



Does purchased goodwill create shareholder value?

Testing the representation of purchased goodwill and its interindustry differences using long-term post-acquisition returns.

Abstract

In this paper we examine the relationship between the purchased goodwill proportion (PGP) and the long-term stock performance of US acquirers and how this relationship is moderated by industry classification. Our final sample consists of 676 M&As in the period 2007-2017. We apply the Buy-and-Hold Stock Return (BHSR) approach, two variations of the Buy-and-Hold Abnormal Return (BHAR) approach as well as the Fama French Three Factor Model. First, we find that acquirers with higher PGP, on average, experience lower incremental post-acquisition stock performance. Second, we find that the effect of PGP on company performance depends on what industry the acquirer belongs to. Our study consequently confirms the general perception that higher purchased goodwill is “bad”, indicating that it is not a reliable measure of synergies and the going concern. Furthermore, certain industries such as Healthcare, Industrials, Financials, and Media and Entertainment significantly moderate PGP’s relation to stock performance negatively compared to our baseline industry Consumer Products and Services.

Keywords: purchased goodwill, mergers and acquisitions, long-term performance, industries

Authors: Philip Eriksson & Edith Wolff

Supervisor: Henrik Andersson

Table of Contents

1. Introduction.....	3
2.1 Theory	5
2.1.1 Goodwill.....	5
2.1.2 Mergers and Acquisition (M&A) Motives.....	7
2.1.3 Impairing Goodwill	8
2.1.4 Goodwill and Industry Differences	10
2.2 Literature Review.....	12
2.2.1 Overview	12
2.2.2 Goodwill Impairment Studies	14
2.2.3 M&A and Long-Term Abnormal Returns.....	16
2.2.4 Contribution.....	17
2.3 Hypotheses	18
3. Methodology	19
3.1 Research Design.....	19
3.1.1 Company Performance Proxies	19
3.1.2 Biases.....	21
3.1.3 Addressing Biases	23
3.2 Variables	24
3.3 Sample Selection and Data Collection	25
4. Findings and Analysis	27
4.1 Description of Data.....	27
4.2 Hypothesis Testing	29
4.3 Summary of Results	30
4.3.1 Hypothesis One	30
4.3.2 Hypothesis Two.....	36
5. Discussion.....	38
5.1 Discussion of Results	38
6. Conclusion.....	42
7. Bibliography	43
8. Appendix.....	45
8.1 Complementary Findings	45
8.1.1 Fama French Three Factor Model.....	45
8.1.2 Short-Term BHSR.....	48
8.1.3 Outlier Exemption Robustness Check.....	48
8.1.4 Three BCa Bootstrap Tables	50

1. Introduction

During periods of economic turbulence, such as the current one, high-profile acquisitions made during more favorable market conditions may face significant consequences. The Financial Times reported on January 15th, 2023, that several notable corporations are grappling with substantial write-downs on recent acquisitions due to increased economic uncertainty and rising interest rates, which have followed a surge of deal-making. In some instances, the losses incurred have been significant enough to entirely negate a company's quarterly profits. A noteworthy case in point is Teladoc Inc., which recorded goodwill valued at over \$14.5 billion in 2021 as a result of numerous acquisitions, representing almost 82% of its total assets (Teladoc Health, Inc, 2022/2023). However, in the following year, Teladoc experienced a write-down of over \$13.4 billion, accompanied by a share price reduction of over 30% after the news was released. This raises the question as to what the purchased goodwill truly represented and why it was subject to a later write-down.

Purchased goodwill is the premium an acquirer pays above the fair value of a target firm's net assets. The reasons why an acquirer pays this premium include, but are not limited to, building more valuable technology, gaining more market share, creating cost advantages, and expanding its services. In this sense, mergers and acquisitions (M&A) are a way for acquiring firms to invest their resources now to create shareholder value in the future. However, one factor for the success of an acquisition, like any investment, is dependent upon the price that is paid. Therefore, the purchased goodwill premium may also entail value-destructing elements like overbidding, overvaluation, and fair value mismeasurements that may cause the acquiring firm to later impair its purchased goodwill. Teladoc's management justifies the acquisitions that were made: "These [impairment] losses and accumulated deficit reflect the substantial investments we have made to expand our business and scope of services, acquire new clients and members, build our proprietary network of healthcare providers, and develop our technology platform.". Teladoc's investments may have been susceptible to a variety of factors that impeded the accuracy of its purchased goodwill, including overly optimistic growth projections and synergy effects, inadequate assessment of potential risks and uncertainties, or influence from market conditions prevailing at the time of acquisition (Boennen & Glaum, 2014).

In this paper, we aim to empirically study the representation of purchased goodwill as we attempt to answer the following questions: What is the relationship between the purchased goodwill proportion (PGP) and long-term stock returns, and how does this relationship vary between industries? In order to answer these, we conduct two tests. First, we use an OLS regression to test whether our variable of interest, PGP, has a relationship with the long-term stock performance of the acquiring firm. Second, we investigate if the effect of PGP on company performance depends on what industry the acquirer belongs to by introducing an industry dummy and PGP interaction term. We conduct this study on a sample of 676 acquiring firms listed in the U.S. between 2007-2017. The post-acquisition long-term returns of interest are one-, two-, three-, four-, and five-year cumulative returns, which we test separately. We focus on the U.S. market for several reasons. U.S. stock exchanges are some of the most developed in the world with a long history of M&A activity across a diverse range of industries. Therefore, the study can be applicable to a larger audience of investors, managers, analysts, and other stakeholders involved in U.S. markets. Furthermore, our sample of stock returns spans from 2008 to 2021, and all of our samples comply with the SFAS 141 "Business Combinations," and SFAS 142 "Goodwill and Other Intangible Assets" standards that were introduced in the early 2000s to better reflect the value of purchased goodwill over time (Burger & Wen, 2021). The goal of our results is subsequently to test if purchased goodwill is, on average, value-creating or value-destroying, and whether industry-specific factors influence this relationship. Finally, we will discuss what implications this has for investors as well as accounting regulators.

The paper proceeds as follows: we begin with relevant theory, including a discussion of goodwill, impairment standards, and mergers and acquisitions. Next, we discuss previous literature on goodwill impairment predictions and M&A long-term abnormal returns as they pertain to our study. Then we present our hypotheses followed by the methodology section which describes the various methods that we implement to test our hypotheses. This section also outlines the variables we consider, the sample selection and the data collection. In the findings and analysis section, we provide a description of our data and test our hypotheses. The results are then summarized, followed by a discussion of the findings.

2. Theory and Literature

2.1 Theory

2.1.1 Goodwill

Today goodwill is one of the most important, yet vague assets for users of financial statements to evaluate. Boennen & Glaum (2014) highlight that goodwill is at times the single largest balance sheet item and can therefore have large effects on the financial statements and financial ratios, which may be of interest to not only academics but also to investors as well as executives. We also find goodwill to be a meaningful component of the balance sheet of our sample firms. For example, between 2007-2016 the average goodwill-to-asset ratio was 17.7%, the highest year being 2016 with a ratio of 20.6%.

In general, goodwill can be divided into two main groups: internally generated goodwill and goodwill that arises as a result of business combinations (purchased goodwill). Bloom (2009) clearly points out that internally generated goodwill is not, and cannot be, accounted for on the balance sheet of a company, chiefly due to the inherent limitations of double-entry bookkeeping and historical cost-based accounting. In other words, if we give rise to internally generated goodwill on the asset side of the balance sheet then what will be the respective liability to match it? Furthermore, will we update the internally generated goodwill every year with its ‘fair value’, while the rest of the balance sheet is recorded at its cost-basis? While internally generated goodwill is thoroughly interesting and important for questioning the current accounting regime, in this paper we will focus on the second group: purchased goodwill. In brief, purchased goodwill amounts to the difference between the purchase price for the target and the fair value of the target’s net assets:

$$\text{Purchased Goodwill} = \text{Purchase Price}_{\text{Acquirer}} - (\text{Fair Value of Assets}_{\text{Target}} - \text{Fair Value of Liabilities}_{\text{Target}})$$

At the time of an acquisition, public firms must disclose how they have allocated their purchase price in a so-called purchase price allocation (PPA). The assets that the acquirer usually

purchases comprise of cash and cash equivalent, property plant and equipment, account receivables, and other identifiable intangible assets. Purchased liabilities usually comprise of debt, accounts payable, pension and postretirement-related benefits, and deferred taxes. Purchased goodwill therefore represents the *unidentifiable* intangible assets that are left, after accounting for the *identifiable* assets and liabilities.

Johnson & Petrone (1999) argue that purchased goodwill should, theoretically, purely represent the future cash flow that management expects will arise as a result of the acquisition. These cash flows arise for two main reasons, also known as the core goodwill components:

1. The going concern element of the acquired target, meaning that the company is more valuable than its assets separately. The organized collection of net assets can generate a higher rate of return when there are high barriers to entry in a particular industry or geographic location, which can lead to the emergence of monopoly-like profits.
2. The synergies from consolidating the acquirer and the target. Revenue-, cost-, and financial synergies are usually found to motivate purchased goodwill (Bruner & Perella, 2004). After a merger, the acquirer may benefit from its larger size and scope, for instance through cross-sales or improved customer engagement, economies of scale, and lower interest rates for loans supported by a larger balance sheet and larger cash flows than prior to the merger.

The problem though, the authors point out, is that other elements will also be brought up in the goodwill component because of current accounting standards. These elements include:

3. Mismeasurement or misstatement of assets acquired. The complexity of assessing certain assets sometimes cause inaccurate amounts being recorded in the financial statements.
4. The omission of assets or liabilities in the PPA that are difficult to measure or identify can have an impact on the calculation of goodwill. In particular, the fair values of certain acquiree assets or liabilities that were not recognized at the time of acquisition can lead to the over- or underestimation of goodwill.

5. Overvaluation/undervaluation of the consideration paid by the acquirer. One common measurement issue that arises in acquisition deals occurs when the acquirer uses stock to settle the transaction, and the value of the stock fluctuates between the announcement date and the effective date. Another example is when the bidder overvalues the target's future cash flows by setting a discount rate that is not representative of the inherent risk of the target during the DCF-valuation process.
6. Overpayment/underpayment by the acquiror. Overpayment may occur for instance when the price is driven up by bidding competition, whereas an underpayment may occur if the target is experiencing financial distress.

Due to these limitations, it is difficult to discern what the acquired goodwill truly represents. We believe that by measuring the relation between the PGP and long-term stock returns we will gain insight into how to think about the premium acquirers pay for target firms. We calculate the PGP in the following manner, based on an acquirer's PPA:

$$\text{PGP} = \text{Goodwill} / \text{Total Deal Value}$$

If purchased goodwill is value-creating, then we can expect that incremental increases in the PGP will have a positive relation to the acquiring firm's stock performance. On the other hand, if purchased goodwill is value-destroying, such that an incremental increase in the PGP result in decreased stock performance, then we expect other negative non-core elements to be present such as overbidding and overpayment.

2.1.2 Mergers and Acquisition (M&A) Motives

To better understand why purchased goodwill has both value-creating and value-destroying aspects, we must investigate firm motives behind M&As. In theory, mergers and acquisitions are considered investment strategies for firms, where the board of directors and management use deployable cash or stock to create shareholder value. Therefore, the success of a merger is usually measured by the increased value of the newly combined firm. Although M&A is often in earnest intended to create shareholder value, the investment strategy is subject to inadequate target selection, costly integration processes, biases, and an underestimation of risk. Since

management tends to focus on the total price of the acquisition rather than the premium paid, many M&A could result in overbidding (Mayrhofer, Hassan, & Ghauri, 2018). Not only do mistakes such as these naturally occur, but firms may also conduct M&A for reasons that are weakly associated with increased shareholder value. Trautwein (1990), for example, highlights empire-building as an M&A motive in which managers seek to maximize their own utility as opposed to the shareholders’.

Tarun et. al (2007) tests three overarching theories that explain management motives for conducting M&A: improved economic performance, personal benefits for managers, and increased market power (Mukherjee, Kiymaz, & Baker, 2004). The authors found, consistent with their expectations, that the most important motive for acquiring another company is the resulting synergy effects (37.3%), and the second highest-ranked motive was diversification (29.3%). Of the firms seeking synergy effects, 89.9% were seeking synergies from the operating economics of the combined firms. But, the authors also found a mismatch of cash flows and bad practice for measuring discount rates in their survey, which implied an overvaluation of target firms. A prime example of this is the persistence of using the buyer's WACC¹ to value the target's equity cash flow. Even though synergy effects were the most important motive for M&A, Goold & Campbell (1998) found that synergy-producing mergers tend to overpay for their targets since the risks and biases of the acquisition may be particularly prevalent when management has high synergy expectations.

2.1.3 Impairing Goodwill

Impairments occur when the carrying value of goodwill becomes less than the recoverable amount in a company’s cash generating unit (CGU)². This occurrence is caused by lower cash flow projections, higher discount rates, or both. Therefore, when management expects less discounted cash flows arising from an acquisition, then the “impairment-only approach” clearly communicates to investors that the company will deliver less-than-expected future cash flows.

