standard set of deal and acquirer characteristics used in previous studies. We measure the forecasting ability of CAR and the benchmark characteristics model in-sample and out-of-sample by decade, industry, deal, and firm type.

In the second set of tests, we relate CAR to other “ex-ante” outcomes. Because our benchmark characteristics model correlates reasonably well with ex-post outcomes, we consider whether CAR captures other value-relevant information known at the time of the announcement by relating CAR to predicted outcomes based on the benchmark model.

Another way to check whether value-relevant information exists at the time of the announcement and whether this information is incorporated in CAR is to construct a trading strategy. We estimate a trading strategy that buys acquirers that CAR expects to do better (i.e., CAR predicts more favorable acquisition outcomes) and sells acquirers that CAR expects to do worse. We replicate this test using the benchmark measure. Further, since some acquisitions are canceled rather than completed, we discuss selection and feedback effects. We consider whether “listening” to CAR (completing deals with positive CAR and withdrawing deals with negative CAR) creates an abnormal return spread relative to “not listening” to CAR. We again replicate this test using the benchmark measure.

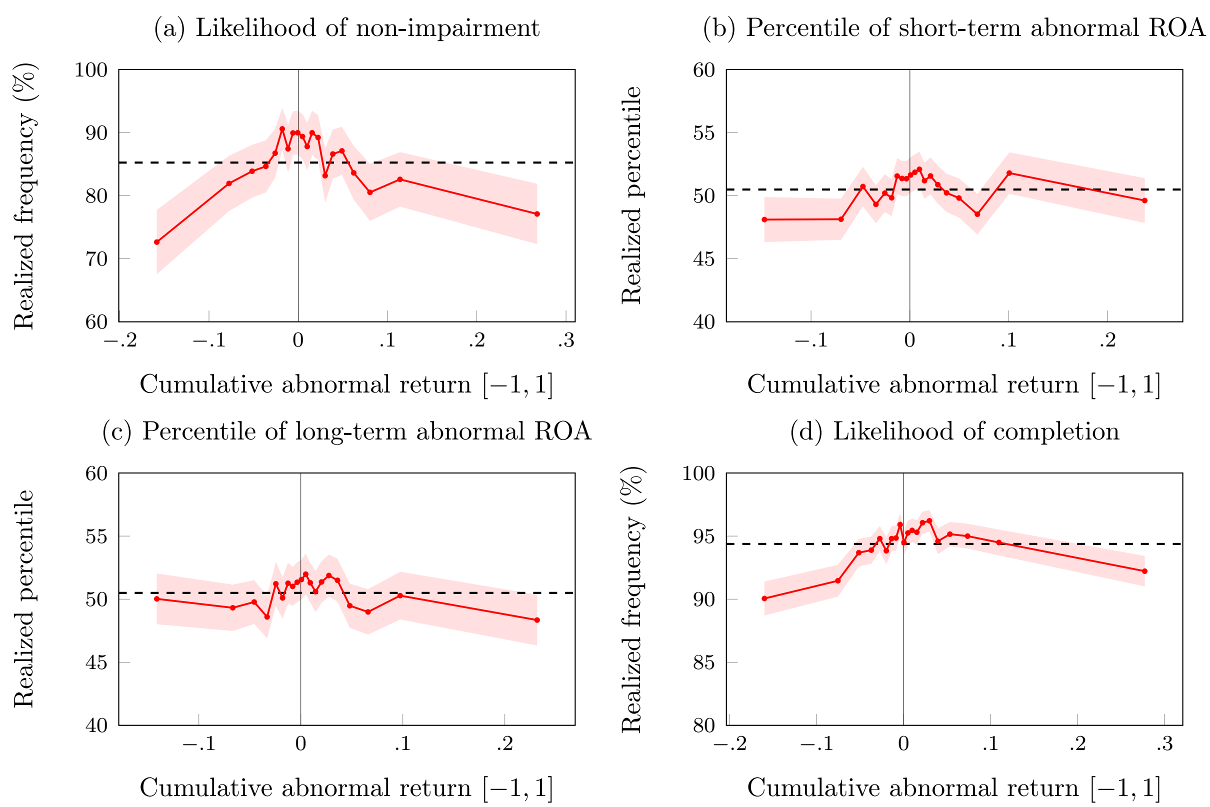
Our final tests consider inferences generated by CAR relative to those generated by our benchmark measure on the “types” of deals associated with ex-post value creation. We consider whether the types of transactions (e.g., cash, private target) predicted to create or destroy value by CAR, align with the types of transactions that do well ex-post and the types of transactions predicted to do well by our benchmark measure.

3.1 Visual Examination

We begin by examining the unconditional relation between transaction- and acquirer-level outcomes and CAR. The implicit assumption behind using CAR to estimate value creation is that CAR is positively correlated with ex-post outcomes. The results of the visual examination are presented in Figure 1. We sort CAR$[-1, 1]$ into 20 equally-sized bins and present the related outcome statistics. In Panel (a), we present the fraction of transactions without transaction-level impairment. The panel shows little correlation between the realized likelihood of non-impairment and CAR: impairment outcomes vary little across CAR vigintiles.

Figure 1: **CAR $[-1, 1]$ and Ex-post Outcomes**

Observations are sorted into 20 equally-sized bins based on their CAR$[-1, 1]$. Panel (a) plots the percentage of transactions without impairment for each acquirer's CAR$[-1, 1]$ vigintile (the solid red line). The horizontal black dashed line represents the unconditional likelihood of not failing in our sample. Panels (b), (c), and (d) present the average realized percentile of short-term abnormal ROA, long-term abnormal ROA, and the likelihood of completion, respectively, for each vigintile of CAR. The light red shading indicates 95% confidence intervals.



Panels (b) and (c) show firm-level outcomes related to ROA. Panel (b) presents the relation between the average realized percentile of short-term abnormal ROA (percentiles within the sample) and CAR vigintiles, and Panel (c) shows the relation between long-term abnormal ROA and CAR vigintiles. Neither chart shows any meaningful correlation between firm-level outcomes and CAR. Panel (d) presents the relation between the fraction of transactions completed (rather than withdrawn) and CAR. For the lowest CAR bins, completion rates are particularly low. However, for the remaining bins, there is little relation between completion rates and CAR, and even a reduction in completion rates for the very highest CAR bins.

Overall, a first visual pass reveals no meaningful association between transaction- and acquirer-level outcomes and CAR.

3.2 In-sample Tests: CAR versus Characteristics

Next, we explore the correlation between the various outcome variables and CAR in a regression framework. Table 2 reports regressions with acquisition outcome measures as the dependent variables and acquirer CARs over multiple windows surrounding the deal announcement as the key independent variables of interest. Panel A reports the results of ordinary least squares (OLS) regressions that model the probability of no goodwill impairment within five years of the deal's effective date. Panels B and C report the results of OLS regressions with short- and long-term abnormal ROA as the dependent variable, respectively, and Panel D reports the results of OLS regressions with the probability of completion as the dependent variable. Some regressions include the following acquirer and deal characteristics as controls: the log of market capitalization, leverage, free cash flow scaled by previous-year assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals.

Columns (1)–(3) of the first three panels of Table 2 show that the explanatory power of

CAR is minuscule: the highest R^2 is 0.20%. For non-impairment, short-term and long-term abnormal ROA outcomes, 3-day and 11-day CARs are never statistically significant. CAR is statistically significant in two regressions when a longer CAR window is used, but the explanatory power is weak. Since the mean time to close a deal is 74 days, CAR measured over announcement to completion, on average, includes more than 10 weeks for which other information (related to ROA or impairment likelihood) may be released.[19]

In Column (4), we add characteristics known at the time of the announcement, and in Column (5), we further saturate the model with year and industry-fixed effects. Again, in these two columns, across all six regressions in Panels A–C, CAR is statistically significant at the 10% level (and the correct sign) in only one regression. With the inclusion of controls, the explanatory power of the dependent variables increases. For example, for short-term abnormal ROA, the R^2 increases from 0.00% (Column (1)) to 2.6% when characteristics are added (Column (4)) to 7.8% when characteristics and year and industry fixed effects are added (Column (5)). Columns (6) and (7) show similar R^2s of 2.6% and 7.8% when CAR is not included in the regression, indicating the explanatory power comes entirely from the controls and not from CAR.[20]

Panel D of Table 2 reports the results of regressions of deal completion on acquirer CAR. Similar to the results reported in Kau et al. (2008) and Luo (2005), we find that CAR correlates with completion outcomes: the coefficient is the correct sign and is significant at the 10% for the majority of the specifications. The results indicate that some managers respond to signals generated from CAR. However, CAR has little economic significance:

[19]Part of the information about the expected value created by the acquisition may already be impounded in the price before the announcement due to leakage or anticipation of the acquisition (e.g., Schwert, 1996; Bhattacharya, Daouk, Jorgenson, and Kehr, 2000; Mitchell et al., 2004; Betton et al., 2008; Edmans, Goldstein, and Jiang, 2012; Offenberg and Officer, 2012; Wang, 2018; Bennett and Dam, 2019; Irani, 2020). In Internet Appendix Table IA.B1, we follow (Schipper and Thompson, 1983b) and extend the measurement period of CAR to begin 41 days before the announcement and end one day following the announcement. The results show that extending the window does not change our inference about CAR's lack of predictive ability.

[20]In Internet Appendix Table IA.B2, we show that the results are robust to using two alternative definitions of non-impairment that classify transactions that lack information as non-impaired (Panel A) or impaired (Panel B) and to using industry-adjusted ROA rather than abnormal ROA (Panel C).

Table 2: **Acquirer CAR and Acquisition Outcomes**

This table reports results from regressions of acquisition outcome measures on acquirer cumulative abnormal returns (CAR). The dependent variable is a non-impairment dummy (Panel A), short-term abnormal ROA (Panel B), long-term abnormal ROA (Panel C), and a completion dummy (Panel D). In Columns (1)–(3), CAR is the only independent variable. In addition to CAR, Column (4) includes characteristics, and Column (5) includes year and industry fixed effects, as well as characteristics as independent variables. Column (6) only includes characteristics, and Column (7) includes year and industry fixed effects, and characteristics as independent variables. The characteristics used in the controls include the log of market capitalization, leverage and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

CAR window:	$[-1,1]$	$[-5,5]$	$[\mathrm{A}-2,\mathrm{C}+2]$	$[-1,1]$		n.a.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Probability of Non-impairment ($N = 6,128$)							
Dependent variable:	Non-impairment Dummy						
CAR	−0.020	−0.008	0.089**	0.082	0.080	Controls	Controls
	(0.105)	(0.081)	(0.040)	(0.083)	(0.089)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.002	0.036	0.088	0.036	0.088
Panel B: Short-term Abnormal ROA ($N = 28,710$)							
Dependent variable:	Short-term Abnormal ROA						
CAR	0.004	−0.001	0.009**	0.019	0.019*	Controls	Controls
	(0.012)	(0.006)	(0.003)	(0.011)	(0.010)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.000	0.026	0.078	0.026	0.078
Panel C: Long-term Abnormal ROA ($N = 22,577$)							
Dependent variable:	Long-term Abnormal ROA						
CAR	−0.015	−0.004	0.004	0.001	−0.002	Controls	Controls
	(0.014)	(0.007)	(0.003)	(0.012)	(0.012)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.000	0.034	0.109	0.034	0.109
Panel D: Probability of Completion ($N = 39,585$)							
Dependent variable:	Completion Dummy						
CAR	0.048*	0.031*	—	0.050**	0.041**	Controls	Controls
	(0.023)	(0.014)		(0.019)	(0.016)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	—	0.147	0.153	0.147	0.153

taking Column (4), which has the largest point estimate in the panel, as the example, for every one percentage point reduction in CAR, the probability of withdrawal increases by 0.05%. Compared with the 94.4% unconditional probability of completion for the sample, the point estimate represents less than a one-basis-point increase. [21]

To summarize, our in-sample tests show that CAR is uninformative about acquisition outcomes. In most specifications, the relation between CAR and the acquisition outcome is not statistically different from zero. When CAR is statistically significant, the magnitude of the variation explained is minute.

In addition to announcement returns, we also consider the predictive properties of other ex-ante measures, specifically, deal and firm characteristics that are also known at the time of the transaction announcement. These ex-ante characteristics serve as a relative benchmark that allows us to better gauge CAR's performance in predicting acquisition performance realizations.[22]

We now discuss the in-sample explanatory power of these deal and acquirer characteristics. Table 2, Column (7), shows that year and industry controls and deal and firm characteristics alone can explain 8.8% of the variation in non-impairment (Panel A), 7.8% and 10.9% of the variation in short- and long-term abnormal ROA (Panels B and C), and 15.3% of the variation in completion rates (Panel D). In contrast, CAR, at best, explains 0.20% of the variation (Columns (1)–(3) across all four panels). Notice also that when comparing the R^2 of Column (7) to Column (5) in all four panels, CAR increases the explanatory power of outcomes in a model with industry controls and deal and firm characteristics by (at most) a meager 0.01%.

To summarize, if the market reaction to the announcement provides additional informa-

[21] In Panel D of Internet Appendix Table IA.B2, we include deals that still may be pending or for which the outcome is unknown as the outcome variable, and we find the relation between CAR and completion is no longer statistically significant.

[22] We consider the standard characteristics utilized in the M&A literature: the logarithm of market capitalization, leverage, free cash flow scaled by previous-year assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and indicators for stock-only consideration, mixed payment, diversifying acquisition, hostile deal, competing bidders, and public targets.

tion related to deal value creation over and above the information contained in the deal and firm characteristics, then the CAR-alone model should perform well (Columns (1)–(3) in all panels in Table 2). It does not. In addition, a model that combines CAR and the characteristics model (Column (5)) should significantly outperform the characteristics-only model (Column (7)). It does not. We conclude that deal and firm characteristics, also known at the time of the deal announcement date, dominate CAR as predictors.

We also test whether CAR is better at predicting short-term or long-run outcomes, e.g., impairment within the first year as opposed to within five years, and ROA the year following the completion date versus ROA five years following the completion date. We rerun the earlier regressions (as in Table 2) but define the dependent variable as the outcome within a particular period relative to the deal's effective date (up to five years) for non-impairment and ROA, and relative to the announcement date for completion. In Internet Appendix Figure IA.B1, we plot the coefficients on CAR (Panels (a), (c), and (e)) and the adjusted R^2 (Panels (b), (d), and (f)). In addition, to provide a benchmark, we add to the latter set of panels the R^2 from the standard regression of deal and acquirer characteristics (without industry or year fixed effects).

Panels (a) and (c) of Internet Appendix Figure IA.B1 show that CAR performs better on some short-term outcomes. Specifically, CAR's coefficient is statistically significant when considering one-year non-impairment and one- and two-year ROA. Yet the practical impact of the coefficients remains minimal: a one-standard-deviation shift in CAR (7.2% and 6.6% for Panels (a) and (c), respectively) correlates with a minute increase in the short-term probability of non-impairment by 0.02 standard deviations, and a similar increase in ROA of 0.03 standard deviations. Furthermore, the R^2s in these regressions are virtually zero. In sharp contrast, deal and acquirer characteristics known at the time of the announcement produce R^2s of 4.2% to 6.2%.

These results show that even though CAR performs better for short-term outcomes, it still is an ineffective predictor of value creation. In particular, CAR's meager economic significance and lack of explanatory power for short-term outcomes render it an uninformative indicator of value creation, relative to the explanatory power of standard deal and acquirer characteristics available at the time of the announcement.

3.3 In-sample Tests: By Subsample

Given that CAR has no material explanatory power over outcomes in the universe of acquisition announcements, we try to find the "golden subset," i.e., a subsample in which CAR has a stronger correlation with acquisition outcomes.

3.3.1 Subsamples by Decades

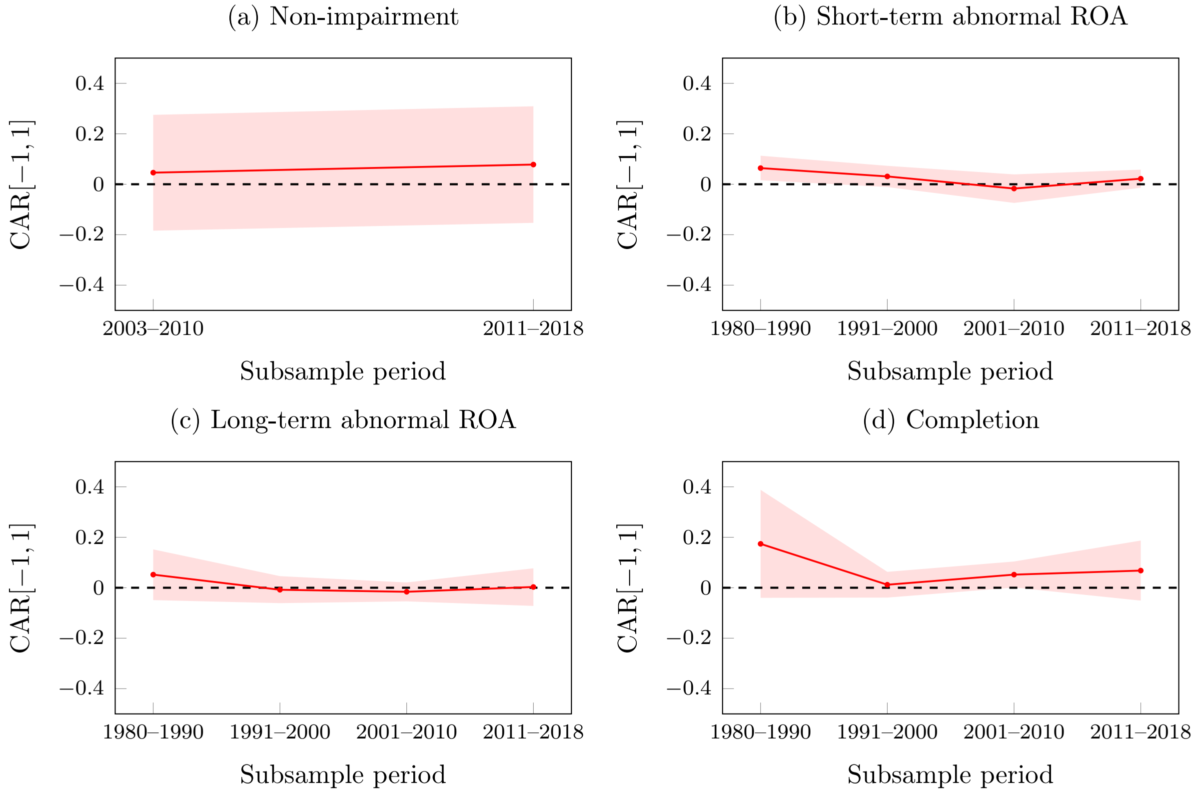
Figure 2, Panels (a)–(d) show the coefficient and 95% confidence intervals for regressions of outcomes on CAR based on the specification in Table 2, Column (5), for each of the four decades in our sample. Panel (a) shows that the coefficient on CAR in regressions of non-impairment on CAR is insignificant in the 2003–2010 and 2011–2018 periods. Panel (b) shows that when short-term abnormal ROA is the outcome variable, CAR is significant (and the correct sign) at the 5% level for the 1980–1990 period; however, it is not statistically significant (and in some periods has the wrong sign) in the 1991–2000, 2001–2010, and 2011–2018 periods. Panel (c) shows that when long-term abnormal ROA is the outcome variable, CAR is not statistically significant (and in two periods has the wrong sign) for all four subperiods. Similarly, Panel (d) shows that CAR correlates with completion for only one of the four subperiods. This result contrasts with the statistically significant (at the 5% level) and positive (but economically weak) relation between CAR and completion reported in Table 2.

3.3.2 Subsamples by Industries

We also split the sample by industries. Internet Appendix Figure IA.B2 replicates Table 2, Column (5), by Fama French 12 industry classifications and reports the coefficients on CAR

Figure 2: **Acquirer CAR and Acquisition Outcomes: By Decade**

This figure plots the coefficients and 95% confidence intervals for regressions of outcomes on CAR based on the specification in Table 2, Column (5), for each of the four decades in our sample (except for impairment, which we can only determine for two decades due to data limitations). Panels (a), (b), (c), and (d) use non-impairment, short-term abnormal ROA, long-term abnormal ROA, and completion, respectively, as the key independent variables. The red dots represent the point estimates, and the light red shading represents 95% confidence intervals.



and 95% confidence intervals. Across 48 regressions (4 outcome variables $\times$ 12 industries), the coefficient on CAR is the correct sign and statistically significant at the 5% level for only four regressions. Although CAR correlates with some outcomes in a few select industries, importantly, for these select industries, CAR does not correlate with all outcomes. Like the time period results, the correlation between CAR and completion rates (Panel (d)) is the correct sign and statistically significant for only one of the 12 industries. Again, this result indicates that the relationship between the withdrawal decision and CAR is economically weak.

