NBER WORKING PAPER SERIES

PRODUCT MARKET SYNERGIES AND COMPETITION IN MERGERS AND ACQUISITIONS: A TEXT-BASED ANALYSIS

Gerard Hoberg
Gordon M. Phillips

Working Paper 14289
http://www.nber.org/papers/w14289

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2008

This paper was previously circulated as "Product Market Synergies and Competition in Mergers and Acquisitions." Hoberg can be reached at ghoberg@rhsmith.umd.edu and Phillips can be reached at gphillips@rhsmith.umd.edu. We thank Michael Faulkender, Kathleen Hanley, Nagpurnanand Prabhala, David Robinson and seminar participants at Duke University, Ohio State University, the Securities Exchange Commission, and the University of Maryland for helpful comments. All errors are the authors alone. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Product Market Synergies and Competition in Mergers and Acquisitions: A Text-Based Analysis
Gerard Hoberg and Gordon M. Phillips
NBER Working Paper No. 14289
August 2008, Revised February 2012
JEL No. G3,G34

ABSTRACT

We examine how product differentiation influences mergers and acquisitions and the ability of firms to exploit product market synergies. Using novel text-based analysis of firm 10K product descriptions, we find three key results. (1) Firms are more likely to enter restructuring transactions when the language describing their assets is similar to all other firms, consistent with their assets being more redeployable. (2) Targets earn lower announcement returns when similar alternative target firms exist. (3) Acquiring firms in competitive product markets experience increased profitability, higher sales growth, and increased changes in their product descriptions when they buy target firms that are similar to them and different from rival firms. Our findings are consistent with similar merging firms exploiting synergies to create new products and increase their product differentiation relative to ex-ante rivals.

Gerard Hoberg
Robert H. Smith School of Business
University of Maryland
4423 Van Munching Hall
College Park, MD 20742
ghoberg@rhsmith.umd.edu

Gordon M. Phillips
Marshall School of Business
University of Southern California
Citigroup Center
Los Angeles, CA 90089
and NBER
Gordon.Phillips@marshall.usc.edu

It has long been viewed that product market synergies and competition are key drivers of mergers.[1] One important dimension of synergies is the ability of merging firms to create new products and differentiate themselves from rivals when merging firms have complementary assets. This dimension of synergies is frequently cited as being important for mergers but has not been documented as being an important factor for mergers.[2] We examine whether product market measures of similarity and competition affect firms' propensity to merge and whether firms use mergers and acquisitions to increase profits and introduce new products given asset complementarities.[3]

In competitive markets, mergers are a quick way to potentially increase product offerings if synergies arise from asset complementarities. Thus, firms may have incentives to merge with firms that have different skills or technologies that increase the ability to introduce new products. Another important concern to acquirers is to introduce new products that differentiate them from existing rival firms. There is thus a potential tension between merging with a firm whose products are very similar, and a firm whose skills or technologies are different enough from rivals to help differentiate the acquirer.

We explore this "similar but different" tension in examining the likelihood of mergers and ex post merger outcomes. Our paper provides direct evidence on asset complementarities and synergies as a source of gains from merging. We provide evidence on how similar firms are with respect to their existing products and also measure new product creation and the extent that product differentiation from rival firms increases post merger.

[1]See Andrade, Mitchell, and Stafford (2001) and Betton, Eckbo, and Thorburn (2007) for two surveys on different motives for mergers.

[2]Rhodes-Kropf and Robinson (2008) model asset complementarity and synergies as a motive for mergers and provide evidence that merging firms have similar market-to-book valuations. However, given they do not have product data or descriptions, they are unable to provide direct evidence that firms with asset complementarities combine to exploit potential synergies. Healy, Palepu, and Ruback (1992) and Andrade, Mitchell, and Stafford (2001) have documented increased industry-adjusted cash flows following mergers. However, the literature has not been able to identify if asset complementarities allowing firms to introduce new products are responsible for gains in cash flows.

[3]Chamberlain (1933) and Hotelling (1929) famously show that the notion of product differentiation is fundamental to theories of industrial organization, with product differentiation increasing the profitability of firms. The Hotelling model allows for a concept of distance and makes a distinction between "close" rivals versus "distant" rivals, a concept that we exploit. In Chamberlain, product differentiation exists but all rivals are equally close. See also the survey by Eaton and Lipsey (1989).

We employ a novel empirical design using text-based analysis of firm product descriptions obtained from web-crawling algorithms that read and process 49,408 10-K filings from the SEC Edgar website. From firm product descriptions we calculate measures of product similarity and product market competition. Our product based measures capture the degree to which products are different from rivals - while independently measuring the similarity between the acquirer and target. Our framework also permits analysis of product descriptions in time series, and we can thus analyze likely synergies in the form of new product introductions.

We test how competition and product similarity affect the likelihood of mergers and acquisitions, the size of stock-market gains, and which transactions are most successful in the long-run. Our similarity measures allow us to directly address the dual role of competition and relatedness in merger decisions and outcomes, and permit us to test whether new product synergies are closely tied to the positive real outcomes we report. Our hypotheses relating ex post outcomes of mergers to product market competition and product similarity cannot be tested using similarity measures based on traditional SIC codes, given that they are too granular and do not capture similarity both within and across industries.

We report three major findings. First, we find that the distribution of rivals around a given firm strongly predicts the likelihood that it will merge. Second, transactions in competitive product markets with similar acquirer and target firms experience increased stock returns and real longer-term gains including higher profitability and increased changes in their product descriptions. Third, these findings are especially strong when the target is less similar to the acquirer's closest rivals. These results are robust to the inclusion of measures of horizontal and vertical relatedness using SIC codes and the product input-output matrix. We now summarize each of these findings in more detail.

Regarding transaction likelihood, we find that firms having products that are more broadly similar to all other firms in the economy are more likely to restructure, and firms having more very close rivals are less likely to restructure. We interpret the first effect as an "asset deployability effect", as firms with more broadly similar products likely have assets that can be more easily deployed in other product markets, and hence face a larger opportunity set of possible transactions. We interpret the second effect as a "competitive effect", as firms with very near rivals must compete for any available restructuring opportunities. These effects are significant, economically and statistically, and for large and small firms alike.

We also find that mergers and acquisition transactions experience higher market reactions and positive long-term real outcomes (higher profitability and evidence of new product introductions) when acquirers reside in ex-ante competitive product markets, and when the transaction increases the acquirer's product market differentiation relative to its close rivals. These findings are both statistically and economically significant. Our results suggest that product market competition can be influenced by merging. In particular, firms residing in ex-ante competitive industries can use restructuring to reduce competition. We also find that gains from this strategy are larger when the target and the acquirer are more pairwise similar, consistent with mergers involving firms with related products experiencing significant gains. Gains are also larger when there is evidence of barriers to entry, suggesting that rivals will not be able to replicate the new products.

We differ from the existing literature in three major ways. First, the literature has not focused on how product differentiation and competition affect the probability of restructuring and subsequent ex-post performance and product development. Second, we present an innovative new methodology based on text analysis that allows us to separately measure an acquirer's similarity to its rivals, the location of the target relative to the acquirer's rivals, and the target and acquirer pairwise similarity. Existing studies have relied on SIC code based variables to measure product market competitiveness and firm similarity. Most recently, Fan and Goyal (2006) show that mergers that are vertically related using the input-output matrix have positive announcement wealth effects in the stock market. While partially informative, SIC code based measures cannot measure the degree to which firms are similar within and across industries. Our continuous measure of similarity is important economically and statistically even after controlling for both horizontal and vertical measures of relatedness using SIC codes. We document that the product similarity of merger pairs is highly diverse *both within and across industries*. Our measure also captures how close merging firms are to existing rival firms. Our results show that our measure of similarity and competition significantly helps explain ex post changes in cash flows, sales, and new product introduction. Lastly, we also measure barriers to entry by examining the frequency of the word root "patent".

Our research contributes to existing strands of literature using text analysis in finance. Hanley and Hoberg (2008) examine prospectus disclosures on the SEC Edgar website and construct section size measures and document similarity measures to address theories of IPO pricing. Outside of corporate finance, other studies that use text based analysis to study the role of the media in stock price formation include Tetlock (2008), Macskassy, Saar-Tsechansky, and Tetlock (2004), Li (2006a), and Boukus and Rosenberg (2006).

The remainder of the paper is organized as follows. A discussion of merger strategies and the incentives to merge is in section I. The data, methodology and summary statistics are presented in section II. Section III introduces our new empirical measures of product market similarity, and Section IV examines key predictions regarding their relationship to firm profitability. Determinants of the likelihood of restructuring transactions are presented in section V. A discussion of ex-post outcomes upon announcement and in the long term is in section VI. Section VII concludes.

I Incentives to Merge and Post-Merger Performance

This section first discusses existing literature on mergers and, second, develops our hypotheses of how potential changes to product differentiation and competition may affect a firm's decision to merge and its post-merger performance.

A Relation to Previous Literature on Mergers

Our paper focuses on how high ex ante product market competition creates incentives to merge, and how mergers may create profits through synergies and subsequent product differentiation. The central idea is that firms may wish to merge with partners with complementary assets who expand their range of products through new product introductions (enabling them to differentiate their products from rival firms) - while also picking partners that are related enough so that they can skillfully manage the new assets.

This rationale for mergers is distinct from the existing motives in the finance literature.[4] Rhodes-Kropf and Robinson (2008) also consider synergies through asset complementarities as a motive for mergers - but do not consider how competition impacts this motive, nor do they empirically measure ex post new product introduction. We also use text based methods to directly measure asset complementarity of merger partners rather than inferring it using market-to-book ratios. Overall, the existing literature does not consider how product market competition affects the incentives to choose targets that can increase product market differentiation from close rivals as in Hotelling (1929). Perhaps more importantly, using new techniques, our paper can measure potential synergies through asset complementarities and how similar merging firms are from rivals and also document ex post changes in product differentiation post merger.

In addition to increases in product differentiation, key to understanding ex post performance after mergers are measures of relatedness of the target and acquirer. As emphasized by many authors, related acquisitions have the potential to perform better as the acquirer is likely to have existing skill in operating the target firm's assets. Kaplan and Weisbach (1992) show that related mergers are less likely to be divested subsequently by the acquirer - although they do not find any difference in the performance of diversifying versus non-diversifying mergers. Maksimovic, Phillips, and Prabhala (2008) find that acquirers with more skill in particular industries are more likely to maintain and increase the productivity of the assets they acquire and keep in related industries. Fan and Goyal (2006) show that mergers that are vertically related using the input-output matrix have positive announcement wealth effects in the stock market. While partially informative, SIC code based measures cannot

[4]Existing reasons for mergers include technological industry shocks and excess industry capacity (Morck, Shleifer, and Vishny (1988), Jensen (1993), Mitchell and Mulherin (1996), Andrade and Stafford (2004), Harford (2005)), reduction of agency problems (Jensen (1993)), agency and empire building (Hart and Moore (1995)), demand shocks and efficiency (Maksimovic and Phillips (2001), Jovanovic and Rousseau (2002), Yang (2008)) and the industry life cycle (Maksimovic and Phillips (2008)).

measure the degree to which firms are similar within and across industries. Our study shows that many merging firms are highly related using text-based measures even though they have different two-digit SIC codes that are not related. Most importantly, discrete measures of relatedness cannot capture how related rivals are to the merging firms both within and across industries.

Research in industrial organization has also studied mergers in industries with existing differentiated products. Baker and Bresnahan (1985) theoretically model how mergers can increase pricing power by enabling post-merger firms to increase prices. The prescribed policy is to merge with firms whose products are close substitutes to yours, when other non-merging firms only produce distant substitutes. Pricing power is enhanced by merging because post-merger firms face an increased steepness in their residual (inverse) demand curves. Empirical studies in this literature have focused on estimating own and cross-price elasticities of demand, as well as the effect on post-merger prices in specific markets including the ready-to-eat cereal market (Nevo (2000)).

Our study has a different focus. Instead of taking the extent of product differentiation as given, we examine how producers in competitive markets with similar products may be able to use mergers to increase their differentiation relative to rival firms by using synergies to enhance and develop new products. The flexibility of our text based measures, especially our ability to measure product similarity across arbitrary groups of firms and in time series, allows us to test for evidence of new product introductions.

B Product Differentiation and Changing Competition

A simple example illustrates how managers can reduce competition via mergers when they reside in ex-ante competitive product markets, and illustrates why a refined notion of product differentiation and competition is important. First consider a manager producing product X in a competitive industry with 100 homogeneous firms. Here, mergers offer little help in reducing competition because an oligopoly of 99 is little different than an oligopoly of 100 firms. However, consider the situation if the 100 firms are somewhat differentiated, with some firms having complementary assets. We assume that firms with complementary assets can merge and the two firms can jointly produce a new product Z at a lower expected cost (equivalently a higher probability of success).[5] Gains from a merger are possible if product Z is somewhat different from existing products such that the new combined firm faces less competition and enjoys higher profit margins. Hence, firms can use mergers to substantially alter the degree of competition they face when they reside in differentiated product markets. Only an empirical design capable of measuring the degree of product similarity both within and across industries can test these theories.

C Hypotheses

We now develop specific hypotheses based on the aforementioned theories of merger pair similarity and industrial organization. Our first two hypotheses concern the probability that a given firm will become part of an acquisition. Our last three hypotheses formulate predictions regarding ex post outcomes.

Hypothesis H1: Asset Redeployability: *Firms with assets that are similar to large numbers of firms are more likely to enter into restructuring transactions.*

Hypothesis H2: Competition for Targets: *Firms with high local product market competition are less likely to be targets and enter restructuring transactions given the existence of multiple substitute target firms. This competition not only reduces the likelihood that any given firm will merge, but will also reduce the premium realized by a target when it does merge.*

Key to testing H1 and H2 is the degree of similarity. The effects of H2 are likely to be more severe when firms are very similar. For example, identical firms would have to compete in order to merge with a third firm offering new synergies requiring their technology. In contrast, the effects of H1 are more likely to hold when firms are moderately similar, as such firms are less viable substitutes. Also, some degree of heterogeneity is likely required in order for new product synergies to be possible.

[5]It is also necessary that contracting costs and ex post holdup prevent a contract from being signed between the firms.

Our remaining hypotheses focus on long term outcomes.