¹ The weighted average cost of capital is the most commonly used discount factor companies use in discounted cash flow models.

² Cash generating units are defined as the smallest identifiable group of assets that generates largely independent cash inflows (IFRS IAS36).

Under previous standards of the US-GAAP and IFRS, goodwill underwent amortization throughout its useful life and was solely tested for impairment when an indication suggested that the carrying value was irrecoverable. In 2001 and 2004, however, the new accounting standards, SFAS 141 "Business Combinations," and SFAS 142 "Goodwill and Other Intangible Assets," popularly recognized as the "impairment-only approach," were introduced. Subsequently, goodwill impairment testing has become a mandatory procedure to perform at least once a year for public firms, with the frequency increasing if there is a basis to presume that an impairment may be necessary. The novel approach aimed to enhance the representational power of goodwill by providing financial information in its present state. This approach seems to resonate with investors as Burger & Wen (2021) found that SFAS 142 was successful in increasing the value relevance of goodwill in comparison to the previous amortization techniques.

FASB and IASB, however, have expressed a concern that acquired goodwill may be confused with internally generated goodwill in the new approach, which may cause write-downs to be hidden by other cash-generating factors (Boennen & Glaum, 2014). This may enable acquirers to avoid recognizing a write-down of goodwill by appointing it to a larger CGU. This CGU aggregation risk is found to vary across industries. Research by Carlin & Finch (2011) suggests that companies in the food and beverage, media, utility, and transportation sectors are particularly prone to consolidating fewer CGUs than business segments, while firms in the commercial services segment are more likely to define more CGUs than business segments. Nonetheless, when aggregated, sample firms define on average 0.91 CGUs per business segment. If management uses the CGU aggregation as a strategy for managing impairment risks and timing, it poses a risk of overstating earnings and net assets, understating leverage, and reducing reporting transparency, as cautioned by the authors. The impairment-only approach has another negative implication, namely that goodwill can easily become biased to the downside. Once goodwill is written down, it remains so indefinitely, causing it to potentially become undervalued on a company's balance sheet over time which makes it difficult for investors to assess its true value. Consequently, larger goodwill posts, such as Teladoc's, may inevitably be impaired during market downturns or moments of lower-than-expected performance, but will never be recovered when market conditions ease and operating performance increases.

2.1.4 Goodwill and Industry Differences

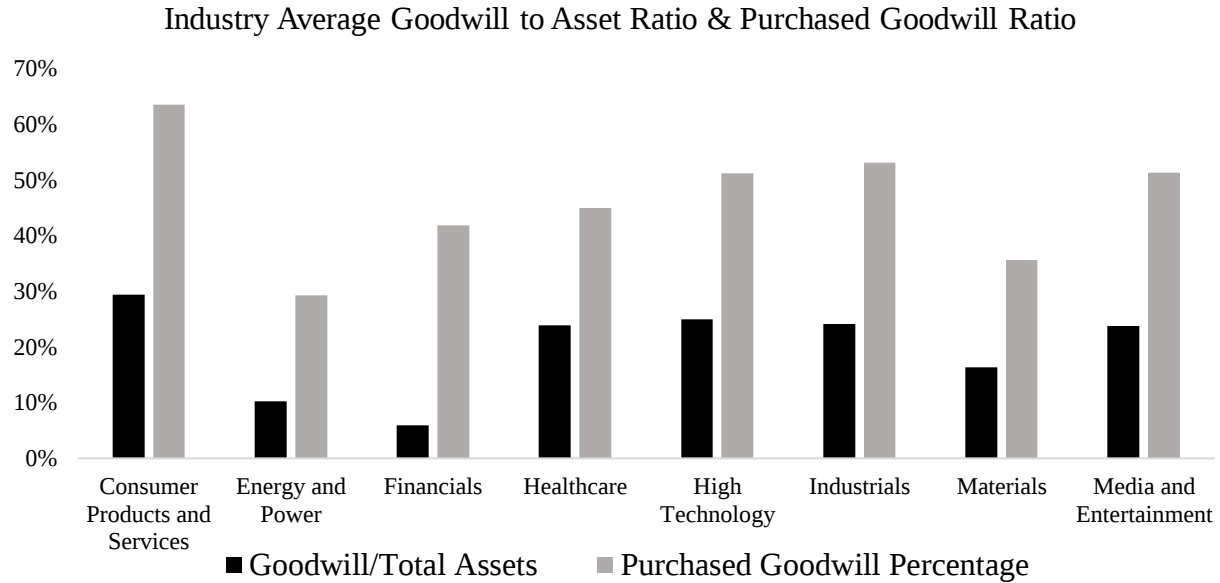
The amount of purchased goodwill a firm has may depend heavily on the industry it is operating within, despite all industries adhering to the same SFAS accounting standards for business combinations. Jarva (2009) finds that goodwill write-offs are unevenly distributed across industries. For example, in his sample 21.1% of write-offs occur in business services, 15% in electronic and other electrical equipment, but only 4% in engineering and management services. Given that a higher PGP at acquisition leads to a higher probability of experiencing a goodwill write-down, this could indicate that there are interindustry differences in the PGP (Olante, 2013).

We, indeed, find in [Figure 1](#) that there are stark differences in both the PGP and goodwill-to-asset ratio between industries. One explanation for the difference in PGP is the varying balance sheet composition between industries. For example, in our sample this can be seen in the Energy and Power industry, where tangible assets such as property, plant and equipment comprise the majority of the balance sheet values. Specifically, oil and gas target companies may own valuable land used for oil extraction in which case the acquirer tends to buy the target and realize little to no goodwill. Companies within asset-light industries, on the other hand, may be valuable due to their internally generated goodwill that is not presented on the balance sheet. Specifically, an industry such as High Technology does not generally have capital tied up in tangible assets, such as land, rather it derives value from *identifiable* intangible assets such as customer lists, trade names, and acquired technology as well as from their *unidentifiable* assets including brand image, market share, high growth rate, key personnel, and more. It is often such intangible and unidentifiable assets that generate large amounts of purchased goodwill. For example, Teladoc purchased HealthiestYou in 2016 with a 90.7% PGP, Facebook Inc purchased WhatsApp in 2014 with a 89.2% PGP, and Delta Airlines purchased Northwest Airlines Corp in 2008 with a 138.1% PGP.

The differences observed in PGP and goodwill-to-asset ratios among various industries indicate that certain industries may have a higher or lower tendency to recognize goodwill compared to others. The PGP and the goodwill-to-asset ratio may be linked to unique growth expectations in certain industries, as higher anticipated synergies and growth frequently result in higher PGPs. To measure industry growth, the book-to-market ratio (BM) can be utilized, with a lower BM indicating a high growth premium and a higher BM indicating a low growth premium (Fama &

French, 1995). [Table 1](#) displays the average growth premium of our acquirers based on their industry classification.

Figure 1



The average purchased goodwill proportion is calculated using our full sample of 676 observations. The goodwill-to-asset ratio is based on the 632 of 676 observations that were retrieved from COMPUSTAT. It is measured as the year-end total goodwill divided by total assets for the year when our sample firms acquired their target.

Table 1

Industry	Book-to-market ratio
Consumer Products and Services	0,44
Energy and Power	0,66
Financials	0,74
Healthcare	0,47
High Technology	0,45
Industrials	0,50
Materials	0,51
Media and Entertainment	0,49

We calculate the BM ratio using our sample firm's average BM ratios at the time of their acquisitions. Although the BM ratios may not be fully indicative of the population, we believe that they give a good representation since we acquired a large sample for each industry.

Target firms within industries of high growth premiums receive a high valuation for cash flows that can be realized far out in the future, as opposed to in the short term. And since targets are usually valued based upon discounted future cash (DCF) projections, then companies belonging to industries with high growth premiums will have cash flows and synergies that are highly uncertain, which should be reflected in increased discount rates in the valuing of the firm. One common fallacy, however, is that bidders fail to acknowledge the inherent risk and hence overvalue the target by underestimating the discount rate. This may be especially prominent when the target is deemed riskier in nature, having less certain cash-flows, more liabilities and operating in uncertain markets. This will, with the same reasoning as before, likely underestimate the risk profile of the target. As discount rates unproportionally affect the terminal value in DCF models in comparison to the forecasted years, an inaccurate risk adjustment can have significant effects on the final valuation. Therefore, high growth industries with higher uncertainty may potentially be more subject to this discount bias, which in turn increase the likelihood of goodwill impairments in colder business climates and a higher sensitivity to post-acquisition stock performance.

Another related explanation for the industry differences in PGP is that certain industries may have stronger competitive dynamics than others, which can potentially lead to overbidding. Porter (1979) explains that “As an industry matures, its growth rate changes, resulting in declining profits and (often) a shakeout”. In this way, as the market saturates, market share becomes increasingly important, and consolidation may become vital for growth. This may give rise to heightened bidding competition and lead to a target becoming overpriced. The airline industry has strong competitive dynamics surrounding flight routes so although Delta Airlines motivated its high goodwill premium with the substantial revenue synergies it would receive from Northwest Airlines, the purchase could have also been prone to a bidding war in which the premium may not be justified.

2.2 Literature Review

2.2.1 Overview

Table 2 presents the key findings and information derived from benchmark studies relevant to our research. We group the studies into two categories: goodwill impairment studies and M&A studies. After the table we review the literature in more detail, discuss the limitations of the previous studies as they pertain to our study, and conclude with our contribution.

Table 2

Study	Data	Contribution	Methodology	Select Results
Hayne & Hughes (2006)	1276 U.S. publicly traded firms, 1988-1998	Compare the difference in premium paid between a write-off and a non-write-off group	Multiperiod binary logit model	Premium paid, the goodwill proportion, and paying with stock instead of cash all significantly increased the probability of a write-down occurring.
Olante (2013)	929 U.S. publicly traded firms, 1999-2007	Compare the difference between a write-off and a non-write-off group under SFAS 141 & 142 standards	Logistic Regression	Risk of impairment is 81% higher when the acquirer had more than 67% in PGP compared to acquirer with 46% in PGP
Jarva (2009)	327 goodwill write-off instances, 2002-2005	Test whether goodwill impairments arise due to earnings management or genuine cash flow expectations	OLS regression and logistic model	Write-offs have a significant positive predictive ability for expected one and two-year-ahead cash flows

Loughran & Vijh (1997)	947 U.S. acquisitions, 1979-1989	Measure pre- & post-acquisition long term abnormal returns	Control firm return approach	Moderately significant -25.0 % five-year abnormal return for stock mergers and significant 61.7% return for cash tenders
Rau & Vermaelen (1998)	3517 NYSE /AMEX/Nasdaq acquisitions, 1980-1991	Measure post-acquisition returns for ‘glamour’ firms vs. ‘value’ firms	Control firm bootstrapping approach	Value firms earn significant positive abnormal returns while glamour firms earn significantly negative abnormal returns
Dutta & Vijay Jog (2009)	1300 Canadian acquisitions, 1993–2002	Compare and test robustness between different long-term buy and hold abnormal return methodologies	Reference portfolio, control firm, and Calendar-time portfolio approach	Control firm and calendar-time approach are superior to BHAR approach.

2.2.2 Goodwill Impairment Studies

Hayne and Hughes’s (2006) and Olante’s (2013) findings are relevant to our study because they find empirical evidence that the PGP is a significant leading indicator of goodwill write-offs. Furthermore, Olante (2013) finds that goodwill impairments are most likely to occur two to three years following the acquisition given the new SFAS 142 standard while impairments were most likely to occur four to five years post-acquisition before the new standard was enacted. Hayne and Hughes (2006), on the other hand, find that most impairments occur six to eleven years

following the acquisition which may be indicative of their sample being more representative of the accounting standards before SFAS 142. Furthermore, Jarva (2009) lays an empirical foundation for what goodwill impairments represent. He confirms that write-offs, in accordance with SFAS 142, are indeed associated with expected future cash flows, rather than earnings management. Specifically, he finds that the impairment charge is predominantly associated with a company's subsequent one-to-two-year cash flows. The implication of the three studies, put together, is that a higher PGP at acquisition will increase the probability of impairment in the second and third years following the acquisition. The impairment, in turn, is found to predict following negative one- and two-year cash flow performance, which in this context would affect the observed acquirers' fourth- and fifth-year cash flows after the effective date. This is a phenomenon that could also be prevalent in our study, as it could explain poor stock performance.

While this group of studies is focused on the impairment of goodwill, we are focused on a holistic view of company performance. We believe that stock performance is the best proxy for company performance because it will incorporate both the positive and negative aspects of purchased goodwill. Tuomo Vuolteenaho (2002) decomposed the stock return of a firm into changes in cash-flow expectations and discount rates, which implies that stock performance should capture positive synergy and going concern news as well as negative impairment news. Impairment prediction studies will not fully capture all cash-flow expectations due to two main limitations. For example, a goodwill component generating less than expected future cash flows can be concealed by other cash-generating factors and hence avoid being impaired, while an efficient stock market on the other hand will price in any negative news that may relate to the acquisition. This means that poor performance may be priced into the stock price when it could otherwise be masked in terms of impairment tests. Limitations also apply toward the upside. As previously mentioned, once goodwill is written down it is done so indefinitely which implies that recoveries in company performance will not be captured in impairment studies. Albeit this rebound in performance is likely to be reflected in the stock price as priced-in future cash flows increase again. Subsequently, stock performance reflects both positive and negative cash flow news which is another reason why it serves as a suitable holistic proxy for company performance. But, given that the breadth of cash flow news is so wide, the indicator will also

capture cash flow news that are not related to the PGP and the acquisition event. This limitation arises from the use of stock performance, as it can exhibit misleading statistical relationships, emphasizing the necessity of a reliable asset pricing model.