3.3.3 Subsamples by Deal and Acquirer Characteristics

We further consider whether a particular set of deal or acquirer firm characteristics drive the lack of relation between outcomes and CAR. For example, the existing literature has discussed anticipation (e.g., serial acquirers), new information on acquiring firm valuation (e.g., stock or diversifying deals), difficulty in assessing value creation due to lack of information (e.g., private and high-tech targets, and small deals), or price pressure from merger arbitrageurs (e.g., public targets) as potential explanations for the lack of relation.

In Internet Appendix Table IA.B3, we replicate Table 2, Column (5), for 29 subsamples based on the deal and acquirer characteristics used in Table 2, Column (5), as well as indicators for serial acquirers and high-tech targets.[23] Across 116 regressions (4 outcome variables $\times$ 29 subsamples), the coefficient on CAR is the correct sign and statistically significant at the 5% level for only 17 regressions. Of more importance is whether CAR's performance improves systematically in particular subsamples: in only three subsamples does CAR achieve statistical significance for two of the four outcome variables, and in no subsample does CAR achieve statistical significance for three or more outcome variables. Overall, the results indicate, to the extent that characteristics correlate with potential explanations for CAR's lack of explanatory power, these particular subsamples do not drive the result.[24]

[23]The deal characteristics we consider are the target's public status, form of payment (stock, cash, mix), diversifying, competitive, hostile, relative size, and high-tech. The acquirer characteristics we consider are serial acquirer, market capitalization, Tobin's Q, past returns, free cash flow, and leverage, defined using either an indicator variable or above/below median values. We include serial acquirers who made more than one deal in a five-year window to capture potential anticipation. We include a high-tech indicator as Luo (2005) finds that the relation between CAR and completion is related to high-tech industry classification. We obtain the high-tech indicator from SDC.

[24]Goodwill impairment tests are performed at the reporting unit level. When several targets operate under a single reporting unit, operating performance improvements by one target may obscure the poor operating performance of the failed acquisition, thereby stalling goodwill impairment. For acquisitions that are large relative to the acquirer's size, it is less likely that other businesses can hide value reductions in the target. In Internet Appendix Table IA.B3, the coefficient on CAR remains insignificant when we zoom in on the sample of acquisitions that are relatively large in size. Further, we focus on extreme impairments, and such large value destruction is difficult to mask.

3.3.4 Brute Force Subsamples by Characteristic Combinations

Particular combinations of deal or acquirer firm characteristics may drive the lack of relation between outcomes and CAR. We, therefore, allow the interaction of characteristics. Following the same approach as in the previous subsection, we create the following 10 indicator variables based on the characteristics: the log of market capitalization, leverage, free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, cash-payment, diversifying, serial, and public target deals. If the characteristic is continuous, we create the indicator variable by splitting the sample at the median. We then form subsamples based on all of the unique interactions of these variables and retain subsamples with at least 30 observations. We then split the sample into two time periods, and, for each subsample and time period, we regress outcomes on CAR$[-1, 1]$ and record the corresponding t-statistic. The results are reported in Internet Appendix Table IA.B4. We report the number of transactions with a t-statistic greater than two, less than minus two, or with an absolute value less than two for both periods. Taking Panel A as an example, for non-impairment, we run 22,298 regressions and find that only 5% of transactions (1,091/22,298) have the correct sign and a t-static of at least two in the first period, and only 3% (735/22,298) do in the second period. Furthermore, only 0.26% (59/22,298) have the correct sign and statistical significance in *both* periods. We draw similar conclusions using the other three outcome variables.

To summarize, even in an extensive data search, we cannot locate a sample for which CAR consistently captures outcomes. We conclude that the lack of relation between outcomes and CAR is systematic and not driven by a particular time period (e.g., the financial crisis), industry, or combination of deal and acquirer characteristics.

3.4 Out-of-sample Tests: CAR versus Characteristics

Next, we compare the ability of CAR versus characteristics-based models to predict deal and acquirer outcomes in out-of-sample settings.

To conduct our out-of-sample tests, we use the following approach. We estimate a CAR-only OLS regression model that uses our transaction- and firm-level ex-post outcome measures as the dependent variable and CAR as the independent variable. We also estimate a characteristics-only OLS regression model that uses our transaction- and firm-level ex-post outcome measures as the dependent variable and the characteristics used in Column (7) of Table 2 as independent variables. (Note that we do not include industry and year controls.) For regressions with non-impairment as the outcome variable, we use the first half of the sample, 2003–2010, as a fit period to estimate coefficients. Then, we use the parameter estimates from this first period to predict outcomes in the second half of the sample, 2011–2018 (i.e., the imputed probability of transaction impairment within five years of the deal's effective date). For regressions with ROA and completion as the outcome variable, we use the first half of the sample, 1980–2000, as a fit period to estimate coefficients. Then, we use the parameter estimates from this first period to predict outcomes in the second half of the sample, 2001–2018. Our analysis examines the ability of characteristics and CAR to predict outcomes in the second period. This is an out-of-sample test because this later period was not used to estimate the model's parameters.

We then compare the quality of the predictions made by CAR and the characteristics-based model out-of-sample. We present the results in Table 3. Panel A shows that the predicted outcome by CAR is not correlated with the realizations of non-impairment, short-term abnormal ROA, or long-term abnormal ROA (Columns (1), (3), and (5)). In contrast, the predicted outcome by the characteristics-based model is positive (correct direction) and significant at the 1% confidence level for all three outcomes (Columns (2), (4), and (6)).

Similar to the results reported in Table 2, CAR predicts completion outcomes better than impairment and ROA outcomes. The coefficient on CAR in Column (7) is the correct sign and is significant at the 5% confidence level. However, the result is economically weak, with an R^2 of 0.03%. When the probability of completion predicted by CAR goes from the 25^{th} percentile to the 75^{th} percentile, the likelihood of completion increases by 0.3%. The

Table 3: **Out-of-sample: Predicted versus Realized Outcomes**

We first estimate OLS regressions of deal outcome measures on CAR$[-1, 1]$ only and characteristics only using only the first half of transactions in each sample as a fit period to estimate coefficients. We then use the parameter estimates from this first half to predict outcomes in the second half of each sample. In Panel A, we assess the correlation between realized outcomes and predicted outcomes produced by the CAR-only model (Columns (1), (3), (5), (7)) and the characteristics-only model (Columns (2), (4), (6), (8)). In Panel B, we assess the correlation between the predicted outcome by the characteristics-only model and acquirer CAR. Standard errors are reported in parentheses below the coefficients. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Predicted versus Realized Outcomes

Dependent variable:	Realized Outcome							
	Non-impair		ST abROA		LT abROA		Completion	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted based on CAR	26.039		−0.855		6.573		1.090**	
	(30.177)		(1.031)		(4.985)		(0.404)	
Predicted based on characteristics		0.708***		1.338***		1.188***		0.995***
		(0.138)		(0.368)		(0.228)		(0.059)
Observations	2,862	2,862	14,358	14,358	10,713	10,713	18,014	18,014
Adjusted R^2	0.000	0.027	0.000	0.039	0.000	0.041	0.000	0.148

Panel B: Is CAR Correlated with the "Predictable" Component of Outcomes?

Dependent variable:	Predicted Outcome by a Characteristics Model							
	Non-impair		ST abROA		LT abROA		Completion	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CAR$[-1, 1]$	−0.104***		−0.014***		−0.010***		−0.008	
	(0.005)		(0.001)		(0.003)		(0.017)	
CAR$[-5, 5]$		−0.078***		−0.009***		−0.008***		−0.011
		(0.009)		(0.002)		(0.003)		(0.008)
Observations	2,862	2,862	14,358	14,358	10,713	10,713	18,014	18,014
Adjusted R^2	0.010	0.009	0.005	0.006	0.001	0.002	0.000	0.000

coefficient on the prediction based on characteristics is statistically significant at the 1% level, achieving an R^2 of 14.8%. When the predicted probability of completion by characteristics goes from the 25th percentile to the 75th percentile, the likelihood of completion increases by 6.0%.

Our analysis so far has identified a set of characteristics that are useful in predicting acquisition outcomes out-of-sample. When acquisitions are announced, is the announcement CAR correlated with the out-of-sample characteristics-based prediction (which we already know is reliable)? We investigate this issue in Panel B of Table 3, which reports results for regressions of the predicted outcome by the characteristics-only model on acquirer CAR. Results show that acquirer CAR in the later sample is either not correlated with the *predictable* part of acquisition outcomes (Columns (7)–(8)) or has the wrong sign (Columns (1)–(6)).

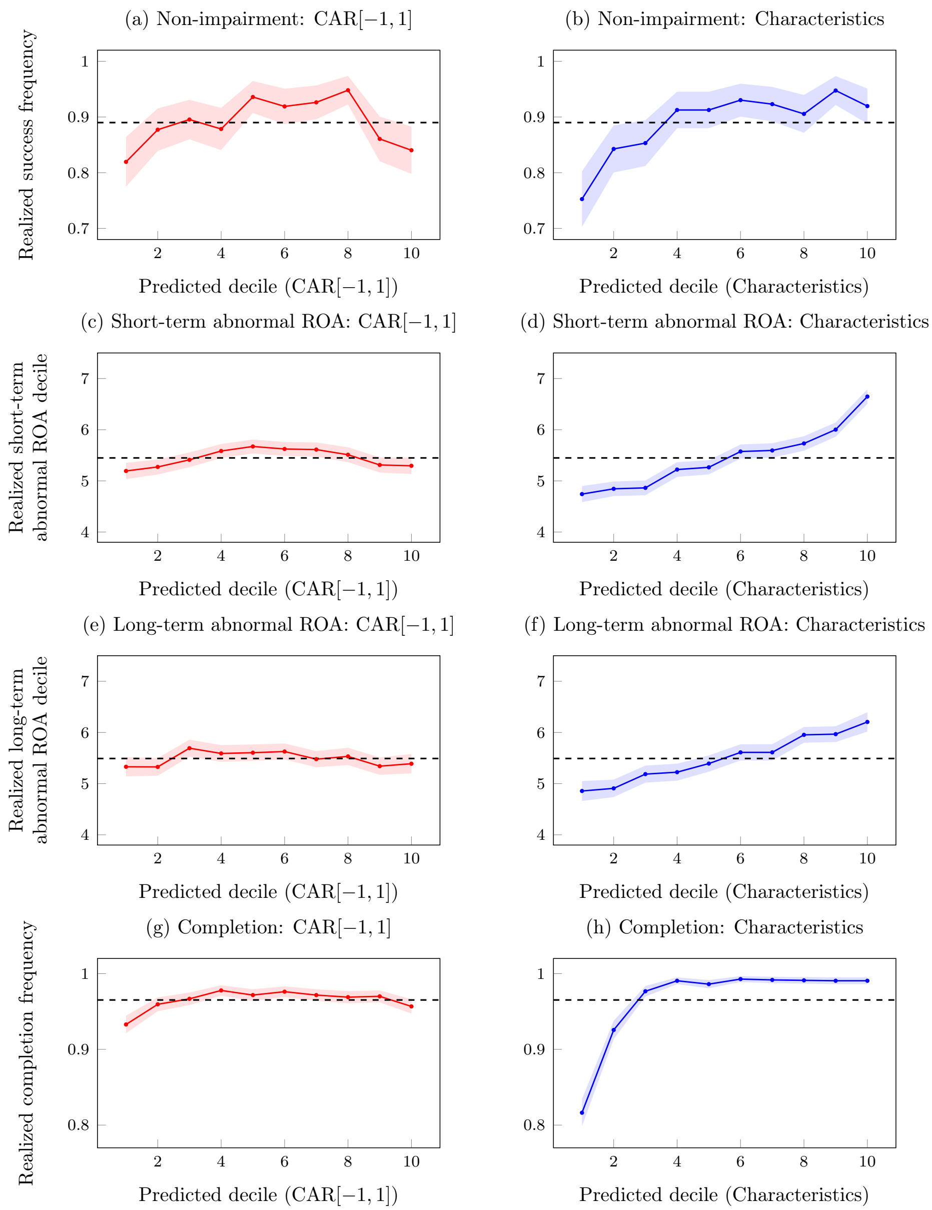
In Figure 3, we present out-of-sample tests graphically, similar in spirit to the tests reported in Table 3. We estimate OLS outcome models on CAR or characteristics for the transaction-level failure measures (non-impairment and completion). We then use the coefficients estimated in the first half of the sample to estimate the predicted probability *decile* in the second half of the sample. Then, we report the fraction of transactions with non-impairment and deal completion rates for each predicted probability decile. Similarly, for the abnormal ROA outcome variables, we estimate OLS models of outcomes on CAR or characteristics. We again use the coefficients estimated in the first half of the sample to estimate the predicted outcome *decile* in the second half of the sample. Then, we report the realized outcome decile for each predicted outcome decile.

Focusing first on non-impairment, if the model has predictive power, then the realized non-impairment rate should increase monotonically as we move from decile 1 (low predicted probability) to decile 10 (high predicted probability). Alternatively, if the model lacks predictive power, the realized non-impairment rate should be close to 90.9% (the unconditional non-impairment rate in the second half of the sample) for all deciles. In Panel (a), we see little evidence of significant predictive power for the CAR-only model. The realized non-impairment rate is non-monotonic as we move from decile 1 to 10. Moreover, realized non-impairment rates are close to 90.9% for many deciles, and there are reduced non-impairment rates for the highest CAR deciles. In contrast, Panel (b), the characteristics-only model, shows a stable positive upward trend, indicating that deciles with higher predicted non-impairment are associated with a higher fraction of realized non-impairment rates.

The results for the firm-level ROA outcome variables are generally similar. In Panels (c) and (e)—the CAR-only model—realized outcome deciles vary little from the unconditional

Figure 3: **Out-of-sample: Predicted versus Realized Outcomes**

These figures report out-of-sample results. We use the first half of each sample to fit OLS regressions of non-impairment, short- and long-term abnormal ROA, and deal completion as outcome variables. Panels (a), (c), (e), and (g) include only acquirer CAR$[-1, 1]$ as an independent variable. Panels (b), (d), (f), and (h) include only deal and firm characteristics as the independent variables. Using the estimates, we obtain predicted outcome deciles for the second half of each sample. For our transaction-level measures, for each predicted probability decile, we report the fraction of transactions with realized success or completion. We report the realized outcome decile for our firm-level measures for each predicted decile. The dashed line indicates the unconditional realized frequency (for success and completion) and the unconditional realized outcome decile (for short- and long-term abnormal ROA) for the second half of each sample. The shaded portion represents the 95% confidence interval.



average decile in the second half of the sample (as indicated by the dashed line) across predicted outcome deciles. In contrast, Panels (d) and (f)—the characteristics-only model—show an upward trend in realized outcome deciles as we move from low predicted to high predicted deciles.

In terms of completion (Panels (g) and (h)), for the characteristics-only model, realized completion is 99.0% for the highest predicted completion decile, and 81.6% for the lowest; for the CAR-only model realized completion for the highest decile is 95.7% and for the lowest decile is 93.3%.

In sum, the out-of-sample tests reiterate the conclusion from the earlier in-sample tests: CAR has only very weak predictive power about acquisition outcomes, whereas acquisition outcomes can be predicted relatively well by characteristics known at the time of the announcement. These results relate to Ellahie, Hshieh, and Zhang (2022), who develop an implied return-on-equity improvement measure (IRI) serving as an ex-ante acquisition quality index. Their measure is constructed using multiple aspects of the transaction, including deal, target, and acquirer information, and it incorporates inputs generated from stock, accounting, and transaction-related data. The authors find that acquirers with high IRI (created from characteristics) are associated with worse accounting and stock outcomes in the years following the acquisition.

3.5 Long-term Stock Returns: CAR versus Characteristics

We further substantiate our conclusion about CAR's lack of predictive ability by relating outcomes predicted by CAR and outcomes predicted by characteristics to long-term returns (DGTW-adjusted BHAR) following the acquisition announcement. For each announcement year in our sample, we estimate OLS regressions of the probability of non-impairment, short-term and long-term abnormal ROA, and the likelihood of completion, respectively, on CAR or characteristics. Using each transaction's imputed outcome, we sort predicted values into 10 outcome deciles based on the announcement year. We then formulate a trading strategy in which we buy the top 30% of acquirers based on the predicted outcome and sell the bottom 30% of acquirers. The positions are held for 60 months starting the month before the announcement date. The trading strategy is not implementable as it requires perfect foresight of CAR and deal characteristics the month before the deal announcement. Still, as a look-back strategy, it is informative about CAR's ability to capture long-term returns.

Table 4: **Trading Strategy Based on CAR and Characteristics**

This table reports 60-month equal-weighted DGTW-adjusted portfolio returns computed beginning the month-end of the deal announcement date. In Columns (1), (3), (5), and (7), we estimate yearly OLS regressions of the probability of impairment, short-term and long-term abnormal ROA, and the probability of completion, respectively, on CAR$[-1, 1]$. We then compute the imputed outcome for each year, and sort predicted values into 10 outcome deciles. We report the equal-weighted 60-month DGTW-adjusted BHAR for acquirers in the bottom-three and top-three deciles and the p-value for the difference test between the two portfolios. Columns (2), (4), (6), and (8) are computed analogously, except we use the characteristics model to predict outcomes.

Predicted variable:	Non-impair		ST abROA		LT abROA		Completion	
Prediction model:	CAR	Char	CAR	Char	CAR	Char	CAR	Char
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	60-month DGTW-adjusted BHAR							
Buy top 3 deciles	−6.9%	−2.7%	−10.9%	−5.1%	−8.9%	−3.8%	−10.9%	−5.0%
Sell bottom 3 deciles	−6.9%	−12.0%	−13.7%	−15.8%	−8.1%	−13.2%	−13.7%	−12.9%
Difference	0.1%	9.2%	2.8%	10.7%	−0.8%	9.4%	2.8%	7.9%
p-value	0.986	0.005	0.158	0.000	0.726	0.000	0.089	0.000

We summarize the results of these trading strategies in Table 4. For example, Column (1) shows that buying a portfolio with the highest predicted non-impairment likelihood by CAR (top-three deciles) yields abnormal returns of −6.9% over five years. The portfolio with the lowest predicted non-impairment likelihood by CAR (bottom-three deciles) yields similar abnormal returns of −6.9%. These two abnormal returns are not statistically different. In contrast, in the characteristics model in Column (2), the portfolio with the highest predicted non-impairment likelihood yields −2.7% after five years, and the portfolio based on the lowest likelihood yields −12.0%. The performance difference between these portfolios is 9.2% and is statistically different at the 1% level.

Columns (3) and (5) show similar results for the short- and long-term abnormal ROA measures. The difference between buying the top three and selling the bottom three deciles predicted by CAR yields the wrong sign on one (Column (5)) and a positive but not statistically significant difference in the other (Column (3)). In Columns (4) and (6), the difference between the top and bottom three deciles predicted by characteristics are 10.7% and 9.4%, respectively, and both are statistically different at the 1% level.

We next focus on completion outcomes. We first estimate outcomes on CAR and characteristics for each announcement year to compute imputed outcomes. We sort predicted values into 10 outcome deciles based on the announcement year, then retain only completed deals. Like other columns, we report returns for the top and bottom three deciles. Since this analysis is conditional on completion, the bottom three deciles can be interpreted as those deals that are completed despite signals from CAR (characteristics) that the deal is creating less value relative to other deals announced in the same year.[25]

In Column (7), the difference in the DGTW-adjusted returns from buying and selling the top three deciles predicted by CAR is positive (2.8%), and weakly significant at the 10% level. However, in Column (8), the difference in the DGTW-adjusted returns from buying and selling the top three deciles predicted by characteristics is large and positive (7.9%) and statistically significant at the 1% level.

Across all four outcome variables, the return spread generated by characteristics outperforms CAR by orders of magnitude. Thus, characteristics not only correlate with ex-post outcomes (e.g., non-impairment, ROA, completion), but predicted ex-post outcomes based on characteristics also correlate with long-term returns around the announcement period. This result further confirms that CAR fails to capture other information known during the announcement. Additionally, it validates our acquisition value creation proxies since they correlate with long-term returns.

[25] We do not include canceled deals in this analysis as there is no clear prediction on returns: acquirers with canceled negative CAR transactions may indeed generate positive returns following the transaction process.

3.6 Selection and Feedback Effects

The analyses that test the correlation of ex-post outcomes (e.g., non-impairment and abnormal ROA) with CAR are conditional on deal completion, which is information not known at the announcement. As such, these analyses implicitly assume that completed deals are a random sample of those announced and that ex-post outcomes are unaffected by management who heeds announcement returns.

If announcement returns reflect market expectations of deal value absent any managerial response, they need not correlate with ex-post outcomes that do reflect managerial response. If managers learn from CAR and take corrective action, it would affect the correlation between CAR and ex-post outcomes. This effect could present itself via a selection bias (elimination of withdrawn bids) and a "feedback effect" (Edmans et al., 2012, 2015). For example, managers who observe a negative CAR may cancel the transaction or allocate more resources to improve the potential for deal success. Conversely, following a positive CAR, managers may allocate more resources to completing and integrating the combined entity or they may become overconfident and do the opposite.