Hypothesis H3: Differentiation from Rivals: *Acquirers in competitive product markets experience positive outcomes when they buy target firms that help them to increase product differentiation relative to their nearest ex-ante rivals.*

Hypothesis 3 follows from two ideas. First, basic industrial organization (as in Hotelling (1929)) suggests that firms that differentiate themselves and generate local product monopolies should be more profitable. Second, profits from this strategy can be leveraged through new product introductions, especially when synergies allow products in markets with lower competition from rivals.

Hypothesis H4: Synergies through Asset Complementarities: *Acquirers buying targets similar to themselves are likely to have asset complementarities and experience future higher profitability and product differentiation.*

The reasons why acquirer and target pairwise similarity should result in positive outcomes as in Hypothesis 4, are numerous as discussed earlier in section I. Our primary rationale is that products developed using technology from similar targets with complementary (similar) assets are more likely to succeed as in Rhodes-Kropf and Robinson (2008).

The merger strategies underlying Hypotheses 3 and 4 might be especially relevant in competitive product markets because the gains from product differentiation are likely to be at a maximum when competition is reduced (Hypothesis 3), and highly similar targets are more likely to exist (Hypothesis 4). Although they might seem at odds, these hypotheses are not mutually exclusive. Key to maximizing gains is finding a target that is both somewhat distant from the acquirer's closest rivals, yet not too distant from the acquirer itself - "similar to self and different than rivals".

[**Insert Figure 1 Here**]

We illustrate these hypotheses using two examples: the Symantec and Veritas (SV) merger, and the Disney and Pixar (DP) merger. We base our discussion on actual similarity data (described fully in Section II). Figure 1 displays the SV merger, and the ten nearest rival firms surrounding both Symantec and Veritas. Firms with a label

"S" are among Symantec's closest rivals and firms with a "V" are among Veritas's closest rivals. Both firms offer products that are indeed related to each other (Veritas is Symantec's 18th closest firm and Symantec is Veritas's 37th closest), but also quite different. Symantec focuses on anti-virus software and Veritas focuses on internet security and authentication. The figure shows that Symantec faced a fair amount of competition from its rivals including Cyberguard, McAfee, and Watchguard, for example.

The SV merger is consistent with Symantec choosing Veritas because it is similar enough to permit successful managerial integration and new product synergies (Hypothesis 4), perhaps in the form of new security products defending joint PC and internet applications. This merger also might help Symantec to differentiate itself from its rivals and introduce new products that will face little in the way of initial competition (Hypothesis 3), which in turn will improve profit margins. Hence, Symantec might use the merger to change the degree of product market competition it faces, a strategy that might be especially effective if Veritas owns key patents or trade secrets preventing rivals from launching similar products (Hypothesis 5, see below). Importantly, this notion of similar but different is underscored by comparing the chosen target to a hypothetical merger between Symantec and one of its other near rivals such as McAfee. Because the products offered by the firms in this hypothetical pair are closer to being perfect substitutes, both in terms of technology and end consumer usage, gains from Hypotheses 3 and 5 are likely not possible, explaining why Veritas might be a more appropriate target.

[**Insert Figure 2 Here**]

The DP merger in Figure 2 shows by example that the same logic applies to larger firms with many products. Disney is Pixar's sixth closest rival, and Pixar is Disney's ninth closest rival. Only one other firm, Newscorp, is a member of both firms' nearest ten rivals. Pixar's nearest two rivals are Dreamworks and IMAX, consistent with Pixar's focus on computer generated motion pictures. Disney is ranked ninth on Pixar's list likely because it offers many products including television programming and amusements that are related to, but not exactly substitutes for motion pictures.

Disney's nearest two rivals are Newscorp and and CBS, and Six Flags is its fourth closest rival, suggesting that Disney has a significant television and print media presence, as well as an amusement park presence. Pixar's appeal to Disney might be due to both its asset similarity, but also its moderately different product offerings, as this merger might permit seamless integration and new product introductions that can enhance pricing power in niche markets.

If Disney were to merge with its closest rival, it would choose to merge with NewsCorp, or perhaps with one of the other top ranked firms. This transaction would score highly on pairwise product similarity (Hypothesis 4), but it might not be different enough from existing rivals to enable the introduction of new products in less competitive product markets (Hypothesis 3). In contrast, Pixar has the benefit of still being pairwise similar to Disney (its ninth closest firm), but it also has the virtue of being distant from Disney's other close rivals. This point is underscored by the fact that only one firm other than Disney and Pixar themselves are in both firms' nearest ten. Hence, Pixar scores highly on both dimensions and is both similar to the acquirer (Hypothesis 4) and distant from Disney's other closest rivals (Hypothesis 3). This right mix of "similar to self but different than rivals" should increase product differentiation and profitability for Disney.

It is also relevant to consider the reaction of Disney's rivals to mergers fitting this profile. The hypothesized gains might be transient, for example, if the target's assets and product lines are not protected by high replication costs or patents. Rival firms could develop similar products and erase any pricing power generated by the transaction. Our next hypothesis is based on this intuition.

Hypothesis H5: Barriers to Entry: *Gains from restructuring will be even larger if barriers to entry such as patents exist, as rivals would be less able to replicate the acquirer's strategy.*

II Data and Methodology

A Data Description

The key innovation of our study is that we use firms' 10-K text product descriptions to compute continuous measures of product similarity for every pair of firms in our sample (a pairwise similarity matrix). We construct these text-based measures of product similarity (described later in this section) using firm product descriptions that we obtain directly from 10-K filings on the SEC Edgar website during the period electronic 10K records are available. We merge these product similarity measures to the COMPUSTAT/CRSP database using the tax identification number (also known as the employee identification number), and we then link this firm-level database to the SDC Platinum database of mergers and acquisition of asset transactions. In order to be in our firm-level database, for both firms involved in mergers and acquisitions and all other firms, a firm must exist in both the CRSP and COMPUSTAT databases in the given year of analysis. 52,013 firm years pass this initial screen.

We electronically gather 10-Ks by searching the Edgar database for filings that appear as "10-K", "10-K405", "10KSB", "10KSB40". Prior to 1997, the Edgar database is somewhat sparse as electronic filing was not required until 1997. Of the 52,013 firm years that are present in both CRSP and COMPUSTAT, we are are able to match and read 49,408 filings (95%) associated with fiscal years ending in 1997 to 2005. Note we do not include 10-K filings from investment trusts, tracking stocks, and inactive firms.[6] These 49,408 filings match with the COMPUSTAT database based on their tax ID number, and most match directly without any intervention. However, a small number of firms (roughly 1% to 2% of our sample) experienced changes in their tax ID number during our sample period, and we hand correct links for these firms.[7]

[6]Although we consider fiscal year endings through 2005, we extract documents filed through December 2006, as most of the filings in 2006 are associated with fiscal years ending in 2005. Hence, our main database ends in 2005, as we consider the fiscal year end to be the unit of observation, as is the case in the standard COMPUSTAT database. Regarding the start date of our main database, we use observations associated with 1996 fiscal year endings only for the purposes of computing lagged variables, and hence our main database begins in 1997. To ensure maximum coverage, we search for filings beginning in January 1996 as some firms with fiscal years ending in the earlier part of 1996 file their 10-K in calendar year 1996.

[7]The COMPUSTAT tax ID number is only available as a header variable and this variable thus reflects the firm's most recent tax ID number. Because some firms maintain the same COMPUSTAT

We extract the product description section from each linked 10-K. This section appears as Item 1 or Item 1A in most 10-Ks. We utilize a combination of PERL web crawling scripts, APL programming, and human intervention (when documents are non-standard) to extract and summarize this section. The web crawling algorithm scans the Edgar website and collects the entire text of each 10-K annual report, and APL text reading algorithms then extract its product description and tax number. This latter process is extensively supported by human intervention when non-standard document formats are encountered. This method is highly reliable and we encountered only a very small number of firms (roughly 100) that we were not able to process because they did not contain a valid product description or because the product description had fewer then 1000 characters. These firms are excluded from our analysis.

Our database of 49,408 filings in our sample years of 1997 to 2005, and 6485 additional observations for fiscal years ending in 1996 (these latter observations are used solely for computing the values of lagged variables) represents 95.0% of the eligible CRSP and COMPUSTAT database. We can also report that our database is well balanced over time, as we capture 95.6% of the eligible data in 1997, and 94.3% in 2005, and this annual percentage varies only slightly in the range of 93.6% in 2000 to 95.9% in 2003. Because we do not observe any time trends in our data coverage, and because database selection can be determined using ex-ante information (ie, the 10-K itself), we do not believe these requirements induce any bias. Our final sample size is 47,394 rather than 49,408 because we additionally require that key COMPUSTAT data items, (sales, assets and operating cash flow), are populated before observations can be included in our analysis.

gvkey and CRSP permno even as their Tax ID number changes, these hand corrections ensure that our database has uniform coverage over all of our sample years. These hand corrections are mainly based on comparing the firm names at the time of filing to those listed in the CRSP historical names database. Although fuzzy matching is used to suggest possible corrections, all final corrections are hand verified.

B Product Similarity

For each firm i and j, we measure product similarity using a method based on each firm's empirical distribution of words used in its product description. This method results in a real number in the interval (0,1) to capture the similarity of words used for each pair of firms. Details are discussed in Appendix 1. The main idea is that firms having more common word usage are scored as having a higher degree of product similarity. The method uses a normalization to avoid over-scoring larger documents, and a simple adjustment to exclude very common words. Based on these pairwise similarities, we compute the following measures local and broad similarity. We also compute a measure of barriers to entry based on usage of the word root "patent".

Product Similarity (All Firms - 10): For a given firm i, this variable is the average pairwise similarity between firm i and all other firms j in the sample - excluding i's 10 closest rivals.

Product Similarity (10 Nearest): For a given firm i, this variable is the average pairwise similarity between firm i and its ten closest rivals j. The closest rivals are the ten firms having the highest similarity relative to i.

% Neighbor Patent Words: For each firm, we first compute the percentage of all of the words in its product description having the word root "patent". Because our focus is on whether a product market is protected by barriers to entry (hypothesis H5), we then compute this variable's average over each firm's ten nearest rivals.

Last Year 10 Nearest Fraction Restructured: For firm i, this variable is equal to the fraction of its ten closest rivals that were either a target or an acquirer in the previous year according to the SDC Platinum database.

Target + Acquirer Product Similarity: For a given merger pair (target and acquirer), this is the product similarity of the two firms.

Gain in Product Differentiation: This variable is the target's average distance from the acquirer's 10 nearest rivals minus the acquirer's average distance from its 10 nearest rivals. Pairwise distance is one minus pairwise similarity. This variable measures the degree to which an acquirer gains product differentiation from its rivals by purchasing the given target.

Product similarities are most intuitive when firms have only one segment. Importantly, our computer-based algorithms are not able to separate the text associated with each segment of conglomerate firms.[8] However, we believe that similarities measured relative to conglomerates are still informative regarding the competition faced

[8] Multiple segments appear in product descriptions, but automated text reading algorithms cannot separate the text attributable to each due to the high degree of heterogeneity in how segment descriptions are organized.

by the firm in all of its segments. To the extent that multiple segments might add noise to our measures, we do not see this as a problem as it would only bias our results away from finding significant results. However, to ensure our inferences do not change, we also rerun all of our tests using the subsample of single segment firms. Although not reported to conserve space, our results change little in this subsample.[9]

C Other Control Variables

Past studies seeking to measure product differentiation have been forced to rely on industry definitions based on SIC codes. We thus control for the following standard measures of industry competition and merger similarity based on SIC codes.

Sales HHI (SIC-3): We use the two step method described in Hoberg and Phillips (2008) to compute sales-based Herfindahl ratios for each three-digit SIC code. This method uses data based on both public and private firms to compute the best estimate of industry concentration given the limited data available on private firm sales.

Last Year SIC-3 % Restructured: For a given firm i, we compute the percentage of firms in its three-digit SIC code that were involved in restructuring transactions in the previous year according to the SDC Platinum database.

Same SIC-3 Industry Dummy: For a given merger pair, this is variable is one if the two firms are in the same three-digit SIC code.

Vertical Similarity Dummy: For a given merger pair, this is variable is one if the two firms are at least 5% vertically related. We use the methodology described in Fan and Goyal (2006) to construct this variable. In particular, based on four digit SIC codes of both the target and the acquirer, we use the Use Table of Benchmark Input-Output Accounts of the US Economy to compute the fraction of the inputs that are from the other firm's SIC industry. If this percentage exceeds 5% for either firm, then the dummy is set to one. For robustness we also replace this dummy variable with the percentage shipped from merger partner's industry i to merging partner's industry j.

Although Herfindahl indices based on SIC codes are informative in many applications, we present evidence that SIC codes provide only a weak measure of competition. This is primarily due to their granularity. We include both SIC code based measures and text-based measures of similarity throughout our analysis, and we document the benefits of each. We also include controls for the following variables at the transaction level.

[9]In some isolated cases, significance levels decline from the 1% level to the 5% or 10% level due to reduced power.

Target Relative Size: The pre-announcement market value of the target divided by the sum of the pre-announcement market values of the target and the acquirer.

Merger Dummy: A dummy variable equal to one if the given transaction is a merger and zero if it is an acquisition of assets. We only examine these two transaction types.

Merger x Relative Size: The cross product of the above two variables.

Log Total Size: The natural logarithm of the sum of the pre-announcement market values of the target and the acquirer.

III Mergers and Product Market Similarity

We begin by exploring how mergers relate to our product market similarity measure. To show their uniqueness relative to SIC codes, Table I lists all restructuring pairs in 2005 that are very similar (the top percentile of similarity from all firm pairings) despite the fact that they reside in different two-digit SIC codes. The list suggests that the high degrees of similarity are, in fact, due to real product similarities. For example, petroleum and pipeline firms are related. Newspapers and radio are also related, and can be viewed as substitute sources of advertising despite their being in different two-digit SIC codes.

To further illustrate how the algorithm rated these firms, Table II displays the full list of words that were common for the first ten of these related transactions. Appendix 2 presents the word lists for the remaining transactions in Table I. Table II further illustrates the limitations of using SIC codes as an all or nothing classification of merger pair similarity. The word lists suggest that the similarity calculations are indeed driven by product market content, consistent with our interpretation of the similarity measures as representing product market similarity. Key to this result is our focus on non-common words, as our similarity calculations are based only on words that appear in no more than 5% of all 10-Ks in the given year. This eliminates templates, legal jargon and other non-product related content. The list of similar words for each pair is also substantial, indicating that our basis for defining similarity is informative.