2.2.3 M&A and Long-Term Abnormal Returns

Bruner (2005), in a review of existing mergers and acquisition studies, exemplifies that the outcome of M&A activities is incredibly mixed. Nonetheless, he notices that the well-accepted notion that M&As are value-destroying is in general false. Shareholders of acquiring firms generally earn the required rate of return on investment, while acquisitions with certain characteristics can even lead to increased shareholder value. For example, M&A that involve private targets, cash payments as opposed to stock, or are structured as earnouts have shown to increase shareholder value. Since a company's stock returns reflect the market's expectations of the potential benefits or risks associated with an acquisition, a large portion of M&A studies use stock returns as a proxy for company performance (Loughran & Vijh 1997, Rau & Vermaelen 1998, Dutta & Vijay Jog 2009). But despite the large number of M&A studies, none have considered the purchased goodwill proportion as a factor that could be related to post-acquisition returns. From a theoretical perspective, this may be because the PGP variable is likely not priced into the stock market since it can be obscure and difficult to find. When an acquisition is announced the deal value and purchase method are normally disclosed by the acquirer, but the PPA that the acquirer makes may not be finalized for several quarters after the announcement date. Furthermore, companies disclose the PPA in SEC filings only and some companies don't disclose PPAs for individual acquisitions, but they rather group several acquisitions into one PPA. These factors make it difficult for investors to evaluate the PGP and may be a reason for why previous studies may have neglected this variable.

Instead, previous studies have primarily focused on the relatedness, mode, geographical proximity, and form of payment of acquisitions. Loughran and Vijh (1997) test and confirm the theory that acquirers tend to conduct mergers with stock when management believes their own company is overvalued whereas they conduct cash tenders if their stock is undervalued. A similar finding is discussed by Rhodes-Kropf, Robinson, & Viswanathan (2015) who conclude that "[m]erger intensity is highly positively correlated with short-run deviations in valuation

from long-run trends, especially when stock is used as the method of payment.”. They even argue that economic shocks to individual industries as well as the to the general market might be the catalysts for merger activity, but the effect of market misvaluations influence which companies acquire which, and how the transaction is financed.

Rau & Vermaelen (1998) also find overly optimistic management to be the most likely cause of negative post-acquisition abnormal returns. They found that value (high BM) firms obtain a statistically significant positive abnormal return of 8% in mergers and 16% in tender offers over a three-year post-acquisition period. On the other hand, glamour (low BM) firms earn a statistically significant negative abnormal return of -17% in mergers and a positive insignificant abnormal return of 4% in tenders over the same period. Dutta & Jog (2009) find that stock-financed deals as well as relatively large acquisitions significantly underperform with regards to the three-year buy and hold abnormal returns (BHAR). Furthermore, they find that acquirers with director ownership above 25% significantly improve BHAR whereas CEO ownership of between 5%-25% lends significantly negative BHAR. The reason high director ownership leads to higher BHAR may be due to an increased alignment of motives with shareholders, coupled with checks and balances between directors which leads to more prudent and long-term dealmaking. Although the authors believe that the BHAR method is biased, they find similar results regarding director ownership and the form of payment using the Fama-French Three Factor Model.

2.2.4 Contribution

Our study contributes to the existing literature in several ways. The purchased goodwill proportion is relatively unexplored, and hence the availability of structured, accurate data of this variable is not available in any database known to us. Therefore, we believe that the 961 data points we manually collected for U.S. acquirers’ purchased goodwill proportion between 2007-2016 can serve as a foundation for further research on goodwill theory. While previous studies have tested the relationship between the PGP variable and goodwill write-downs, no study has yet tested the PGPs relationship with a holistic firm-performance measure: stock performance. Because PGP is found to relate to goodwill write-downs and since goodwill has become an increasingly sizeable asset of a company’s balance sheet, we believe this factor is worth

examining in more detail. Finally, interindustry differences in the sensitivity to purchased goodwill are as of yet unexplored and our test can provide insight into the reason for why purchased goodwill arises and why certain industries might be more or less sensitive to the PGP.

The overarching aim of our study is to provide investors and accounting regulators with empirical insights into how the markets truly value purchased goodwill in comparison to what theory states. Our focus on the representation of goodwill on its own as well as between industries can be useful for investors who are evaluating M&As on the public markets. Furthermore, we believe that our findings will be relevant for evaluating whether accounting standards for business combinations provide the best representation for what an acquirer pays for during an acquisition. In summary, this study holds the potential to provide valuable insights not only for future research but also for various stakeholders utilizing financial statements, including individual investors and accounting regulators responsible for business combination accounting standards.

2.3 Hypotheses

We would first like to establish an understanding of what purchased goodwill, on average, represents. Based on Johnson and Petrone's (1999) theory, goodwill can represent positive elements such as synergies, and the going concern cash flows; goodwill can also represent negative aspects such as overvaluation, overbidding, and fair value mismeasurements. For this reason, we test how the PGP impacts publicly listed U.S. acquirers' post-acquisition stock returns. We recognize that assessing a target's value is a difficult procedure subject to biases and risks of overbidding, which may lead to negative stock returns. Nonetheless, purchased goodwill should, in its ideal state, only represent the positive going concern and synergy elements that subsequently correspond to the expected future cash flows arising from the acquisition. In order to test whether purchased goodwill is value-creating or value-destroying, our null hypothesis is that the PGP has no impact on acquirer post-acquisition returns. If the PGP has a significant positive relationship with post-acquisition returns, then purchased goodwill will on average be value-creating rather than value-destroying and vice versa.

H0: The purchased goodwill proportion has no impact on the long-term stock performance of acquirers.

Second, we are interested in whether the effect of PGP on company performance depends on what industry the acquirer belongs to. Specifically, we are interested to find if industry classifications *alter* PGP's relation to company performance, compared to the absolute impact itself. We notice that the PGP, goodwill-to-asset ratio and book-to-market averages markedly differ between industries, but since the topic of interindustry differences and their moderating effect on purchased goodwill is an uncovered area in prior literature, we do not anticipate a significant difference in sensitivity to the PGP across industries.

H0: The purchased goodwill proportion's effect on long-term stock performance does not differ across industries.

3. Methodology

3.1 Research Design

This section is dedicated to presenting our distinct approaches for testing each hypothesis. We will explain the potential biases that may arise during the application of these approaches and conclude with the measures we take to address these biases.

3.1.1 Company Performance Proxies

To test our hypotheses, we use ordinary least squares (OLS) regression models, each complemented with a bootstrapping approach. These models will serve to test the average impact of the PGP on post-acquisition returns and if there is an interaction effect between the PGP and the acquirer industry. The OLS regressions explain the one, two-, three-, four-, and five-year post-acquisition performance by including our variable of interest as well as various control variables described later in the Variables section. We use three different measures as a proxy for company performance:

1. Buy-and-hold stock returns (BHSR)

2. Buy-and-hold abnormal returns ($BHAR_{BMS}$), book-to-market and size reference portfolio
3. Buy-and-hold abnormal returns ($BHAR_I$), industry reference portfolio

While we use the BHSR approach to test both our first and second hypotheses, we use the BHAR approach only to test the first hypothesis. We also perform a complementary method to our first hypothesis, the Fama French Three Factor Model, to test abnormal returns between three PGP groups (Low, Medium, High) while considering the risk profile for each of the groups. The methodology and findings for the complementary model are presented in the [appendix](#). Below is a detailed description of how we measure the stock performance for each of the approaches.

BHSR

The first method we use to calculate post-acquisition performance is a buy-and-hold stock return whereby we compute the raw stock returns of our acquiring firms. The raw return (R_{it}) for a given security (i) is measured each year during periods $t=1$ to $t=5$. The return is calculated using the stock price from the date that the acquisition is effective, plus one day.

$$R_{it} = (Stock\ Price_{it} - Stock\ Price_{i,t=0}) / (Stock\ Price_{i,t=0}),$$

BHAR_{BMS}

The second method we use to measure post-acquisition performance is a buy-and-hold abnormal return, where we match and subtract each of our sample firm's BHSR with the returns of a corresponding BM and size reference portfolio. The $BHAR_{BMS}$ accounts for the opportunity cost of holding a given security in our sample based on its BM and size (Barber, Lyon, and Tsai 2003). In order to match each of our observations with a corresponding reference portfolio, we begin by splitting our sample into six groups in the same fashion as Kenneth French. The sample is first split between big and small firms, based on the median market size of the sample. Then each of the two groups is split into three separate groups based on the 30th percentile and the 70th percentile of the sample's BM ratios. Firms with a BM ratio in the lowest 30th percentile are classified as Low, firms in the 30-70th percentile are Neutral, and firms in the 70th percentile are High. As a result, this creates six separate portfolios with the same number of observations in each: small-low, small-neutral, small-high, big-low, big-neutral, and big-high.

$$AR_{it(BMS)} = R_{it} - R_{BMS,t}$$

$AR_{it(BMS)}$ is the $BHAR_{BMS}$ for acquirer (i) in period (t) and is calculated as the difference between a sample firm's BHSR (R_{it}) and the return of a portfolio with a matching BM and size ($R_{BMS,t}$) in the corresponding period (t).

$BHAR_t$

The third method we use to measure post-acquisition performance is a buy-and-hold abnormal return based on our sample firms' industry classification. We implement a matching procedure that corresponds with the $BHAR_{BMS}$ method. We first classify each of our firms into one of the Fama-French 49 industries and proceed to match each of our firm's returns with the returns of the corresponding industry portfolio.

$$AR_{it(I)} = R_{it} - R_{I,t}$$

3.1.2 Biases

In this section, we will discuss biases that may arise from using the BHSR and BHAR approaches, whereas a similar discussion about the Fama-French Three Factor Model is found in the appendix. Finally, we discuss the ways in which we respond to these limitations.

The BHSR and BHAR method share four chief biases. The first bias is having a *bad model of asset pricing*. This bias occurs when an asset pricing model excludes relevant factors or includes irrelevant factors that explain stock returns. This can lead to a degree of misestimation of the expected returns and/or the relationship between the variables of interest and the dependent variable, which means that the beta coefficients of our independent variables PGP and PGP x Industry in the regression models may fail to capture a true relationship. Second, gathering an M&A sample over a long period of time will result in the sample including the same acquirer several times. There are stock-listed companies that make many acquisitions, sometimes referred to as serial acquirers. If all acquisitions of the same acquirer are included in the sample that will lead to a *cross-sectional dependence* bias. Third, the methodologies are subject to a *skewness bias*. Raw stock returns as well as abnormal stock returns are positively skewed, which is a result of some stocks significantly outperforming the rest of the market. As a result stock returns do not

follow a normal distribution, rather the distribution has positive long tails (Barber, Lyon, & Tsai, 2003). The fourth bias, which we have identified in our data is a *survival bias*. Since we are measuring stock performance over a five-year period, some of our observations become delisted over time. Delisting could be an indication of poor company performance leading up to bankruptcy or a target turnaround opportunity for another firm if the company is mismanaged, but it could also be caused by a strategic decision of a well-performing venture. Bessler et.al (2022) conducted a study on why firms exit the German stock exchange Prime Standard and found that out of their 136 exiting public firms 61 firms merged, 53 were insolvent, and 22 firms went private. Similarly, we might expect our surviving firms to induce a positive bias in the later return periods leading to abnormal returns that are not representative of the total population of acquirers.

Since the BHAR methodology involves portfolio construction, it has three additional biases. The first is the *new listing bias*. This refers to the fact that a randomly chosen sample of acquirers tends to have been listed for some time while a benchmark portfolio may include newly listed firms without pre-event return data which, on average, underperform the market. Subsequently, our sample likely consists of firms that have a track record before their acquisition, while the reference portfolios include companies that recently began trading. Second, the BHAR method is subject to a *calendar clustering bias* when multiple acquirers share the same date of acquisition. This could create a bias where several firms in our sample are matched with the same reference portfolio and hence have the same ‘opportunity cost’ even though they may differ completely. Third, our BHAR method is subject to a *rebalancing bias* that arises because the Kenneth French portfolios are rebalanced monthly based on a given firm’s new BM and size or yearly based on industry classification, whereas our sample is not rebalanced. When a firm gets a new BM or size profile, the Kenneth French 2x3 portfolio is rebalanced. Furthermore, when a firm changes the industry it operates within, then the rebalanced Kenneth French portfolio will account for that. However, our portfolios are only classified into BM and size, as well as industry portfolios at the time the acquisition was effective. This implies that the abnormal returns may not be representative over time, as firm classifications may change. We would like to note, though, that the $BHAR_i$ is not as subject to the rebalancing bias as the $BHAR_{BMS}$.