Feedback effects imply a flattening of the relation between CAR and deal outcomes for very negative CAR and an ambiguous relation for very positive CAR. We do not observe the flattening for very negative CAR in Figures 1 and 3. Internet Appendix Table IA.B5 replicates Table 2, Column (5), but removes CAR's extreme top and bottom deciles. The lack of correlation between CAR and outcomes remains for the remaining non-extreme eight deciles.

Selection effects imply a correlation between CAR and withdrawal rates, which we document both in-sample and out-of-sample, indicating that truncation bias may partially explain CAR's failure to capture outcomes. However, we caution that this relation is economically weak with low explanatory and predictive power (Tables 2 and 3), lacking consistent results across periods (Figure 2) and subsamples (Internet Appendix Table IA.B3).

Although we cannot isolate the counterfactual (outcomes that do not reflect managerial action), we conduct two tests to explore how truncation bias and feedback effects might account for the lack of correlation between CAR and ex-post outcomes.

The first test involves the following steps. We conduct out-of-sample tests that rely on the insight that the likelihood of canceling a deal is predictable using acquirer and deal characteristics (e.g., Luo, 2005; Betton, Eckbo, and Thornburn, 2009; Wang, 2017). Our Table 3 and Figure 3 show that characteristics predict deal completion reasonably well out-of-sample.

Using the first half of the sample, we regress the completion indicator on characteristics. We then predict the cancellation probability for transactions in the second half of the sample. We sort transactions based on their completion probabilities into three terciles, then repeat the Table 2 tests for both the lowest tercile (low withdrawal probability) and the highest tercile (high withdrawal probability). Internet Appendix Table IA.B6 shows that CAR does not perform better for the sample of transactions with a low cancellation probability—for the 21 regressions in Panels B, D, and F, the coefficient on CAR is statistically significant for only one—than it does for the sample of transactions with a high cancellation probability. Thus, while selection bias is present in our completed deal analyses, the lack of relation between CAR and outcomes for deals, even with a low cancellation probability, indicates that selection may not be the driver of the failure of CAR.

The second test checks whether the feedback provided by CAR is useful. If that is the case, "listening" to CAR from a long-term return perspective would be beneficial. We estimate OLS regressions of the completion probability on CAR (or deal characteristics) using the early years of the sample before 2000. We then compute the imputed outcome for years after 2001. For the samples of completed and withdrawn deals, we sell the bottom three deciles based on CAR (or characteristics) and buy the top three deciles based on CAR (or characteristics). We then consider the returns to listening to the sell signal (negative CAR) by withdrawing versus completing transactions in the bottom three deciles, and listening to the buy signal (positive CAR) by completing versus canceling transactions in the top three

Table 5: **"Listening" to CAR**

This table reports 60-month equal-weighted DGTW-adjusted portfolio returns computed beginning the month-end of the deal announcement date. In Columns (1) and (2), we estimate OLS regressions of the withdrawal probability on deal characteristics using the early years of the sample before 2000. We then compute the imputed outcome for years after 2001. Columns (3) and (4) are computed analogously, except we use the CAR[−1, 1] to predict outcomes. In Panels A and B, we limit the sample to completed and withdrawn deals, respectively, and then sort predicted values into outcome deciles. We report the equal-weighted 60-month DGTW-adjusted mean (median) BHAR for acquirers in the bottom-three and top-three deciles and the p-value (Wilcoxon rank sum) for the difference test between the two portfolios. Panel C reports the differences between the buy and sell signals and the overall return if one consistently "listened" to CAR.

	60-month DGTW-adjusted BHAR			
Prediction model:	Characteristics		CAR	
	Mean	Median	Mean	Median
	(1)	(2)	(3)	(4)
Panel A: Completed Deals				
Buy top 3 deciles	1%	−7%	−1%	−21%
Sell bottom 3 deciles	−12%	−28%	−15%	−26%
	13%	21%	14%	5%
Two-sided test: (t-test/Wilcoxon rank sum)	0.000	0.000	0.000	0.000
Panel B: Withdrawn Deals				
Buy top 3 deciles	−17%	−30%	−16%	−28%
Sell bottom 3 deciles	−9%	−25%	−35%	−50%
	−8%	−5%	19%	22%
Two-sided test: (t-test/Wilcoxon rank sum)	0.460	0.558	0.101	0.004
Panel C: "Listening to CAR"				
Sell signal: Listened (cancelled) vs. not	3%	3%	−20%	−24%
Buy signal: Listened (completed) vs. not	18%	23%	15%	7%
Consistently listened	21%	26%	−5%	−17%

deciles.

We report the results in Table 5. The results indicate that listening to CAR results in losses. Panel C tells us that listening to sell signals generated by CAR (withdrawing vs.completing results in mean losses of 20% (-35% vs. -15%) and median losses of 24%, whereas listening to sell signals produced by characteristics results in positive returns of 3%. CAR performs better for buy signals: completing versus withdrawing generates returns of

7%–15%; however, signals generated from characteristics produce higher returns of 18%–23%. The net effect of listening to CAR is −5% to −17%, while the net effect of listening to characteristics is 21% to 26%.

Overall, although selection and feedback effects are present, they are likely not the primary (or only) driver of the lack of correlation between CAR and ex-post outcomes. Until now, we have used four ex-post outcomes. Though each has its own strengths and weaknesses, in Section 5, we will present tests based on CAR's second moment (that do not rely on ex-post proxies for deal quality). We will find that CAR's second moment shows properties indicating that CAR is not a reliable measure of NPV.

3.7 Which Deals Create Value?

Another way to investigate the forecasting ability of CAR is to consider the quality of inferences regarding deal quality generated from announcement returns relative to ex-post measures. To do so, we consider the "types" of transactions (defined by deal, target, or acquirer characteristics) that CAR predicts will create or destroy the most value. We then relate these deal types to realized outcomes. For example, if CAR for the average public target is negative, one might infer that acquiring a typical public target destroys value. How reliable are these inferences? We address this question in multiple ways.

3.7.1 Univariate Tests: One Characteristic at a Time

We run 65 univariate regressions. There are five dependent variables: CAR, non-impairment, short- and long-term abnormal ROA, and completion. Each of these are regressed against one of these thirteen independent variables: various deal and firm characteristics. So we obtain 65 coefficients (13 $\times$ 5 outcomes). All acquirer characteristics are computed before the announcement. Leverage, free cash flows, assets, and Tobin's Q are computed in the year before the announcement. Past returns are computed in the quarter and month before the announcement.[26]

We standardize the 65 coefficients and present them in Figure 4. The coefficients are sorted by characteristics that predict the lowest CAR (large acquirer, public target, large deal, stock-only) to those that predict the highest CAR (large relative size, mixed payment, high leverage).[27] In general, the relations between CAR and characteristics that we document match those found in earlier studies exploring this relation, although often in different periods and using different samples.

Two important inferences can be drawn from Figure 4. First, the coefficients of the four ex-post outcomes correlate despite originating from different sources. This result implies that characteristics associated with a high likelihood of success (e.g., large acquirer size) are also associated with high ex-post performance measures, as indicated by low impairment outcomes, high short-term and long-term abnormal ROA, and high completion rates. Similarly, characteristics associated with a low likelihood of success (e.g., diversifying and large relative size) are also associated with poor ex-post performance measures. This fact provides further validation of our ex-post proxies for acquisition quality.

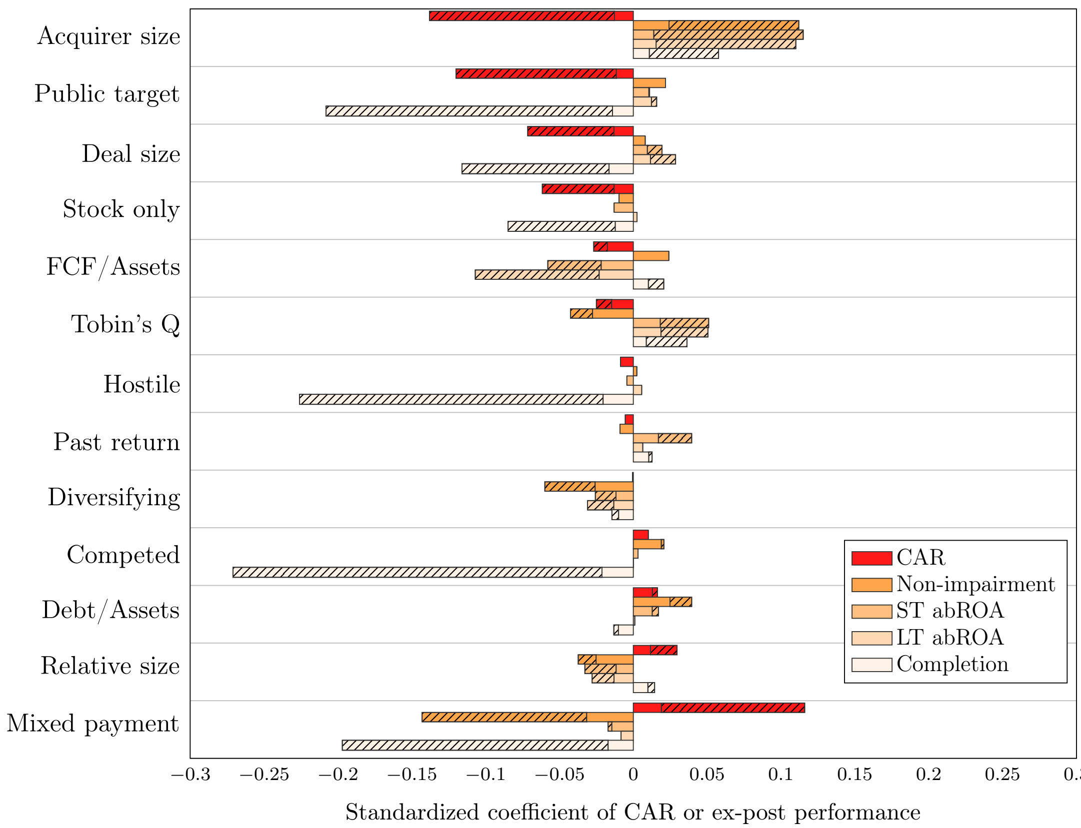
Second, and strikingly, Figure 4 shows no association (in terms of sign and relative importance) between the characteristics for which CAR predicts failure or success and the characteristics that are associated with failure or success ex-post. For example, transactions with large acquirer size are associated with low CARs but are not associated with an increased rate of impairment, withdrawal, or low abnormal ROA; transactions with large relative size are associated with high CARs but are not associated with higher non-impairment or completion or abnormal ROA.

[26] This test also helps address an errors-in-variables critique. Specifically, our main tests regress ex-post outcomes on announcement returns. Standard regression analysis assumes that regressors are observed without noise. CAR, however, could be noisy and hence may lead to attenuated coefficients—an econometric issue often referred to as errors-in-variables in the literature. However, in this section, CAR is the dependent rather than the independent variable.

[27] As discussed in Section 2, the sample size varies across ex-post outcome measures due to data availability. We report the coefficients for regressions when CAR is the dependent variable using the short-term abnormal ROA sample. The results are nearly identical when we use the samples associated with our three other outcome variables.

Figure 4: **Correlation of CAR and Outcomes with Characteristics**

The bar chart shows the standardized coefficients for regressions in which the dependent variable is CAR, non-impairment, short- and long-term abnormal ROA, or completion on the various deal and firm characteristics. Each characteristic enters each regression individually (univariate regressions). The red bars indicate the standardized coefficients from regressions in which CAR is the dependent variable, and the four lighter bars indicate regressions for which non-impairment, short- and long-term abnormal ROA, and completion are the dependent variables. The patterned portion of the bars indicates a coefficient larger than 1.96 standard errors of the standardized coefficient, i.e., statistically significant at least at the 5% level. All acquirer characteristics are computed before the announcement. Leverage, free cash flows, assets, and Tobin's Q are computed the year before the announcement. Past returns are computed in the quarter before the announcement.



Overall, on a univariate basis, there is often a mismatch between the types of deals and acquirers predicted to do well or to destroy value by CAR and the ex-post realizations of these deal types. The results in this section show that the inferences about the quality of acquisition decisions generated by CAR are inconsistent with those generated from ex-post measures.

3.7.2 Combining CAR-based Inferences into a Single Predictor

We further consider the combination of characteristics often used in the M&A literature. Earlier studies have found that announcement returns are persistently associated with particular characteristics; hence, researchers have concluded that deals with certain characteristics create value for acquirers, on average, while others destroy value.

We construct a single measure of CAR-predicted deal success based on characteristics. We first predict CAR by regressing acquirer CAR on characteristics. The in-sample predicted CAR summarizes the associations of CAR with all of the regressors. We then sort the predicted CAR into deciles such that the top (bottom) deciles contain transactions that have characteristics associated with high (low) CAR, implying that, on average, they should predict high (low) NPV transactions.

Our analyses use these predictive regressions to explore whether high-NPV transactions, according to CAR, are indeed associated with better ex-post outcomes. In the four panels on the left-hand side of Figure 5, Panels (a), (c), (e), and (g), we present the ex-post outcomes for predicted CAR deciles. Panel (a) shows that the rate of no goodwill impairment does not vary with the first eight deciles and actually declines for the highest deciles of the combined CAR predictor. In Panels (c) and (e), the sign is clearly wrong—realized abnormal ROA declines from the lowest to highest deciles of the combined CAR predictor. In Panel (g), realized completion rates are lower for the lowest and highest deciles, but do not vary across combined CAR predictor deciles 3–9.

Next, we test whether characteristics do a better job of predicting ex-post outcomes. We utilize the coefficients from regressions of ex-post outcomes on characteristics to obtain in-sample predicted non-impairment, short-term abnormal ROA, long-term abnormal ROA, and completion. Then, we sort predicted values into deciles. On the right-hand side of Figure 5, for each predicted decile, we report realized non-impairment frequency (Panel (b)), average realized abnormal ROA (Panels (d) and (f)), and average realized completion rates (Panel (h)). All four panels show a clear positive slope, suggesting that characteristics are good predictors of ex-post outcomes.

Overall, our results indicate that CAR is not directly or indirectly associated with outcomes via characteristics. These results contrast with the moderate ability of characteristics to predict transaction- and firm-level acquisition outcomes.

3.8 Following the Literature's Advice

We next zoom in on the most common determinants of acquisition quality that have been discussed in the literature (and taught in the classroom): the form of payment, the target's status as public or private, acquirer size, and relative transaction size.[28]

We form 16 combinations of these characteristics (in their binary forms) and calculate average CAR and average ex-post outcomes for transactions classified in each combination. Table 6 presents the results. The combinations are sorted by their average CARs.[29] To facilitate interpretation, statistics within each column are color-coded from red (signifying the worst performance) to green (signifying the best performance) for each measure.

The table shows no positive association between announcement returns and ex-post outcomes. If anything, the association is often negative. The transactions ranked as having the best performance according to CAR (2.8%) have the following acquisition characteristics: not limited to cash, private target, small acquirer, and large relative size. However, their ex-post outcomes are the worst: only 75% of them do not impair (versus a sample mean of 85.2%), and their average short- and long-term abnormal ROA is -0.8% (versus a sample mean of -0.2%). In contrast, the bottom half of the characteristic combinations according to CAR (indicated by red shading) are often ranked in the top half of non-impairment and

[28] For research that links announcement returns to these four characteristics, see, e.g., Travlos (1987), Eckbo, Giammarino, and Heinkel (1990a), Morck, Shleifer, and Vishny (1990), Chang (1998), Andrade, Mitchell, and Stafford (2001), Fuller et al. (2002), Moeller, Schlingemann, and Stulz (2004), Moeller, Schlingemann, and Stulz (2005), Faccio, McConnell, and Stolin (2006), Officer (2007), Bayazitova, Kahl, and Valkanov (2012), Harford, Humphry-Jenner, and Powell (2012), Eckbo, Makaew, and Thorburn (2018), and Schneider and Spalt (2022).

[29] As discussed in Section 2, the sample size varies across ex-post outcome measures due to data availability. We report average CARs based on the short-term abnormal ROA sample. The results are nearly identical when we sort CAR using the samples associated with our three other outcome variables.

Figure 5: **CAR-based Predictors versus Characteristics-based Predictors**

We utilize the coefficients from a regression of CAR on characteristics to obtain an in-sample predicted CAR for the sample of completed transactions, i.e., a summary of what CAR would be given the set of deal and acquirer characteristics. We then sort the predicted CAR into deciles. On the left-hand side of the figure, for each predicted CAR decile, we report (solid red line) realized non-impairment frequency (Panel (a)), average realized short- and long-term abnormal ROA (Panels (c) and (e), respectively), and realized completion frequency (Panel (g)). The red shading indicates a 95% confidence interval. Similarly, we utilize the coefficients from regressions of ex-post outcomes on characteristics to obtain in-sample predicted non-impairment, short-term abnormal ROA, long-term abnormal ROA, and completion, and then sort predicted values into deciles, on the right-hand side of the figure. For each predicted decile, we report (solid blue line) realized non-impairment frequency (Panel (b)), average realized short- and long-term abnormal ROA (Panels (d) and (f), respectively), and realized completion frequency (Panel (h)). The blue shading indicates the 95% confidence interval.

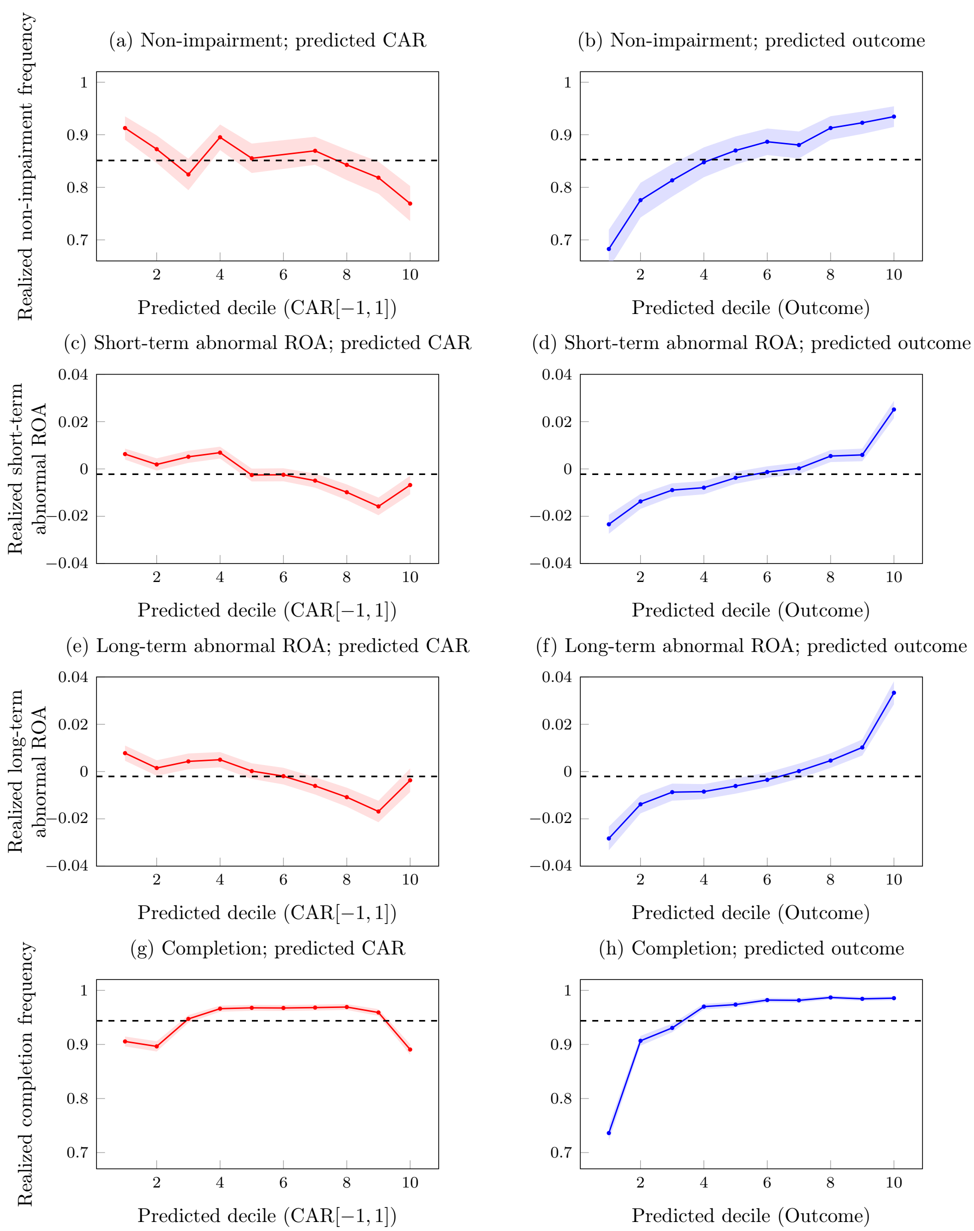


Table 6: **Acquisition Outcomes and CAR, Grouped by Characteristics**

This table reports the average of the acquisition outcome variables and CAR for acquisitions grouped by the characteristics identified in the extant literature as being correlated with CAR. *Rank* is the average rank of the four outcome variables. To facilitate interpretation, statistics within each column are color-coded from red (signifying the worst performance) to green (signifying the best performance) for each measure.