Table III displays summary statistics for our key variables based on our firm and transaction level databases. 15.1% of the firms in our firm-level database were targets either of a merger or an acquisition of assets, and 28.2% were acquirers. These numbers are somewhat larger than some existing studies because (1) our sample includes more recent years in which transactions were more common, (2) these figures include both mergers and partial acquisitions, and (3) transactions are included if the counterparty is public or private. We next report the fraction of targets and acquirers by transaction type, and find that mergers (4.3% targets and 10.4% acquirers) are less common than asset acquisitions (10.8% targets and 17.7% acquirers). However, both transaction types are sufficiently common to permit statistical analysis.

All product similarities are bounded in the range (0,1). The average product similarity between randomly chosen firms (excluding ten closest rivals) is .017 (or 1.7%). The average similarity between a firm and its ten closest neighbors is considerably higher at 15.9%. The average sales based HHI for firms in our sample is .048. The average percentage of rival firm 10-K words having the word root "patent" is 0.147%.

In Panel B, we report summary statistics for the transaction level database. Importantly, we measure ex-post changes in profitability, sales growth, and expenses of the acquiring firm starting from the year after the transaction becomes effective (ie, year t+1).[10] For example, a three year change in profitability is equal to profitability in year t+4 minus that in year t+1. The average acquirer experienced an announcement return very close to zero (0.4%), and the average target experienced an announcement return of 9.5%. The average merger pair is 9.3% similar, and the average target is 8.8% less similar to the acquirer's rivals than the acquirer itself is (ie, an average 8.8% potential gain in acquirer product differentiation). Panel C shows that the average acquiring firm experiences very little change in profitability, and 15.9% to 27.0% sales growth over the one to three year horizon following the transaction.

Table IV displays Pearson correlation coefficients between our measures of product differentiation and other key variables. Product similarity relative to all firms (excluding nearest ten) is 60.1% and 52.0% correlated with similarity measured relative

[10] Although this is conservative and reduces the power of our tests (some performance gains might accrue immediately), this avoids potential bias due to the challenges associated with computing the ex-ante performance given that two firms exist prior to the transaction, and because most transactions are partial acquisitions.

to the one hundred and ten nearest neighbors, respectively.[11] We also find that the sales HHI variable is roughly -10% correlated with the product similarity variables. This suggests that firms in concentrated industries have somewhat lower product similarities, which is consistent with higher product similarity and lower HHI both being associated with product market competition. However, this correlation is modest and both measures contain much distinct information. Overall, we conclude that most correlations are small and that multi-collinearity is not likely to be an issue in our analysis. We confirm this later using formal tests for multicollinearity.

A The Similarity Measure

Figure 3 displays the distributional properties of our pair-wise similarity measure. The uppermost graph displays the distribution of similarities for all randomly chosen firm pairs (ie., we do not condition on restructuring). The vertical axis is the frequency and the horizontal axis is the pairwise similarity expressed as a percentage. Randomly chosen firms generally have similarity percentages ranging from zero to four, but a relatively fat tail also stretches beyond scores of ten percent. The second graph displays similarities for firms entering into restructuring transactions. Our broad conclusion is that restructuring pairs are highly similar relative to randomly selected firms, and that merger pair similarities are quite diverse with considerable mass attached to values ranging from zero all the way to thirty.

[**Insert Figure 3 Here**]

Existing studies measure merger pair similarity by asking if the target and acquirer reside within the same SIC code or in vertically related industries. The third and fourth graphs confirm that product similarities are very high for merger pairs in the same two and three-digit SIC codes. However, the high diversity of similarities illustrates that the granular nature of SIC codes fails to capture much product heterogeneity. Also striking, however, is the relatively high level of merger pair similarity observed for merger pairs residing in different two-digit SIC codes in the bottom most

[11] Although we focus on the ten nearest neighbors in most of our analysis, our results change little if we instead focus on the 100 nearest neighbors.

graph. Firms in this sample are in different two-digit SIC codes, and hence any study assessing merger pair similarity on the basis of SIC groupings would label these pairs as being dissimilar. We find that the overwhelming majority of firms in this group have relatively high levels of similarity compared to the randomly chose firms in the top most graph. Even if we exclude firms that are in vertically related industries we still find striking evidence that merging firms in unrelated industries are in many cases very similar. We conclude that product similarity adds information not contained in SIC codes and allows us to measure the degree of similarity within industries and also across dissimilar industries.

IV Product Similarity, Product Market Competition and Profitability

Before beginning our examination of mergers and acquisitions, we test whether our similarity variables constructed from the 10-K product descriptions explain the profitability of firms in cross section. We examine the relation between our similarity variables and profitability to provide verification that our similarity measures are valid measures of product characteristics for all firms. The theoretical link is that our similarities are based on an empirical multi dimensional Hotelling grid that underlies the product space of all firms (the grid has as many dimensions as there are words in firm 10-K product descriptions, and a firm's location is determined by its word usage).

Theories of industrial organization predict that firms residing in locations where fewer close rivals exist - thus firms with more product differentiation - should enjoy higher profit margins in equilibrium. Hence, if our empirical measure of product market competition is valid, there should be a negative relationship between a firm's similarity relative to its ten nearest rivals, and its profitability.

Table V displays the results of OLS regressions in which one observation is one firm in one year. The dependent variable is profitability defined as operating income scaled by sales (Panel A) and scaled by assets (Panel B). We present results in groups of three: all firms, large firms (above median book assets) and small firms. In all specifications, we include three digit SIC industry fixed effects and we report $t-$statistics that are adjusted to account for clustering at the year and three digit SIC industry level.

[Insert Table V Here]

Table V strongly confirms that our product similarity variable is negatively related to profitability. This finding is significant both statistically (at better than the 1% level) and economically. Where economic impact is defined as a regression coefficient multiplied by the given variable's standard deviation, and based on the standard deviations from Table III and the coefficients from Table V, rows one and three imply that the economic impact of product differentiation on profitability is 2.6 percentage points. When profitability is scaled by assets rather than sales (Panel B), this number is still large at 1.7 percentage points. These numbers are roughly ten percent of the total standard deviations of the dependent variables (operating income scaled by sales and assets have standard deviations of 35.4 and 23.7 percentage points, respectively). We conclude that our 10-K based measures of product market competition are indeed consistent with key predictions of industrial organizational theories, lending further support to our assumption that the ten nearest similarity variable is a valid proxy for local product market competition and product differentiation.

V Merger and Asset Acquisition Likelihood

In this section, we test our first two hypotheses and examine the likelihood of restructuring transactions and its link to measures of product market competition and similarity. Table VI displays the results of logistic regressions in which one observation is one firm in one year. All reported figures are marginal effects, and t-statistics are reported in parentheses. The dependent variable is a dummy variable equal to one if the given firm is a target of a restructuring transaction in a given year (Panel A) and a dummy equal to one if the given firm is an acquirer (Panel B). In all specifications, we report $t-$statistics that are adjusted to account for clustering at the year and industry level.

Consistent with Hypothesis 1, the table shows that a firm is more likely to be an acquirer or a target (especially an acquirer) if its overall product similarity to all firms is high. This result is highly significant for acquirers at the 1% level regardless of specification. However, it is somewhat weaker for targets, as the product similarity relative to all firms is significant only when the similarity relative to ten nearest neighbors is included in the model. Although this result is weaker, tests we conduct indicate that the observed significance in the joint model is robust as a conditional result and not due to multicollinearity.[12] We interpret these results in terms of asset deployment. A firm whose product is broadly similar likely owns assets that can be deployed in many different product markets. These firms, especially acquirers, engage in more restructuring transactions. We refer to this as the "Redeployment effect".

Second, consistent with Hypothesis 2, firms in highly competitive local product markets are less likely to restructure both as targets and acquirers. The product similarity relative to a firm's ten nearest rivals significantly negatively predicts transaction likelihood at the 1% level for both targets and acquirers. This finding does not depend on the specification. We interpret this as a "competitive effect", as firms having very similar rivals must compete for restructuring opportunities. Later in this section, we confirm that both the redeployment effect and the competitive effect are economically significant.

Table VI also shows that firms using the word "patent" in their product descriptions are more likely to restructure, both as targets and as acquirers. A likely explanation is that patents serve as barriers to entry, and hence a firm that needs a technology protected by patents has few options to acquire it outside of merging with the patent holder. For example, patents might effectively preclude entry via organic investment.

The table also shows that the SIC-based sales HHI variable is negatively related to restructuring for acquirers, and weakly positively linked to restructuring for tar-

[12]The two variables are only 51% correlated and the highest variance inflation factor (VIF) is 1.5, which is considerably below the critical value of 2.5, and far below the value of 10 deemed necessary to declare the presence of multicollinearity. The results are also not driven by functional form as they change little in an OLS linear probability model. We also randomly divide the sample into two and find similar results on the resulting subsets of data.

gets. A likely explanation for the strong negative link for acquirers is that firms might anticipate federal regulations that block acquiring firms in concentrated industries. It is also possible that HHIs load on the redeployment effect more than the competitive effect. We conclude that concentration measures and product similarity measures should indeed be jointly considered in studies of product market competition as they contain much distinct information. The table also shows that recent restructuring activity predicts future restructuring for both SIC-based and product similarity based measures, confirming that SIC codes and product similarities contain distinct information.

Although we do not report the results to conserve space, we separately reproduce the tests in Table VI for small and large firms. Our goal is to examine robustness, and to determine which group of firms is most responsible for our findings. In each year, we divide firms into a big firm sample and small firm sample based on whether their assets are above or below the median. We then reproduce the logistic regressions in Table VI for both groups. Both the redeployment effect and the competitive effect are especially robust for large firms, and results for acquirers are also robust for small firms. The results suggest that the likelihood that small firms will be targets is less related to similarity variables. The link between patents and restructuring is also stronger for larger firms. The importance and strongest robustness of our findings for large firms leads us to conclude that our results are important, as the restructuring decisions of these firms have a greater impact on the overall utilization of assets.

In Table VII, we reproduce the tests in Table VI for mergers and acquisition of assets transactions. The competitive effect is stronger for acquisition of asset transactions in Panel B than for mergers in Panel A. One reason is that acquisition of assets are more common, as noted earlier. Furthermore, it is possible that asset deployment and competition for assets might be easier to measure when the matter of control and managerial job loss are less an issue. Although the competitive effect retains its negative sign for acquirers in mergers, it becomes positive for targets of mergers. This feature might reflect consolidating industries, where complete mergers to reduce high competition might be the best strategy to restore profitability. It is also possible that the matter of target control drives this finding. The redeployment effect remains highly positive and robust for both mergers and acquisition of assets transactions.

A Economic Magnitudes

In this section, we summarize the economic magnitude of our findings regarding transaction likelihood. We examine the effect of changing one of our three key variables (product similarity 10 nearest, product similarity all firms excluding the ten nearest firms, and the % neighbor patent words) on the probability of a given transaction. In later sections, we also examine economic magnitudes related to announcement returns and real outcomes. Because some of our models are logistic models, and others OLS based, we adopt a general framework based on predicted values. We first compute a model's predicted value when all of the independent variables are set to their mean value. We then set one of our three variables to its expected 10%ile value, and recompute the predicted value holding all other variables fixed. We repeat this procedure for the 90%ile. We are thus able to report how a given dependent variable changes when a key independent variable moves from its 10%ile value, to its mean value, and to its 90%ile value. Key benefits of this approach include its generality and its ability to show not only impact, but also the magnitude of the dependent variable itself.

Table VIII confirms that our findings regarding economic magnitudes for transaction incidence are economically relevant. The effect of changing similarity (10 nearest) from the 10%ile to the 90%ile changes acquirer incidence from 19.4% to 10.8%. The economic magnitude of our "competitive effect" is thus quite substantial, and it is especially large for big firms in row 4. This effect is somewhat smaller, but still large, for target incidence at 16.6% to 13.5%. The overall similarity variable, the "redeployment effect", is also large and moves from 12.5% to 17.7% for acquirer incidence. The redeployment effect is considerably smaller for target incidence. The patent variable also has a rather substantial economic impact as it changes acquirer incidence from 12.7% to 17.5%. Interestingly, this variable has just as much impact on both target and acquirer incidence. Its importance is consistent with acquisitions being a necessity when firms need to acquire patent protected technology as organic investment is less feasible. Acquisitions should be more common when they are the only option.

VI Ex-post Outcomes

We now examine ex post outcomes, including combined announcement returns and longer-term changes in cash flows, sales and product descriptions. Our tests relate ex post outcomes to the similarity between the acquirer and target and also the similarity between the target and the acquirer's existing rival firms.

A Announcement Returns

This section tests Hypotheses H3, H4, and H5 by examining the returns of the combined acquirer and target firms preceding and surrounding transaction announcements. Importantly, these hypotheses predict that total value creation will be larger when mergers are more likely to permit new product synergies in markets facing little competition. Although these hypotheses thus have strong predictions regarding the combined firm's returns, they are silent on how the gains would be split between the target and the acquirer. Hence, we focus our analysis on the combined firm.

Table IX reports OLS regressions with the acquirer and target's combined abnormal announcement return as the dependent variable. We consider one to eleven day event windows ending on the announcement date (from day t=-10 to day t=0), and we adjust standard errors to reflect possible clustering at the industry and year level. The combined firm's raw return is the total market capitalization of both firms (in dollars) at the end of the event window minus the original market capitalization (in dollars), all divided by the original market capitalization. Hence, this is a simple value weighted return for the combined firm. The abnormal return is equal to this raw return less the return of the CRSP value weighted market index over the same event window. Because the many of our transactions are partial acquisitions, these returns are noisy measures of the transactions's true return, and their economic magnitude can vastly understate the true impact of the transaction when the transaction is smaller. Given we wish to also control for partial anticipation of the deals, we examine event windows starting at ten days before announcement.[13] Note that our

[13]To measure deal anticipation further, we also considered a variable summarizing the fraction of a firm's ten nearest rivals that were involved in restructuring in the past year. Consistent with anticipation, this variable negatively predicts announcement returns, but only at the 10% level in

sample of transactions for this test is somewhat smaller than our main transaction sample because we must further require that the target firm is both publicly traded and that it has available CRSP stock returns on the event date.