3.1.3 Addressing Biases

First, we have addressed the possibility of having a *bad asset pricing model* by implementing control variables that are highly relevant to post-acquisition returns, which have also been widely tested in previous studies. Furthermore, by testing our variable of interest with the three methods mentioned above, as well as the Fama-French Three Factor Model, we can draw more robust conclusions since we are not limited to the results of one method. Second, we have completely alleviated the *cross-sectional dependence bias* of stock returns by removing duplicate acquirers from our sample. Third, in order to reduce the *skewness bias*, we run each regression through a BCa bootstrapping procedure. By using the estimate for the original data and adjusting for skewness in the bootstrap distribution, this approach is superior to the Percentile and the Residual bootstrapping approaches for our data since we cannot assume linearity, normality, stationarity, and time independence of our observations. Therefore, we believe this bootstrapping approach with 1999 replications, by offering an alternative to such problems associated with standard *t*-tests over long horizons, can give more prudent results. Fourth, we do not have a specific mitigation method for the *survival bias*. But, since we measure the PGP in relation to post-acquisition returns for multiple periods, the later yearly returns may only be affected by this bias, whereas three-month and one-year returns, for example, will not be affected.

Although we were unable to eliminate recently listed firms or adjust for the variability of firm inclusion and exclusion in the databases used to generate rebalanced reference portfolios and sample data, we were still able to use the BHSR method to compare our variable of interest against raw returns. As a result, we were able to address our hypothesis with a comparable level of precision while sidestepping the potential issue of the *new listing bias* in our BHAR sample. Moreover, we have minimized the *calendar clustering bias* by dividing our sample into six Fama-French BM and size groups as well as groups based on the Fama French 49 industries. Creating more concentrated groups reduces the number of overlapping observations matched to the same benchmark groups but does not alleviate the bias completely. Finally, while previous studies (Rau & Vermaelen 1998; Dutta & Jog 2009) have implemented monthly rebalanced portfolios to alleviate the *rebalancing bias*, we do not do so in the scope of this study. This bias impacts the BM and size BHAR as well as the Fama French Three Factor Model, but industry BHAR_{*i*} to a large extent avoids this. The Kenneth French industries are rebalanced yearly based

on a firm's SIC code but since firm classifications are sticky, the amount of rebalancing made will be limited. For example, none of our sample firms were reclassified into a new industry during the period 2007-2016.

3.2 Variables

The dependent variable in our study are variations of stock return which were presented earlier in this section. To be clear, the stock returns are measured strictly as a capital gain and do not represent the total gains that an investor may experience since dividends are not included.

Our independent variable of interest is the purchased goodwill proportion (PGP). To complement this independent variable, we include industry and year fixed effect variables as well as several control dummies in our Full model. The eight industries in our sample are Consumer Products and Services, Energy and Power, Financials, Healthcare, High Technology, Industrials, Materials, Media and Entertainment. Furthermore, we include year-fixed effects for each year in our sample period 2007-2016. We believe accounting for differences in years is especially appropriate during this time interval as it includes a bear market following 2007 and the culmination of a bull market in 2021.

We have chosen the majority of our remaining control variables in accordance with Loughran & Vijh (1997), Rau & Vermaelen (1998), and Dutta & Jog (2009). The control dummies are Related/Unrelated, Merger/Tender, Cross-Border/Domestic, Cash/Mix/Shares. If the acquirer purchased a company within the same industry that it operates in, then the acquisition is considered related, otherwise, it is considered unrelated. Moeller, Schlingemann, & Stulz (2004) find that conglomerate firms generally have lower post-acquisition performance, thus in a similar fashion, we expect that firms that acquire targets within the same industry have better estimates of the valuation of the target and can better measure the potential synergies compared to when the target is unrelated. Furthermore, the authors compare bidder gains between cross-border and domestic acquisitions and find that acquiring cross-border targets relative to domestic targets lead to significantly lower announcement stock returns by approximately 1%. Subsequently, we believe the effect may also hold for long-term stock performance. We control for this impact by including the Cross-Border/Domestic dummy, in which a Domestic acquisition occurs when the

acquiring firm purchases a firm that has its headquarters in the United States and is considered Cross Border if otherwise.

Consideration Type is coded as either Tender or Merger. A tender offer is defined as a public offer made by an acquiring company to purchase a significant stake in a target company directly from the target company's shareholders. In a tender offer, the acquirer typically bypasses the target company's management and board of directors, which is different from a merger or disclosed value deal where the acquirer negotiates with the target company's management and board of directors to agree on the terms of the deal. Loughran and Vijh (1997) found that mergers experience a significant abnormal return of about -15.9 percent, whereas tender offers earn 43.0 percent during a five-year period after acquisition. Lastly, in our study, the Consideration Category can take on three different variables: Cash, Stock, and Mixed. Cash is coded for acquiring firms that pay solely with cash, Stock is coded for firms that acquire using their own stock, whereas Mix represents firms that pay with both cash and stock.

3.3 Sample Selection and Data Collection

In this study, we use Refinitiv Eikon to gather a specific set of acquisition data from January 1, 2007, to December 31, 2016. The deal screener includes various filters such as the disclosed dollar value of the deal, a market capitalization four weeks before the acquisition, and the acquirer's ultimate parent nation that must be the United States. Additionally, we retrieve both the acquirers' macro industry as well as the acquirers' industry based on their SIC codes. The Real Estate and the Government Agency industries are excluded, and the form of the deal is limited to mergers and acquisitions. The shares acquired must be at least 90% of the target's total outstanding shares, and the acquirer must have a public status. The deal value must be equal to or greater than \$0, and only cash consideration, common/ordinary shares, or a combination of both were considered, excluding all other Consideration Offered Categories. The acquirer's ultimate parent must be listed on one of the following exchanges: AMEX, NASDAQ, NYSE, NYSE Alter, NYSE Arca, NYSE Amex, NYSE MKT, and only Deal Type 'Disclosed Value' & 'Tender Offers' are included. Lastly, the acquirer must own 0% of the target's shares six months before the transaction occurred. These filters result in an initial sample of 2,330 acquisitions. Filtering

and excluding data points is done to reduce the noise in our data set caused by abnormal merger characteristics or outlier acquirers that behave drastically differently than the general market.

After exporting the data to Excel, we add three additional requirements. First, we remove duplicates keeping only the oldest observation³ which results in a loss of 748 observations, leaving a sample of 1,582. Second, acquisitions with a value of less than 5% of the acquirer's market capitalization are excluded, resulting in a further loss of 495 observations, leaving 1,098 observations. By doing so, non-material deals were excluded from the data set, while the ones that were likely to affect the acquirer, in the long run, are kept. Third, industries with less than 30 observations are excluded since the sub-samples are considered too small to serve as a fair proxy for the entire industry population.

Additional information on the acquirer is then manually added to the data set. The goodwill arising from the acquisition as well as the deal value is recorded, found in the PPA of the SEC filing of the acquirer. When available, we use the SEC filing of the year after the acquisition as it contains revised information. Otherwise, we use the preliminary PPA from the year of acquisition. This resulted in a loss of 137 observations for companies that did not report the PGP of the acquisition of interest, leaving 961 observations. Furthermore, we retrieve stock prices over the following five years after the acquisition, and BM ratios and market values through the Intrinio and CapitalIQ APIs and add them to the Excel list. A total of 181 observations were excluded due to missing stock price data over all the five years and missing BM ratios, resulting in 779 observations. Finally, we exclude observations that were listed on OTC markets since many of them lost 100% of their value over the subsequent years of observation, which did not occur for stocks listed on the primary exchanges. For this reason, we believe that the firms on the OTC markets are not representative of our full sample. This resulted in a final sample of 676 observations. It should also be noted that not all firms in our data set were listed for all five years after the acquisition. This implies that in year one retained our full sample, year two retained 648 observations, year three retained 622 observations, year four retained 596 observations, and year five retained 583 observations.

³ This filter has been widely used for collecting acquisition observations (Loughran & Vijh, 1997; Kooli & L'Her, 2004; Dutta & Jog 2009).

Finally in order to match our sample firms' post-acquisition returns, we use the Kenneth R. French data library to retrieve historical archives of daily 2x3 book-to-market and size portfolio returns, monthly three-factor data, as well as 49 industry portfolio returns. The Kenneth French data set uses COMPUSTAT data of firms listed on the NYSE, AMEX, and NASDAQ which are the same exchanges that our sample firms are listed on.

4. Findings and Analysis

4.1 Description of Data

Table 3 reports the annual number of acquisitions in our sample, as well as the aggregate dollar value of these acquisitions from the beginning of 2007 to the year-end of 2016. During these ten years, the aggregate dollar value of all acquisitions was \$685.2 billion in which the majority (52%) of the value comprised of purchased goodwill. Furthermore, most of the sample acquisitions were mergers (93.7%) rather than tender offers, and the majority were related (76.5%).

Table 3

Number of Acquisitions by Calendar Year													
Year	Deals	Deal Value	Purchased Goodwill	Average PGP	Merger	Tender	Stock	Cash	Mixed	Related	Unrelated	Domestic	Cross Border
2007	128	117,30	60,25	51%	124	4	20	54	54	103	25	116	12
2008	82	64,13	41,86	65%	71	11	16	33	33	66	16	75	7
2009	31	10,35	4,61	45%	28	3	7	15	9	26	5	27	4
2010	59	37,98	24,54	57%	52	6	11	31	16	45	13	48	10
2011	56	31,13	13,63	44%	55	1	7	33	16	37	19	45	11
2012	59	58,70	26,34	45%	55	4	11	33	15	41	18	52	7
2013	58	39,58	18,70	47%	55	3	7	30	21	42	16	49	9
2014	85	84,48	49,98	59%	82	3	15	33	37	67	18	74	11
2015	78	114,05	51,94	46%	72	6	13	28	37	61	17	70	8
2016	41	130,80	67,46	52%	40	1	5	15	21	29	12	35	6
Total	676	685,20	354,39	52%	634	42	112	305	259	517	159	591	85

The data is retrieved from Refinitiv Eikon, Deal Screener. The table portrays the most important characteristics of our sample on a yearly basis. Deal value and purchased goodwill are presented in (\$) billions. The average PGP is defined as the total Purchased Goodwill divided by the total Deal Value for a given year. The remaining columns in the table show the number of transactions within each category for each given year.

Figure 2 is a visualization of the yearly number of acquisitions and the yearly deal value in our sample. The graph reveals a strong downward trend in both M&A activity and deal value from 2007 until 2015. Interestingly, the number of M&A transactions never recovered from 2007 highs, while a full recovery in deal value wasn't realized until 2016. The year 2007 was the culmination of a bull market that is reflected in our data by the apparent merger wave. Out of the 128 deals in 2007, 124 of them were classified as mergers. After the merger wave, The Great Financial Crises plagued the following ten-year period with uncertainty that was offset by quantitative easing policies enacted by the United States Federal Reserve (Timiraos, 2022). The five-year period following 2016 subsequently ended with all-time market highs. With this in mind, the average five-year stock return of our 583 observations amounted to 61.4%.