Acquisition Characteristics					Ex-post Outcomes				
Cash Only	Public Target	Large Acquirer	Large Relative Size	Avg CAR$[-1,1]$	Non-impairment	ST abROA	LT abROA	Completion	Avg Rank
			Y	0.028	0.753	−0.008	−0.008	0.935	14
Y			Y	0.022	0.818	−0.007	−0.010	0.973	11
		Y	Y	0.018	0.834	−0.006	−0.005	0.950	7
Y	Y		Y	0.018	0.791	−0.015	−0.010	0.818	16
Y		Y	Y	0.014	0.856	0.002	0.002	0.975	5
Y				0.010	0.874	−0.008	−0.010	0.989	6
Y	Y			0.008	0.925	−0.002	−0.013	0.876	10
Y	Y	Y	Y	0.007	0.823	0.002	0.005	0.804	8
				0.004	0.821	−0.013	−0.012	0.981	12
Y		Y		0.002	0.913	0.012	0.011	0.986	1
		Y		0.002	0.919	0.004	0.004	0.979	3
Y	Y	Y		0.001	0.888	0.020	0.018	0.931	4
	Y	Y		−0.003	0.953	0.008	0.009	0.938	2
	Y			−0.004	0.949	−0.014	−0.024	0.892	12
	Y		Y	−0.011	0.849	−0.012	−0.011	0.789	15
	Y	Y	Y	−0.026	0.837	0.000	0.004	0.843	8

ROA outcomes (as indicated by green shading). For completion, there does not appear to be a negative relation between CAR and non-withdrawal rates, but high withdrawal rates (indicated by red shading) appear throughout the CAR distribution.

Notably, the types of deals indicated to be successful by our four ex-post measures are largely correlated (particularly for non-impairment and ROA). Cash, public target, not large acquirer, and large relative size deals destroy value based on these outcomes. No cash only, private target, large acquirer, and not large relative size create value based on these outcomes.

Overall, these results reiterate our earlier findings that CAR is not a reliable indicator of acquisition quality, and is beaten by a simple characteristics model where the characteristics are known at the time of announcement.

4 What Does CAR Measure?

A firm's attempt to acquire another firm is one of its most important decisions. These decisions are not random; something triggers this action at a particular time. We introduce a simple model in which acquisition decisions are endogenous, and we assess the information communicated to the market at the time of the announcement. Earlier studies discuss the implications of endogeneity of corporate actions (e.g., Shleifer and Vishny, 2003; Viswanathan and Wei, 2008; Savor and Lu, 2009), and others recognize that CAR includes other value-relevant signals beyond the NPV of the specific deal (e.g., Hietala et al., 2003; Grinblatt and Titman, 2002; Bhagat et al., 2005; Pan et al., 2016; Gokkaya, Liu, and Stulz, 2024). Here, we formally assess the information impounded in CAR at the announcement.

4.1 Model Framework

Consider two equity-financed firms: an acquirer and a target. Equity prices are efficient, reflecting all available information. From the acquirer's perspective, the NPV of a deal equals the sum of the present values of the target's cash flows and the synergies minus the price paid to obtain the right over these cash flows. There is no agency problem: the acquirer's manager engages only in positive-NPV deals.

The model operates over two periods: $t = 0$ and $t = 1$. Before $t = 0$, acquiring was not viable because the deal's NPV was not positive. Had it been viable, the acquirer would have already acquired the target.

At $t = 0$, a trigger may occur that turns the NPV of the deal positive. If this trigger is activated, the manager announces the deal at $t = 1$. The trigger could emanate from within the firm (e.g., the outcome of an R&D project, change in management) or from outside (e.g., appearance of an attractive target, change in the competitive landscape, change in regulation, climate change). The trigger affects the acquirer's value by Z. We define X as the component of Z that is orthogonal to deal value. X could be positive, zero, or negative.

An example of a trigger associated with positive value X is the hiring of a talented CEO who brings value to the acquirer and, in addition, can turn a zero-NPV target into a positive-NPV one. An example of a trigger associated with negative X value is the failure of an internal R&D project that destroys value within the acquirer and simultaneously turns a zero-NPV target into an attractive acquisition, as it possesses the desired technology. An example of a trigger with $X = 0$ is simply the serendipitous discovery of a hidden gem of a target.

4.2 Model Analysis

Suppose that at $t = 1$, investors observe an acquisition announcement. Based on this announcement, they would infer that a trigger prompted the manager of the acquiring firm to make a merger announcement. The trigger must surprise investors; otherwise, they would have already incorporated the NPV from the projected deal in the acquirer's price before the announcement. This implies that every acquisition announcement contains two signals, which, by the market efficiency and orthogonality of X assumptions, sum up to the observed \$CAR (where \$CAR is defined as the market capitalization of the acquirer multiplied by CAR at the announcement at $t = 1$). Therefore,

$$\$CAR = NPV + X. \tag{2}$$

A more general version of the above formula that incorporates many situations is

$$\$CAR = (1 - \alpha)(1 - \beta)NPV + (1 - \beta)X + Y. \tag{3}$$

We have four situations in mind:

1. Pre-announcement anticipation by investors. Here, β is the degree of information leakage at $t = 0$. $\beta = 1$ suggests that the news of the occurrence of a trigger and the subsequent decision to acquire a target leaked out at $t = 0$.

2. Uncertainty about the likelihood of merger completion. Here, α is the probability of non-completion after the announcement.

3. Merger arbitrage, which generates noise Y that is orthogonal to the trigger.

4. Behavioral biases (e.g., Roll, 1986), which generate noise Y that is orthogonal to the trigger.

Equation (2) tells us that, even if we make the most classical assumptions, $\$CAR$ combines the two signals NPV and X. Therefore, extracting the NPV signal by simply observing $CAR alone is not possible. This conceptual problem worsens in the more general case because now $\$CAR$ combines several signals, as in Equation (3).

As a practical matter, however, the above conceptual problem may not be that important if the magnitude of X and Y are small compared to the magnitude of NPV. We investigate this point in the next section.

5 Is the Variance of CAR Reasonable?

We use a novel methodology to estimate whether the magnitude of X is small compared to the magnitude of NPV. In other words, this methodology helps us evaluate the extent to which NPV can be extracted from $CAR. Specifically, we ask whether the observed variability of $CAR reflects variability in NPV that would be considered plausible.

An advantage of this test is that it measures the informational content of $CAR via its variability, not by its link with ex-post outcomes. As such, these tests do not rely on measuring ex-post outcomes (or the availability of these variables) and are robust to the feedback and selection issues discussed earlier.

5.1 The Plausible Range of NPV Values

We start with the simplest formula we developed in Section 4:

$$\$CAR = NPV + X,$$

which implies that

$$Var(\$CAR) = Var(NPV) + Var(X). \tag{4}$$

If \$CAR is primarily determined by NPV, i.e., X is relatively small, then the magnitude of the dollar value created or destroyed—as implied by \$CAR—should be within reasonable NPV bounds.

What is a plausible range of NPV values? NPV is a value construct (measured in dollars) likely to be related to the deal amount. It is improbable that the acquirer loses more than it invested. Therefore, NPV is unlikely to be lower than $-\$DealSize$; i.e., a \$1 billion investment is unlikely to result in a value destruction of more than \$1 billion. A reasonable upper bound of NPV also exists. The target's shareholders are unlikely to sell their firm at a deep discount, say -50%; hence, NPV is unlikely to exceed $\$DealSize$. Thus, reasonable bounds on NPV are $\pm\$DealSize$.[30]

The above rationale can be illustrated with an example of Microsoft's acquisitions. Let us assume that CAR is a good measure of NPV. One of Microsoft's most significant acquisitions was Activision Blizzard for \$68 billion in 2022. Based on the above mentioned bounds, we expect CAR to reflect NPV creation or destruction of up to \$68 billion. Given Microsoft's market capitalization of circa \$2 trillion, CAR should be within the $\pm3.5\%$ range. In 2022, Microsoft also acquired CyberX for \$165 million. Applying the same logic, we expect CAR to have a narrower $\pm0.0083\%$ range. Gross deviations from these ranges would indicate that CAR likely contains other information unrelated to NPV.

[30]In practice, the range of \$CAR realizations could be tighter than the bounds on NPV if information leaked before the announcement or investors doubt the deal will be completed ($\alpha > 0$ in Equation (3)).

After all, whether CAR moves within the range of plausible NPV values is an empirical question. In the context of our model, $\$CAR = NPV + X$. If \$CAR is primarily determined by NPV, i.e., X is small, the magnitude of the value created or destroyed—as implied by \$CAR—should be related to the size of the deal. In this case, one can learn about NPV from the information in \$CAR. If, however, X is sizeable, \$CAR will have only limited association with the deal size, and the information about NPV would be masked by X.

We now aim to assess the importance of X within \$CAR.

5.2 \$CAR Realizations Compared to NPV's Plausible Range

We begin the study of the relationship of \$CAR and deal sizes by simply plotting the frequency of transactions with respect to deal sizes and the value created or destroyed. Our sample includes the entire set of 47,543 deal announcements for which CAR$[-1, 1]$ is available, irrespective of deal completion and the availability of other variables.

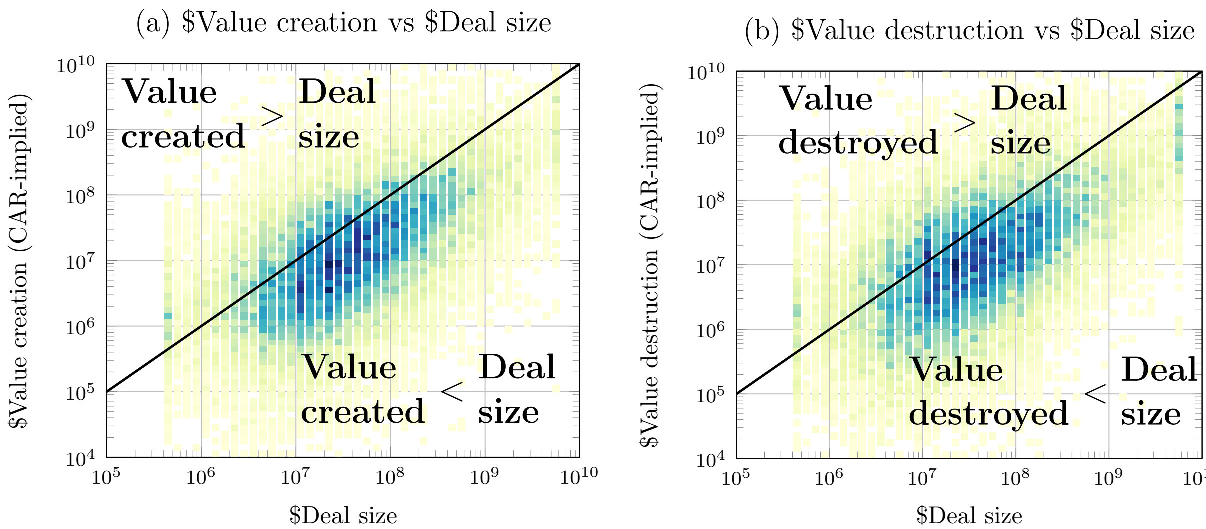
Figure 6 presents the transaction frequency as a function of \$CAR and \$deal size. For ease of presentation, we split the sample by CAR sign. Panel (a) includes only deals with positive CAR, whereas Panel (b) includes only deals with negative CAR. Darker colors represent a higher concentration of deals. The black lines represent $\$CAR = \$DealSize$. In Panel (a), deals above the black line indicate that the \$CAR-implied value created exceeds the amount paid. If \$CAR indeed measures NPV, then the conclusion from Panel (a) is that in 16.3% of the deals that created value, the target's shareholders sold their firm at a discount deeper than 50%. Similarly, the conclusion from Panel (b) would be that in 26.8% of value-destroying deals, the value destroyed was greater than the amount invested by the acquirer.

The alternative explanation would be that these violations of the reasonable bounds indicate that the non-deal information X, contained in \$CAR, is materially large. As a result, \$CAR has extreme tail values.

We delve deeper into the reasonableness of CAR variance by adding more structure.

Figure 6: **$CAR and $Deal Size**

The figures plot the frequency of transactions, presented as a function of $CAR and $deal size. Panel (a) uses the subsample of positive $CAR deals, and Panel (b) uses the subsample of negative $CAR deals. Darker colors represent a higher concentration of deals. The black lines represent $\$CAR = \$DealSize$.



Suppose that X is insignificant; then $CAR would be a good measure of NPV. Therefore, CAR (measured as a fraction of market capitalization) can be expressed as

$$\begin{aligned} CAR &= \frac{NPV}{MktCap} \\ &= \frac{NPV}{DealSize} \times \frac{DealSize}{MktCap} \\ &= NPVratio \times RelativeSize, \end{aligned} \tag{5}$$

where $NPVratio$ is the deal's NPV scaled by its size and $RelativeSize$ is the deal's size relative to the acquirer's market capitalization.

Next, we derive CAR's variation. Internet Appendix D provides this derivation. If we divide the $RelativeSize$ of acquisitions into percentiles, where within a percentile, $RelativeSize$ is homogeneous, the above derivation simplifies a lot. Within a percentile group k, several terms in the derivation become negligible, e.g., $Var(RelativeSize_k) \to 0$. The standard deviation of CAR within group k can then be approximated as

$$\widehat{\text{Std}}(CAR_k) \approx \text{Std}(NPVratio_k) \times \text{E}(RelativeSize_k). \tag{6}$$

Equation (6) tells us that the link between the variability in CAR and the variability in the NPV ratio depends on the expected relative size of the deal in each percentile.

We mentioned earlier that NPV is likely bound by $\pm\$DealSize$. This implies that a reasonable bound for $NPVratio$, as defined above, is ± 1. Further, assuming $NPVratio$ is normally distributed with zero mean and knowing that 95% of the distribution lies within two standard deviations from the mean, we can put the following approximate upper bound on the standard deviation of $NPVratio$: $+0.5$.

Figure 7 explores the variability in CAR for different relative deal size bins. In Panel (a), we notice the variability in CAR largely does not depend on the relative size of the deal. We see further that if we use the 0.5 upper bound for the standard deviation of $NPVratio$, deals below the 70^{th} percentile of $RelativeSize$ violate this reasonable upper bound. The CARs in these deals imply unreasonable NPV values.

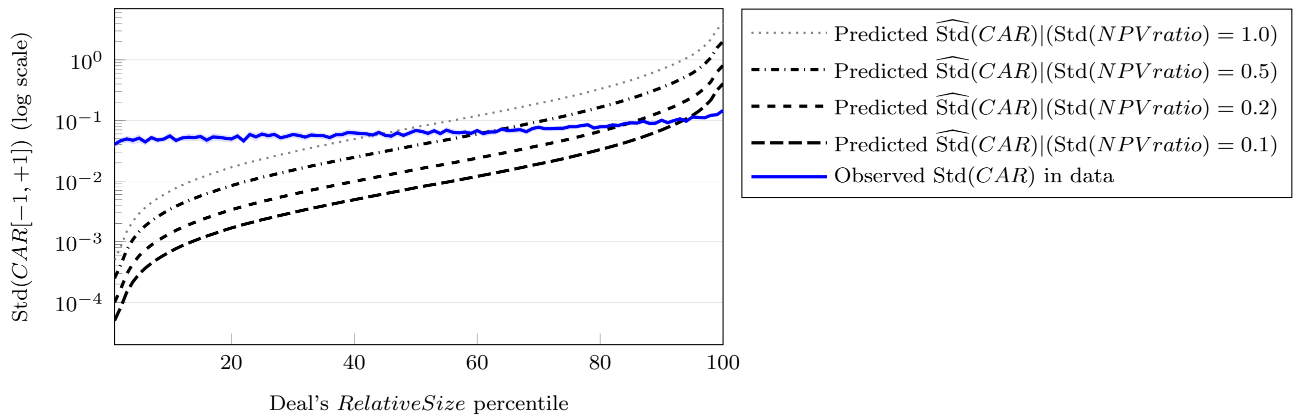
Another way to see the above point is to use Equation (6). Knowing CAR's variation for different relative deal sizes, we can back out the distribution of the implied NPV ratio for various values of relative deal size. This analysis is shown in Panel (b) of Figure 7. At low $RelativeSize$ values, the distribution of the inferred NPV ratios has a wide range, implying that it frequently violates the NPV ratio bounds of ± 1. For instance, most deals at the 10^{th} percentile of $RelativeSize$ have NPV ratios below -3 or above 3, indicating expected value creation or destruction at least threefold of their original investment. Conversely, transactions with high $RelativeSize$ percentiles show a narrow range of NPV ratios.

We conclude that the variance of CAR does not seem reasonable if CAR is a good measure of NPV. Relative to a reasonable range of variability in NPV ratios, about 70% of deals have a far too volatile CAR. This pattern can only happen because X exists and is large.

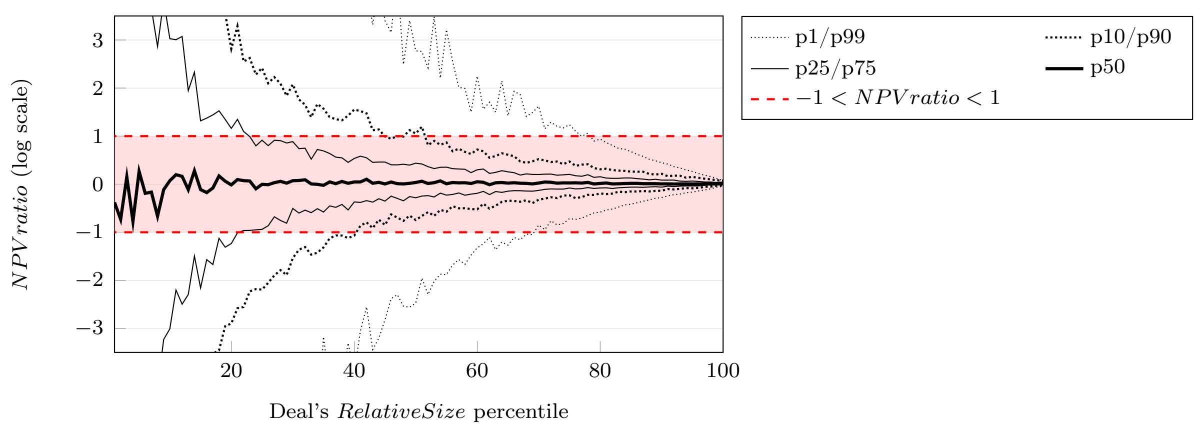
Figure 7: **Estimated and Predicted Variability in CAR, per Relative Size**

The figure presents charts plotting the statistical properties of $CAR[-1,1]$, $RelativeSize$, and $NPVratio$ with respect to the percentiles of $RelativeSize$ (deal value scaled by the acquirer's market capitalization). In all panels, the x-axis measures percentiles of $RelativeSize$. Panel (a) shows the observed standard deviation of CAR ($\mathrm{Std}(CAR)$) and the predicted range of volatilities ($\widehat{\mathrm{Std}}(CAR)$) based on plausible NPV ratios ($\widehat{\mathrm{Std}}(CAR)|(\mathrm{Std}(NPVratio) = m)$) for various values of m. The shaded area around $\mathrm{Std}(CAR)$ represents a 95% confidence interval (calculated using a bootstrap procedure within each $RelativeSize$ percentile, with 1,000 repetitions). Note that the width of the shaded area is in the order of $\pm 0.1 \times \mathrm{mean}(CAR)$, and therefore is barely noticeable in the chart. Panel (b) presents the raw distribution of $NPVratio$. The red dashed lines mark the $NPVratio \in (-1,1)$ bounds.

(a) Observed and Predicted $\mathrm{Std}(CAR)$



(b) Distribution of Implied $NPVratio$



5.3 What Determines the Variability in $CAR?

We now explore the variability in $CAR further and ask whether it is affected more by the acquirer or deal size. As discussed earlier, if $CAR measures value creation, then the variation in the dollar magnitudes it produces should be primarily related to the size of the deal. However, if the non-NPV component X is material, the magnitude of \$CAR is expected to covary with the acquirer's size.

We implement this analysis by regressing the variability in \$CAR on both the acquirer and deal sizes as well as other controls. The results are presented in Table 7. The sample is at the transaction level in Columns (1)–(3). The dependent variable is the logged absolute value of \$CAR. The independent variables are the logged market capitalization of the acquirer and the logged deal size. The ratio of the coefficients indicates that the acquirer's market capitalization is about 6× more important than deal size when it comes to explaining the magnitude of \$CAR. In Columns (4)–(7), we group transactions in bins and use the logged standard deviation of \$CAR within each bin as our dependent variable. In Columns (4) and (5), bins are defined by the logged acquirers' market capitalization and logged deal size.[31] In Columns (6) and (7), acquisitions are binned by acquirers. Hence, all Microsoft deals are placed in the same bin. This allows us to assess the degree to which the economic magnitude of CAR varies with targets' size for the *same acquirer*.