[Insert Table IX Here]

We find strong support for the conclusion that more value is created around the announcement date when the acquirer is in a relatively competitive product market, and it is buying target assets in a product market that is less competitive. Rows 1, 3, 5, and 7 support this conclusion at the five percent level or better in all but one event window (the acquirer result is not significant for the t=-1 to t=0 event window in row 3). This rather strong evidence supports the notion that the market rewards acquirers that buy assets permitting new product synergies in less competitive markets.

In rows 2, 4, 6, and 8, we replace the basic product market competition variable with two other variables to more directly test H3 and H4. Namely, we include the gain in product differentiation relative to the acquirer's rivals and the merger pair similarity. The results broadly support H3 for all horizons, as the gain in product differentiation is positive and significant at the 5% level or the 1% level in all rows. We also see some support for H4 over longer event windows, as the pairwise similarity variable is positive and significant at the 5% level in row 8 for the t=-10 to t=0 event window. These results suggest that significant gains accrue consistent with H3 and H4 despite the low degree of power in these tests, and also that these gains appear both on the event day, and in the form of leakage prior to the event day. Although the % neighbor patents variable is positive in every row, it is not statistically significant. Hence, we find only very weak support for H5 in this test.

Lastly, we also find that announcement returns are larger when the transaction is a merger involving a target firm that is large relative to the acquirer, and smaller when the firms are larger unconditionally. The larger returns when bigger targets are involved in mergers likely reflects the leveraged gains that should be observed in the combined firm's returns when transactions involve a larger fraction of the combined

some specifications, or less in others. We omit this variable as it does not alter any inferences and to conserve space.

firm (mergers, unlike asset acquisitions, involve all of the target's assets).

B Real Performance

Although we observe evidence of financial value creation consistent with our hypotheses in Section A, it is important to examine the additional predictions of H3, H4, and H5. In particular, this value increase should be accompanied by real post-transaction gains in both sales and profitability. Also, we should observe evidence consistent with new product synergies. This section presents the real long-term outcomes of acquiring firms as a function of their ex-ante competitive environment.[14]

An important challenge faced by researchers studying ex-post restructuring performance is that two separate firms exist ex ante, and one or two firms might exist ex post depending upon the transaction type. The matter of measuring ex-ante profitability or sales is especially confounding for partial asset purchases. We avoid this issue entirely by only considering the acquirer's post-effective change in performance measured relative to the first set of numbers available after the transaction's effective date. Our hypotheses thus assume that profitability and sales growth accrue over time, as should be the case when new product development drives gains given issues related to the time to build. We examine changes from year t+1 to year t+2 or t+4 (one and three year horizons). For this reason, the sample of transactions used in this section is somewhat smaller than our sample in section A because we must further require that the acquirer has valid COMPUSTAT data at least two years after a given transaction closes. By examining post-effective changes only, we bias our analysis toward not finding results due to lost power, but we avoid biases associated with attempts to measure year t=-1 performance, and complications due to changes in accounting methods following the transaction. As documented by Maksimovic, Phillips, and Prabhala (2008), many mergers also involve selling off divisions at the time of transaction and hence t-1 assets may no longer be owned by either firm by year 3. Our results thus likely understate the true relationship between our key variables.

[14]Although not displayed, we also find that the same variables used to predict changes in ex post profitability and sales growth also positively predict long-term abnormal stock returns, although these findings are only significant at the 10% level. These tests confirm that our long-term findings are not unique to accounting data.

We consider three measures of ex post performance: the change in operating income divided by sales, change in operating income divided by assets, and sales growth. All measures are SIC-3 industry adjusted. The first is COMPUSTAT item 13 divided by item 12. The second is COMPUSTAT item 13 divided by item 6. To mitigate the effect of outliers, we truncate both profitability variables to lie in the interval (-1,1). Changes are computed from year t+1 to year t+4. To reduce survivorship issues, we assign any missing values for a given horizon the value of the last known horizon (for example, if three year sales growth is missing, we populate the given observation with two year sales growth, or one year sales growth).[15]

Table X reports the results of OLS regressions where the ex post change in performance (horizons noted in column two) is the dependent variable. In each panel, we first present two rows including acquirer product similarity (one- and three-year windows, respectively). In the third and fourth regressions of each panel (regressions (3) and (4) in Panel A, regressions (7) and (8) in Panel B, and regressions (11) and (12) in Panel C) we replace this variable with two other variables to more directly test H3 and H4. Namely, we include the gain in product differentiation relative to the acquirer's rivals and the merger pair similarity.

We find evidence that acquirers residing in highly competitive product markets experience positive improvements in profitability and sales growth. The gains in profitability in Panel A accrue over three years, and gains in sales in Panel C accrue more quickly. These results are significant at the 1% level or better for sales growth, and the three year profitability result is significant at the 5% level. These results are also economically significant, as we report later in this section. Acquirers in competitive product markets thus appear able to influence the degree of competition they face by restructuring and developing new products, which is consistent with the growth in sales. The potential for these new products to generate product differentiation can explain why increased profitability accompanies the higher sales. The weaker results in Panel B for profitability normalized by sales rather than assets is likely due to the high sales growth documented in Panel C that coincides with the profitability growth noted in Panel A, as sales is in the denominator of the profitability variable in Panel

[15] Our results are robust to simply discarding these observations rather than using the last value.

B. The strong results for sales growth are especially consistent with the development of new products.

Rows (3) and (4) in Panel A, and rows (11) and (12) of Panel C, show strong support for Hypothesis 3. In particular, the gain in product differentiation coefficient is positive and significant. Hypothesis 4 is also supported by the positive and significant target+acquirer similarity coefficient. As before, we do not see significant coefficients in rows (7) and (8) in Panel B for these variables because Panels A and C suggest that predictions regarding operating income scaled by sales are ambiguous.

We view the especially strong support for sales growth to be consistent with new product development being a key feature of merger strategies. This result helps to separate our hypotheses from the pure pricing power hypothesis advocated by Baker and Bresnahan (1985). Our more direct evidence supporting new product development presented later in this section further illustrates this distinction. The increased profitability noted in Panel A supports both hypotheses. In the case of new product development, managers should favor developing new products in markets where rivals do not yet exist, which will generate increased pricing power.

We do not find support for Hypothesis 5 on this table (the % target neighbor's patent words is generally insignificant, and even negative in some specifications). However, some of the weakness in testing H5 might be related to industry adjusting the dependent variables that we use in this analysis. In particular, if patents matter at the industry level, the theoretical predictions of H5 would still obtain, however they would not be visible in our regression design.

C Product Descriptions

In this section, we consider the prediction of Hypotheses 4 and 5 that new product development will accompany positive real outcomes. New product development is most likely when the target firm is more similar to the acquirer itself due to complementary assets (H4), and when barriers to entry protect the target's uniqueness (H5).

To test these predictions, we consider the time series of the product descriptions, and examine how their size varies over time. We proxy for new product development by examining whether firms experience growth in the size of their product descriptions in the years following the merger's effective date. This first description fully integrates the initial state of the post merger firm. We define "product description growth" as the logarithmic growth in the number of words used in the product market description from year t+1 to either year t+2 or t+4. We then explore whether the same set of variables used to predict ex-post real performance also predict product description growth. In our analysis, we use an OLS specification in which all standard errors are adjusted for clustering at the year and SIC-3 industry level. Note the sample of transactions used in this section is smaller than our sample in section B because we were able to obtain COMPUSTAT data going out to a longer horizon than our available 10-K data, which is needed to construct the dependent variable in this section.

Table XI presents the results of these tests. We find support for hypotheses H4 to H5. Rows (1) to (3) show that product descriptions increase dramatically when acquiring firms reside in ex-ante competitive industries (a key prediction of both hypotheses). This result is significant at the 1% level. Rows (4) to (6) show that this result can be explained by two key factors. Product development is most aggressive when (1) the acquirer and target are more similar, and (2) barriers to entry protect the target's technologies. Evidence supporting H4 is particularly strong, as pairwise similarity is significant at better than the 1% level. For the three year horizon, the % neighbor patent words variable is positive and significant at the 1% level.

The Table also shows that vertical mergers, as described in Fan and Goyal (2006) also experience ex-post growth in the size of product descriptions, consistent with these mergers also permitting the introduction of new products. These results support the broader conclusions of Fan and Goyal (2006), who show that vertical mergers are indeed value creating.

The table also shows that pairwise similarity measured using SIC codes is too granular to produce similar inferences. Hence, understanding the relationship between pairwise similarity and product differentiation relies on the researcher's ability to measure the degree of product similarity, and also to measure similarities relative to different groups of firms. The table also documents that product descriptions also have a tendency to mean revert over time, a feature that is likely due to writing style and a long term preference for brevity. We control for this feature of the data in addition to the other key variables discussed above.

An alternative explanation for our findings is that repeat acquisitions explain why product descriptions expand. We test this hypothesis and compute dummy variables indicating whether a given firm completes a future acquisition during the same horizon over which these calculations are performed. We do not find a statistically significant link between the incidence of repeat acquisitions and our key product market competition and pairwise similarity variables. We also reproduce these tests using the subsample of firms that do not experience any future acquisitions and our results remain robust. We conclude that repeat acquisitions do not explain our findings.

D Economic Magnitudes for Ex Post Outcomes

In this section, we summarize the economic magnitude of two key variables (product similarity 10 nearest, and the % neighbor patent words) on announcement returns and real outcomes. We use the same generalized method used to examine transaction incidence in section V.

Table XII displays the results of these calculations. The economic magnitudes in Panel A regarding announcement returns are modest relative to our incidence results, at least in nominal terms. The competitive effect we document for the combined firm increases event day announcement returns by 0.3%, and longer horizon (11-day) event returns by 0.8%. This spread is modest in nominal terms, but is not trivial given that the mean combined event day firm announcement return is just 0.5% (standard deviation of 4.2%), and the eleven day announcement return averages 0.6% (standard deviation of 7.9%). Hence, over the longer window, the 0.8% is 10% the total standard deviation, and this variable moves the announcement return by more than the mean itself. The relatively small nominal size of these returns is consistent with the bulk of our database being partial acquisitions, and with target firms being smaller than acquiring firms in general. These features make it very hard to measure the impact of a transaction when all we observe is an aggregated return that includes changes in the value of the non-acquired assets. The patent variable has a smaller magnitude consistent with its low statistical significance in Table IX. Because transactions can be anticipated months or even years in advance, we believe that additional gains likely accrue to the combined firm outside of the traditional event windows that we examine.

The results in Panel B show that the economic magnitude of our findings regarding real outcomes are large, especially regarding sales growth. The acquirer's product similarity (10 nearest) increases the change in profitability from -0.9% to -0.2% (1 year) and from -2.3% to -0.9% (3 years). Noting that this variable is a change (hence its mean is near zero) that has a standard deviation of roughly 10% (See Table III), we conclude that this predicted spread is economically relevant, especially for the three year horizon where the predicted spread is nearly 20% of the standard deviation of the dependent variable. Put differently, profitability does not change very much year to year, and a 1.4% shift in profitability can be quite substantial. The table also shows that acquirer product similarity generates a sales growth spread from 11.9% to 19.9% (one year) and 20.0% to 33.9% (three years). These spreads are substantial and economically significant both in nominal and relative terms.

Panel C reports the economic impact of our variables on the ex post growth in the size of the product description. The spread for the acquirer product similarity variable is from -2.4% to 9.1% (results similar for one and three year horizons). This spread is economically large. We conclude that our findings regarding incidence, real performance, and product description growth are economically large. Our most economically significant findings relate to transaction incidence, sales growth, and product description growth. Our findings regarding announcement returns are modest in nominal terms, but are larger when compared to the dependent variable's narrow distribution.

E Robustness

In this section, we examine key predictions of some alternative theories and summarize additional robustness tests.

One main alternative for higher cash flows ex post is that firms cut expenses post-merger. This could arise if the combined firm experiences economies of scale or if acquirers deliberately purchase less efficient targets in order to improve efficiency. We use methods similar to those of Table X and measure changes in the cost of goods sold (scaled by sales), selling and administrative expenses (scaled by sales), and capital expenditures (scaled by assets) starting from year t+1 to year t+2 or t+4 (one year or three year performance). We find that our key variables are weakly negatively related to changes in these expense ratios, but none of the coefficients are statistically significant. We conclude that cost savings and economies of scale likely cannot explain our previous results. Despite this finding, it is important to note that expense-based merger strategies are not mutually exclusive to our product market based strategies, and the best mergers might generate gains along both dimensions.

Second, we also examine whether our results are driven by measures of vertically related mergers, as in Fan and Goyal (2006). Although we find independent support for the conclusion that vertical mergers have positive attributes, we also find that our similarity measures correlate very little with vertical similarity measured using the input-output tables (correlation less than 10%). We also control for vertical similarity in our analysis and find that our results are unchanged regardless of whether we include or exclude controls for vertical similarity. Our measures thus capture additional information that is economically and statistically significant in explaining ex post changes in merger profitability and new product introduction.

Third, we test whether our results are driven by the technology boom of the late 1990s. We perform several tests to exclude this possibility. Throughout our study, we control for time and industry effects. We also run an unreported robustness test where we exclude all technology firms from our sample (technology firms defined as in Loughran and Ritter (2004)). Our results are almost unchanged, and in some cases, become stronger in this test. Our results are thus not driven by technology firms.

Fourth, our results might be driven by multiple segment firms, and in particular, our broad measure of product similarity might be measuring firm diversification rather than asset redeployability. We test this hypothesis by rerunning our tests after excluding all multiple segment firms, where multiple segment firms are identified using the COMPUSTAT segment tapes. Our results change little in this test, and the relevant broad product similarity variable's coefficient declines by only 5% to 10% in various tests. Although diversification might play a small role, our results are not driven by firm diversification.

Fifth, our product similarity variables might be driven by corporate culture, and our results thus driven by similar cultures being more conductive to innovation and the introduction of new products. We do not believe our results support this hypothesis for two reasons. First, the word lists that drive our similarity measure (see Table II) fit a product market interpretation, and not a corporate culture interpretation. Second, although corporate culture predicts that outcomes might be related to target and acquirer pairwise similarity, it is silent on whether the target's distance from the acquirer's nearest rivals (gain in differentiation) will matter to ex post outcomes thus corporate culture likely cannot explain our results.