Figure 2

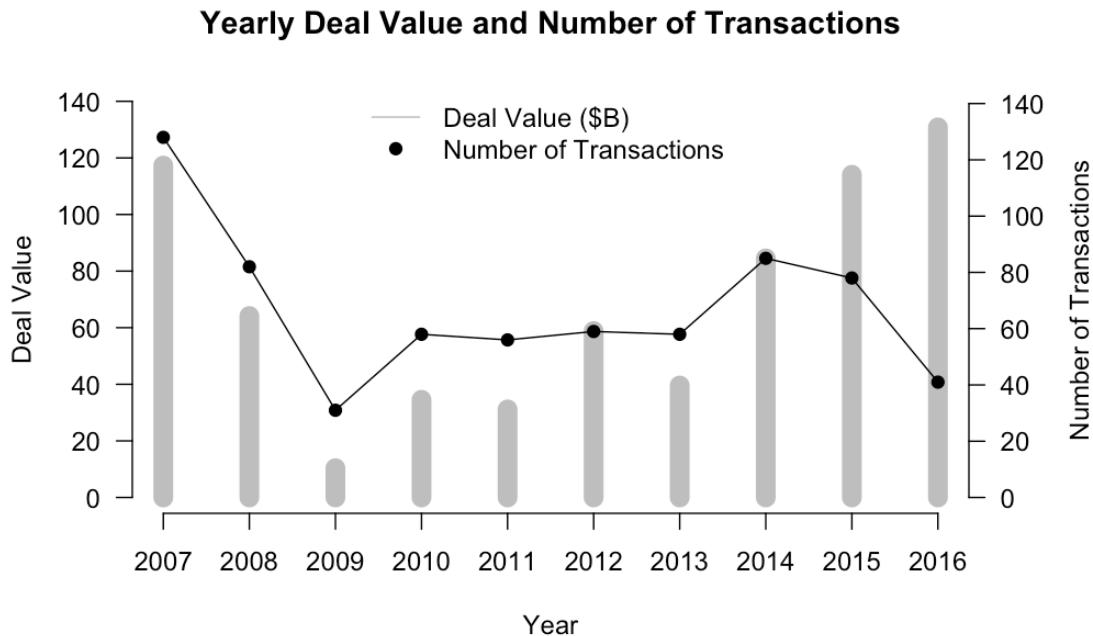
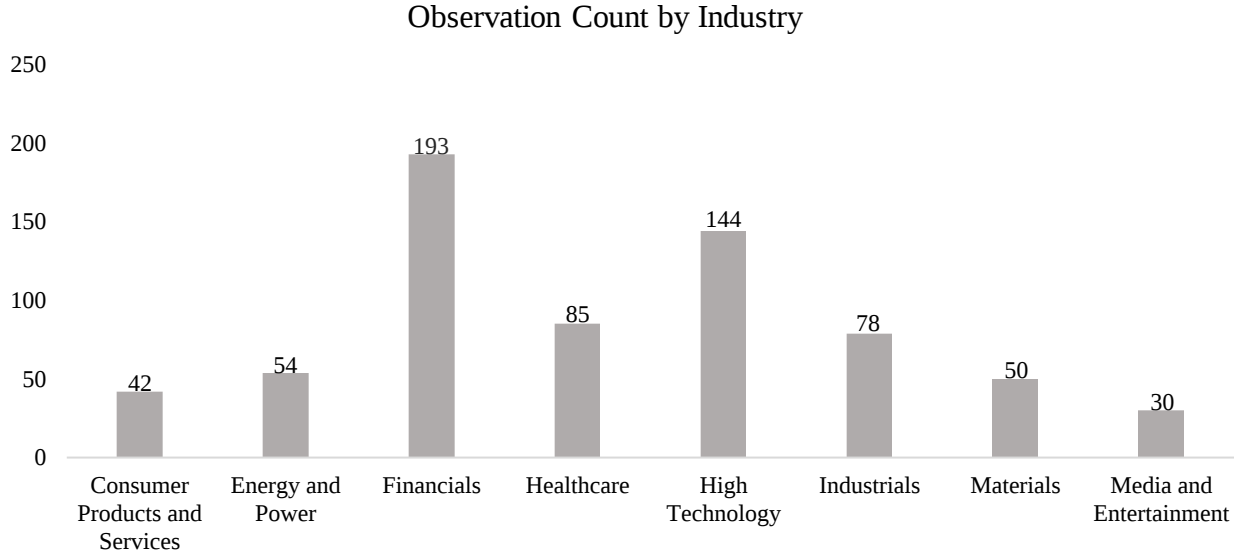


Figure 3 displays the sample size grouped by industry. The Financials and High Technology industries have the highest number of observations in comparison to the other industries. The predominance of samples in the Financials industry may be due the large number of acquisitions

in the banking sector during The Great Financial Crises. Out of the total 193 acquisitions in the Financial industry, 23% of them occur in 2007 and 41% occur in years 2007-2009.

Figure 3



4.2 Hypothesis Testing

We have two main regression models that we use in this study: the IV model and the Full model. We iterate our first regression by changing the dependent variable Stock return (R_{it}) to account for BHSR and the two BHAR methodologies as well as the year in which we measure the returns. In the IV model we simply regress company return on the independent variable PGP. The Full model formula we use for our two tests are exemplified below.

$$R_{it}/AR_{it} = \alpha_{it} + B_1 PGP_{it} + B_2 \text{Related dummy}_{it} + B_3 \text{Consideration Type dummy}_{it} + B_4 \text{Target Location dummy}_{it} + B_5 \text{Payment Type}_{it} \text{ dummy} + \sum \text{Industry FE}_{it} + \sum \text{Year FE}_{it}$$

$$R_{it} = \alpha_{it} + B_1 PGP_{it} + B_2 PGP * \text{Industry Dummy}_{it} + B_3 \text{Related dummy}_{it} + B_4 \text{Consideration Type dummy}_{it} + B_5 \text{Target Location dummy}_{it} + B_6 \text{Payment Type}_{it} \text{ dummy} + \sum \text{Industry FE}_{it} + \sum \text{Year FE}_{it}$$

4.3 Summary of Results

In this section we will begin by presenting the results from our first hypothesis test with the results structured in the same order as they are presented in our methodology. Thereafter, we will present the results of our second hypothesis. We will begin by giving a short introduction to our findings for each hypothesis before displaying the regression and mentioning our bootstrapped results. Lastly, for each test we will give a short summary about our complementary findings when applying the Fama-French Three Factor Model as well as the Outlier Exemption Robustness Check. Complementary results are found in the [appendix](#).

4.3.1 Hypothesis One

To begin, we test if the purchased goodwill proportion has an impact on the long-term stock performance of acquirers. For each method, we first present a simple regression with our independent variable (IV), and then a Full model including our fixed effects and dummy variables. For the sake of simplicity, we choose not to present the coefficients of the fixed effects in our regression tables, except for the ones included in the year four bootstrapping tables that can found in the appendix.

BHSR

In our IV model we found that the PGP has a significant negative relationship with one-, two-, and four-year post-acquisition [BHSR](#). The one- and two-year post-acquisition returns decrease, on average, with -0.195 (-19.5%) and -0.256 (-25.6%) for every 1 unit (100%) increase in the PGP respectively and the four-year post-acquisition returns decrease further by a total of -0.395 (-39.5%). This implies that the largest proportion of the PGPs negative impact on BHSR occur in the first two years post-acquisition. This is in-line with the three-month post-acquisition returns which, although insignificant, display that three-month post-announcement returns decrease with -10.6% for every 100% increase in the PGP⁴.

In our Full model, however, the PGP no longer explains the returns of the first couple of years but still significantly explains the four-year post-acquisition return; every one unit (100%)

⁴ The table with the full results for the three-month BHSR regressions can be found in the appendix.

increase in the PGP at acquisition, on average, results in a -0.358 unit decrease in (-35.8%) four-year stock returns at the 5% level. Furthermore, Mix acquirers underperform Cash acquirers by -0.067% after the first year of acquisition. In addition, acquirers that buy within the United States earn 17.7% higher two-year return after the acquisition in comparison with Cross-Border acquirers at the 5% level. All year fixed effects are significant for each post-acquisition year and this variable makes up most of the Full model's explanatory power.

When bootstrapping our IV model we find that year one and two are still significant, whereas year four is not. As for our Full model, we no longer find support for the PGP variable's impact on the four-year returns at a 5% assertion level through a bootstrapped method. Furthermore, the underperformance of Mix acquirers relative to Cash acquirers lacks evidence in the bootstrapped version. We can confirm, however, that Domestic acquirers do indeed perform better than Cross-Border acquirers at the 5% significance level in both tests. Also, the bootstrapped output confirms that acquisitions made in 2007 performed significantly worse than every other year and that acquisitions in 2016 performed the best (+0.841) relative to 2007, followed by acquisitions made in 2011 (+0.812) and in 2009 (+0.749). An acquirer that belongs to any industry besides Healthcare, on average, performs worse than our base industry Consumer Products and Services. However, the Energy and Power industry is the only significant industry both in the original regression as well as the bootstrapped version, with a coefficient of -0.587. This implies that the Energy and Power industry earned about -57.8% lower returns in comparison to Consumer Products and Services four years after acquisition. One last and interesting finding from the bootstrapped approach is that acquirers buying targets belonging to an industry different to its own, classified as Unrelated, significantly overperform the alternative, Related, in the fourth year even though this control dummy is insignificant in the original regression.

Method 1: Buy and Hold Stock Returns (BHSR)

Year	(1)	(2)	(3)	(4)	(5)
IV MODEL					
PGP	-0.195*** (0.058)	-0.256*** (0.093)	-0.113 (0.107)	-0.395** (0.172)	-0.344 (0.254)
Constant	0.099*** (0.031)	0.221*** (0.051)	0.296*** (0.059)	0.583*** (0.094)	0.754*** (0.139)
Adjusted R ²	0.015	0.010	0.0002	0.007	0.001

FULL MODEL

PGP	-0.093 (0.058)	-0.126 (0.093)	-0.048 (0.109)	-0.358** (0.179)	-0.379 (0.265)
Unrelated	-0.021 (0.040)	-0.015 (0.063)	0.055 (0.075)	0.107 (0.124)	0.070 (0.183)
Tender	-0.022 (0.069)	-0.033 (0.111)	-0.130 (0.134)	-0.238 (0.220)	-0.331 (0.327)
Domestic	0.055 (0.049)	0.177** (0.077)	0.151 (0.092)	0.204 (0.149)	0.317 (0.221)
Mix	-0.067* (0.038)	-0.066 (0.060)	-0.113 (0.072)	-0.054 (0.119)	-0.123 (0.176)
Shares	-0.002 (0.048)	-0.081 (0.076)	-0.111 (0.090)	0.130 (0.148)	0.099 (0.218)
Constant	-0.154 (0.099)	-0.339** (0.155)	-0.121 (0.186)	0.111 (0.304)	0.159 (0.451)
Observations	676	648	622	596	583
Adjusted R ²	0.149	0.162	0.115	0.078	0.068
Year & Industry FE	Yes	Yes	Yes	Yes	Yes

Standard error (SE) is shown in parentheses

*p<0.1; **p<0.05; ***p<0.01

The table portrays the variables investigated on a yearly basis, starting from year 1 until year 5. In the IV model, returns are regressed on PGP, and in the Full model all of the control variables are included.

BHAR_{BMS}

In both our IV and Full model, applying the [BHAR_{BMS}](#), we find that PGP has a significant negative relationship with acquirer abnormal returns in year four of -0.292 and -0.363 respectively. Besides our variable of interest, we see that companies defined as Domestic significantly outperform the alternative by over 16% in the second and third year. One interesting finding is that acquisitions that took place in 2009, 2011, 2013, 2014 and 2016 all significantly outperform the baseline year 2007 for certain yearly post-acquisition returns, the most significant outperformance is found in the three-year abnormal stock returns of acquisitions made in 2011 with a coefficient of 0.346 at a 1% assertion level. This is seemingly unexpected as the BHAR methodology involves matching our firms' yearly returns with portfolio returns that have similar BM and size characteristics, therefore yearly variations in returns should be accounted for. Therefore, this result implies that acquirers in general had better three-year post-acquisition returns than their respective benchmark portfolios during these years, on average. Similar to the BHSR findings, as we bootstrap the four-year BHAR_{BMS} for both the IV and Full model, we no longer find support for the PGP variable's impact on the four-year returns at a 5% assertion level. Furthermore, the outperformance of Domestic relative to Cross-Border acquirers

lacks evidence in the bootstrapped version. Again, we see that an acquirer that belongs to any industry besides Healthcare, on average, performs worse than our base industry Consumer Products and Services, however, the Energy and Power and Materials industries are the only statistically significant results, with a coefficient of -0.583 (-57.8%) and -0.482 (-48.2%) respectively. Lastly, year 2011 and 2014 remain significant in the bootstrapped version.

Method 2: BM, Size Buy & Hold Abnormal Returns

Year	(1)	(2)	(3)	(4)	(5)
IV MODEL					
PGP	-0.072 (0.051)	-0.097 (0.084)	-0.012 (0.103)	-0.292* (0.168)	-0.256 (0.247)
Constant	-0.007 (0.028)	-0.011 (0.046)	-0.078 (0.056)	0.085 (0.092)	0.122 (0.135)
Adjusted R ²	0.002	0.001	0	0.003	0.0001
FULL MODEL					
PGP	-0.074 (0.055)	-0.107 (0.091)	-0.052 (0.109)	-0.363** (0.177)	-0.377 (0.263)
Unrelated	-0.028 (0.038)	-0.022 (0.062)	0.044 (0.075)	0.104 (0.123)	0.043 (0.182)
Tender	0.006 (0.066)	-0.021 (0.108)	-0.111 (0.134)	-0.210 (0.219)	-0.307 (0.324)
Domestic	0.056 (0.047)	0.162** (0.076)	0.168* (0.092)	0.220 (0.148)	0.331 (0.219)
Mix	-0.058 (0.036)	-0.053 (0.059)	-0.119 (0.072)	-0.058 (0.118)	-0.126 (0.175)
Shares	-0.021 (0.045)	-0.081 (0.075)	-0.131 (0.090)	0.119 (0.147)	0.104 (0.216)
Constant	0.025 (0.094)	-0.062 (0.152)	-0.123 (0.186)	-0.059 (0.302)	-0.137 (0.448)
Observations	676	648	622	596	583
Adjusted R ²	0	0.012	0.026	0.039	0.030
Year & Industry FE	Yes	Yes	Yes	Yes	Yes

Standard error (SE) is shown in parentheses

*p<0.1; **p<0.05; ***p<0.01

The table portrays the variables investigated on a yearly basis, starting from year 1 until year 5. In the IV model, returns are regressed on PGP, and in the Full model all of the control variables are included.

$BHAR_t$

Like the previous two methods, we find that the PGP has a significant negative relationship with the abnormal returns in the fourth year using an industry-related benchmark. For the IV model and Full model, every 1 unit (100%) increase in the PGP at acquisition, on average, results in decreases in four-year [BHAR_t](#) stock returns of -0.331 (-33.1%) and -0.358 (-35.8%),

respectively. Besides our variable of interest, we again see that Domestic acquirers significantly outperform the alternative with 17% in the second year and 18% in the third year. Similar to the BHAR_{BMS} approach, we find that acquisitions that took place in 2009, 2011, 2012, and 2016 all significantly outperformed the baseline year 2007 for certain yearly post-acquisition returns, and yet again the most significant outperformance is found in three-year abnormal stock returns of acquisitions made in 2011 with a coefficient of 0.336 at a 1% assertion level.