Table 7 shows that regardless of the sample aggregation level and bin definition, the sensitivity of \$CAR variability to acquirer size is 7 to 13 times higher than its sensitivity to target size (i.e., log(deal size)). The dramatic importance of acquirer size relative to target size is even present when considering acquisitions by the same acquirer.

These findings imply that CAR contains more information about the acquirer than the target. As X is likely to be related primarily to acquirer characteristics, and NPV is likely to be related primarily to deal characteristics, this implies that X likely dominates NPV when analyzing CAR.

[31] To be consistent with Figure 6, we form the bins using log with basis 10, where transactions are placed in bins based on 0.2 increments of their logged acquirers' market capitalization and logged deal size.

Table 7: **Determinants of the Variability in CAR**

This table reports regressions of the variability in $CAR on both the acquirer's market capitalization and deal size. Columns (1)–(3) present regressions using a sample based on individual acquisitions. The dependent variable is the logged absolute value of $CAR. $CAR is the change in the acquirer's market capitalization around the announcement event (window of $[-1, 1]$). In Columns (4)–(7), observations are defined over bins of acquisition announcements. In Columns (4) and (5), binning is based on the interaction of logged deal size and logged acquirer market capitalization (both rounded to the nearest 0.2). In Columns (6) and (7), acquisitions are binned by acquirers. In Columns (4)–(7), the dependent variables are the logarithms of the standard deviation of acquisitions' $CAR within each bin. We require bins to have at least five acquisitions. All regressions include an intercept, which is not reported. Deal characteristic controls include leverage and free cash flow scaled by lagged assets, Tobin's-Q, previous-quarter market-adjusted stock returns, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Dependent variable:	log(abs($\$CAR$))			log(Std($\$CAR$))		log(Std($\$CAR$))	
Sample:	All			Acquirer MktCap-Deal Size Bins		Acquirer-based Bins	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(Acquirer market cap)	0.808***	0.788***	0.852***	0.858***	0.809***	0.786***	0.743***
	(0.018)	(0.012)	(0.022)	(0.007)	(0.008)	(0.015)	(0.022)
log(Deal size)	0.118***	0.135***	0.139***		0.097***		0.056***
	(0.011)	(0.010)	(0.008)		(0.005)		(0.010)
Acquirer FE	No	No	Yes	No	No	No	No
Deal characteristic controls	No	Yes	Yes	No	No	No	No
Observations	47,543	41,958	38,039	1,273	1,273	2,997	2,997
Adjusted R^2	0.672	0.693	0.716	0.975	0.981	0.738	0.740

6 Conclusion

Whether CAR around acquisition announcements is a reliable measure of NPV has important implications for corporate finance scholars, the judicial system, and the economy. If CAR were a reliable barometer of the value created in executive decision-making, it should be harnessed to improve economic efficiency. Here are some examples. Executives' incentive pay and promotion could be tied directly to the value created in specific deals they worked on. When investors indicate value destruction via negative CAR, firms' directors could use it as a cause to dismiss the executive team. The judgment of antitrust investigators should be questioned if legal actions by the Department of Justice's antitrust division are uncorrelated with the information conveyed in CAR (Gao, Peng, and Strong, 2017).

Our tests reveal that CAR is not meaningfully correlated with ex-post outcomes. We use four measures of ex-post acquisition outcomes: two transaction-level measures—goodwill impairment and deal completion—and two acquirer-level measures of ex-post performance, short-term and long-term abnormal ROA. Despite capturing different aspects of acquisition performance, these measures are correlated. We first document that CAR has no meaningful correlation with transaction-specific outcomes or measures of the acquirer's future performance, implying that CAR is a poor measure of value creation or destruction. We show that a standard list of deal and acquirer characteristics known at the time of the announcement can, unlike CAR, predict acquisition outcomes reasonably well.[32] We use this superior predictability to assess the relation between CAR and the predictable component (using these characteristics) of acquisition outcomes, but we find no relation. Thus, announcement returns fail to reflect all information available at the acquisition announcement and are likely unable to capture expected acquisition outcomes sufficiently. We show that the poor performance of CAR results in unreliable inferences regarding the types of transactions (i.e., stock vs. cash deals, public vs. private targets, or large vs. small acquirers) that create or destroy value.

Why does CAR not measure NPV? Though we do not claim that we have a definitive answer to this question, we argue that a likely possibility is that CAR could be dominated by non-NPV information. Specifically, since acquisition decisions are endogenous, their announcement must reveal information about their triggers. Our empirical investigation shows that the variability in CAR is often larger than one would expect had it just measured NPV. Furthermore, this variability in CAR itself is correlated with acquirer size much more than deal size, which implies that CAR contains more information about the acquirer than the target. As NPV is related primarily to deal characteristics, the trigger (often associated with the acquirer) likely dominates NPV in determining CAR. Therefore, one cannot easily extract information about the value created in the announced transaction.

[32] CAR is beaten by even the simplest of models like ours; CAR will be trounced by non-linear models as shown in Campbell et al. (2024).

Overall, our results indicate that CAR lacks construct validity (Trochim and Donnelly, 2001): there is a disconnect between the empirical metric of `CAR` and the theoretical construct `Value Creation` that economists intend to measure.

References

Ahern, Kenneth R., and Denis Sosyura, 2014, Who writes the news? Corporate press releases during merger negotiations, *Journal of Finance* 69, 241–291.

Andrade, Gregor, Mark Mitchell, and Erik Stafford, 2001, New evidence and perspectives on mergers, *Journal of Economic Perspectives* 15, 103–120.

Asquith, Paul, Robert F. Bruner, and David W. Mullins, 1983, The gains to bidding firms from merger, *Journal of Financial Economics* 11, 121–139.

Bayazitova, Dinara, Matthias Kahl, and Rossen I. Valkanov, 2012, Value creation estimates beyond announcement returns: Mega-mergers versus other mergers, Working paper, University of Colorado and University of California San Diego.

Beatty, Anne, and Joseph Weber, 2006, Accounting discretion in fair value estimates: An examination of SFAS 142 goodwill impairments, *Journal of Accounting Research* 44, 257–288.

Ben-David, Itzhak, Michael S. Drake, and Darren T. Roulstone, 2015, Acquirer valuation and acquisition decisions: Identifying mispricing using short interest, *Journal of Financial and Quantitative Analysis* 50, 1–32.

Bennett, Benjamin, and Robert A. Dam, 2019, Merger activity, stock prices, and measuring gains from M&A, Working paper, Tulane University.

Bens, Daniel A., Wendy Heltzer, and Benjamin Segal, 2011, The information content of goodwill impairments and SFAS 142, *Journal of Accounting, Auditing & Finance* 26, 527–555.

Berger, Philip G., and Eli Ofek, 1996, Bustup takeovers of value-destroying diversified firms, *Journal of Finance* 51, 1175–1200.

Betton, Sandra, B. Espen Eckbo, Rex Thompson, and Karin S. Thornburn, 2014, Merger negotiations with stock market feedback, *Journal of Finance* 69, 1705–1745.

Betton, Sandra, B. Espen Eckbo, and Karin S. Thorburn, 2008, Corporate takeovers, in *Handbook of empirical corporate finance*, 291–429 (Elsevier).

Betton, Sandra, B. Espen Eckbo, and Karin S. Thornburn, 2009, Merger negotiations and the toehold puzzle, *Journal of Financial Economics* 91, 158–178.

Bhagat, Sanjai, Ming Dong, David Hirshleifer, and Robert Noah, 2005, Do tender offers create value? New methods and evidence, *Journal of Financial Economics* 76, 3–60.

Bhattacharya, Utpal, Hazem Daouk, Brian Jorgenson, and Carl-Heinrich Kehr, 2000, When an event is not an event: The curious case of an emerging market, *Journal of Financial Economics* 55, 69–101.

Brav, Alon, and John Brian Heaton, 2015, Event studies in securities litigation: Low power, confounding effects, and bias, *Washington University Law Review* 93, 583–614.

Brealey, Richard A., Stewart C. Myers, Franklin Allen, and V. Sivarama Krishnan, 2006, *Corporate finance*, volume 8 (McGraw-Hill/Irwin Boston et al.).

Campbell, John L., Erik T. Elfrink, Ashley Fu-Hsien Huang, and Hsin-Min Lu, 2024, What is the deal? Predicting M&A outcomes with machine learning, Working paper, University of Georgia.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University press).

Chang, Saeyoung, 1998, Takeovers of privately held targets, methods of payment, and bidder returns, *Journal of Finance* 53, 773–784.

Chen, Changling, Mark J. Kohlbeck, and Terry Warfield, 2004, Goodwill valuation effects of the initial adoption of SFAS 142, Working paper, University of Wisconsin.

Chen, Xia, Jarrad Harford, and Kai Li, 2007, Monitoring: Which institutions matter?, *Journal of Financial Economics* 86, 279–305.

Cowan, Arnold, Cynthia Jeffrey, and Qian Wang, 2019, Does writing down goodwill imperil a CEO's job?, Working paper, Iowa State University.

Cunningham, Colleen, Florian Ederer, and Song Ma, 2021, Killer acquisitions, *Journal of Political Economy* 129, 649–702.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Dasgupta, Sudipto, Jarrad Harford, and Fangyuan Ma, 2023, EPS-sensitivity and mergers, *Journal of Financial and Quantitative Analysis* forthcoming.

Dong, Ming, David Hirshleifer, Scott Richardson, and Siew Hong Teoh, 2006, Does investor misvaluation drive the takeover market?, *Journal of Finance* 61, 725–762.

Eckbo, B. Espen, Ronald M. Giammarino, and Robert L. Heinkel, 1990a, Asymmetric information and the medium of exchange in takeovers: Theory and tests, *Review of Financial Studies* 3, 651–675.

Eckbo, B. Espen, Tanakorn Makaew, and Karin S. Thorburn, 2018, Are stock-financed takeovers opportunistic?, *Journal of Financial Economics* 128, 443–465.

Eckbo, B. Espen, Vojislav Maksimovic, and Joseph Williams, 1990b, Consistent estimation of cross-sectional models in event studies, *Review of Financial Studies* 3, 343–365.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2012, The real effects of financial markets: The impact of prices on takeovers, *Journal of Finance* 67, 933–971.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2015, Feedback effects, asymmetric trading, and the limits to arbitrage, *American Economic Review* 105, 3766–97.

Ellahie, Atif, Shenje Hshieh, and Feng Zhang, 2021, Measuring the quality of mergers and acquisitions, Working paper, University of Utah.

Ellahie, Atif, Shenje Hshieh, and Feng Zhang, 2022, Measuring the quality of mergers and acquisitions, Working paper, University of Utah.

Elliott, John A., and Wayne H. Shaw, 1988, Write-offs as accounting procedures to manage perceptions, *Journal of Accounting Research* 26, 91–119.

Faccio, Mara, John J. McConnell, and David Stolin, 2006, Returns to acquirers of listed and unlisted targets, *Journal of Financial and Quantitative Analysis* 41, 197–220.

Fama, Eugene F., and Kenneth R. French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153–193.

Foster, Benjamin P., Robin Fletcher, and William D. Stout, 2003, Valuing intangible assets, *CPA Journal* 73, 50.

Francis, Jennifer, J. Douglas Hanna, and Linda Vincent, 1996, Causes and effects of discretionary asset write-offs, *Journal of Accounting Research* 34, 117–134.

Fu, Fangjian, Leming Lin, and Micah S. Officer, 2013, Acquisitions driven by stock overvaluation: Are they good deals?, *Journal of Financial Economics* 109, 24–39.

Fuller, Kathleen, Jeffry Netter, and Mike Stegemoller, 2002, What do returns to acquiring firms tell us? Evidence from firms that make many acquisitions, *Journal of Finance* 57, 1763–1793.

Gao, Ning, Ni Peng, and Norman Strong, 2017, What determines horizontal merger antitrust case selection?, *Journal of Corporate Finance* 46, 51–76.

Gokkaya, Sinan, Xi Liu, and René M Stulz, 2024, Is there information in corporate acquisition plans?, Working paper, The Ohio State University.

Grinblatt, Mark, and Sheridan Titman, 2002, *Financial markets and corporate strategy* (McGraw Hill).

Gu, Feng, and Baruch Lev, 2011, Overpriced shares, ill-advised acquisitions, and goodwill impairment, *Accounting Review* 86, 1995–2022.

Harford, Jarrad, Mark Humphry-Jenner, and Ronan Powell, 2012, The sources of value destruction in acquisitions by entrenched managers, *Journal of Financial Economics* 106, 247–261.

Harford, Jarrad, and Kai Li, 2007, Decoupling CEO wealth and firm performance: Evidence from acquisitions and capital expenditures, *Journal of Finance* 62, 917–949.

Hayn, Carla, and Patricia J. Hughes, 2006, Leading indicators of goodwill impairment, *Journal of Accounting, Auditing & Finance* 21, 223–265.

Healy, Paul M., Krishna G. Palepu, and Richard S. Ruback, 1992, Does corporate performance improve after mergers?, *Journal of Financial Economics* 31, 135–175.

Henning, Steven L., Barry L. Lewis, and Wayne H. Shaw, 2000, Valuation of the components of purchased goodwill, *Journal of Accounting Research* 38, 375–386.

Henning, Steven L., and Toby Stock, 1997, The value-relevance of goodwill write-offs, Working paper, Southern Methodist University.

Hietala, Pekka, Steven Kaplan, and David Robinson, 2003, What is the price of hubris? Using takeover battles to infer overpayments and synergies, *Financial Management* 32, 5–31.

Irani, M. Vahid, 2020, The real merger gains: Correcting for partial anticipation, Working paper, University of South Carolina.

Jacobsen, Stacey, 2014, The death of the deal: Are withdrawn acquisition deals informative of CEO quality?, *Journal of Financial Economics* 114, 54–83.

Jennings, Robert H., and Michael A. Mazzeo, 1991, Stock price movements around acquisition announcements and management's response, *Journal of Business* 64, 139–163.

Jensen, Michael C., and Richard S. Ruback, 1983, The market for corporate control: The scientific evidence, *Journal of Financial Economics* 11, 5–50.

Jenter, Dirk, and Katharina Lewellen, 2021, Performance-induced CEO turnover, *Review of Financial Studies* 34, 569–617.

Johnson, L. Todd, and Kimberly R. Petrone, 1998, Is goodwill an asset?, *Accounting Horizons* 12, 293–303.

Kaplan, Steven N., and Michael S. Weisbach, 1992, The success of acquisitions: Evidence from divestitures, *Journal of Finance* 47, 107–138.

Kau, James B., James S. Linck, and Paul H Rubin, 2008, Do managers listen to the market?, *Journal of Corporate Finance* 14, 347–362.

Lehn, Kenneth M., and Mengxin Zhao, 2006, CEO turnover after acquisitions: Are bad bidders fired?, *Journal of Finance* 61, 1759–1811.

Li, Kevin K., and Richard G. Sloan, 2017, Has goodwill accounting gone bad?, *Review of Accounting Studies* 22, 964–1003.

Li, Zining, Pervin K. Shroff, Ramgopal Venkataraman, and Ivy Xiying Zhang, 2011, Causes and consequences of goodwill impairment losses, *Review of Accounting Studies* 16, 745–778.

Luo, Yuanzhi, 2005, Do insiders learn from outsiders? Evidence from mergers and acquisitions, *Journal of Finance* 60, 1951–1982.

Malatesta, Paul, and Rex Thompson, 1985, Partially anticipated events: A model of stock price reaction with an application to corporate acquisitions, *Journal of Financial Economics* 14, 237–250.

Malmendier, Ulrike, Enrico Moretti, and Florian S. Peters, 2018, Winning by losing: Evidence on the long-run effects of mergers, *Review of Financial Studies* 31, 3212–3264.

Mitchell, Mark, Todd Pulvino, and Erik Stafford, 2004, Price pressure around mergers, *Journal of Finance* 59, 31–63.

Mitchell, Mark L., and Kenneth Lehn, 1990, Do bad bidders become good targets?, *Journal of Political Economy* 98, 372–398.

Mitchell, Mark L., and Erik Stafford, 2000, Managerial decisions and long-term stock price performance, *Journal of Business* 73, 287–329.

Moeller, Sara B., Frederik P. Schlingemann, and René M. Stulz, 2004, Firm size and the gains from acquisitions, *Journal of Financial Economics* 73, 201–228.

Moeller, Sara B., Frederik P. Schlingemann, and René M. Stulz, 2005, Wealth destruction on a massive scale? A study of acquiring-firm returns in the recent merger wave, *Journal of Finance* 60, 757–782.

Morck, Randall, Andrei Shleifer, and Robert W. Vishny, 1990, Do managerial objectives drive bad acquisitions?, *Journal of Finance* 45, 31–48.

Offenberg, David, and Micah S. Officer, 2012, Anticipation and returns in event studies, Working paper, Loyola Marymount University.

Officer, Micah S., 2007, The price of corporate liquidity: Acquisition discounts for unlisted target, *Journal of Financial Economics* 83, 571–598.

Pan, Yihui, Tracy Yue Wang, and Michael S. Weisbach, 2016, CEO investment cycles, *Review of Financial Studies* 29, 2955–2999.

Papadakis, Vassilis M., and Ioannis C. Thanos, 2010, Measuring the performance of acquisitions: An empirical investigation using multiple criteria, *British Journal of Management* 21, 859–873.

Parrino, Robert, 1997, CEO turnover and outside succession a cross-sectional analysis, *Journal of Financial Economics* 46, 165–197.

Powell, Ronan G., and Andrew W. Stark, 2005, Does operating performance increase post-takeover for UK takeovers? A comparison of performance measures and benchmarks, *Journal of Corporate Finance* 11, 293–317.

Ramanna, Karthik, and Ross L. Watts, 2012, Evidence on the use of unverifiable estimates in required goodwill impairment, *Review of Accounting Studies* 17, 749–780.

Roll, Richard, 1986, The hubris hypothesis of corporate takeovers, *Journal of Business* 59, 197–216.

Savor, Pavel G., and Qi Lu, 2009, Do stock mergers create value for acquirers?, *Journal of Finance* 64, 1061–1097.

Schipper, Katherine, and Rex Thompson, 1983a, Evidence on the capitalized value of merger activity for acquiring firms, *Journal of Financial Economics* 11, 85–119.

Schipper, Katherine, and Rex Thompson, 1983b, The impact of merger-related regulations on the shareholders of acquiring firms, *Journal of Accounting Research* 21, 184–221.

Schneider, Christoph, and Oliver Spalt, 2022, Bidder and target size effects in M&A are not driven by overconfidence or agency problems, *Critical Review of Finance* forthcoming.

Schoenberg, Richard, 2006, Measuring the performance of corporate acquisitions: An empirical comparison of alternative metrics, *British Journal of Management* 17, 361–370.

Schwert, G. William, 1996, Markup pricing in mergers and acquisitions, *Journal of Financial Economics* 41, 153–192.

Shleifer, Andrei, and Robert W. Vishny, 2003, Stock market driven acquisitions, *Journal of Financial Economics* 70, 295–311.

Travlos, Nickolaos G., 1987, Corporate takeover bids, method of payment, and bidding firms' stock returns, *Journal of Finance* 52, 943–963.

Trochim, William M.K., and James P. Donnelly, 2001, *Research methods knowledge base*, volume 2 (Atomic Dog, Macmillan, New York).

Viswanathan, S., and Bin Wei, 2008, Endogenous events and long-run returns, *Review of Financial Studies* 21, 855–888.

Wang, Baolian, 2017, Probability weighting and asset prices: Evidence from mergers and acquisitions, Working paper, University of Florida.

Wang, Wenyu, 2018, Bid anticipation, information revelation, and merger gains, *Journal of Financial Economics* 128, 320–343.

Weisbach, Michael S., 1995, CEO turnover and the firm's investment decisions, *Journal of Financial Economics* 37, 159–188.

The (Missing) Relation Between Announcement Returns and Value Creation

Internet Appendix

A Sample Construction

Our sample of mergers and acquisitions comes from the Thomson Reuters Securities Data Company (SDC) Domestic Merger and Acquisition database. Our sample begins in 1980 and ends in 2018; so we can track acquisition outcomes in the five years after the transaction. We include transactions that satisfy the following criteria: (a) the merger or acquisition was announced on or after January 1, 1980, and completed by December 31, 2018; (b) the acquirer is a U.S. company; (c) the acquirer is a publicly-traded firm; (d) the deal is not classified as a leveraged buyout, spinoff, repurchase, self-tender, recapitalization, privatization, stake purchase, or acquisition of partial or remaining interest; (e) the percentage of shares acquired (or sought for not completed deals) is at least 50%; (f) the percentage of shares held by the acquirer six months before the announcement is less than 50%; (g) Compustat has accounting data on the bidder and the Center for Research in Security Prices (CRSP) database has stock data for the month of the deal announcement; and (h) the deal value is non-missing in the SDC database. These requirements result in an initial sample of 47,543 deals, of which 42,354 are completed, 2,227 are withdrawn (the deal outcome is known in these cases), and 2,962 are not completed but not withdrawn (e.g., the transaction may be pending, or the outcome is unknown, which we exclude from the analysis). Internet Appendix Table IA.A1 below lists the steps and number of deals remaining after each filter.