Sixth, we examine whether our results are driven by repeat-acquiring firms by rerunning our tests after excluding firms that were involved in an acquisition in the past year. Our results change little in these tests. We conclude that our findings are most consistent with product market competition and similarity influencing merger incidence and ex post changes in profitability and product offerings.

VII Conclusions

Using novel text-based measures of product similarity between firms, we analyze how similarity and competition impacts the incentives to merge and whether mergers with potential product market synergies through asset complementarities add value. Our 10-K based text measures of product market competition, differentiation, and new product development have advantages over simple SIC code based measures as they can capture similarity both within and across product markets, and changes in product descriptions over time.

We find that firms with higher economy-wide product market similarity are more likely to enter into restructuring transactions, consistent with their assets being more broadly deployable. Firms with higher localized product market competition are less likely to enter into restructuring transactions. These results are consistent with local competition for restructuring opportunities. We find that firms protected by barriers to entry are also more likely to enter into restructuring transactions. These results are consistent with acquirers of firms in protected product markets obtaining patent protected technologies through mergers or asset purchases instead of developing the technologies or products within the existing firm.

Examining post-merger outcomes, we find that value creation upon announcement and long-term profitability and sales growth are higher when acquirers purchase targets that 1) Increase the acquirer's product differentiation relative to its nearest rivals and 2) Have high similarity relative to the acquirer's own products. These gains are larger when barriers to entry exist, making it harder for rival firms to replicate the given strategy. We also find evidence that product descriptions for the combined acquirer and target firm increase when acquirers are in competitive product markets and transacting firms have high product similarity.

Overall, our results are consistent with merging firms exploiting asset complementarities to create value through sales growth and new product introductions to increase their product differentiation. More broadly, our results suggest that firms can actively reduce their product market competition by engaging in strategic restructuring transactions to increase product differentiation through new product introductions.

Appendix 1

This Appendix explains how we compute the "product similarity" and "product differentiation" between two firms i and j. We first take the text in each firm's product description (from its 10-K) and construct a binary vector summarizing its usage of English words. The vector has a length equal to the number of unique words used in the set of all product descriptions. For a given firm, a given element of this vector is set to one if the word associated with the given element is in the given firm's product description. It is set to zero otherwise. Because we wish to measure product descriptions, we seek to restrict the number of words in this vector to those that are less commonly used across all product descriptions. Very common words are likely to be articles, conjunctions, personal pronouns, abbreviations, or other words that do not identify the firm's product offerings. Hence, we restrict attention to words that appear in fewer than five percent of all product descriptions in the given year. For each firm i, we thus have a binary vector P_i, with each element taking a value of one if the associated word is used in the given firm's product description and zero otherwise.

We next define the normalized frequency vector V_i, which normalizes the vector $P_{x,i}$ to have unit length.

$$V_i = \frac{P_i}{\sqrt{P_i \cdot P_i}} \tag{1}$$

To measure the degree of similarity between the products of firms i and j, we simply take the dot product of the two normalized vectors, a quantity we define as "product similarity". We utilize this measure throughout this study.

$$Product\ Similarity_{i,j} = (V_i \cdot V_j) \tag{2}$$

To measure product differentiation across firms i and j, we simply take one minus the product similarity.

$$Product\ Differentiation_{i,j} = 1 - (V_i \cdot V_j) \tag{3}$$

Because all normalized vectors V_i have a length of one, product similarity and product differentiation both have the nice property of being bounded in the interval (0,1). This normalization is important because it ensures that product descriptions with fewer words are not penalized excessively relative to those using more words. Intuitively, the differentiation between two products is zero if they are the same, and can never exceed one if they are entirely different.

Appendix 2 (Table II Cont)

Word lists of merging firms with high similarity.

This appendix presents the common words for the mergers from Table I that are not already presented in Table II. Merging firms are both (1) in different two-digit SIC codes and (2) have a merger pair similarity in the highest percentile in 2005 (the most recent year in our sample).

Acquirer (Industry) + Target (Industry): list of common words
Enterprise Products Partners (SIC3=131, Crude Petroleum & Natural Gas) + El Paso Corp-Natural Gas (SIC3=492, Natural Gas Transmission): abandonment, acreage, anadarko, basin, basins, border, bound, ceiling, coal, coastal, compression, compressor, condensate, connects, continent, deepwater, depleted, differs, discontinuation, downstream, exact, ferc, gulf, gulfterra, horsepower, hurricanes, inch, indian, inlet, interconnections, interconnects, interruptible, intrastate, juan, justify, kerr, liquids, mainline, mcgee, midstream, minerals, mmcf, morgan, mountains, northeastern, onshore, paso, permian, pipelines, reservoir, rockies, rocky,
Express Scripts Inc (SIC3=641, Insurance Agents, Brokers & Service) + Priority Healthcare Corp (SIC3=512, Wholesale-Drugs & Druggists Sundries): accreditation, admit, advancepcs, alert, alerts, asthma, beneficiaries, bids, broadened, calculations, caremark, cited, compiled, counseling, deadline, differently, dispense, dispensing, enact, exclusion, false, filling, formularies, formulary, fraudulent, freedom, harbors, hipaa, hmos, implicate, induce, inspector, interactions, journal, kickback, knowingly, marked, medco, medication, medications, medicines, nurses, nursing, outpatient, paths, pbms, pharmacies, pharmacist, pharmacists, pharmacy, phoenix,
First Advantage Corp (SIC3=738, Services-Miscellaneous Business Services) + Credit Information Group (SIC3=636, Title Insurance): corefacts, credco, dedicate, diversifying, eviction, fadv, forensics, here, insuring, investigative, justify, landlords, malfeasance, misuse, mitigation, multifamily, nonetheless, omega, paperwork,
General Dynamics Corp (SIC3=373, Ship & Boat Building & Repairing) + Anteon International Corp (SIC3=737, Services-Computer Programming): agrees, allied, appropriated, armed, army, attack, ballistic, battle, cargo, civilian, combat, combatant, command, commanders, corps, deployments, destroyer, fighter, littoral, missile,
H&R Block Inc (SIC3=720, Services-Personal Services) + American Express Tax & Bus (SIC3=619, Finance Services): accessibility, advantaged, affirmed, annuities, annuity, applicability, attest, attracted, attrition, captured, censure, charging, clearing, contacting, custodian, delinquency, delinquent, franchisee, franchisees, franchising, hsbc, inappropriate, iras, join, nasd, partnering, planners, preempt, professionally, ranked, rewards,
Hampshire Group Ltd (SIC3=225, Knitting Mills) + Kellwood Co-David Brooks Bus (SIC3=233, Women's, Misses, and Juniors Outerwear): apparel, casual, dockers, juniors, knit, pants, quota, skirts, sportswear, styles,
Hewlett-Packard Co (SIC3=357, Computer & of fice Equipment) + Peregrine Systems Inc (SIC3=737, Services-Computer Programming): alignment, allocating, americas, answer, architectures, comparing, continuity, corrections, deliverables, descriptions, desk, diego, emea, geographies, heterogeneous, hewlett, infrastructures, laptop, lifecycle, metrics, middleware, networked, packard, pdas, prevents, printer, redundancies, resides,
Highland Hospitality Corp (SIC3=679, Miscellaneous Investing) + Hilton Boston Back Bay Hotel (SIC3=701, Hotels & Motels): airport, convention, courtyard, embassy, franchised, franchisees, franchisor, garden, guest, hilton, homewood, hospitality, hotel, knew, lodging, omaha, parking, renovation, repositioning, reservation,
IRIS International Inc (SIC3=382, Laboratory Apparatus & Furniture) + Quidel Corp-Urinalysis Bus (SIC3=283, Medicinal Chemicals & Botanical Products): bacteria, bayer, bench, characterize, classify, clearances, cleared, diagnostics, exploit, glucose, infection, infections, nitrite, proposition, protein, reagents, seizure, specimen,
Journal Communications Inc (SIC3=271, Newspapers: Publishing Printing) + Emmis Comm-Radio Stations(3) (SIC3=483, Radio Broadcasting Stations): absent, adult, advertiser, advertisers, affiliation, affirmed, allotted, appropriations, arbitron, asking, assignments, assigns, attribution, audience, audiences, broadcaster, broadcasters, broadcasting, broadcasts, brokered, carriage, circulation, circulations, classic, clusters, compelling, conclusion, contemporary, denial, dividing, egregious, electing, equitable, failing, frequencies,
Kinder Morgan Energy Partners (SIC3=492, Natural Gas Transmission) + Terra Nitrogen Co-Blytheville (SIC3=287, Agricultural Chemicals): accidents, acre, ammonia, analogous, anhydrous, barge, barges, basin, beaumont, blend, carbon, cars, cercla, coincide, creek, crop, crops, depleted, dioxide, drainage, environmentally, exported, exports, farm, feed, feedstock, fertilizer, fertilizers, gallons, gasoline, grain, groundwater, gulf,
LaSalle Hotel Properties (SIC3=679, Miscellaneous Investing) + Hilton San Diego Resort,CA (SIC3=701, Hotels & Motels): accommodations, airline, airport, bankrupt, contrary, convention, diego, earthquake, franchisor, guests, hilton, hospitality, hotel, indianapolis, instructions, insuring, knew, leisure, lodging, luxury, omaha, pines,
Leggett & Platt Inc (SIC3=251, Household Furniture) + Foamex International-Rubber (SIC3=308, Miscellaneous Plastics Products): bedding, carpet, cushioning, cushions, fabricated, fibers, foam, leggett, mattress, mattresses, mexican, molded, pillows, platt, quilting, scrap, seat, seating, shape, springs, upholstered
Lionbridge Technologies Inc (SIC3=737, Services-Computer Programming) + Bowne Global Solutions (SIC3=275, Commercial Printing): adapting, bowne, cultural, culturally, freelance, globalization, instructions, languages, leverages, lifecycle, linguistic, localization, multimedia, shorten, standardizes, translated, translation,

Appendix 2 Continued

This appendix presents the common words for the mergers from Table I that are not already presented in Table II. Merging firms are both (1) in different two-digit SIC codes and (2) have a merger pair similarity in the highest percentile in 2005 (the most recent year in our sample).

Acquirer (Industry) + Target (Industry): list of common words
MarkWest Energy Partners LP (SIC3=131, Crude Petroleum & Natural Gas) + Javelina Gas Processing (SIC3=492, Natural Gas Transmission): accidental, acreage, anadarko, appalachian, basin, basins, cogeneration, compression, compressor, condensate, condensed, counterparty, diminished, disciplined, exact, ferc, grouped, haul, horsepower, inch, injected, inlet, insignificant, intrastate, lateral, liquids, midstream, mmcf,
McDATA Corp (SIC3=366, Telephone & Telegraph Apparatus) + Computer Network Technology (SIC3=357, Computer & office Equipment): backbone, backup, broadened, brocade, campus, cisco, dell, depended, diagnose, disk, disruptive, encompassing, escon, extensibility, fabric, fibre, ficon, forecasting, fractional, heterogeneous, hewlett, hitachi, infrastructures, interoperability, iscsi, mainframe, mcdata, merits, migrate,
Nektar Therapeutics (SIC3=283, Medicinal Chemicals & Botanical Products) + AeroGen Inc (SIC3=384, Surgical & Medical Instruments & Apparatus): absorbed, absorption, activated, adult, aerogen, aerosol, aerosols, aeruginosa, alkermes, antibiotic, antibiotics, aradigm, bachelor, battelle, biologic, biology, bloodstream, breath, buccal, carbon, cgmp, clinically, collaborations, commercialized, conceived, cystic, diabetes, diabetic, dioxide, dosage, dose, doses, dosing, fibrosis, filling, formulated, formulations, founder, glucose, harvard, inconvenience,
NRG Energy Inc (SIC3=491, Electric Services) + West Coast Power LLC (SIC3=131, Crude Petroleum & Natural Gas): absent, acid, ahead, allegedly, approves, asbestos, attainment, balancing, baseload, basin, bids, bilateral, bilaterally, blanket, boiler, btus, cabrillo, cair, capped, caps, carbon, cdwr, cercla, coal, combustion, commences, compel, consumed, cooling, cooperatives, cpuc, crisis, crude, curve, defines, depressed,
Pacific Energy Partners LP (SIC3=461, Pipe Lines (No Natural Gas)) + Valero LP- Terminal & Pipeline (SIC3=291, Petroleum Refining): barges, barrels, benchmark, blend, blended, blending, cercla, complements, connects, conocophillips, crude, deepwater, denver, diesel, distillate, dock, exxonmobil, feedstock, feedstocks, futures, gasoline, grades, grandfathered, groundwater, gulf, heavier, hydrocarbon, inch, oils, paso, pipe,
Polo Ralph Lauren Corp (SIC3=232, Men's & Boys Work Clothg & Allied Garments) + Ralph Lauren Footwear Co Inc (SIC3=302, Rubber & Plastics Footwear): accessory, amortize, apparel, appearances, athletic, atmosphere, attitude, boys, caribbean, casual, catalogs, chile, classic, clubs, coincide, colombia, compensatory, contemporary, designer, distinctive, dress, eyewear, famous, flagship, footwear, girls, gloves, golf, gucci, hilfiger, hosiery, importing, jeans, knit, launches, lauren, leather, lifestyle, madrid, malaysia, message, mills, newest,
Renaissance Learning Inc (SIC3=737, Services-Computer Programming) + AlphaSmart Inc (SIC3=357, Computer & office Equipment): administrators, bundled, classroom, curricula, districts, educators, english, grammar, handles, instructional, learning, literacy, math, multimedia, principals, quizzes, renaissance, sessions,
RH Donnelley Corp (SIC3=731, Services-Advertising) + Dex Media Inc (SIC3=274, Miscellaneous Publishing): accountable, advertise, advertisement, advertiser, advertisers, affiliation, amdocs, bell, billboards, bottom, bound, bundled, circulation, citysearch, cmrs, column, completeness, decentralized, delinquent, directional, directories, directory, donnelley, enduring, english, fingers, george, golf, google, guides, households, incumbent, listings, logos, longest, newspaper, newspapers, official, permission, phrase, portals, postal, premise, proposition,
Stonemor Partners LP (SIC3=650, Real Estate) + Service Corp Intl-Cemeteries (SIC3=720, Services-Personal Services): bronze, burial, casket, caskets, cemeteries, cemetery, closings, cremation, crypts, funeral, funerals, gardens, heritage, interment, lawn, lots, mausoleum, memorial, memorials, openings, receptacles, spaces,
Sunstone Hotel Investors Inc (SIC3=679, Miscellaneous Investing) + Renaissance Hotels-Hotel (SIC3=701, Hotels & Motels): accommodations, affiliation, airline, airports, amenities, booked, bookings, contagious, courtyard, cured, earthquakes, floods, franchisees, franchisor, franchisors, guest, guests, hospitality, hotel, indemnities, intermediaries, justify, laundry, leisure, lodging, marriott, parking, portugal, qualifies, renovation,
Sunstone Hotel Investors Inc (SIC3=679, Miscellaneous Investing) + Renaissance Washington DC (SIC3=701, Hotels & Motels): accommodations, affiliation, airline, airports, amenities, booked, bookings, contagious, courtyard, cured, earthquakes, floods, franchisees, franchisor, franchisors, guest, guests, hospitality, hotel, indemnities, intermediaries, justify, laundry, leisure, lodging, marriott, parking, portugal, qualifies, renovation,
Titan Tire Corp (SIC3=331, Steel Works, Blast Furnaces & Finishing Mills) + Goodyear Tire & Rubber-North (SIC3=301, Tires & Inner Tubes): australian, belts, bridgestone, carbon, earthmoving, exports, farm, goodyear, haul, highway, luxembourg, michelin, mills, pounds, rubber, textile, tire, tires, titan, tread, wheel

References

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

Andrade, Gregor, and Erik Stafford, 2004, Investigating the economic role of mergers, *Journal of Corporate Finance* 1, 1–36.