Again, as we bootstrap the original four-year BHAR_I regression of our IV and Full model, we no longer find support for the PGP variable's impact on the four-year returns at a 5% assertion level. Furthermore, the outperformance of Domestic acquirers relative to the alternative lacks evidence in the bootstrapped version. Moreover, we see that an acquirer that belongs to any industry besides Healthcare, on average, performs worse than our base industry Consumer Products and Services, however, none of the industries are statistically significant. Like in the previous method, year 2011 significantly explains returns. One final and interesting finding with regards to the bootstrapped approach is that Tender acquirers significantly underperform Mergers by -0.254, although this control dummy is insignificant in the original regression.

Method 3: Industry Buy & Hold Abnormal Returns

Year	(1)	(2)	(3)	(4)	(5)
IV MODEL					
PGP	-0.074 (0.050)	-0.128 (0.082)	-0.030 (0.097)	-0.331** (0.162)	-0.326 (0.238)
Constant	0.001 (0.027)	0.00000 (0.045)	-0.074 (0.053)	0.097 (0.089)	0.131 (0.130)
Adjusted R ²	0.002	0.002	0	0.005	0.002

FULL MODEL

PGP	-0.059 (0.054)	-0.118 (0.089)	-0.041 (0.104)	-0.358** (0.173)	-0.397 (0.254)
Unrelated	-0.014 (0.037)	-0.019 (0.061)	0.036 (0.071)	0.095 (0.119)	0.036 (0.176)
Tender	0.005 (0.064)	-0.014 (0.106)	-0.146 (0.127)	-0.254 (0.213)	-0.343 (0.314)
Domestic	0.064 (0.045)	0.170** (0.074)	0.180** (0.088)	0.225 (0.144)	0.319 (0.212)
Mix	-0.051 (0.035)	-0.053 (0.058)	-0.106 (0.069)	-0.023 (0.115)	-0.091 (0.169)
Shares	-0.024 (0.044)	-0.070 (0.073)	-0.100 (0.086)	0.138 (0.143)	0.120 (0.209)
Constant	0.015 (0.091)	-0.041 (0.149)	-0.104 (0.178)	0.002 (0.294)	-0.015 (0.434)

Observations	676	648	622	596	583
Adjusted R ²	0	0.007	0.015	0.029	0.023
Year & Industry FE	Yes	Yes	Yes	Yes	Yes

Standard error (SE) is shown in parentheses

*p<0.1; **p<0.05; ***p<0.01

The table portrays the variables investigated on a yearly basis, starting from year 1 until year 5. In the IV model, returns are regressed on PGP, and in the Full model all of the control variables are included.

In summary, we find that the PGP factor has a significant negative correlation with post-acquisition returns using both BHSR and BHAR methodologies. In our BHSR IV model, the PGP was significantly related to one-, two-, and four-year BHSR, and year one and two were still significant using the bootstrapping robustness check. In our Full model regressions, the PGP was only significantly related to four-year post-acquisition returns in all methodologies, but this significance was lost when applying the BCa bootstrapping robustness check.

Complementary Findings

Using the Fama-French Three Factor Model, we observe that the Low PGP portfolio has the highest compounded monthly abnormal returns (alpha) of 0.010%, followed by the Medium PGP portfolio returning 0.006%, and the High PGP portfolio returning 0.002%. Only the Medium PGP portfolio exhibited a significant alpha at a 10% level, while the other two groups did not. The High PGP group has the highest market beta of 1.473, significantly correlated with monthly portfolio returns. The SMB factor was positive for all portfolios, indicating a high proportion of small firms. The Low PGP group had the highest proportion of small firms, followed by the Medium PGP group, while the High PGP group had the least. The HML factor provides an idea of the book-to-market ratio of the portfolios, with the Low PGP group having predominantly growth firms, the Medium PGP group having predominantly value firms, and the High PGP group being balanced between growth and value, with none of the HMLs being significant.

Based on our Outlier Exemption approach we can confirm that the PGP has a significant negative effect on company performance in year one, two and four of -0.148, -0.214 and -0.194 respectively in the IV model, however with the introduction of various controls, this relationship is lost. Furthermore, the robustness check reconfirms that Domestic acquirers significantly

outperform Cross-Border acquirers. The payment method also plays a role in explaining stock performance. Both Mix and Shares acquirers significantly underperform Cash acquirers in all five years investigated.

4.3.2 Hypothesis Two

Our results from testing our second hypothesis show that the effect of PGP on company performance is indeed moderated by the industry the acquirer belongs to. The top row for each industry shows the original regression with the betas and the standard errors in the parenthesis, while the bottom row shows the bootstrapped version with the BootMed and the significance interval in the parenthesis. The model includes fixed effects as well as the remaining control variables used in our previously presented regressions.

There are no significant results in the first year, however, Financials and High Technology are significant in the second year with coefficients of about -0.659 and -0.854 respectively, and the Products and Services, Healthcare and Industrials become significant in the third year. The significance seems to increase over time for most industries except for the Materials and Energy and Power that remain insignificant until a modest significance in year 5 at a 10% level. In general the results indicate that all industries have a negative impact on PGP's relation to stock returns compared to the baseline Consumer Products and Services, with the most negative being Healthcare, followed by Industrials, and Media and Entertainment. This can be interpreted as these industries are more sensitive to increases in the PGP.

Although there are no significant results in the first year of the original regression, the bootstrapped versions find significance for the Consumer Products and Services as well as the Industrials at a 5% level. In year two and three on the other hand there is no evidence of cross-industry differences in how PGP affects returns. Yet again Industrials is significant in year four as well as Healthcare. Finally, in the fifth year, Industrials, Healthcare, Financials and Media and Entertainment are significant.

Cross Industry Sensitivity (PGP x Industry FE)

Year	(1)	(2)	(3)	(4)	(5)
Consumer Products and Services (Baseline)	0.091 (0.226)	0.435 (0.354)	0.816** (0.413)	1.421** (0.658)	3.006*** (0.966)
	0.092 (-0.547:-0.015)	0.419 (-0.384:1.604)	0.801 (-0.633:1.819)	1.364 (-0.366:8.042)	2.933 (-0.216: 6.165)
Energy and Power	0.381 (0.316)	-0.072 (0.497)	-0.733 (0.578)	-1.264 (0.937)	-2.378* (1.441)
	0.355 (-0.299:1.042)	-0.04806 (-1.186:1.064)	-0.660 (-2.225:0.871)	-1.223 (-6.829: 1.024)	-2.197 (-5.775: 1.189)
Financials	-0.267 (0.245)	-0.659* (0.384)	-1.081** (0.448)	-1.688** (0.715)	-3.303*** (1.049)
	-0.266 (-0.812:0.061)	-0.636 (-1.829:0.201)	-1.039 (-2.258:0.295)	-1.622 (-7.986: 0.112)	-3.188 (-6.332: -0.007)
Healthcare	-0.285 (0.285)	-0.646 (0.455)	-1.246** (0.534)	-3.923*** (0.852)	-5.930*** (1.261)
	-0.301 (-0.913:0.224)	-0.625 (-1.820:0.393)	-1.166 (-2.797:0.272)	-3.854 (-9.607:-1.287)	-5.642 (-10.792:-2.542)
High Technology	-0.224 (0.266)	-0.854** (0.419)	-0.576 (0.491)	-1.192 (0.792)	-2.588** (1.163)
	-0.231 (-0.865:0.272)	-0.747 (-3.234:0.219)	-0.491 (-1.942:0.715)	-1.140 (-7.501:0.627)	-2.439 (-5.914:0.569)
Industrials	-0.438 (0.291)	-0.705 (0.455)	-1.182** (0.531)	2.894*** (0.8459)	-5.913*** (1.238)
	-0.441 (-1.031:-0.059)	-0.682 (-1.908:0.259)	-1.124 (-2.438:0.153)	-2.822 (-9.610:-0.290)	-5.336 (-17.308:-1.823)
Materials	0.070 (0.309)	-0.236 (0.485)	-0.421 (0.565)	-0.939 (0.899)	-2.220* (1.316)
	0.082 (-0.592:0.642)	-0.186 (-1.432:0.724)	-0.337 (-1.593:0.947)	-0.900 (-7.115:1.023)	-2.059 (-5.514:1.076)
Media and Entertainment	-0.105 (0.378)	-0.365 (0.590)	-0.640 (0.723)	-1.914 (1.206)	-3.981** (1.763)
	-0.144 (-0.940:0.535)	-0.370 (-1.831:0.963)	-0.598 (-3.053:1.718)	-2.040 (-6.412:0.553)	-3.842 (-8.854:-0.667)
Observations	676	648	622	596	583
Adjusted R ²	0.155	0.164	0.121	0.117	0.115
Year & Industry FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Standard error (SE) is shown in parentheses

* p<0.1; ** p<0.05; *** p<0.01

In the first column, titled PGP x Industry FE, the various interaction between industries and PGP is presented. The Consumer Products and Services industry represents the baseline in which a one-unit change in PGP results in a 0.088 change in stock returns for the first year, an effect that increases over time to 2.946 in the fifth year. By subtracting the coefficient of the baseline industry and the interaction coefficient between the industry of interest and PGP, we can evaluate the modifying effect that each specific industry has on PGP's relation to the acquirer returns over a 5-year period.

Complementary Findings

For the scope of our second hypothesis the Fama-French Three Factor model is not applicable. Based on our Outlier Exemption approach, however, we do not find general support for the PGP being moderated by industry classifications. We find that the only significant moderating effect is for Energy and Power in year one, where the industry performed better than our baseline, Consumer Products and Services by 0.640 or 64%.

5. Discussion

5.1 Discussion of Results

The results from testing our first hypothesis suggest that purchased goodwill is on average value-destroying as we found that incremental increases of the PGP at acquisition are negatively associated with short term, and long term BHSR and BHAR. This finding aligns with the general view that higher goodwill arising at acquisition is ‘bad’ and does not exclusively represent the theoretically ideal goodwill, that is, the future value creation from synergy effects and the going concern, but rather non-core elements such as overbidding and overvaluation as suggested by Johnson & Petrone (1999). So, although goodwill has seen an increase in value relevance since the adoption of the “impairment-only approach”, this study provides evidence that the valuation of goodwill is yet vague and the way in which it is accounted for may lead investors astray.

The beta coefficient of the PGP was negative in all BHSR/BHAR tests and in every year observed, but this negative relationship was only statistically significant in year one and two of the BHSR IV model, and in year four of the BHSR and BHAR Full models. The complementary Fama-French Three Factor Model reinforces this association as we found a stepwise decreasing abnormal return (alpha) from the low PGP group to the highest PGP group. These findings suggest that M&A, on average, are susceptible to overoptimism about future synergies, bidding competitions, and inaccurate valuation of net assets or discount rates which is reflected in the purchased goodwill premium paid by the acquirer. One explanation for the negative post-acquisition returns of acquirers paying a higher PGP premium is the potential goodwill impairments that these acquirers may experience. Olante (2013) found that goodwill impairments are most likely to occur two to three years post-acquisition which is in line with our results that show that the PGP has a significant negative relationship with post-acquisition returns in year

one, two, and four. Although we have not identified which of our samples experienced goodwill impairments, a higher PGP increases the probability of an impairment occurring (Olante 2013). This, in turn, can have adverse effects on stock performance, which Teladoc in our introduction clearly exemplifies. When combining our results with theory and prior literature, we can with fair accuracy conclude that investors should be wary of M&A with a high purchased goodwill premium. But we cannot emphasize enough the limitations of measuring long-term post-acquisition performance that is subject to clear biases. When alleviating numerous biases using a bootstrapping robustness check, for example, we find that four-year post-acquisition returns cannot be significantly explained by the PGP in neither our IV models, nor our Full models.

Furthermore, by investigating the moderating effect of industry classification on the PGP, through our second hypothesis, we find that there are indeed interindustry differences in the PGP's relationship with post-acquisition returns. Theoretical considerations prompted us to explore the possibility of interindustry differences in the sensitivity to the PGP. First, we reasoned that interindustry differences in the average PGP could explain the moderating effect, yet we find no support for this. For example, High Technology and Industrials have similar average PGPs of 51% and 53% respectively, but for every incremental increase in PGP Industrials experiences on average more than two times lower BHSR than High Technology.

Second, we reasoned that acquisitions in industries with a higher growth premium (BM) will experience more difficulties in forecasting future cash flows and hence may be subject to a discount bias. In general, this could be true for some of our industries. Energy and Power, Financials, and Materials have the lowest growth premiums coupled with a moderating effect coefficient of -2.378, -3.03, and -2.220 in year five, respectively. In comparison, Healthcare and Industrials with high growth premiums have the lowest coefficients of -5.930 and -5.913, respectively. When comparing these groups of industries, the theory seems to hold. But the rest of our industries do not follow the same trend. For example, Consumer Products and Services has the highest growth premium while also having post-acquisition returns that are positively correlated with the PGP, which is counterintuitive based on this theory. Furthermore, High Technology has the second highest growth premium of 0.45 and has a coefficient in year five of -

2.558 which is even lower than Financials and only slightly higher than Energy and Power and Materials.