For each transaction, we compute acquirer announcement returns. We estimate daily abnormal returns using a market model and a CRSP value-weighted index (r_{mt}). The market model parameters, α_i and β_i, are estimated from 361 to 61 trading days before the deal announcement day. CARs are then computed by summing the daily abnormal returns over various event horizons. We estimate CARs over three days $[-1, 1]$ and an 11-day period $[-5, 5]$ surrounding each acquisition announcement, and over the entire acquisition process beginning two days before the announcement and ending two days following the deal completion $[\text{Announcement} - 2, \text{Close} + 2]$. CAR may understate absolute value expectations if the probability of deal completion is uncertain; using the long window that includes the deal

Table IA.A1: **Sample Construction**

This table reports the filters applied to the Thomson Reuters Securities Data Company (SDC) Domestic Merger and Acquisition database.

Step	Filter description	# of Deals		
		Completed	Not Completed	Total
1	Date announced: 1/1/1980–12/31/2018			
2	Acquirer country: U.S.			349,687
3	Acquirer is public	117,566	48,504	166,070
4	Eliminate leveraged buyouts, spinoffs, repurchases, self-tenders, recapitalization, privatization, stake purchase, acquisitions of partial or remaining interest	103,015	23,443	126,458
5	Percent of shares acquired ("sought" for deals not completed) in the transaction: 50 to Hi	99,939	21,538	121,477
6	Percent of shares held by acquirer six months before the announcement: 0 to 49	99,881	21,527	121,408
7	Drop duplicate deals in terms of the announcement and effective date, acquirer and acquirer parent name, deal value, target and acquirer sic code, and % of stock as method of payment	97,745	21,116	118,861
8	Require match to CRSP	81,086	13,850	94,936
9	Require match to Compustat	80,444	13,630	94,074
10	Require CAR $[-1, 1]$ measure to be non-missing	78,406	13,256	91,662
11	Deal value is non-missing	42,354	5,189	47,543

completion date overcomes this issue as the probability of completion has moved toward one.

We construct transaction- and firm-level proxies for acquisition outcomes to assess the core relation between announcement returns and value creation. Due to differences in data availability across outcome measures, the sample sizes vary for each measure. In the next two subsections, we provide further detail on sample filters and the number of observations for the various outcome variables.

A.1 Abnormal ROA and Deal Completion Samples

Internet Appendix Table IA.A2 below presents the additional filters to obtain the short-term abnormal ROA (Panel A), long-term abnormal ROA (Panel B), and deal completion (Panel C) samples. We require the particular outcome measure to be non-missing and all firm-level control variables to be non-missing, including the log of market capitalization, leverage, and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals.

Table IA.A2: **Abnormal ROA and Deal Completion Samples**

This table reports the filters applied to obtain the short-term abnormal ROA (Panel A), long-term abnormal ROA (Panel B), and deal completion (Panel C) samples.

Filter Description	# of Deals
Panel A: Firm-level Short-term abROA Sample	
Short-term abROA to be non-missing	31,266
Controls non-missing	28,710
Panel B: Firm-level Long-term abROA Sample	
Long-term abROA to be non-missing	24,497
Controls non-missing	22,577
Panel C: Transaction-level Deal Completion Sample	
Deal withdrawn to be non-missing	44,825
Controls non-missing	39,585

A.2 Goodwill Impairment Data

In an acquisition, the acquirer exchanges consideration (cash, stock, or both) for the target's stock or assets. In most cases, the acquirer pays more than the value of the identifiable assets of the target. As such, on the acquirer's balance sheet, the value of the target is recorded as a combination of the value of the identifiable assets and goodwill. Goodwill is the account on the acquirer's balance sheet that captures the difference between the consideration paid in the acquisition and the value of the identifiable net assets:

$$\mathrm{Goodwill_i} = \mathrm{Price_i} - \mathrm{Value(Identifiable\ Assets)_i}. \tag{7}$$

From an economic point of view, goodwill can include the value of (a) a standalone going-concern element, which reflects the higher value of a collection of assets over assets held independently; (b) a synergy element, which reflects the value from combining the acquirer and target businesses; and (c) any overpayment or overvaluation of the stock consideration (Johnson and Petrone, 1998; Henning, Lewis, and Shaw, 2000).

Accounting rules require occasional downward adjustments to the goodwill account (goodwill write-downs or impairments). The impairment of goodwill can arise because of the following factors: overvaluation of existing target assets, overestimated synergies, or the inability to realize synergies due to firm, industry, or economy-wide shocks.

The Financial Accounting Standards Board (FASB) published a new financial accounting standard, SFAS 142, effective December 2001, intending to increase transparency and generate goodwill balances that better reflect the underlying economic value of the acquisition on an ongoing basis (Foster, Fletcher, and Stout, 2003). SFAS 142 introduced four significant changes to the existing rules. First, goodwill assignment and impairment tests must be conducted at the "reporting unit" level (an operating segment or one component level below a segment), making it easier to identify the goodwill recorded and the source of future impairments at the transaction level. Second, acquirers can "write up" the target's assets to fair value at the time of the acquisition.[33] Third, goodwill is no longer amortized but is considered an asset that can stay on the firm's balance sheet indefinitely.[34] Fourth, firms must conduct impairment tests following "material" events for reductions in the value of goodwill, and for many years in our sample annual impairment tests were conducted. If the appraised value is less than the recorded value, then a goodwill "impairment" occurs. The amount of goodwill is reduced on the balance sheet, and an impairment expense is incurred on the income statement as a component of income from continuing operations. In September 2011, FASB modified SFAS 142 so that formal valuations to produce comparisons of fair value and carrying value of a reporting unit are only required when certain qualitative indicators of impairment exist; thus, impairment tests are no longer required to be conducted annually.[35]

[33] Identifiable intangible assets, such as patents and customer lists, are no longer included in goodwill balances.

[34] Before SFAS 142, acquisition goodwill was amortized over a maximum of 40 years.

[35] Before the 2001 rule change, SFAS 121 prescribed only non-routine impairment tests following certain triggering events that indicated that goodwill might no longer be recoverable. Under SFAS 142, the impairment amount must be determined using a fair value approach based on a two-step impairment test. In the

Goodwill reflects the premium paid over the identifiable assets in nominal terms. Thus, the impairment of goodwill indicates that the remaining value of the target is lower than the nominal value paid a few years earlier at acquisition. A large goodwill impairment, therefore, likely captures value destruction. Due to the increased precision and timeliness of goodwill reporting required by SFAS 142, we can construct goodwill balances and impairment at the transaction level, yielding a direct and quantifiable representation of transaction-specific acquisition failure.

To construct the goodwill impairment sample, we start with the 42,354 completed deals described in Internet Appendix Table IA.A1. To align with SFAS 142 roll-out, we retain transactions announced between 2003 and 2018. We include additional filters that are not imposed on our samples that use ROA and completion data. We require the transaction value to exceed \$10 million and to be at least 5% of the acquirer's market capitalization at the end of the fiscal year before the deal was announced. This filter allows a more precise measure of impairment. For very small deals (both in dollar and relative terms), it is difficult to determine the source of the impairment and, in many instances, the amount of goodwill originally produced from the transaction. These filters yield 8,367 transactions.

Next, we link sample firms to Compustat goodwill data and identify all acquirers with firm-level goodwill impairments. In this step, we exclude transactions with missing assets in the year of deal close and transactions with missing or zero goodwill in both the year of and the year after close. This yields 6,767 transactions.

To identify goodwill impairments in the data, we follow Bens et al. (2011). We initially screen for potential goodwill impairments by flagging instances in which the Compustat variable "Impairments of Goodwill Pretax" (item 368 or GDWLIP) is at least 5% of previous-

first step, the reporting unit's fair value is compared to the book value. A second step is performed if the fair value is less than the book value. In the second step, the fair value of the unit's (non-goodwill) net assets is determined, and the fair value of goodwill is the difference between the unit's fair value and the fair value of the unit's identifiable net assets. The impairment amount is the excess of the book value of goodwill and the newly assessed fair-value estimate of goodwill. Firms often use a weighted combination of discounted cash flow, public comparable company multiples, and precedent acquisition transaction multiples valuation as techniques to determine fair value.

Table IA.A3: **Sample Construction for Goodwill Impairments**

The table shows the sample construction. Panel A includes transactions from SDC that were announced from January 2003 and completed by December 2018. Sample filters are described in the text. Panel B describes the classification of the "potentially impaired" transactions. For this sample, we read through the 10-K Notes and Factiva to identify the target(s) that triggered the impairment. Panel C shows the final sample composition.

Panel A: Sample Construction

# Deals	6,767
# Transactions without acquiring firm-level impairment within 5 years of deal effective date	5,229
# Transactions "potentially impaired" with acquiring firm-level impairment within 5 years	1,538

Panel B: Classification of "Potentially Impaired" Transactions

Deals classified in goodwill impairment sample	
Impairment linked directly to target and exact impairment amount can be identified	543
Impairment linked directly to target, other targets in firm or segment also linked	126
Target is in impaired segment, target goodwill > 20% of segment goodwill	277
Total (% of deals potentially impaired)	946 (62%)
Deals classified in no goodwill impairment sample	
Impairment is not in target's segment or 10-K specifies another target as a source of impairment	262
Total (% of deals potentially impaired)	262 (17%)
Deals excluded from sample: cannot classify as impaired or not impaired	
Target is in the impaired segment, but target goodwill is < 20% of segment goodwill	159
No information on goodwill created from acquisition	136
No information on the source of impairment	17
No goodwill created from acquisition	18
Total (% of deals potentially impaired)	330 (21%)

Panel C: Final Goodwill Impairment Sample Summary

Impairment sample	946
Non-impairment sample	5,491
Total	6,437
Controls non-missing	6,128
Final impairment sample	906
Final non-impairment sample	5,222

year total acquirer assets in any year between the year of the acquisition close and 5 years following. This requirement ensures that the impairment event has detectable valuation effects. Of the 6,767 transactions in the sample, 1,538 deals are associated with a firm-level impairment within 5 years of the deal's effective date. This is summarized in Internet Appendix Table IA.A3.

The Compustat goodwill and impairment data are based on aggregate firm-level data, so it is not directly possible to identify transaction-specific measures. To identify the amount of goodwill recorded for each transaction in our sample, we read through the Notes to Consolidated Financial Statements in the first 10-K filing following the deal's effective date. Following an acquisition, the notes include an "Acquisitions" or "Business Combinations" section that presents the preliminary allocations of the aggregate purchase price based on the assets and liabilities estimated at fair values to line items, such as net tangible assets, identifiable intangible assets, and goodwill. In this step, we also determine the recording unit for which the goodwill has been allocated.

For the years with indicated firm-level impairment, we use the Notes to Consolidated Financial Statements to determine whether and how much of the impairment is due to the specific transaction in our sample. We also read through news articles and press releases in Factiva if more information is required.

In many instances, the source and the amount of the impairment assigned to each target are straightforward. In the most uncomplicated scenarios, the targets with goodwill impairment and the amount of target-level impairment are directly listed in the Notes section of the 10-K, or the firm writes off the entirety of its goodwill balance. In other cases, the Notes list the reporting unit(s) that suffered the loss. We search the 10-K, the Notes, and Factiva in the year of the goodwill allocation to determine the reporting unit(s) to which the target's goodwill is allocated. If target goodwill is 100% of the impaired reporting unit goodwill, the impairment attributable to the target is straightforward. For 543 transactions in the potentially impaired sample of 1,538, we can link the impairment directly to the target and determine the exact impairment amount.

In 126 other instances, the target is listed as impaired in the Notes, but the impairment amount is unknown due to other targets also triggering the impairment. If the impairment is at the reporting-unit level, we set target impairment equal to unit impairment $\times$ (target goodwill/unit goodwill). If the impairment is reported at the consolidated firm level, we set target impairment equal to total impairment × (target goodwill / total goodwill). Note that our variable of interest is the occurrence of an impairment, which will be unaffected by errors in the estimated impairment size.

For some transactions, we are uncertain as to the source and amount of the impairment. If the target is in the impaired segment and target goodwill is at least 20% of segment goodwill, we conclude that it is reasonably likely that the target has been impaired and include these 277 transactions in the impairment sample. We estimate the size of the impairment using the relative size of the target goodwill as described above. Therefore, of the 1,538 "potentially impaired" deals, we can classify 543 + 126 + 277 = 946 as "impaired deals."

For 262 transactions flagged as potentially impaired, we determine that the impairment is not in the target's segment or that other targets have been listed as the source of the impairment. These transactions are included in the non-impairment sample.

We cannot reasonably classify some transactions as impaired or not impaired, and thus, they are excluded from the sample. We exclude 17 deals where the 10-K provides no details on the source of the impairment, 159 deals where the target is in the impaired segment but target goodwill is less than 20% of segment goodwill. (We run robustness tests in Internet Appendix Table IA.B2 and show that our results are unaltered if these deals are included in the sample and classified as either impaired or not impaired.) We exclude 136 deals that lack information in the Notes on the amount of goodwill created from the particular acquisition and 18 deals where goodwill was not created from the acquisition.

Internet Appendix Table IA.A3, Panel B, shows that we could successfully link impairment events to specific transactions. Of 1,538 transactions flagged as potentially impaired, we can credibly classify 62% as impaired, and 17% as not impaired, and we cannot classify 21% of transactions. Moreover, for 71% ((543+126)/946) of the transactions classified as impaired, we know unambiguously the source of the impairment. To our knowledge, we are the first to construct a comprehensive data set that includes transaction-specific goodwill balances and transaction-specific impairment outcomes in the post-SFAS 142 period. Hayn and Hughes (2006) also trace initial goodwill balances and subsequent impairments at the transaction level, but they exclude 55% of transactions due to insufficient information. Overall, they focus largely on the pre-SFAS 142 period, when the disclosure of initial goodwill and the source of the impairment was generally less comprehensive.

Internet Appendix Table IA.A3, Panel C, shows that the sample (6,437) is further reduced when we require announcement returns to be non-missing (6,435) and controls to be non-missing (6,128). Thus, our final sample for goodwill impairment analyses is 6,128 transactions, of which 906 are classified as impaired, and 5,222 are classified as not impaired.

Internet Appendix Table IA.A4 shows summary statistics for this sample. We find that 14.8% of transactions are impaired by year 5 and that, conditional on impairment, 79% of the impairments occur by year 3. In Internet Appendix Table IA.A4, Panel B, we report goodwill and impairment statistics for the sample of 906 deals associated with transaction-level impairment. The dollar values of goodwill impairments are large. On average, acquirers write down 83% of the original goodwill allocated to the deal, and the impairment size is about 11% of the acquirer's assets.

Table IA.A4: **Sample Statistics**

This table provides summary statistics. Panel A shows sample statistics for the percentage of transactions with goodwill impairment within five years of the deal's effective date. Panel B shows statistics for the 906 transactions in the impairment sample.

Panel A: Transaction-level Impairment Percentages

	%	N
Year 0–1	5.5%	339
Year 2–3	6.1%	374
Year 4–5	3.1%	193
Impaired by year 5	14.8%	906
Not impaired by year 5	85.2%	5,222
Total completed deals	100.0%	6,128

Panel B: Transaction-level Impairment Statistics

	Mean	Std dev
$ Goodwill ($m)	422.3	1,252.8
Goodwill/Net purchase price	69%	73%
Goodwill/Total assets	14%	12%
Impairment $ loss ($m)	−242.1	643.7
Impairment/Goodwill	83%	35%
Impairment/Purchase price	57%	76%
Impairment/Total assets	11%	11%

B Predicting Outcomes: Additional Tests

We consider whether our Section 3 tests that indicate a lack of correlation between outcomes (e.g., non-impairment, ROA, and completion) and CAR are robust to alternative definitions of CAR and outcomes and alternative measurement periods for the outcome variables of interest, across industries, deal types, and firm characteristics.

Internet Appendix Table IA.B1 below replicates Table 2, Columns (1), (4), and (5), but redefines CAR using a longer event window $[-41, 1]$ to capture potential leakage or deal anticipation (e.g., Betton et al., 2014; Ahern and Sosyura, 2014). The results in all four panels confirm the lack of relation between CAR and outcomes.

Panels A and B in Internet Appendix Table IA.B2 replicate Table 2, Panel A, but redefines the non-impairment dummy in two alternative ways. We define the goodwill impairment sample in Internet Appendix A.2. For 159 deals in which the target is in the impaired segment but the target goodwill is less than 20% of segment goodwill, we cannot reasonably classify these deals as impaired or not impaired, so we exclude them in our main tests. In Panel A, we retain these deals and assume they did not result in impairment; in Panel B we retain these deals and assume they resulted in goodwill impairment. Internet Appendix Table IA.B2, Panel C, replicates Table 2, Panel B, but uses industry-adjusted ROA (measured as ROA minus the median Fama-French 12 industry ROA and averaged over the three years after deal completion) rather than short-term abnormal ROA. In Panel D, we include deals that still may be pending or for which the outcome is unknown as the outcome variable. The results in all four panels confirm the lack of relation between CAR and alternative definitions of ex-post outcomes.

We next define our ex-post outcomes at various periods following the deal completion date. We again replicate Table 2, yet redefine the dependent variable each year relative to the deal's effective date (up to five years) for non-impairment and ROA, and relative to the announcement date for completion. In Internet Appendix Figure IA.B1, we plot the coefficients on CAR and 95% confidence intervals based on the specification in Table 2

Table IA.B1: **Acquirer CAR $[-41, 1]$ and Acquisition Outcomes**

This table reports results from regressions of acquisition outcome measures on acquirer cumulative abnormal returns (CAR), measured over $[-41, 1]$. The dependent variable is a non-impairment dummy (Panel A), short-term abnormal ROA (Panel B), long-term abnormal ROA (Panel C), and a completion dummy (Panel D). In Column (1), CAR is the only independent variable. In addition to CAR, Column (2) includes characteristics. Column (3) further includes year and industry fixed effects as independent variables. The characteristics used as controls include the log of market capitalization, leverage, and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

CAR window:		$[-41, 1]$	
	(1)	(2)	(3)
Panel A: Probability of Non-impairment ($N = 6,128$, DV: Non-impairment Dummy)			
CAR	0.039	0.067	0.066
	(0.061)	(0.059)	(0.061)
Controls	—	Char	Year, Ind, Char
Adjusted R^2	0.000	0.036	0.089
Panel B: Short-term Abnormal ROA ($N = 28,710$, DV: ST abROA)			
CAR	0.001	0.005*	0.004
	(0.003)	(0.003)	(0.003)
Controls	—	Char	Year, Ind, Char
Adjusted R^2	0.000	0.026	0.078
Panel C: Long-term Abnormal ROA ($N = 22,577$, DV: LT abROA)			
CAR	0.004	0.006	0.004
	(0.004)	(0.005)	(0.004)
Controls	—	Char	Year, Ind, Char
Adjusted R^2	0.000	0.0347	0.109
Panel D: Probability of Completion ($N = 39,585$, DV: Completion Dummy)			
CAR	−0.003	0.009	0.007
	(0.005)	(0.006)	(0.006)
Controls	—	Char	Year, Ind, Char
Adjusted R^2	0.000	0.056	0.111

(Panels (a), (c), and (e)) and the adjusted R^2 (Panels (b), (d), and (f)). In addition, to provide a benchmark, we add to the latter set of panels the R^2 from the standard regression of deal and acquirer characteristics (without industry or year fixed effects).