Baker, Jonathan, and Timothy Bresnahan, 1985, The gains from merger or collusion in product differentiated industries, *Journal of Industrial Economics* 33, 427–444.

Betton, Sandra, Espen Eckbo, and Karin Thorburn, 2007, Corporate takeovers, forthcoming Handbook of Corporate Finance: Empirical Corporate Finance: Elsevier/North-Holland.

Boukus, Ellyn, and Joshua Rosenberg, 2006, The information content of fomc minutes, Yale University working paper.

Chamberlain, EH, 1933, *The Theory of Monopolistic Competition* (Harvard University Press: Cambridge).

Davis, James, Eugene Fama, and Kenneth French, 2000, Characteristics, covariances, and average returns: 1929-1997, *The Journal of Finance* 55, 389–406.

Eaton, Curtis, and Richard Lipsey, 1989, *Product Differentiation, Handbook of Industrial Organization* (Elsevier).

Fan, Joseph, and Vidhan Goyal, 2006, On the patterns and wealth effects of vertical mergers, *Journal of Business* 79, 877–902.

Hanley, Kathleen, and Gerard Hoberg, 2008, Strategic disclosure and the pricing of initial public offerings, Working Paper, Securities Exchange Commission and the University of Maryland.

Harford, Jarrad, 2005, What drives merger waves?, *Journal of Financial Economics* 77, 529–560.

Hart, Oliver, and John Moore, 1995, Debt and seniority: an analysis of the role of hard claims in constraining management, *American Economic Review* 85, 567–85.

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

Hoberg, Gerard, and Gordon Phillips, 2008, Real and financial industry booms and busts, Working Paper, University of Maryland.

Hotelling, H., 1929, Stability in competition, *Economic Journal* pp. 41–57.

Jensen, Michael, 1993, The modern industrial revolution, exit, and the failure of internal control systems, *The Journal of Finance* 48, 831–80.

Jovanovic, Boyan, and P Rousseau, 2002, The q-theory of mergers, *American Economic Review* 92, 198–204.

Kaplan, Steve, and Michael Weisbach, 1992, The success of acquisitions: evidence from divestitures, *The Journal of Finance* 47, 107–138.

Li, Feng, 2006a, Account report readibility, current earnings and earnings persistence, University of Michigan Working Paper.

Loughran, Tim, and Jay Ritter, 2004, Why has ipo underpricing changed over time?, *Financial Management* 33, 5–37.

Macskassy, Sofus, Maytal Saar-Tsechansky, and Paul Tetlock, 2004, More than words: Quantifying language to measure firms' fundamentals, Forthcoming, Journal of Finance.

Maksimovic, Vojislav, and Gordon Phillips, 2001, The market for corporate assets: Who engages in mergers and asset sales and are there efficiency gains?, *The Journal of Finance* 56, 2019–2065.

——— , 2008, The industry life-cycle, acquisitions and investment: Does firm organization matter?, *The Journal of Finance* 63.

——— , and Nagpurnanand Prabhala, 2008, Post-merger restructuring and the boundaries of the firm, University of Maryland Working Paper.

Mitchell, Mark, and Harold Mulherin, 1996, The impact of industry shocks on takeover and restructuring activity, *Journal of Financial Economics* 41, 193–229.

Morck, Randall, Andrei Shleifer, and Robert Vishny, 1988, Management ownership and market valuation: an empirical analysis, *Journal of Financial Economics* 20, 293–315.

Nevo, Aviv, 2000, Mergers with differentiated products: the case of the ready to eat cereal industry, *Rand Journal of Economics* 31, 395–421.

Rhodes-Kropf, Matthew, and David Robinson, 2008, The market for mergers and the boundaries of the firm, Forthcoming, Journal of Finance.

Tetlock, Paul, 2008, Giving content to investor sentiment: the role of media in the stock market, *The Journal of Finance* 62, 1139–1168.

Yang, Liu, 2008, The real determinants of asset sales, Forthcoming, Journal of Finance.

Table I: Merging firms in 2005 with highest percentile similarity but different two-digit SIC codes

This table presents a list of firms that are both (1) in different two-digit SIC codes and (2) have a merger pair similarity in the 100th percentile in 2005 (the most recent year in our sample).

quirer	Target	Acquirer SIC-3
is-Chalmers Energy Inc	RPC Inc-Technical Services	SIC3=735, Equipment Rental + Leasin
nericredit Financial Services	Bay View Acceptance Corp	SIC3=619, Finance Services
as Pipeline Partners LP	Energy Transfer Partners	SIC3=492, Natural Gas Transmission
o Corp	WUPL-TV, New Orleans	SIC3=271, Newspapers: Publishing + I
ckeye GP LLC	Atlas Oil Co-Refined Petroleum	SIC3=461, Pipe Lines
reerStaff Unlimited Inc	ProCare One Nurses LLC	SIC3=805, Services-Nursing + Persona
evronTexaco Corp	Unocal Corp	SIC3=291, Petroleum Refining
rrectional Properties Trust	Geo Group Inc-Lawton	SIC3=679, Miscellaneous Investing
gle Hosp Prop Trust Inc	Hilton Glendale	SIC3=679, Miscellaneous Investing
ICORE Corp	JDS Uniphase Corp	SIC3=355, Special Industry Machinery
terprise Products Partners	El Paso Corp-Natural Gas	SIC3=131, Crude Petroleum + Natural
press Scripts Inc	Priority Healthcare Corp	SIC3=641, Insurance Agents, Brokers
st Advantage Corp	Credit Information Group	SIC3=738, Miscellaneous Business Ser
neral Dynamics Corp	Anteon International Corp	SIC3=373, Ship + Boat Building + Rep
R Block Inc	American Express Tax & Bus	SIC3=720, Services-Personal Services
mpshire Group Ltd	Kellwood Co-David Brooks Bus	SIC3=225, Knitting Mills
wlett-Packard Co	Peregrine Systems Inc	SIC3=357, Computer + of fice Equipm
ghland Hospitality Corp	Hilton Boston Back Bay Hotel	SIC3=679, Miscellaneous Investing
S International Inc	Quidel Corp-Urinalysis Bus	SIC3=382, Laboratory Apparatus + Fu
rnal Communications Inc	Emmis Comm-Radio Stations	SIC3=271, Newspapers: Publishing Pri
nder Morgan Energy Partners	Terra Nitrogen Co-Blytheville	SIC3=492, Natural Gas Transmission
Salle Hotel Properties	Hilton San Diego Resort	SIC3=679, Miscellaneous Investing
ggett & Platt Inc	Foamex International-Rubber	SIC3=251, Household Furniture
nbridge Technologies Inc	Bowne Global Solutions	SIC3=737, Services-Computer Progran
rkWest Energy Partners LP	Javelina Gas Processing	SIC3=131, Crude Petroleum + Natural
DATA Corp	Computer Network Technology	SIC3=366, Telephone + Telegraph App
ktar Therapeutics	AeroGen Inc	SIC3=283, Medicinal Chemicals + Bot
G Energy Inc	West Coast Power LLC	SIC3=491, Electric Services
cific Energy Partners LP	Valero LP- Terminal & Pipeline	SIC3=461, Pipe Lines (No Natural Gas
o Ralph Lauren Corp	Ralph Lauren Footwear Co Inc	SIC3=232, Men's + Boys Work Clothg
naissance Learning Inc	AlphaSmart Inc	SIC3=737, Services-Computer Progran
Donnelley Corp	Dex Media Inc	SIC3=731, Services-Advertising
nemor Partners LP	Service Corp Intl-Cemeteries	SIC3=650, Real Estate
nstone Hotel Investors Inc	Renaissance Hotels-Hotel	SIC3=679, Miscellaneous Investing
nstone Hotel Investors Inc	Renaissance Washington DC	SIC3=679, Miscellaneous Investing
an Tire Corp	Goodyear Tire & Rubber-North	SIC3=331, Steel Works, Blast Furnace

Table II: Common words for merging firms from Table I with high similarity

This table presents the common words for the first 10 mergers from Table I. Firms have both (1) in different two-digit SIC codes and (2) have a merger pair similarity in the highest percentile in 2005 (the most recent year in our sample). The word lists for the remaining mergers from Table I are presented in Appendix II as Table XIII.

Acquirer (Industry) + Target (Industry): list of common words
Allis-Chalmers Energy Inc (SIC3=735, Equipment Rental & Leasing) + RPC Inc-Technical Services Bus (SIC3=138, Drilling Oil & Gas Wells): accessing, bore, casing, coiled, depths, diameter, downhole, drill, drilled, economical, fluids, forklifts, geological, gulf, halliburton, hammer, hevi, hydraulic, hydrocarbons, inject, injected, laydown, motors, mountains, nitrogen, oilfield, onshore, permeability, pipe, pipelines, pumps, reservoirs, retrieve,
Americredit Financial Services (SIC3=619, Finance Services) + Bay View Acceptance Corp (SIC3=602, National Commercial Banks): advise, aforementioned, applicants, captive, dealership, dealerships, defenses, delinquencies, depress, depressed, discriminating, discriminatory, earns, franchised, obligors, recession,
Atlas Pipeline Partners LP (SIC3=492, Natural Gas Transmission) + Energy Transfer Partners-Elk (SIC3=131, Crude Petroleum & Natural Gas): abandonment, analogous, atlas, basin, basins, bbls, cercla, compressing, condensate, cubic, diameter, discrimination, discriminatory, ebitda, ferc, fractionation, gallon, gather, gathered, gatherer, geothermal, grange, grievances, hazard, hydrocarbon, hydrocarbons, intrastate, liquids, midstream,
Belo Corp (SIC3=271, Newspapers: Publishing or Publishing & Printing) + WUPL-TV, New Orleans, LA (SIC3=483, Radio Broadcasting Stations): advertiser, advertisers, affiliation, affirmed, assignments, attribution, audience, audiences, austin, broadcaster, broadcasters, broadcasting, broadcasts, carriage, charlotte, compulsory, contests, digitally, distant, dmas, duopolies, duopoly, empowers, fare, finds, flag, frequencies, george, hispanic, households, indecency, indecent, informational, inquiry, insulated, king, louis, magazines, morning, multichannel, necessity, netratings, newspaper, newspapers, nielsen, norfolk, orleans, paramount, passive, phoenix, piracy, portland, primetime, providence, purport, random, rank, ranked, reauthorization,
Buckeye GP LLC (SIC3=461, Pipe Lines) + Atlas Oil Co-Refined Petroleum (SIC3=590, Retail-Miscellaneous Retail): allentown, barge, cercla, citgo, crude, diesel, fuels, gasoline, harrisburg, haven, legality, newark, philadelphia, propane, refineries, refiners, spite, superfund, trigger, unitholder, unitholders
CareerStaff Unlimited Inc (SIC3=805, Services-Nursing & Personal Care Facilities) + ProCare One Nurses LLC (SIC3=874, Services-Management Services): accrue, acted, acuity, admissions, adult, arranging, arthritis, basket, beds, beneficiaries, careerstaff, contest, disabled, disabling, divested, elderly, encouraged, ethical, exclusion, false, hipaa, hospice, incrementally, induce, inpatient, inspector, intermediaries, kickback, knowing, knowingly, licensure, liquidated, manuals, mental, nurse, nurses, nursing, offenses, outpatient, payor, pediatric, punitive, refinement, rehabilitation, rehabilitative, reimbursable, reimbursements, reimburses, remuneration,
ChevronTexaco Corp (SIC3=291, Petroleum Refining) + Unocal Corp (SIC3=131, Crude Petroleum & Natural Gas): accidental, acreage, alaska, appraisal, argentina, averaged, barrels, basin, basins, bids, border, canyon, caspian, coal, commerciality, concession, concessions, condensate, congo, consortium, consumed, crude, cubic, deepwater, delineation, democratic, discoveries, drill, drilled, exploitation, exploratory, exported, extracted, exxonmobil, farm, fired, formulas, fuels, gasoline, gasolines, geothermal, gulf, hydrocarbon, hydrocarbons, indonesia, java, liquids, malo, megawatts, memorandum, onshore, petrochemical, philippine, philippines, pipelines, ports, producible, progressed, prospect, proved, ramp, refined, refinery, refining, reservoir, reservoirs,
Correctional Properties Trust (SIC3=679, Miscellaneous Investing) + Geo Group Inc-Lawton (SIC3=874, Services-Management Services): adelanto, adult, alcohol, appropriation, appropriations, architects, aurora, awarding, beds, broward, cornell, correctional, corrections, counseling, customs, deems, desert, detainees, detention, falck, george, golden, hobbs, immigration, inmate, inmates, jena, juvenile, karnes, lawton, male,
Eagle Hosp Prop Trust Inc (SIC3=679, Miscellaneous Investing) + Hilton Glendale, Glandale, CA (SIC3=701, Hotels & Motels): accumulates, airport, booked, bookings, concessions, contrary, embassy, franchisees, franchisor, guest, guests, hilton, hospitality, hotel, illiquidity, indoor, instructions, insuring, intermediaries, lodging,
EMCORE Corp (SIC3=355, Special Industry Machinery) + JDS Uniphase Corp-Analog CATV (SIC3=366, Telephone & Telegraph Apparatus): achieves, addressable, agilent, alcatel, amplification, amplifiers, amplitude, attest, band, brightness, broadest, cables, catv, cavity, centralizing, choosing, cisco, closer, complimentary, consumed, contributors, converging, datacom, defect, defective, deferring, dense, deposition, destinations, detectible, deteriorate, diode, diodes, disadvantages, discouraging, disproportionate, disruptive, diverting,

Table III: Summary Statistics

Summary statistics are reported for our sample based on 1997 to 2005. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. We compute product similarities based on the ten nearest firms and based on all firms excluding the 10 nearest. The fraction of nearest neighbors who were involved in restructuring transactions in the past year is computed both based on the firm's ten nearest neighbors and based on three-digit SIC codes. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. Announcement returns are net of the CRSP value weighted index and are measured relative to the three-day window bracketing the announcement date. The gain in product differentiation is the product distance from the target to the acquirer's ten nearest neighbors, less the acquirer's distance to its ten nearest neighbors. We compute three measures of ex-post acquirer real performance. All are based on the first set of accounting numbers available after the transaction is effective (call this the "effective year"), and we consider one to three year changes in performance thereafter (this method avoids bias from trying to measure pre-merger performance of two separate firms). We compute profitability as operating income divided by assets or sales in each year, and we then truncate the distribution at (-1,1) to control for outliers (winsorizing produces similar results). We then compute the change in this variable from the effective year to one to three years thereafter. We compute log sales growth as the natural log of ex-post sales divided by the level of sales in the effective year.