Lastly, although higher bidding competition for maturing industries may be an explanation for interindustry differences, we cannot reasonably test this because the macro classification of our industries is vague and we cannot determine their level of maturity. It is important to note that every PGP x Industry interaction term is statistically significant in year five and this is most likely due to the bias of long-term returns. When using the bootstrapping method in this regression, we find that Healthcare, Financials, Industrials, and Media and Entertainment are the only industries with significant moderating effects. In all, we were not able to find any conclusive evidence for the theoretical framework in regard to the second hypothesis but can conclude that there are interindustry differences regarding the sensitivity of post-acquisition returns on PGP.

Now we would like to gear the discussion towards complementary findings. The Fama French Three Factor Model provides insights into other characteristics, beyond industry, of acquiring firms for different levels of PGP. When applying this methodology, we observed that the high PGP group has the highest and significant market risk, the lowest proportion of small firms, and a balanced mix of growth and value firms. In contrast, the low PGP group has the highest proportion of small firms and the highest number of growth firms. While the high PGP group experienced the lowest abnormal returns, the low PGP group was rewarded with the highest return out of the three groups. This raises the question: do smaller firms tend to be more risk-averse? And is insider ownership more significant for smaller growth firms, leading to goal congruencies and well-motivated acquisitions like they seem to be in Dutta & Jog's (2009) findings? Although our findings may not provide definitive answers, they nonetheless lend insights into other characteristics that are coupled with purchased goodwill.

Next, we will discuss the control variables that significantly explain post-acquisition stock performance in our regressions and explain how they may relate to our variable of interest. Moeller and Schlingemann (2005) found that acquiring cross-border targets, relative to domestic targets, significantly lowers post-announcement stock returns and we find a similar trend for

long-term post-acquisition returns. The size and significance of this relationship are reconfirmed by our outlier elimination robustness check but not when applying the bootstrapping method on our original regression. Acquisitions of foreign targets and acquisitions with a high purchased goodwill premium may share uncertainties. The risks associated with a cross-border target, such as legal and regulatory, political, currency exchange, and due diligence risks, as well as cultural differences and differing market dynamics may be more challenging to address than for a domestic target. Failure to properly calculate the risk-adjusted valuation of the target, to discount future synergies appropriately, or measure the value of the assets and liabilities may all be prevalent non-core goodwill elements belonging to the acquisition of a cross-border target. This may result in an overstated goodwill that is at risk of future impairment.

Furthermore, Mix and Stock acquisitions have significantly lower post-acquisition performance than Cash acquisitions in our original BHSR regression and outlier eliminated BHAR regressions. Similar relationships have been found in prior literature (Dutta & Jog 2009; Loughran & Vijh 1997), and we believe that paying with stock is closely linked to the non-core goodwill component of overvaluation/undervaluation. If an acquirer pays with stock, then the deal value will differ between the announcement date of the acquisition and the date with which it is effective depending upon the acquirer's stock price. Therefore, if the acquirer's stock price increases during this timeframe, then the deal value will increase and as a result the amount of goodwill purchased will also increase. However, if the acquirer's stock price decreases, then so will goodwill. These two components imply that either an overvaluation or an undervaluation are captured in the purchased goodwill component (Johnson & Petrone 1999). Although, we can conclude that goodwill is not solely representative of the core goodwill elements, we cannot conclude which of the non-core goodwill elements that purchased goodwill precisely contain. One interesting topic for future research is hence to investigate if the overvaluation/undervaluation component is linked to the form of payment at acquisition.

Lastly, our tests do not show that the mode of acquisition plays a significant role in explaining post-acquisition returns. However, as we referred to in the literature review, merger waves, like the one in 2007, are found to be positively linked to short-term deviations in valuation from long-term trends, particularly when companies use stock to pay for the mergers (Rhodes-Kropf,

Robinson, & Viswanathan, 2015). Therefore, in theory the mix of high bidding activity and over-optimistic market sentiment could increase the risk of mergers susceptible to overbidding, which implies that the mode of acquisition could be linked to the PGP.

In summary, we cannot conclude in the scope of this study, but have reason to believe, that geography, form of payment, and mode of acquisition could have a significant relationship with purchased goodwill. Future research testing these relationships would bring us closer to understanding the true representation or composition of purchased goodwill.

6. Conclusion

In this study, we find empirical ground showing that purchased goodwill is not only representative of the core goodwill elements of going concern and synergies. As a matter of fact, we find that purchased goodwill is also representative of negative aspects such as overbidding, overvaluation, and fair-value mismeasurements because incremental increases in the purchased goodwill proportion (PGP) is related to lower post-acquisition performance, on average.

Furthermore, we find evidence suggesting that the relationship between the PGP and post-acquisition performance is moderated by the acquiring firm's industry classification. Although our tests cannot explain why interindustry differences in the sensitivity to the PGP exist, we recognize that interindustry differences in balance sheet structure, the intangibility of assets, and growth premiums could in theory play an important role in the realization of purchased goodwill. We have implemented several methodologies, namely the BHSR, BHAR, and Fama-French Three Factor Model coupled with BCa bootstrapping tests and outlier omission tests as robustness checks. By implementing these methodologies as well as robustness checks we have been able to mitigate several biases, but we cannot ascertain that the methods are free from biases. Nonetheless, we believe that future research building on our data collection and our results could be conducted in order to uncover what purchased goodwill represents in more detail. For example, how much of the purchased goodwill is related to overvaluation, overbidding, synergy effects, the going concern, or fair-value mismeasurements? This is a relevant question for future research that can serve to better inform investors and accounting regulators of the true representation of goodwill.

7. Bibliography

- Bagna, E., Ramusino, E. C., & Ogliari, M. (2023). The impact of different goodwill accounting methods on stock prices: A comparison of amortization and impairment-only methodologies. *International Review of Financial Analysis*, 85.
- Bloom, M. (2009). Accounting for Goodwill. *ABACUS, A Journal of Accounting, Finance and Business Studies*, 379-389.
- Boennen, S., & Glaum, M. (2014). Goodwill Accounting: A Review of the Literature. *SSRN*.
- Bruner, R. (2005). Where M&A Pays and Where It Strays: A Survey of the Research. *Journal of Applied Corporate Finance*, 63-76.
- Bruner, R. F., & Perella, J. R. (2004). *Applied Mergers and Acquisitions*. New Jersey: John Wiley & Sons.
- Burger, M., & Wen, H. (2021). The relative and incremental value relevance of goodwill before and after SFAS No. 142. *Journal of Accounting and Public Policy*, 40(6).
- Carlin, T. M., & Finch, N. (2011). Goodwill impairment testing under IFRS: a false impossible shore? *Pacific Accounting Review*.
- Dutta, S., & Jog, V. (2009). The long-term performance of acquiring firms: A re-examination of an anomaly. *Journal of Banking & Finance*, 1400-1412.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 427-465.
- Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *Journal of Finance*, 131-155.
- Fox, J., & Weisberg, S. (2019). *Bootstrapping Regression Models in R*. London: SAGE Publications.
- Hayne, C., & Hughes, P. J. (2006). Leading Indicators of Goodwill Impairment. *Journal of Accounting, Auditing & Finance*, 21(3).
- Jarva, H. (2009). Do Firms Manage Fair Value Estimates? An Examination of SFAS 142 Goodwill Impairments. *Journal of Business & Accounting*, 1059-1086.
- Johnson, L. T., & Petrone, K. R. (1999). Commentary: Is Goodwill an Asset? . *SSRN*.

- Kothari, S., & Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics*, 301-339.
- Loughran, T., & Vijh, A. M. (1997). Do Long-Term Shareholders Benefit From Corporate Acquisitions? *The Journal of Finance*, 1765-1790.
- Lyon, J. D., Barber, B. M., & Tsai, C.-L. (2003). Improved Methods for Tests of Long-Run Abnormal Stock Returns. *The Journal of Finance*, 165-201.
- Mayrhofer, U., Hassan, I., & Ghauri, P. N. (2018). Merger and acquisition motives and outcome assessment . *Thunderbird International Business Review*, 709-718.
- Moeller, S. B., Schlingemann, F. P., & Stulz, R. M. (2004). Firm size and the gains from acquisitions. *Journal of financial economics*, 201-228.
- Mukherjee, T. K., Kiymaz, H., & Baker, H. K. (2004). Merger Motives and Target Valuation: A Survey of Evidence from CFOs. *Journal of Applied Finance*.
- Olante, M. E. (2013). Overpaid acquisitions and goodwill impairment losses - Evidence from the US. *Advances in Accounting*, 243-254.
- Rau, P. R., & Vermaelen, T. (1998). Glamour, value and the post-acquisition performance of acquiring firms. *Journal of Financial Economics*, 223-253.
- Rhodes-Kropf, M., Robinson, D. T., & Viswanathan, b. s. (2015). Valuation waves and merger activity: The empirical evidence. *Journal of Financial Economics*, 561-603.
- Timiraos, N. (2022). *Trillion Dollar Triage: How Jay Powell and the Fed Battled a President and a Pandemic and Prevented Economic Disaster*. Hachette Book Group and Blackstone Publishing.

8. Appendix

8.1 Complementary Findings

8.1.1 Fama French Three Factor Model

We use the Fama French Three Factor Model to obtain risk-weighted abnormal returns for three groups with varying levels of goodwill. This approach complements the BHSR and BHAR methods by comparing the return profiles of low, medium, and high goodwill groups. By considering market risk premiums, book-to-market values, and sizes, the Three-Factor Model uniquely accounts for the risk profile of our sample firms. Each Three Factor portfolio is modeled as follows:

$$R(t) - RF(t) = \alpha + \beta_1[RM(t) - RF(t)] + \beta_2SMB(t) + \beta_3HML(t) + e(t)$$

The regression equation estimates the coefficients, denoted as β_1 , β_2 , and β_3 , using the following variables: $R(t) - RF(t)$ represents the excess return of the portfolio, $RM(t) - RF(t)$ signifies the market risk premium, $SMB(t)$ denotes the excess return of small minus big firms based on their market capitalization, $HML(t)$ represents the excess return of high book-to-market (value) minus low book-to-market (growth) firms, and $e(t)$ is the error term. Lastly, the alpha term that arises in the regression represents everything that cannot be explained by the model, hence known as the abnormal return, and can be modeled as follows:

$$\alpha_i = R(t) - [RF(t) + \beta_1 [RM(t) - RF(t)] + \beta_2SMB(t) + \beta_3HML(t)]$$

In the implementation of this model, we have retrieved the monthly and rebalanced Fama-French data on $RM(t)$, $RF(t)$, $SMB(t)$ and, $HML(t)$ over the 2007-2017 period from the Kenneth R. French database. In this regression, we are interested in creating three portfolios based on the PGP at acquisition. More specifically we find the 30th percentile and the 70th percentile of PGP in our data and group firms with less than 31.935% PGP as low, firms in the 31.935-62.66%

range of the distribution as neutral, and firms with greater than 62.66% of the sample as high PGP. We proceed by calculating the average monthly returns for each group and matching these monthly returns with the average market premium, SMB, and HML. We continue this process for each month, up until 60 months in order to calculate the compounded monthly returns for five years. As previously mentioned, 583 of our 676 firms were still trading on an exchange after five years after its acquisition. Therefore, the average monthly returns in the latter months are measured with a smaller sample size. We present the results below.

Fama French Abnormal Calendar Time Returns

	Acquirer Stock Returns		
	Low PGP Group (1)	Medium PGP Group (2)	High PGP Group (3)
Market Return	0.345 (0.682)	0.484 (0.339)	1.473*** (0.376)
SMB	2.340* (1.242)	1.304* (0.759)	0.322 (0.942)
HML	-1.247 (1.175)	0.088 (0.595)	-0.066 (0.796)
Alpha	0.010 (0.008)	0.006* (0.003)	0.002 (0.004)
Observations	60	60	60
Adjusted R ²	0.017	0.048	0.222

*p<0.1; **p<0.05; ***p<0.01

A similar tendency to what we see in the fourth year using the BHSR and BHAR methodologies, the three-factor model reinforces that higher goodwill has a negative relationship with stock returns – seen by the alpha yielding higher returns for acquirers with lower PGP on average. The Low PGP portfolio experienced compounded monthly abnormal returns of .010%, while the Medium PGP portfolio experienced .006%, and the High PGP portfolio experienced .002%. The three-factor model displays abnormal returns that are adjusted for the portfolio's market risk, size, and book-to-market value. The beta on the market return can be interpreted as the beta of the portfolio. The market has a beta of one, therefore we notice that the high PGP group has the highest market risk of 1.473 and is significantly correlated with the monthly portfolio returns. The small minus big (SMB) factor gives us a sense of the size makeup of the portfolios. When SMB is positive, as it is for all portfolios, it is an indication that the portfolio firms are

predominantly small. Fama & French (1992) found that small firms outperform large firms over the long run. The Low PGP group has the highest proportion of small firms, followed by the Medium PGP group and lastly the High PGP Group, the first mentioned are significant at a 10% level. The high minus low (HML) factor gives an idea of the book-to-market ratio of the different portfolios. A negative HML implies that the portfolio has predominantly low book-to-market ratios meaning that they are growth firms. The low PGP group has an overwhelming amount of growth firms, whereas the medium PGP portfolio consists of higher book-to-market or value firms, and the high PGP portfolio is balanced between growth and value. None of the HMLs are significant. Fama & French (1992) found that value firms outperform growth firms in the long term. Finally, the medium PGP group has a significant excess return, at the 10% significance level. This implies that the group has a significant outperformance that cannot be explained by the three factors of the model.