Table IA.B2: **Acquirer CAR and Acquisition Outcomes: Alternative Definitions**

This table presents regression outcomes linking acquirer CAR to various acquisition performance indicators: alternative non-impairment dummy definition 1 (Panel A), alternative non-impairment dummy definition 2 (Panel B), industry-adjusted ROA (Panel C), and a completion status dummy for deals that includes unknown or pending deals (Panel D). Columns (1)–(3) use CAR as the sole independent variable. Column (4) adds characteristics, and Column (5) adds year and industry fixed effects. Column (6) only includes characteristics, and Column (7) adds year and industry fixed effects to characteristics. The characteristics used as controls include the log of market capitalization, leverage, and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

CAR window:	$[-1,1]$	$[-5,5]$	$[A-2,C+2]$	$[-1,1]$		n.a.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Probability of Non-impairment (Alternative Definition 1, $N = 6,278$)							
Dependent variable:	Non-impairment Dummy						
CAR	−0.002	0.001	0.085**	0.110	0.106	Controls	Controls
	(0.106)	(0.069)	(0.037)	(0.088)	(0.101)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.002	0.027	0.081	0.027	0.081
Panel B: Probability of Non-impairment (Alternative Definition 2, $N = 6,278$)							
Dependent variable:	Non-impairment Dummy						
CAR	−0.023	0.010	0.088**	0.076	0.072	Controls	Controls
	(0.103)	(0.082)	(0.040)	(0.081)	(0.085)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.002	0.036	0.086	0.036	0.086
Panel C: Ind-adj ROA ($N = 30,060$)							
Dependent variable:	Ind-adj ROA						
CAR	−0.038	−0.009	0.015*	0.028	0.030	Controls	Controls
	(0.031)	(0.020)	(0.007)	(0.019)	(0.017)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.001	0.245	0.327	0.245	0.326
Panel D: Probability of Completion ($N = 41,951$)							
Dependent variable:	Completion Dummy						
CAR	−0.017	0.004	—	0.029	0.022	Controls	Controls
	(0.029)	(0.017)		(0.021)	(0.020)	only	only
Controls	—	—	—	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	—	0.062	0.072	0.062	0.072

We next replicate Table 2 for each Fama French 12 industry classification. We report the results in Internet Appendix Figure IA.B2. Panels (a)–(d) show the coefficient and 95% confidence intervals for regressions of outcomes on CAR based on the specification in Table 2, Column (5), for each of the 12 industries. Panel (a) shows that the coefficient on CAR in regressions of non-impairment on CAR is only significant at the 5% level (and the correct sign) for the "other" industry (industry 12). Panel (b) shows that when short-term abnormal ROA is the outcome variable, CAR is significant (and the correct sign) at the 5% level for the "business equipment" (industry 1) and "non-durable" (industry 8) industries, but is not statistically significant in the remaining 10 industries. Panel (c) shows that when long-term abnormal ROA is the outcome variable, CAR is not statistically significant (and in one industry has the wrong sign) for all 12 industries. Similarly, Panel (d) shows that CAR correlates with completion for only one of the 12 industries (i.e., "manufacturing"). This result is in contrast to the statistically significant (at the 5% level) and positive (but economically weak) relation between CAR and completion reported in Table 2. Although CAR correlates with outcomes in a few select industries, importantly, there is no overlap in these industries across outcome variables. The results show that the lack of correlation between CAR and outcomes is persistent across industries.

Internet Appendix Table IA.B3 replicates Table 2 across various deal types and acquirer characteristics. More specifically, we run OLS regressions of each outcome on CAR, year-fixed effects, and industry-fixed effects using 29 subsamples. We split the sample based on deal and firm characteristics, including the log of market capitalization, leverage, and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for serial acquirer, stock-only, mixed-payment, cash-payment, diversifying, competed, hostile, high-tech, and public target deals. In Internet Appendix Table IA.B3, we report the t-statistic of each regression in a given cell. The green shading identifies coefficients that are the correct sign and statistically significant at the 5% level. The results indicate that CAR's performance does not improve systematically in particular subsamples: in only three subsamples does CAR achieve statistical significance for two of the four outcome variables, and in no subsample does CAR achieve statistical significance for three or more outcome variables.

Internet Appendix Table IA.B3 examines deal and firm characteristics individually. We next allow for the interaction of characteristics. We create 10 indicator variables based on deal and firm characteristics, including the log of market capitalization, leverage, free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, cash-payment, diversifying, serial, and public target deals. If the characteristic is continuous, we create the indicator variable by splitting the sample at the median. We then form subsamples based on all the unique interactions of these variables and retain subsamples with at least 30 observations. We then split the sample into two time periods, and, for each subsample and time period, we regress outcomes on CAR$[-1, 1]$ and record the corresponding t-statistic. For both periods, we report the number of transactions with a t-statistic greater than or equal to two, between two and minus two, and less than or equal to minus two. In Internet Appendix Table IA.B4, Panel A, for non-impairment, we run 22,298 regressions for each period and find that only 5% of transactions have the correct sign and a t-static of at least two in the first period; only 3% do so in the second period; and only 0.26% have the correct sign and statistical significance in *both* periods. Similarly, using short-term abnormal ROA, long-term abnormal ROA, and a completion dummy in Panels B–D, respectively, we find no more than 10% of the regression have the correct sign and a t-static of at least two in either period. Only 0.54%, 0.11%, and 1.45% of the regressions have the correct sign and statistical significance in *both* periods for short-term abnormal ROA, long-term abnormal ROA, and a completion dummy, respectively.

Given that extreme announcement returns could also point to feedback effects, in Internet Appendix Table IA.B5 below, we replicate Table 2, Column (5), without extreme CAR, i.e., we eliminate deals with CAR in the top and bottom 10% of the sample. The coefficient on CAR is not statistically significant for any of the outcomes (despite achieving significance

Table IA.B3: **Acquirer CAR and Acquisition Outcomes: Subsamples**

This table reports results from regressions of acquisition outcome measures on acquirer cumulative abnormal returns (CAR) using 29 different subsamples. The dependent variable is a non-impairment dummy (Column (1)), short-term abnormal ROA (Column (2)), long-term abnormal ROA (Column (3)), and a completion dummy (Column (4)). Using subsamples, we run OLS regressions of each outcome on CAR, year-fixed effects, and industry-fixed effects. The subsamples are based on deal and firm characteristics, including the log of market capitalization, leverage, free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for serial acquirer, stock-only, mixed-payment, cash-payment, diversifying, competed, hostile, high tech, and public target deals. We split the sample at the median if the characteristic is continuous. Each cell represents a regression. We report the t-statistic and shade the cell green if the coefficient is statistically significant at or above the 5% level.

	Non-impairment (1)	ST abROA (2)	LT abROA (3)	Completion (4)
Private target	−0.051	1.426	−0.718	−0.011
Public target	2.831	0.249	0.337	3.945
Stock deals	4.460	0.131	−0.956	1.121
Cash deals	0.823	1.026	−0.335	1.942
Mix deals	0.221	2.876	0.490	1.348
Diversifying deals	1.776	0.322	0.106	−0.185
Not diversifying deals	−0.177	3.567	−0.359	3.469
Competitive deals	0.489	3.727	1.159	1.274
Not competitive deals	1.011	1.628	−0.190	1.550
Serial acquirer	0.919	2.726	−0.457	1.531
Not serial acquirer	0.981	1.808	1.188	0.976
Hostile deals	—	1.684	−0.509	0.735
Not hostile deals	0.991	1.871	−0.211	2.491
Large acquirer	2.317	1.667	−0.107	3.402
Small acquirer	−0.677	1.133	−0.110	0.841
Large deal	1.260	3.314	1.107	4.955
Small deal	−0.308	0.148	-1.058	−0.091
Large Tobin's Q	−0.257	2.577	−0.606	0.582
Small Tobin's Q	1.957	0.500	0.252	2.377
High past return	0.593	2.078	-1.394	2.097
Low past return	1.265	1.154	0.938	0.849
High free cash flow	0.302	0.656	0.239	2.871
Low free cash flow	1.321	1.267	−0.502	0.525
High debt	1.277	2.784	0.505	3.232
Low debt	0.546	0.694	−0.526	0.711
High relative size	2.049	1.660	0.822	1.618
Low relative size	0.539	0.139	-2.755	1.616
High tech	−0.754	1.041	0.605	1.511
Not high tech	1.513	3.972	−0.474	2.181

Table IA.B4: **Acquirer CAR and Acquisition Outcomes: Interactions**

This table reports aggregated results from regressions of acquisition outcome measures on acquirer cumulative abnormal returns (CAR), allowing for the interaction of characteristics. The dependent variables are a non-impairment dummy, short-term abnormal ROA, long-term abnormal ROA, and a completion dummy in Panels A–D, respectively. We create ten indicator variables based on the characteristics, including the log of market capitalization, leverage, free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, cash-payment, diversifying, serial, and public target deals. (We create the indicator variable by splitting the sample at the median if the characteristic is continuous.) We then form sub-samples based on all of the unique interactions of these variables and retain sub-samples with at least 30 observations. We split the sample into two time periods, and, for each sub-sample and time period, we regress outcomes on CAR$[-1, 1]$ and record the corresponding t-statistic. For both periods, we report the number of transactions with t-statistic≥ 2, not significant (n.s.), and t-statistic≤ -2. Boldface figures indicate the number of regressions with significant coefficients and the correct sign over the two sample periods.

Panel A: Non-impairment Dummy

Total number of regressions: 49,902

Drop $N \leq 30$ in both periods: 22,298

			First Period		
			t-stat≥ 2	n.s.	t-stat≤ -2
			1,091	20,554	653
Second Period	≥ 2	735	**59**	674	2
	n.s.	20,000	990	18,444	566
	≤ -2	1,563	42	1,436	85

Panel B: Short-term Abnormal ROA

Total number of regressions: 50,103

Drop $N \leq 30$ in both periods: 16,354

			First Period		
			t-stat≥ 2	n.s.	t-stat≤ -2
			1,777	13,932	645
Second Period	≥ 2	1,074	**88**	920	66
	n.s.	14,153	1,584	12,019	550
	≤ -2	1,127	105	993	29

Panel C: Long-term Abnormal ROA

Total number of regressions: 49,614

Drop $N \leq 30$ in both periods: 15,551

			First Period		
			t-stat≥ 2	n.s.	t-stat≤ -2
			796	13,779	976
Second Period	≥ 2	686	**17**	622	47
	n.s.	13,780	721	12,292	767
	≤ -2	1,085	58	865	162

Panel D: Completion Dummy

Total number of regressions: 40,594

Drop $N \leq 30$ in both periods: 16,394

			First Period		
			t-stat≥ 2	n.s.	t-stat≤ -2
			1,186	14,438	770
Second Period	≥ 2	1,067	**238**	818	11
	n.s.	14,601	927	13,008	666
	≤ -2	726	21	612	93

for a few outcomes in Table 2, Column (5)).

Table IA.B5: **Acquirer CAR and Acquisition Outcomes: Trim Extreme Values**

This table replicates Table 2, Column (5), but we eliminate deals with CAR in the top and bottom 10% of the sample. The dependent variable is a non-impairment dummy (Column (1)), short-term abnormal ROA (Column (2)), long-term abnormal ROA (Column (3)), and a completion dummy (Column (4)). CAR$[-1,1]$ is the independent variable. The regressions include year and industry fixed effects, and characteristics as independent variables. The characteristics used as controls include the log of market capitalization, leverage, and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

CAR window: Outcome:	$[-1,1]$ Non-impairment (1)	$[-1,1]$ Short-term abROA (2)	$[-1,1]$ Long-term abROA (3)	$[-1,1]$ Completion (4)
CAR	−0.077	0.003	0.001	0.039
	(0.189)	(0.014)	(0.025)	(0.027)
Controls	Year, Ind, Char	Year, Ind, Char	Year, Ind, Char	Year, Ind, Char
Observations	4,893	22,976	18,063	31,668
Adjusted R^2	0.0843	0.0873	0.124	0.165

We next test whether truncation bias and feedback effects account for the lack of correlation between CAR and ex-post outcomes. Specifically, we rely on the insight that the likelihood of canceling a deal is predictable using acquirer and deal characteristics. Indeed, Table 3 and Figure 3 show that characteristics predict the deal completion reasonably well out-of-sample. To carry out the test, we regress the completion indicator on characteristics using the first half of the sample. We then predict the cancellation probability for transactions in the second half of the sample. We sort transactions based on their completion probabilities into three terciles, then repeat the Table 2 tests for both the lowest tercile (low withdrawal probability) and the highest tercile (high withdrawal probability). Internet Appendix Table IA.B6 shows that CAR does not perform better for the sample of transactions with a low cancellation probability (likely less feedback effects): of the 21 regressions reported in Panels B, D, and F, the coefficient on CAR is statistically significant for only one regression.

Table IA.B6: **Acquirer CAR and Withdrawal Prediction**

This table reports results from regressions of acquisition outcome measures on acquirer cumulative abnormal returns (CAR) in samples with high versus low withdrawal probability. We first regress the completion indicator on characteristics using the first half of the sample. We then predict the cancellation probability for transactions in the second half of the sample. We sort transactions based on their completion probabilities into high withdrawal probability (the highest tercile) and low withdrawal probability (the lowest tercile). The dependent variable is a non-impairment dummy (Panels A and B), short-term abnormal ROA (Panels C and D), or long-term abnormal ROA (Panels E and F). In Columns (1)–(3), CAR is the only independent variable. For Column (4), in addition to CAR, we include characteristics as independent variables, and Column (5) further includes year and industry fixed effects. Column (6) only includes characteristics, and Column (7) includes year and industry fixed effects as well as characteristics as independent variables. The characteristics used as controls include the log of market capitalization, leverage, and free cash flow scaled by lagged assets, Tobin's Q, previous-quarter market-adjusted stock returns, relative deal size, and a set of dummy variables for stock-only, mixed-payment, diversifying, competed, hostile, and public target deals. Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

CAR window:	$[-1,1]$	$[-5,5]$	$[A-2, C+2]$	$[-1,1]$		n.a.	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Probability of Non-impairment – High W/D Probability ($N = 954$, DV: Not impair)							
CAR	−0.019	−0.036	0.011	−0.037	0.063	Controls	Controls
	(0.114)	(0.097)	(0.025)	(0.139)	(0.097)	only	only
Controls	–	–	–	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	−0.001	−0.001	−0.001	0.009	0.033	0.010	0.034
Panel B: Probability of Non-impairment – Low W/D Probability ($N = 954$, DV: Not impair)							
CAR	0.014	0.118	0.113	0.243	0.244	Controls	Controls
	(0.137)	(0.107)	(0.092)	(0.140)	(0.144)	only	only
Controls	–	–	–	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	−0.001	0.000	0.005	0.068	0.128	0.065	0.124
Panel C: ST Abnormal ROA – High W/D Probability ($N = 4,783$, DV: ST abROA)							
CAR	0.012	0.010	0.027**	0.034	0.022	Controls	Controls
	(0.023)	(0.012)	(0.010)	(0.023)	(0.026)	only	only
Controls	–	–	–	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.002	0.047	0.228	0.047	0.228
Panel D: ST Abnormal ROA – Low W/D Probability ($N = 4,786$, DV: ST abROA)							
CAR	−0.024	−0.020	0.000	−0.011	−0.007	Controls	Controls
	(0.030)	(0.012)	(0.006)	(0.028)	(0.025)	only	only
Controls	–	–	–	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.000	0.056	0.111	0.056	0.111
Panel E: LT Abnormal ROA – High W/D Probability ($N = 3,571$, DV: LT abROA)							
CAR	−0.045	−0.034	0.021	−0.015	−0.023	Controls	Controls
	(0.035)	(0.019)	(0.014)	(0.033)	(0.025)	only	only
Controls	–	–	–	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.001	0.001	0.045	0.274	0.045	0.274
Panel F: LT Abnormal ROA – Low W/D Probability ($N = 3,571$, DV: LT abROA)							
CAR	−0.026	−0.022	0.007***	−0.014	−0.007	Controls	Controls
	(0.039)	(0.030)	(0.002)	(0.031)	(0.029)	only	only
Controls	–	–	–	Char	Year, Ind, Char	Char	Year, Ind, Char
Adjusted R^2	0.000	0.000	0.000	0.062	0.133	0.062	0.133

Figure IA.B1: **Predictive Performance of CAR versus a Characteristics-Based Model**

Panel (a) reports the coefficients of OLS regressions of the non-impairment dummy on CAR$[-1, 1]$. Panels (c) and (e) are similar, except the dependent variable is abnormal ROA and a completion dummy, respectively. Panels (b), (d), and (f) report the adjusted R^2 from these regressions of acquisition outcomes on CAR (in red), and also the adjusted R^2 for similar regressions of acquisition outcomes on deal and firm characteristics (in blue). In Panels (a) and (b), in the Year 1 regression, the dependent variable is the non-impairment dummy. In the Year 2 regression, we exclude firms with impaired transactions within one year, and the dependent variable is the non-impairment dummy in Year 2. The Year 3 regression excludes firms with impaired transactions in Years 1 or 2, and the dependent variable is the non-impairment dummy in Year 3. Year 4 and Year 5 regressions are computed similarly. In Panel (c), we measure abnormal ROA at the end of Years 1, 2, 3, 4, and 5 following the deal completion. In Panel (e), we measure deal completion at the end of Years 0, 1, 2, 3, and 4 since the announcement. In Panels (a), (c), and (e), the light-shaded region indicates the 95% confidence interval.

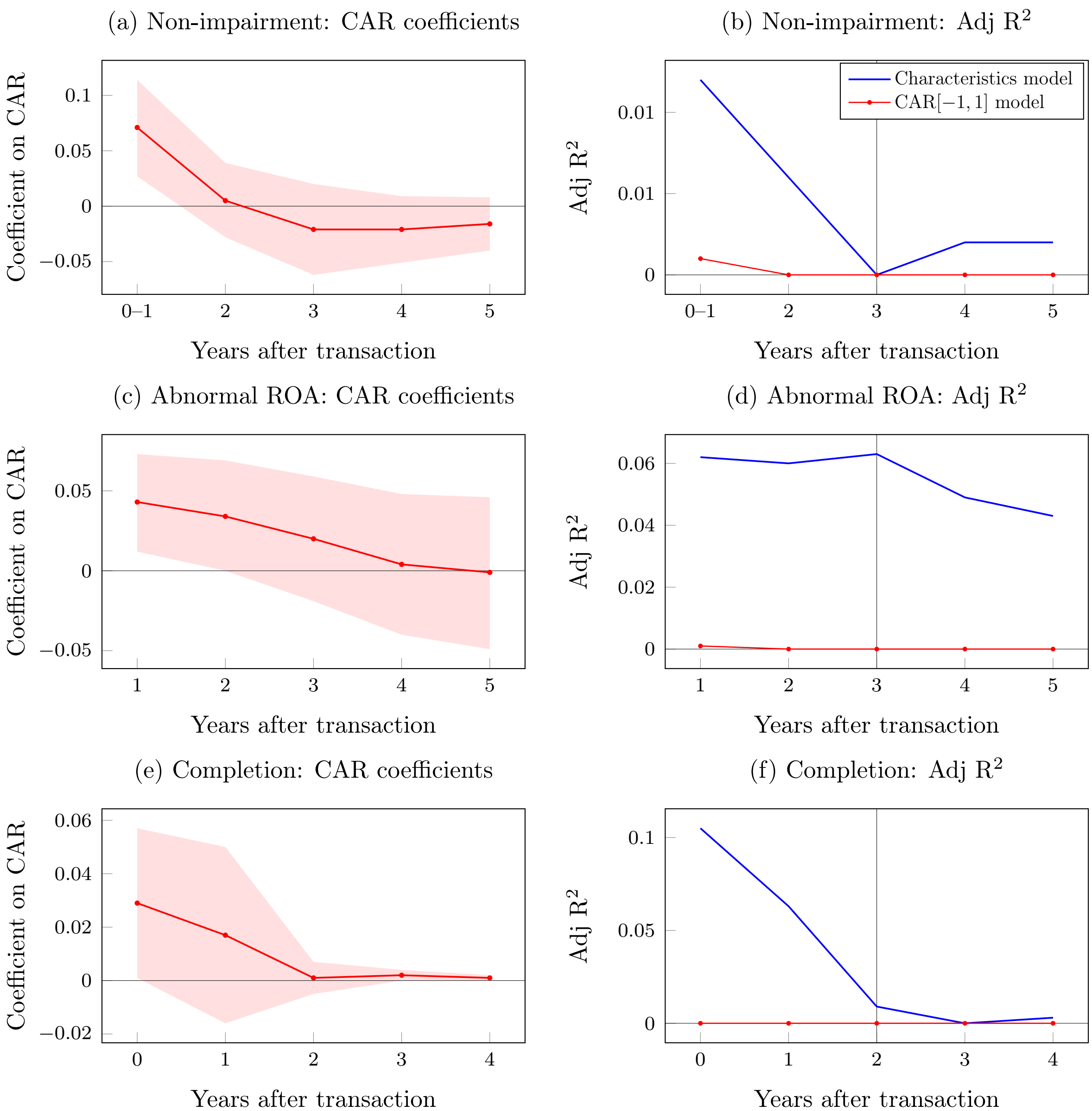


Figure IA.B2: **CAR and Acquisition Outcomes: Fama-French 12 Industry Classification**

This figure plots the coefficients and 95% confidence intervals for regressions of outcomes on CAR based on the specification in Table 2, Column (5), for each of the Fama-French 12 industries. Numbers 1 through 12 on the x-axis correspond to "business equipment," "chemicals and applied products," "consumer durable," "oil, gas, and coal extraction and products," "healthcare, medical equipment, and drugs," "manufacturing," "finance," "consumer nondurables," "wholesale, retail, and some services," "telephone and television transmission," "utilities," and "other," respectively. Panels (a), (b), (c), and (d) use a non-impairment dummy, short-term abnormal ROA, long-term abnormal ROA, and a completion dummy, respectively, as the key independent variable. The red dots represent the point estimates, and the light red shading represents 95% confidence intervals.