Variable	Mean	Std. Dev.	Minimum	Median	Maximum	Obs.
Panel A: Firm Variables						
Target Dummy	0.151	0.358	0.000	0.000	1.000	47,394
Acquirer Dummy	0.282	0.450	0.000	0.000	1.000	47,394
Target of Merger Dummy	0.043	0.202	0.000	0.000	1.000	47,394
Acquirer in Merger Dummy	0.104	0.305	0.000	0.000	1.000	47,394
Target of Acq. of Assets Dummy	0.108	0.311	0.000	0.000	1.000	47,394
Acquirer of Acq. of Assets Dummy	0.177	0.382	0.000	0.000	1.000	47,394
Product Similarity (Overall-10)	0.017	0.005	0.002	0.017	0.055	47,394
Product Similarity (10 nearest)	0.159	0.069	0.028	0.143	0.639	47,394
Fraction 10 Nearest Restructuring	0.377	0.115	0.020	0.380	0.810	47,394
Fraction SIC-3 Restructuring	0.371	0.138	0.000	0.366	1.000	47,394
Log Assets	5.374	2.132	0.000	5.304	13.962	47,394
SIC-3 Industry Sales-based HHI	0.048	0.026	0.000	0.044	0.229	47,394
% Neighbor Patent Words	0.147	0.182	0.000	0.083	0.900	47,394
Panel B: Transaction Level Variables						
Target Ann. Return (-1,+1)	0.095	0.220	-0.878	0.023	4.941	4,932
Acquirer Ann. Return (-1,+1)	0.004	0.079	-0.602	0.000	1.906	5,274
Gain in Product Differentiation	0.088	0.076	-0.355	0.077	0.944	5,274
Merger Pair Similarity	0.093	0.066	0.007	0.080	0.246	5,274
Panel C: Acquirer Ex-Post Real Performance						
1-Year Δ Profitability (scaled by assets)	-0.005	0.088	-0.940	0.000	0.985	4,451
3-Year Δ Profitability (scaled by assets)	-0.016	0.113	-1.117	-0.004	0.985	4,451
1-Year Δ Profitability (scaled by sales)	-0.005	0.126	-1.136	-0.003	1.321	4,451
3-Year Δ Profitability (scaled by sales)	-0.020	0.167	-1.131	-0.013	1.486	4,451
1-Year Log Sales Growth	0.159	0.334	-4.039	0.125	3.559	4,451
3-Year Log Sales Growth	0.270	0.558	-6.090	0.233	7.179	4,451
1-Year Δ COGS (scaled by sales)	0.005	0.088	-0.791	0.003	0.930	4,451
3-Year Δ COGS (scaled by sales)	0.011	0.122	-0.923	0.009	0.962	4,451
1-Year Δ SG+A (scaled by sales)	-0.000	0.079	-0.974	0.000	0.935	4,451
3-Year Δ SG+A (scaled by sales)	0.008	0.120	-0.957	0.002	1.029	4,451
1-Year Δ CAPX (scaled by assets)	-0.000	0.037	-0.381	0.000	0.291	4,451
3-Year Δ CAPX (scaled by assets)	-0.001	0.045	-0.435	0.000	0.304	4,451

Table IV: Pearson Correlation Coefficients

Pearson Correlation Coefficients are reported for our sample based on 1997 to 2005. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. We compute product similarities based on the ten and one hundred nearest firms and based on all firms excluding the 10 nearest. We also include the fraction of nearest neighbors who were involved in restructuring transactions in the past year, as well as a similar fraction based on three-digit SIC code groupings. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent.

Row	Variable	Product Similarity (All Firms-10)	Product Similarity (100 Nearest)	Product Similarity (10 Nearest)	Last Year 10 Nearest % Restructured	Last Year SIC-3 % Restructured	Industry HHI
	Panel A: Correlation Coefficients						
(1)	Product Similarity (100 Nearest)	0.601					
(2)	Product Similarity (10 Nearest)	0.520	0.914				
(3)	% Restructured Last Year (100 Nearest)	-0.060	-0.275	-0.204			
(4)	% Restructured Last Year (SIC-3)	-0.007	-0.095	-0.067	0.547		
(5)	Industry Sales based Herfindahl Index (SIC-3)	-0.097	-0.115	-0.080	0.070	0.084	
(6)	% Neighbor Patent Words	0.015	-0.011	-0.120	-0.431	-0.277	-0.048

Table V: Effect of product similarity on profitability

OLS regressions with profitability defined as operating income divided by sales (Panel A) or Assets (Panel B) as the dependent variable. All specifications include yearly fixed effects and standard errors account for clustering across year and SIC-3 industries. The sample is from 1997 to 2005, and product similarity is based on the word content of the product description section of the 10-K filing. A higher similarity measure implies the firm has a product description more closely linked to those of other firms. We compute product similarities based on the 10 most similar firms. We report Sales HHI (SIC-3) based on the two step fitted method described in Hoberg and Phillips (2008) (accounts for public and private firms). Log assets is the natural log of COMPUSTAT assets. The log book to market ratio is as defined in Davis, Fama, and French (2000) and we use a dummy to indicate when the raw book to market ratio is negative. We define Big (Small) firms as those with above (below) median ex-ante book assets.

Row	Depdendent Variable	Sample	Product Similarity (10 Nearest)	SIC-3 Sales HHI (fitted)	Log Assets	Log Book/ Market	Negative B/M Dummy	Year+ SIC-3 Fixed Effects	Adj R^2	Obs
			Panel A: Profitability scaled by sales							
(1)	oi/sales	All Firms	-0.379	-0.297	0.057	0.015	-0.114	Yes	0.374	46,312
			(-2.85)	(-0.80)	(19.82)	(2.66)	(-8.52)			
(2)	oi/sales	Big Firms	-0.156	-0.343	0.021	-0.027	-0.008	Yes	0.355	23,160
			(-2.66)	(-1.14)	(9.09)	(-5.37)	(-0.71)			
(3)	oi/sales	Small Firms	-0.990	0.099	0.090	0.025	-0.118	Yes	0.314	23,152
			(-4.67)	(0.16)	(21.70)	(3.49)	(-7.14)			
			Panel B: Profitability scaled by assets							
(4)	oi/assets	All Firms	-0.244	0.004	0.037	-0.001	-0.099	Yes	0.252	46,312
			(-3.28)	(0.02)	(14.21)	(-0.11)	(-8.92)			
(5)	oi/assets	Big Firms	-0.099	0.130	0.003	-0.034	0.041	Yes	0.318	23,160
			(-3.48)	(0.85)	(2.82)	(-11.38)	(5.37)			
(6)	oi/assets	Small Firms	-0.700	0.182	0.081	0.009	-0.115	Yes	0.270	23,152
			(-5.82)	(0.43)	(20.89)	(1.52)	(-8.28)			

Table VI: Mergers and Acquisitions and Product Similarity

The table displays marginal effects of logistic regressions where the dependent variable is a dummy indicating whether the given firm is an acquirer of a merger or an acquisition of assets (Panel A) or a target (Panel B). Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. The independent variables include the fraction of nearest neighbors who were involved in restructuring transactions in the past year, as well as the average product similarity of each firm relative to its ten nearest neighbors, and relative to all firms excluding its ten nearest neighbors. We also include the fraction of past-year restructurings based on three-digit SIC code groupings. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The sample is from 1997 to 2005. t-statistics are adjusted for clustering at the year and industry level. Marginal effects are displayed as percentages.

Row	Dependent Variable	Product Similarity (All Firms-10)	Product Similarity (10 Nearest)	Industry HHI (SIC-3)	Last Year 10 Nearest % Restructured	Last Year SIC-3 % Restructured	Last Year % Neighbor Patent Words	Last Year Log Total Assets	Obs
				Panel A: Acquirer Likelihood					
(1)	Acquirer?	2.073	-3.100		6.258		1.747	10.250	47,394
		(7.64)	(-7.65)		(20.25)		(4.61)	(34.92)	
(2)	Acquirer?	0.671			7.133		2.409	9.749	47,394
		(3.10)			(22.28)		(5.96)	(32.18)	
(3)	Acquirer?		-1.863		6.499		1.971	10.239	47,394
			(-6.00)		(20.86)		(5.14)	(35.23)	
(4)	Acquirer?			-1.297		5.287	1.034	10.368	47,394
				(-2.76)		(16.32)	(2.16)	(35.68)	
(5)	Acquirer?	1.942	-3.308	-1.662	4.849	2.891	1.878	10.262	47,394
		(7.66)	(-8.14)	(-4.10)	(14.59)	(10.25)	(4.88)	(34.96)	
				Panel B: Target Likelihood					
(6)	Target?	0.626	-1.281		3.559		1.738	8.289	47,394
		(3.20)	(-4.29)		(14.82)		(6.50)	(44.36)	
(7)	Target?	0.061			3.891		2.013	8.116	47,394
		(0.35)			(16.87)		(7.39)	(43.98)	
(8)	Target?		-0.894		3.632		1.796	8.284	47,394
			(-3.45)		(15.02)		(6.57)	(44.41)	
(9)	Target?			0.566		2.378	1.268	8.465	47,394
				(2.38)		(12.36)	(4.79)	(44.43)	
(10)	Target?	0.657	-1.225	0.420	3.013	1.045	1.878	8.196	47,394
		(3.35)	(-4.05)	(1.92)	(11.75)	(5.22)	(6.83)	(43.80)	

Table VII: M+A and Product Similarity by Transaction Type

The table examines the likelihood of mergers and acquisition of assets. We display marginal effects from logistic regressions. In Panel A, the dependent variable equal to one if the given firm was a target or acquirer in a merger as identified in column two. In Panel B, we focus on acquisition of assets transactions. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. The independent variables include the fraction of nearest neighbors who were involved in restructuring transactions in the past year, as well as the average product similarity of each firm relative to its ten nearest neighbors, and relative to all firms excluding its ten nearest neighbors. We also include the fraction of past-year restructurings based on three-digit SIC code groupings. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The sample is from 1997 to 2005. t-statistics are adjusted for clustering at the year and industry level. Marginal effects are displayed as percentages.

Row	Depdendent Variable	Product Similarity (All Firms-10)	Product Similarity (10 Nearest)	Industry HHI (SIC-3)	Last Year 10 Nearest % Restructured	Last Year SIC-3 % Restructured	Last Year % Neighbor Patent Words	Last Year Log Total Assets	Obs
				Panel A: Mergers Only					
(1)	Acquirer of a Merger	0.994	-0.586		1.594		0.975	4.793	47,394
		(4.11)	(-1.99)		(6.56)		(3.77)	(19.77)	
(2)	Acquirer of a Merger			-2.088		1.546	0.724	5.005	47,394
				(-5.39)		(7.39)	(3.17)	(21.53)	
(3)	Acquirer of a Merger	0.859	-0.868	-2.144	1.176	0.946	0.915	4.910	47,394
		(3.88)	(-3.02)	(-5.45)	(5.03)	(4.78)	(3.75)	(21.96)	
(4)	Target of a Merger	0.276	0.318		0.603		-0.395	1.151	47,394
		(2.64)	(2.43)		(4.76)		(-2.78)	(10.40)	
(5)	Target of a Merger			-0.246		0.334	-0.613	1.301	47,394
				(-1.94)		(3.07)	(-4.27)	(11.86)	
(6)	Target of a Merger	0.267	0.293	-0.150	0.529	0.146	-0.399	1.154	47,394
		(2.53)	(2.24)	(-1.19)	(3.87)	(1.27)	(-2.75)	(10.36)	
				Panel B: Acquisition of Assets Only					
(7)	Acquirer of an Acq of Assets	1.020	-2.660		4.644		0.573	5.251	47,394
		(4.59)	(-8.28)		(16.52)		(1.70)	(21.36)	
(8)	Acquirer of an Acq of Assets			0.469		3.592	0.195	5.185	47,394
				(1.46)		(14.61)	(0.48)	(22.54)	
(9)	Acquirer of an Acq of Assets	1.026	-2.611	0.176	3.706	1.873	0.810	5.142	47,394
		(4.58)	(-8.09)	(0.63)	(12.40)	(8.05)	(2.35)	(20.47)	
(10)	Target of an Acq of Assets	0.419	-1.891		2.980		2.055	7.009	47,394
		(2.42)	(-7.44)		(13.85)		(8.83)	(45.22)	
(11)	Target of an Acq of Assets			0.731		1.986	1.739	7.060	47,394
				(3.86)		(11.46)	(7.21)	(44.74)	
(12)	Target of an Acq of Assets	0.456	-1.816	0.498	2.519	0.877	2.203	6.918	47,394
		(2.64)	(-7.11)	(2.92)	(11.49)	(5.23)	(9.39)	(44.80)	

Table VIII: Economic Magnitudes of Predicting Transaction Incidence

The table displays economic magnitudes associated with various findings reported earlier in this study. All magnitudes are predicted values, and all magnitudes are conditional and thus account for the effects of industry, year and all control variables (based models in earlier tables as noted in panel headers). For each dependent variable being considered (noted in the description column), we first set all control variables to their mean values and compute the model's predicted value. The result of this calculation is the value displayed in the "mean" column in each category. For each independent variable whose economic magnitude we are measuring (product similarity 10 nearest, product similarity overall, and neighbor patent worlds), which is noted in the column headers, we also compute the model's predicted value assuming the given independent variable is expected to be in the 10th and 90th percentile of its distribution, while still holding all control variables fixed at their mean. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). We consider similarity based on a firm's ten closest rivals, and similarity based on all firms in the universe excluding these ten firms. A higher similarity implies that the firm has a product description more closely related to those of other firms. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The sample is from 1997 to 2005.