The Fama-French Three-Factor Model it has been acknowledged for its ability to mitigate cross-sectional dependence among sample firms and reduce misspecification errors that may arise from inadequately specified asset pricing models compared to buy-and-hold methods. Furthermore, the methodology also adds a risk adjusted perspective to our analysis by including the Market Premium, SMB, HML factors. For example, Fama & French (1992) find that firm returns can be explained by their BM and size over time. But the model has been criticized for imperfections not only by later authors but even by its creators who themselves later found that the model failed to explain that stock returns respond to the BM factor in earnings (Fama & French, 1995). Furthermore, the model fails to represent the investor experience, and assumes linearity and no interaction between the three factors: BM, size, and the excess return of the market - although studies have found evidence of the contrary (Lyon, Barber, & Tsai, 2003).

The Fama-French Three-Factor Model it has been acknowledged for its ability to mitigate cross-sectional dependence among sample firms and reduce misspecification errors that may arise from inadequately specified asset pricing models compared to buy-and-hold methods. Furthermore, the methodology also adds a risk adjusted perspective to our analysis by including the Market Premium, SMB, HML factors. For example, Fama & French (1992) find that firm returns can be explained by their BM and size over time. But the model has been criticized for imperfections not only by later authors but even by its creators who themselves later found that the model

failed to explain that stock returns respond to the BM factor in earnings (Fama & French, 1995). Furthermore, the model fails to represent the investor experience, and assumes linearity and no interaction between the three factors: BM, size, and the excess return of the market - although studies have found evidence of the contrary (Lyon, Barber, & Tsai, 2003).

8.1.2 Short-Term BHSR

The table below shows the three models, first the IV model, then we add the industry and last we test our control variables. None of the variables have a significant impact on company performance. Nonetheless, the results indicate a negative PGP effect on returns.

3-Month Buy and Hold Stock Returns

	(1)	(2)	(3)
PGP	-0.106 (0.150)	-0.153 (0.157)	-0.106 (0.152)
Energy and Power		-0.156 (0.242)	
Financials		-0.074 (0.199)	
Healthcare		0.123 (0.219)	
High Technology		-0.073 (0.203)	
Industrials		0.180 (0.221)	
Materials		-0.107 (0.244)	
Media and Entertainment		-0.094 (0.275)	
Unrelated			0.073 (0.106)
Tender			-0.071 (0.187)
Domestic			0.068 (0.135)
Mix			-0.037 (0.100)
Shares			0.068 (0.131)
Constant	.121 (0.082)	0.168 (0.205)	0.052 (0.151)
Observations	674	674	674
Adjusted R ²	0	0	0

Standard error is presented within the parentheses *p<0.1; **p<0.05; ***p<0.01
 Two observations did not have three-month stock return in Refinitiv Eikon, which is why the sample decreased by two. Furthermore, we did not include year fixed effects in this regression due to the short-term horizon of the measurement period. When we did include them, the years did not significantly contribute to explaining the three-month post-acquisition returns.

8.1.3 Outlier Exemption Robustness Check

For the excluding outliers regression, we excluded only outliers in stock returns. We did this using a Z-score method, in which any observations that were three standard deviations higher or lower than the mean, in any year, were excluded. From this process, we removed 27 observations and received a new full sample size of 649 for the one-year stock returns.

BHSR, Outliers Excluded

Year	(1)	(2)	(3)	(4)	(5)
IV MODEL					
PGP	-0.148*** (0.053)	-0.214*** (0.077)	-0.148 (0.091)	-0.194* (0.115)	-0.063 (0.151)
Constant	0.042 (0.028)	0.142*** (0.041)	0.228*** (0.049)	0.340*** (0.062)	0.432*** (0.081)
Adjusted R ²	0.012	0.012	0.004	0.005	0.0003

FULL MODEL

PGP	-0.009 (0.052)	-0.060 (0.075)	-0.048 (0.093)	-0.077 (0.118)	0.038 (0.155)
Unrelated	-0.029 (0.033)	-0.017 (0.047)	0.051 (0.058)	0.102 (0.076)	0.114 (0.099)
Tender	0.035 (0.057)	0.068 (0.081)	-0.001 (0.102)	-0.122 (0.132)	-0.171 (0.173)
Domestic	0.091** (0.041)	0.176*** (0.058)	0.165** (0.073)	0.149 (0.092)	0.231* (0.120)
Mix	-0.090*** (0.031)	-0.112** (0.045)	-0.133** (0.056)	-0.137* (0.073)	-0.170* (0.095)
Shares	-0.068* (0.040)	-0.133** (0.058)	-0.172** (0.071)	-0.271*** (0.092)	-0.323*** (0.120)
Constant	-0.262*** (0.082)	-0.475*** (0.117)	-0.310** (0.147)	-0.292 (0.188)	-0.360 (0.246)
Observations	649	621	595	570	557
Adjusted R ²	0.198	0.223	0.153	0.125	0.128
Year & Industry FE	Yes	Yes	Yes	Yes	Yes

Standard error is presented within the parentheses

*p<0.1; **p<0.05; ***p<0.01

Cross Industry Sensitivity (PGP x Industry FE), Outliers Excluded

Year	(1)	(2)	(3)	(4)	(5)
Consumer Products and Services (Baseline)	0.012 (0.282)	-0.250 (0.399)	-0.436 (0.491)	-0.431 (0.639)	-0.139 (0.838)
	0.016 (-0.409:0.313)	-0.254 (-0.116:0.636)	-0.439 (-1.578:0.621)	-0.454 (-1.446:0.651)	-0.149 (-1.933: 1.529)

Energy and Power	0.640* (0.334)	0.708 (0.474)	0.718 (0.582)	0.974 (0.760)	1.365 (1.022)
	0.642 (0.114:1.274)	0.710 (-0.312:1.821)	0.777 (-0.798:2.098)	0.990 (-0.668: 2.747)	1.355 (-0.912: 3.998)
Financials	-0.143 (0.294)	-0.037 (0.416)	0.135 (0.512)	0.108 (0.666)	-0.183 (0.871)
	-0.146 (-0.473:0.310)	-0.045 (-0.929:0.837)	0.147 (-0.956:1.325)	0.130 (-1.022: 1.157)	-0.166 (-1.916: 1.604)
Healthcare	-0.141 (0.320)	-0.002 (0.459)	0.205 (0.567)	0.064 (0.734)	-0.671 (0.966)
	-0.138 (-0.635:0.396)	0.018 (-1.045:0.925)	0.223 (-1.037:1.458)	0.097 (-1.468:1.455)	-0.628 (-3.010:1.531)
High Technology	0.050 (0.303)	0.499 (0.430)	0.738 (0.531)	0.615 (0.689)	0.519 (0.903)
	0.051 (-0.416:0.534)	0.520 (-0.526:1.401)	0.750 (-0.472: 1.932)	0.641 (-0.608:1.720)	0.540 (-1.306:2.360)
Industrials	-0.247 (0.321)	0.082 (0.453)	0.284 (0.558)	0.414 (0.722)	0.607 (0.943)
	-0.242 (-0.72:0.196)	0.091 (-0.955:1.090)	0.313 (-0.917:1.447)	0.473 (-0.890:1.686)	0.661 (-1.398:2.639)
Materials	0.086 (0.329)	0.327 (0.466)	0.712 (0.573)	0.698 (0.739)	0.700 (0.965)
	0.114 (-0.516:0.698)	0.360 (-0.772:1.322)	0.768 (-0.574:1.927)	0.735 (-0.549:1.909)	0.764 (-1.201:2.531)
Media and Entertainment	0.096 (0.375)	0.455 (0.529)	0.773 (0.672)	0.254 (0.893)	-0.372 (1.160)
	0.068 (-0.602:0.789)	0.430 (-0.948:1.813)	0.699 (-1.402:2.980)	0.188 (-1.744:1.983)	-0.391 (-2.895:1.799)
Observations	649	621	595	570	557
Adjusted R ²	0.213	0.230	0.157	0.124	0.138
Year & Industry FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Standard error (SE) is shown in parentheses				*p<0.1; **p<0.05; ***p<0.01	

8.1.4 Three BCa Bootstrap Tables

Bootstrap 1: Stock Buy & Hold Returns, Year 4

	Estimate	95% CI		BootMed
		2.50%	97.50%	
FULL MODEL				
PGP	-0.358	-0.965	0.024	-0.337
Energy and Power	-0.587	-1.508	-0.098	-0.557
Financials	-0.354	-1.364	0.077	-0.321
Healthcare	0.310	-0.595	0.850	0.339

High Technology	-0.148	-1.101	0.235	-0.123
Industrials	-0.106	-0.962	0.390	-0.087
Materials	-0.457	-1.377	0.014	-0.435
Media and Entertainment	-0.233	-1.243	0.313	-0.199
2008	0.242	0.018	0.543	0.235
2009	0.749	0.287	1.326	0.743
2010	0.741	0.446	1.071	0.737
2011	0.812	0.513	1.175	0.807
2012	0.614	0.302	0.936	0.615
2013	0.517	0.223	0.871	0.518
2014	0.642	0.366	1.082	0.640
2015	0.368	0.131	0.602	0.374
2016	0.841	0.221	2.252	0.784
Unrelated	0.107	-0.129	0.349	0.110
Tender	-0.238	-0.493	-0.021	-0.233
Domestic	0.204	-0.051	0.445	0.204
Mix	-0.054	-0.252	0.181	-0.058
Shares	0.130	-0.220	0.636	0.121
Constant	0.111	-0.338	0.844	0.090

Bootstrap 2: BM, Size Buy & Hold Abnormal Returns, Year 4

		95% CI		BootMed
	Estimate	2.50%	97.50%	
FULL MODEL				
PGP	-0.363	-0.919	0.043	-0.364
Energy and Power	-0.583	-1.396	-0.093	-0.556
Financials	-0.358	-1.315	0.062	-0.336
Healthcare	0.290	-0.587	0.866	0.316
High Technology	-0.146	-1.112	0.240	-0.125
Industrials	-0.104	-0.996	0.376	-0.080
Materials	-0.482	-1.354	-0.025	-0.454
Media and Entertainment	-0.259	-1.236	0.285	-0.238
2008	-0.030	-0.269	0.284	-0.031
2009	0.122	-0.349	0.694	0.116

2010	0.214	-0.068	0.527	0.211
2011	0.364	0.069	0.731	0.364
2012	0.245	-0.051	0.582	0.242
2013	0.190	-0.083	0.556	0.190
2014	0.360	0.071	0.816	0.354
2015	0.130	-0.100	0.380	0.131
2016	0.571	-0.015	1.919	0.518
Unrelated	0.104	-0.124	0.363	0.103
Tender	-0.210	-0.443	0.010	-0.206
Domestic	0.220	-0.054	0.467	0.223
Mix	-0.058	-0.268	0.172	-0.058
Shares	0.119	-0.211	0.631	0.107
Constant	-0.059	-0.506	0.599	-0.074

Bootstrap 3: Industry Buy & Hold Abnormal Returns, Year 4

95% CI				
	Estimate	2.50%	97.50%	BootMed
FULL MODEL				
PGP	-0.358	-0.899	0.034	-0.022
Energy and Power	-0.442	-1.461	0.013	-0.361
Financials	-0.308	-1.351	0.082	-0.403
Healthcare	0.281	-0.629	0.786	-0.261
High Technology	-0.282	-1.217	0.069	0.327
Industrials	-0.160	-1.010	0.315	-0.240
Materials	-0.418	-1.313	0.012	-0.126
Media and Entertainment	-0.315	-1.328	0.180	-0.385
2008	-0.051	-0.282	0.230	-0.270
2009	0.181	-0.249	0.722	-0.047
2010	0.177	-0.111	0.489	0.165
2011	0.303	0.033	0.659	0.173
2012	0.148	-0.147	0.453	0.294
2013	0.056	-0.225	0.412	0.149
2014	0.254	-0.008	0.772	0.055
2015	-0.053	-0.281	0.186	0.231
2016	0.314	-0.313	1.617	-0.051
Unrelated	0.095	-0.116	0.364	0.272

Tender	-0.254	-0.495	-0.044	0.089
Domestic	0.225	-0.009	0.481	-0.250
Mix	-0.023	-0.213	0.209	0.223
Shares	0.138	-0.175	0.589	-0.030
Constant	0.002	-0.432	0.666	0.126