C Validation of Impairment as a Measure of Value Destruction

In this section, we validate goodwill impairment events as a signal of value destruction. To do so, we examine (a) the market's reaction to the news that the goodwill of a past transaction has been impaired; (b) distressed delistings following the impairment announcement; (c) the operating and financial performance of the impaired acquirers after the deal announcement; and (d) management turnover around the impairment announcement.

C.1 Market Response to Impairment News

We test whether investors perceive goodwill impairment as conveying negative news, i.e., they recognize that value has been lost. Our test replicates prior research in the accounting literature documenting that goodwill impairment events are value-relevant.[36]

We use Compustat quarterly data to identify the first quarter in which each transaction in our impairment sample experienced a goodwill write-down and the earnings announcement date for that quarter. Unique earnings announcement dates for an acquirer are included in the sample only once if multiple transactions experience a goodwill impairment announcement for a particular acquirer on the same earnings announcement date. We create three control samples. First, for the non-impairment sample, we generate pseudo-impairment dates on earnings announcements three years following the deal's effective date. (The mean time to impairment is about 3 years.) Our second control sample, "Matched Control Sample 1," comprises firms that announce earnings in the same quarter and have the same fiscal year-end and two-digit SIC code as the impaired firm. Our third control sample, "Matched Control Sample 2," comprises firms that announce earnings in the same quarter, have the same fiscal year-end and two-digit SIC code as the impaired firm, and are in the same market capitalization tercile as the impaired firm. To avoid estimating market model parameters in

[36] In tune with this literature, we interpret this result as a response to a revelation of *past* value destruction (e.g., Henning and Stock, 1997; Chen et al., 2004; Bens et al., 2011; Gu and Lev, 2011; Li et al., 2011).

pre- and post-acquisition periods, we compute market-adjusted returns using the Center for Research in Security Prices (CRSP) value-weighted index.

Table IA.C1 shows the results over four event windows. For the impairment sample, cumulative abnormal returns are negative and statistically different from zero for all four event windows (mean CARs range from −2.8% to −3.7%). The market response to earnings announcements for the three control samples is small and positive for all four event windows (mean CARs range from 0.0% to 1.7%). Importantly, the market response to earnings announcements containing goodwill impairment is statistically lower than the three control samples for all event windows. Although earnings announcements contain other information besides goodwill impairment news, the results suggest that the market considers goodwill impairment events to be bad news.

Table IA.C1: **Market Reaction to Goodwill Impairment News**

This table reports the mean cumulative abnormal returns (CAR) surrounding quarterly earnings announcement dates. For the impairment sample, we focus on the first earnings announcement when a goodwill impairment is announced for a particular transaction. Unique earnings announcement dates for an acquirer are included in the sample only once if multiple transactions experience a goodwill impairment announcement for a specific acquirer on the same earnings announcement date. For the Non-impairment sample, we generate "pseudo" impairment dates three years (the average time to impair) following the deal close date. We also create two matched samples of control firms that did not announce impairment news. "Control1" is a matched sample that includes firms that announce earnings in the same quarter and have the same fiscal year-end and two-digit SIC code as the impaired firm. "Control2" is a matched sample that includes firms that announce earnings in the same quarter and have the same fiscal year-end and two-digit SIC code and are in the same market capitalization tercile as the impaired firm. CARs are based on market-adjusted returns using the Center for Research in Security Prices (CRSP) value-weighted index. The event period is listed in brackets."Difference" refers to the differences between the Impairment and Control samples. Tests for differences are based on the t-test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively. "ns" denotes mean CARs or differences that are not statistically different from zero.

Sample:	Impairment	Non-impair	Control1	Control2	Difference ($t-$test)		
Window:	(1)	(2)	(3)	(4)	(1)−(2)	(1)−(3)	(1)−(4)
CAR[−1, 1]	−2.8% ***	0.3% **	0.2% ***	0.2% ***	−3.1% ***	−3.0% ***	−3.0% ***
CAR[0, 1]	−2.9% ***	0.1% ns	0.0% ns	0.0% ns	−3.0% ***	−2.9% ***	−2.9% ***
CAR[-5,5]	−3.3% ***	0.4% **	0.7% ***	0.8% ***	−3.7% ***	−4.0% ***	−4.1% ***
CAR[−10, 10]	−3.7% ***	0.7% ***	1.5% ***	1.7% ***	−4.4% ***	−5.2% ***	−5.4% ***

C.2 Acquirer's Distressed Delisting

Table IA.C2, Panel A, shows univariate statistics on the number of acquirer firms that exit the public market within 5 years of the deal's effective date. Public market exit data are obtained using the CRSP delisting code. Acquirers are categorized as "Merged/Went private" for delisting codes 200–390 and 573. Acquirers are classified as "Delisted" for delisting codes between 500 and 600 (excluding 573 and 574) and as "Bankrupt/Liquidated" for delisting codes 400–490 and 574. We retain only one observation when an acquirer in the impairment or non-impairment sample announces multiple transactions in the same year.

Table IA.C2, Panel A, shows that firms in the impairment sample are significantly more likely to be delisted and to go through a bankruptcy or liquidation process than firms in the non-impairment sample. In contrast, firms in the non-impairment sample are more likely to merge or go private. These findings imply that impairment is a good proxy for deal failure.

Table IA.C2: **Post-deal Performance for Firms with Goodwill Impairment**

Panel A reports univariate statistics on the number of acquirer firms that exit the public market within 5 years of the deal's effective date. Panel B reports median industry-adjusted accounting performance in the third year after the deal announcement. Tests for differences between samples are based on the Wilcoxon rank test. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Panel A: Post-deal Public Market Exits

Sample:	Impairment		Non-impairment		Difference	
	#	%	#	%		
Merged/Went private	124	15%	1,098	23%	−9.0%	***
Delisted	77	9%	113	2%	6.6%	***
Bankrupt/Liquidated	23	3%	35	1%	1.9%	***

Panel B: Industry-adjusted Accounting Performance During 3 Years After Deal

	Impairment sample	Non-impairment sample	Difference	
Sales growth	−5.3%	0.7%	−6.0%	***
COGS/Sales	2.0%	−1.5%	3.0%	***
SGA/Assets	−0.3%	−1.2%	0.9%	***
PPE Growth	−4.0%	1.5%	−5.5%	***
FCF/Assets	−3.1%	0.9%	−4.0%	***
ROA	−0.1%	1.1%	−1.2%	***
ROE	−7.1%	0.5%	−7.6%	***
Tobin's Q	−26.8%	−0.7%	−26.1%	***
Earnings/Price	−4.7%	0.8%	−5.5%	***

C.3 Acquirers' Long-term Performance

We examine industry-adjusted accounting and stock performance for the three years after the deal announcement. We retain only one observation when an acquirer in the impairment or non-impairment sample announces multiple transactions in the same year. We report the following median performance measures, adjusted by the median Fama-French 48 industry value: sales growth; cost of goods sold (COGS) scaled by sales; selling, general, and administrative expenses (SG&A) scaled by sales; property, plant, and equipment (PPE) growth; free cash flow (FCF) scaled by assets; return on assets (ROA); return on equity (ROE); Tobin's Q; and the earnings-to-price ratio.

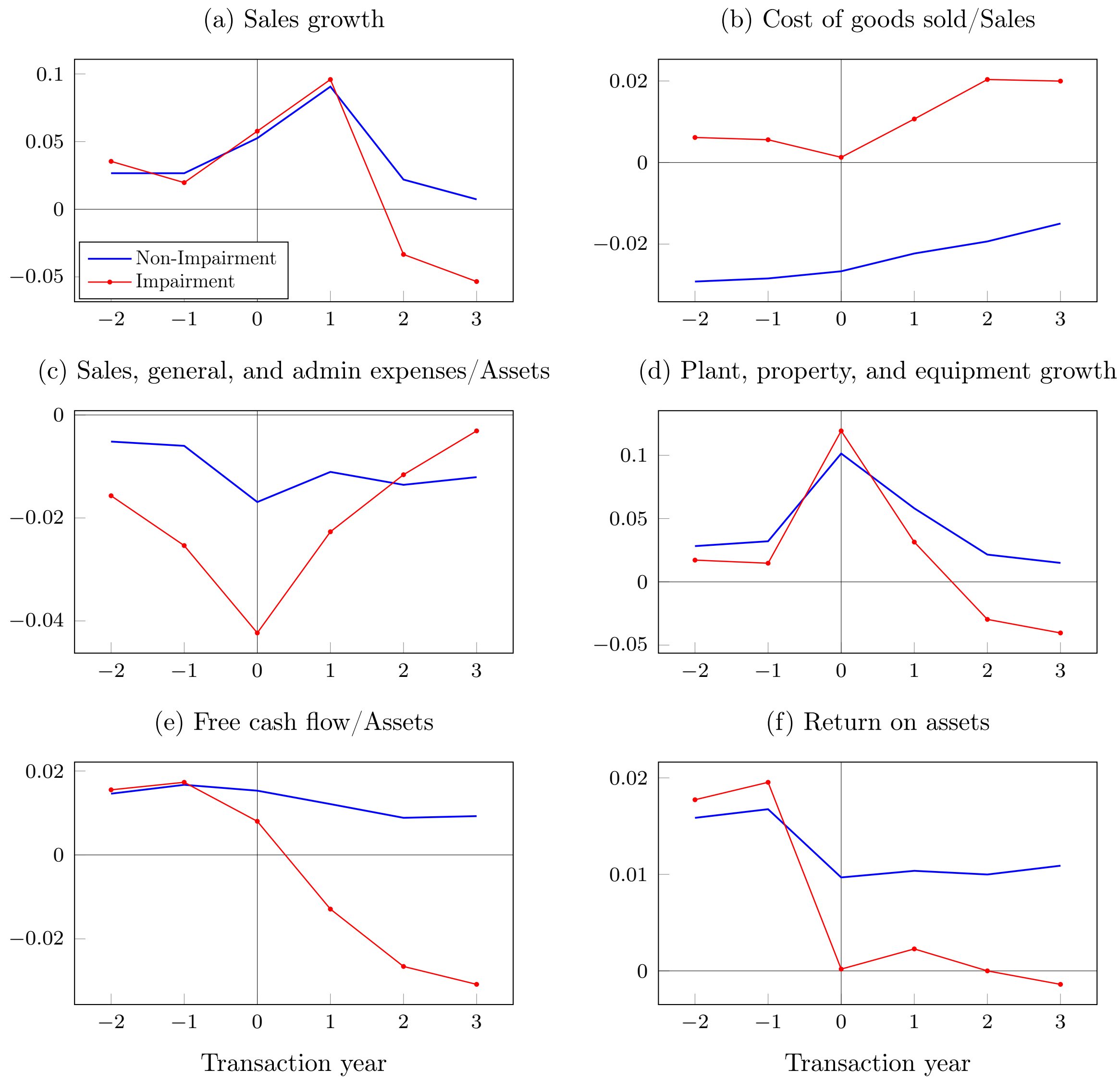
Table IA.C2, Panel B, reports median industry-adjusted statistics and tests of statistical differences between the non-impairment and impairment samples. We observe statistically superior performance for the non-impairment sample relative to the impairment sample for the three years following the acquisition announcement for all nine performance measures.

Figure IA.C1, Panels (a)–(f), show the operating performance from one year before to three years following the acquisition. Across panels, we generally observe that industry-adjusted performance measures begin to materially diverge in the years following the deal announcement for the impairment sample (red lines) and the non-impairment sample (blue lines), indicating that impairment firms encounter significant firm-level adverse shocks in the years following the acquisition. For many of the measures, the divergence begins in the year following the acquisition but widens further two years following the acquisition.

Figure IA.C2, Panels (a)–(d), show the financial performance from two years before to three years after the acquisition. Note here that the gap between the blue and red lines increases not so much before but after the deal announcement. Figure IA.C2, Panel (d), shows that the returns to the realized impairment sample remain relatively flat at the announcement but decline dramatically thereafter. Returns to the realized non-impairment sample continue their steady growth; consequently, the gap between the two subsamples widens.

Figure IA.C1: **Operating Performance and Goodwill Impairment**

The figure shows the industry-adjusted operating performance of acquirers that impaired goodwill (in red) relative to acquirers that did not impair goodwill (in blue). The period begins two years before the acquisition and ends three years after the acquisition. Panel (a) shows sales growth. Panel (b) shows the cost of goods sold/sales. Panel (c) shows sales, general, and administrative expenses/assets. Panel (d) shows plant, property, and equipment growth. Panel (e) shows free cash flow/assets. Panel (f) shows the return of assets.

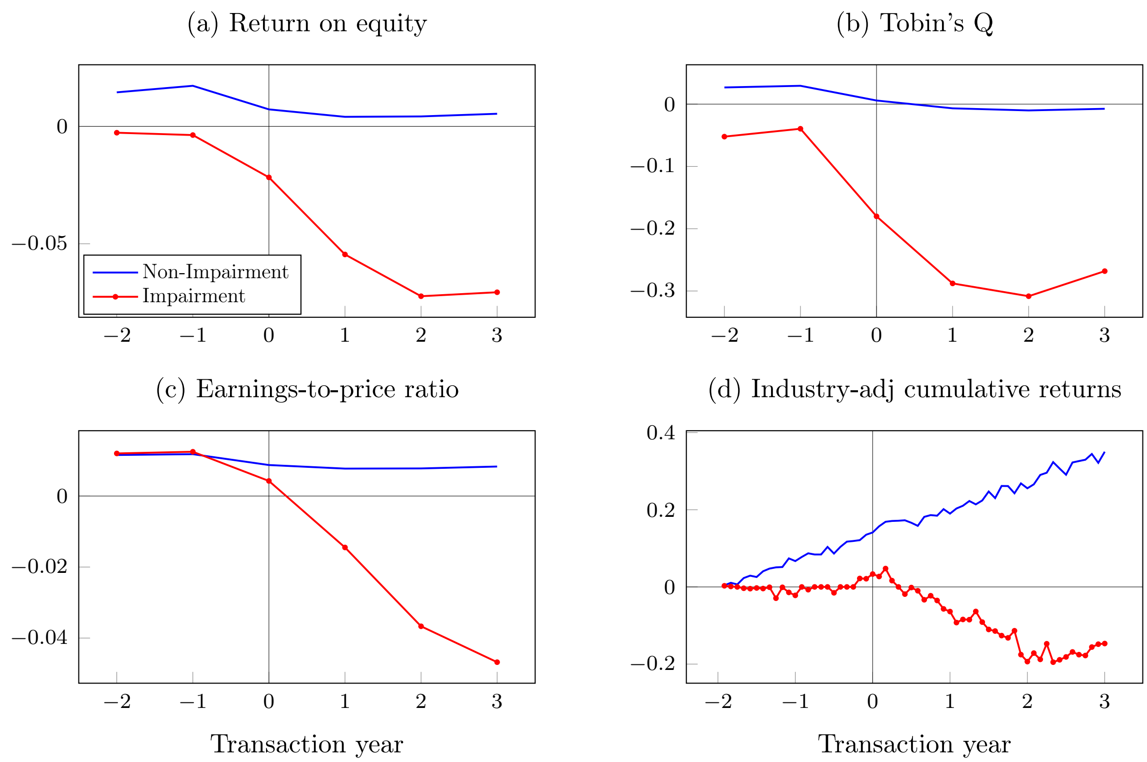


C.4 CEO Turnover Around Goodwill Impairment

We consider both the likelihood of CEO turnover following the deal and the timing of turnover for the impairment sample. In independent work, Cowan, Jeffrey, and Wang (2019) perform a similar analysis and reach the conclusion that goodwill impairment is a good

Figure IA.C2: **Financial Performance and Goodwill Impairment**

The figure shows the industry-adjusted financial performance of acquirers that impaired goodwill (in red) relative to acquirers that did not impair goodwill (in blue). The period begins two years before the acquisition and ends three years after the acquisition. Panel (a) shows the return on equity. Panel (b) shows Tobin's Q. Panel (c) shows the earnings-to-price ratio. Panel (d) shows industry-adjusted buy-and-hold cumulative returns.



indicator of CEO underperformance.

Unlike the previous tests in Appendix Sections C.1, C.2, and C.3, which utilize the full sample of 906 goodwill impairments, because turnover events require manual hand-collected data, we report the results for a smaller subsample of 355 impairments that was utilized in a previous version of the paper. This subsample only includes impairments between 2003 and 2013 and includes more stringent filters than the current version of the paper (e.g., excludes acquisitions by financial firms).

We track turnover events between the deal announcement and four years after the first impairment event. This analysis is conducted at the CEO-impairment level. If a CEO is associated with multiple impairment events, we retain only the transaction with the largest impairment amount. We identify three types of forced CEO turnover: (1) internal turnover (fired by the board), (2) takeover turnover, and (3) bankruptcy turnover. Turnover events are identified using proxy statements, press releases, and news articles in Factiva. We follow Weisbach (1995), Parrino (1997), and Lehn and Zhao (2006) in identifying turnover events. If the CEO is reported as fired, forced from his or her position, or departed due to unspecified policy differences, then the CEO is classified as experiencing an internal turnover event. If the CEO is under the age of 65 and the reason for departure is unrelated to death, poor health, or the acceptance of another position, or if it is announced that the CEO is retiring and yet the announcement is not at least six months before succession, then the CEO is classified as experiencing an internal turnover event. For firms that are acquired, if we cannot find evidence that the CEO retained a role in the acquiring entity, then the CEO is classified as experiencing a takeover turnover event. For firms that enter bankruptcy, if we cannot find evidence that the CEO retained his or her job during the bankruptcy process, then the CEO is classified as experiencing a bankruptcy turnover event.

Table IA.C3 presents results for the full sample of transactions in the impairment sample. We find that 45% of CEOs experience a turnover event between the deal announcement and four years following the impairment, indicating that close to half of the impairment sample CEOs are disciplined by the labor market. To provide a relative comparison, Jenter and Lewellen (2021) show that, unconditional on acquisition activity, on average, 12% of CEOs experience turnover in a given year. For acquiring firms (that may or may not experience impairment), Lehn and Zhao (2006) find a 47% CEO turnover propensity within five years of the deal announcement date.

However, our main interest is the turnover timing, which allows us to assess whether the CEO's departure results from the market's assessment of value destruction at the deal announcement or results from the subsequent impairment event itself. If value destruction is anticipated, CEOs should be more likely to be fired immediately following the acquisition announcement rather than the impairment. We find that 13% of impaired firm CEOs are

Table IA.C3: **Post-deal CEO Turnover for Firms with Goodwill Impairment**

This table reports univariate statistics for CEO turnover for the sample of firms experiencing a goodwill impairment. We track CEO turnover events between the deal announcement and four years after the first impairment event.

Impairment sample	#	%
% Turnover between deal announcement year and impairment year + 4	142	45%
Firms subject to internal turnover	118	38%
Firms subject to takeovers	19	6%
Firms subject to bankruptcy	5	2%
% Turnover year of or year after deal effective year (% of total sample)	19	13%
% Turnover year of or year after impairment year (% of total sample)	58	41%

terminated in the year of or year following the deal effective year, whereas 41% are fired in the year of or year following the impairment year.

To summarize, the results in Table IA.C3 indicate that the majority of turnover events in the impairment sample do not result from anticipated value destruction at the deal announcement but rather because of deal failure that is signaled by goodwill impairment. Specifically, CEO turnover events are three times more likely to occur immediately following the impairment than the deal announcement. This finding implies that the labor market considers impairment to be a proxy for deal failure.

To conclude, the results in Appendix C provide strong evidence that firms in the impairment sample experience all symptoms of deal failure—forced CEO turnover, delistings, bankruptcies, poor accounting, and poor stock performance—supporting our conclusion that goodwill impairment is a good proxy for deal failure.

D Algebraic Derivation of CAR's Variance

CAR (measured as a fraction of market capitalization) can be expressed as

$$\begin{aligned} CAR &= \frac{NPV}{MktCap} \quad (8) \\ &= \frac{NPV}{DealSize} \times \frac{DealSize}{MktCap} \\ &= NPVratio \times RelativeSize. \end{aligned}$$

The expected variance of CAR is as follows:

$$\begin{aligned} Var(CAR) &= Var(NPVratio \times RelativeSize) \quad (9) \\ &= Var(NPVratio)E(RelativeSize)^2 \\ &\quad + Var(RelativeSize) \times [Var(NPVratio) + E(NPVratio)^2] \\ &\quad + Cov(NPVratio^2, RelativeSize^2) - Cov(NPVratio, RelativeSize)^2 \\ &\quad - 2Cov(NPVratio, RelativeSize) \times E(NPVratio) \times E(RelativeSize) \end{aligned}$$

When $RelativeSize$ is constant, the variance of $RelativeSize$ is zero, and the covariance of $RelativeSize$ or $RelativeSize^2$ with any variable is zero. Therefore, Equation (9) simplifies to

$$Var(CAR) = Var(NPVratio) \times E(RelativeSize)^2. \quad (10)$$