		Product Similarity (10 Nearest)			*Product Similarity (All-10)*			*Neighbor Patent Words*		
Row	Description	10 %ile	Mean	90 %ile	10 %ile	Mean	90 %ile	10 %ile	Mean	90 %ile
	Panel A: Target and Acquirer Logit Models (Based on models in Table VI)									
1	All Firms: Target Incidence	16.6%	15.1%	13.5%	14.3%	15.1%	15.9%	12.7%	15.1%	17.5%
2	All Firms: Acquirer Incidence	19.4%	15.1%	10.8%	12.5%	15.1%	17.7%	12.7%	15.1%	17.5%
	Panel B: Large Firms									
3	Big Firms: Target Incidence	23.1%	21.3%	19.5%	19.7%	21.3%	22.8%	16.4%	21.3%	26.2%
4	Big Firms: Acquirer Incidence	41.8%	37.1%	32.4%	34.5%	37.1%	39.7%	31.5%	37.1%	42.7%
	Panel C: Small Firms									
5	Small Firms: Target Incidence	9.3%	8.9%	8.5%	8.9%	8.9%	8.9%	8.7%	8.9%	9.0%
6	Small Firms: Acquirer Incidence	22.6%	19.2%	15.8%	16.1%	19.2%	22.2%	18.8%	19.2%	19.6%
	Panel D: Mergers Only (Based on models in Table VII)									
7	Mergers Only: Target Incidence	3.9%	4.3%	4.7%	3.9%	4.3%	4.6%	4.8%	4.3%	3.7%
8	Mergers Only: Acquirer Incidence	11.6%	10.4%	9.3%	9.3%	10.4%	11.6%	9.2%	10.4%	11.6%
	Panel E: Acquisition of Assets Only (Based on models in Table VII)									
9	Acq Assets Only: Target Incidence	13.1%	10.8%	8.5%	10.3%	10.8%	11.4%	8.0%	10.8%	13.7%
10	Acq Assets Only: Acquirer Incidence	21.1%	17.7%	14.3%	16.3%	17.7%	19.1%	16.7%	17.7%	18.8%

Table IX: Announcement Returns

The table displays panel data regressions in which the dependent variable is the abnormal announcement return of combined target and acquirer. Announcement returns are computed over various windows including day t=-10 to day t=0 (t=0 is the announcement date) as indicated in the event window column. The combined firm's raw return is the total market capitalization of both firms at the end of the event window minus the original market capitalization, divided by the original market capitalization. The abnormal return results after subtracting the return of the CRSP value weighted market index over each event window. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. We compute product similarities based on the ten nearest firms (for both the acquirer and the target). We also compute the pairwise similarity of the target and the acquirer. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The same SIC-3 industry dummy is one if the target and acquirer reside in the same three-digit SIC code. The vertical similarity dummy is one if the target and acquirer are more than 5% vertically related (based on Fan and Goyal (2006)). Target relative size is the ex-ante market value of the target divided by that of the acquirer. The merger dummy is one if the transaction is a merger and zero if it is an acquisition of assets. Log total size is the natural logarithm of the summed ex-ante market values of the two firms. Large targets are defined as those at least 10% of the size of the acquirer. The sample is from 1997 to 2005. *t*-statistics are adjusted for clustering at the year and industry level.

	Event Window	Acquirer Product Simil. to Rivals	Target Product Simil. to Rivals	Gain in Prod. Diff. vs. Rivals	Target + Acquirer Pair Simil.	% Neighbor Patent Words	Same SIC-3 Industry Dummy	Vertical Simil. Dummy	Acquirer Industry HHI (SIC-3)	Target Relative Size	Full Merger Dummy	Merger x Relative Size	Log Total $ Size	R^2	Obs
							Combined Firm Announcement Returns								
(1)	t=0 only	0.017	-0.015			0.006	-0.000	-0.004	0.012	0.000	-0.003	0.021	-0.002	0.022	4,937
		(2.32)	(-2.21)			(1.35)	(-0.20)	(-1.16)	(1.26)	(0.09)	(-1.44)	(3.16)	(-6.56)		
(2)	t=0 only			0.017	0.007	0.007	-0.000	-0.004	0.010	0.000	-0.003	0.021	-0.002	0.022	4,937
				(2.41)	(0.58)	(1.56)	(-0.09)	(-1.16)	(1.02)	(0.05)	(-1.47)	(3.18)	(-6.53)		
(3)	t=-1 to 0	0.015	-0.018			0.006	0.002	-0.004	0.011	0.001	-0.002	0.023	-0.003	0.023	4,937
		(1.66)	(-2.15)			(1.15)	(1.11)	(-0.94)	(1.02)	(0.49)	(-1.02)	(3.06)	(-6.62)		
(4)	t=-1 to 0			0.020	0.002	0.007	0.002	-0.004	0.008	0.001	-0.002	0.023	-0.003	0.023	4,937
				(2.52)	(0.14)	(1.39)	(1.22)	(-0.95)	(0.75)	(0.42)	(-0.98)	(3.08)	(-6.75)		
(5)	t=-5 to 0	0.032	-0.042			0.005	0.002	0.000	0.004	0.003	0.003	0.022	-0.004	0.024	4,937
		(2.31)	(-3.51)			(0.82)	(0.78)	(0.06)	(0.25)	(0.89)	(1.16)	(2.35)	(-6.76)		
(6)	t=-5 to 0			0.026	0.024	0.008	0.001	-0.000	-0.003	0.002	0.003	0.022	-0.004	0.023	4,937
				(2.02)	(1.32)	(1.28)	(0.71)	(-0.01)	(-0.19)	(0.80)	(0.96)	(2.36)	(-6.55)		
(7)	t=-10 to 0	0.041	-0.037			0.003	0.003	0.002	-0.002	-0.001	0.006	0.028	-0.004	0.018	4,937
		(2.42)	(-2.67)			(0.50)	(1.23)	(0.21)	(-0.12)	(-0.15)	(1.89)	(2.79)	(-5.01)		
(8)	t=-10 to 0			0.033	0.047	0.006	0.003	0.001	-0.007	-0.001	0.006	0.028	-0.003	0.018	4,937
				(2.01)	(2.23)	(0.89)	(1.06)	(0.15)	(-0.41)	(-0.20)	(1.65)	(2.79)	(-4.71)		

Table X: Long Term Performance of Acquirers

The table displays panel data regressions in which three year ex-post (after the effective date) changes in industry-adjusted performance measures are the dependent variable. For a transaction that becomes effective in year t, ex-post cashflow change or sales growth is the one to three year change in profitability from year t+1 until year t+2 (one year) or t+4 (three year) as noted in the horizon column. The measure of performance in Panel A is operating income divided by assets, in Panel B, it is operating income divided by sales, and in Panel C, it is log sales growth. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. The acquirer product similarity (10 nearest) is the average similarity between the acquirer and its ten closest rivals. The target and acquirer product similarity is the pairwise similarity between the acquirer and target firms' products. The gain in product differentiation is the product distance from the target to the acquirer's ten nearest neighbors, less the acquirer's distance to its ten nearest neighbors. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The same SIC-3 industry dummy is one if the target and acquirer reside in the same three-digit SIC code. The vertical similarity dummy is one if the target and acquirer are more than 5% vertically related (based on Fan and Goyal (2006)). Target relative size is the ex-ante market value of the target divided by that of the acquirer. The merger dummy is one if the transaction is a merger and zero if it is an acquisition of assets. Log total size is the natural logarithm of the summed ex-ante market values of the two firms. The sample is from 1997 to 2005. t-statistics are adjusted for clustering at the year and industry level.

Row	Horizon	Acquirer Product Simil. (10 Near.)	Gain in Prod. Diff. vs. Rivals	Target + Acquirer Pair Simil.	% Neighbor Patent Words	Same SIC-3 Industry Dummy	Vertical Simil. Dummy	Acquirer Industry HHI (SIC-3)	Target Relative Size	Merger Dummy	Merger x Relative Size	Log Total $ Size	R^2	Obs
						Panel A: Operating Income/Assets								
(1)	1 Year	0.035			0.006	-0.003	-0.006	0.033	0.005	0.002	-0.004	0.000	0.004	4,451
		(1.39)			(0.49)	(-1.01)	(-0.75)	(1.11)	(1.18)	(0.49)	(-0.27)	(0.35)		
(2)	3 Year	0.071			0.001	-0.008	-0.031	0.080	0.017	0.000	0.005	-0.002	0.017	4,451
		(2.47)			(0.04)	(-1.96)	(-2.86)	(1.88)	(2.86)	(0.00)	(0.31)	(-1.39)		
(3)	1 Year		0.034	0.054	0.007	-0.004	-0.006	0.034	0.005	0.001	-0.004	0.001	0.006	4,451
			(1.88)	(2.28)	(0.56)	(-1.12)	(-0.74)	(1.17)	(1.15)	(0.26)	(-0.30)	(0.72)		
(4)	3 Year		0.052	0.085	0.001	-0.009	-0.031	0.083	0.017	-0.001	0.004	-0.001	0.018	4,451
			(2.27)	(2.74)	(0.05)	(-2.05)	(-2.81)	(1.97)	(2.87)	(-0.24)	(0.26)	(-1.00)		
						Panel B: Operating Income/sales								
(5)	1 Year	0.020			0.025	-0.005	-0.002	0.056	0.013	-0.001	-0.006	-0.004	0.008	4,451
		(0.42)			(1.75)	(-1.13)	(-0.21)	(1.60)	(1.86)	(-0.24)	(-0.35)	(-2.42)		
(6)	3 Year	0.049			0.023	-0.010	-0.044	0.077	0.026	-0.003	0.009	-0.006	0.017	4,451
		(0.96)			(0.89)	(-1.78)	(-2.97)	(1.44)	(2.70)	(-0.30)	(0.34)	(-3.19)		
(7)	1 Year		0.020	-0.002	0.024	-0.005	-0.002	0.058	0.013	-0.001	-0.006	-0.004	0.008	4,451
			(0.58)	(-0.05)	(1.68)	(-0.99)	(-0.18)	(1.73)	(1.88)	(-0.15)	(-0.36)	(-2.49)		
(8)	3 Year		0.015	0.032	0.021	-0.010	-0.044	0.082	0.026	-0.003	0.009	-0.006	0.017	4,451
			(0.38)	(0.64)	(0.84)	(-1.69)	(-2.95)	(1.53)	(2.73)	(-0.35)	(0.33)	(-3.18)		
						Panel C: Log Sales Growth								
(9)	1 Year	0.410			0.002	0.009	-0.001	-0.180	0.069	0.014	0.049	-0.015	0.030	4,451
		(5.83)			(0.05)	(0.80)	(-0.04)	(-2.05)	(3.79)	(0.99)	(1.20)	(-4.27)		
(10)	3 Year	0.714			-0.016	0.011	-0.008	-0.022	0.071	-0.006	-0.019	-0.014	0.018	4,451
		(5.31)			(-0.16)	(0.59)	(-0.16)	(-0.16)	(2.36)	(-0.26)	(-0.26)	(-2.63)		
(11)	1 Year		0.275	0.296	-0.005	0.009	0.003	-0.153	0.070	0.012	0.046	-0.014	0.027	4,451
			(4.37)	(3.78)	(-0.11)	(0.83)	(0.10)	(-1.78)	(3.86)	(0.90)	(1.13)	(-3.82)		
(12)	3 Year		0.472	0.488	-0.029	0.012	-0.001	0.026	0.074	-0.008	-0.025	-0.012	0.015	4,451
			(4.05)	(3.56)	(-0.30)	(0.64)	(-0.02)	(0.19)	(2.45)	(-0.33)	(-0.33)	(-2.22)		

Table XI: Ex-post Product Descriptions of Acquirers

The table displays panel data regressions in which three year ex-post (from year t+1 to t+4) logarithmic growth in the size of the firm's product description is the dependent variable. Size of the product description is measured as the number of words. Firms with larger increases in the size of their product description are interpreted as having introduced more products relative to other firms. For a transaction that becomes effective in year t, ex-post product line growth is the one to three year growth in the size of the product description size from year t+1 until year t+2 (one year), t+3, and t+4 (three year) as noted in the horizon column. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). A higher similarity measure implies that the firm has a product description more closely related to those of other firms. The acquirer product similarity (10 nearest) is the average similarity between the acquirer and its ten closest rivals. The target and acquirer product similarity is the pairwise similarity between the acquirer and target firms' products. The gain in product differentiation is the product distance from the target to the acquirer's ten nearest neighbors, less the acquirer's distance to its ten nearest neighbors. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The same SIC-3 industry dummy is one if the target and acquirer reside in the same three-digit SIC code. The vertical similarity dummy is one if the target and acquirer are more than 5% vertically related (based on Fan and Goyal (2006)). Target relative size is the ex-ante market value of the target divided by that of the acquirer. The merger dummy is one if the transaction is a merger and zero if it is an acquisition of assets. Log total size is the natural logarithm of the summed ex-ante market values of the two firms. The initial product description size is the natural logarithm of the total number of words in the firm's initial (year t+1) product description. The sample is from 1997 to 2005. t-statistics are adjusted for clustering at the year and industry level.

Row	Horizon	Acquirer Product Simil. (10 Near.)	Gain in Prod. Diff. vs. Rivals	Target + Acquirer Pair Simil.	% Neighbor Patent Words	Same SIC-3 Industry Dummy	Vertical Simil. Dummy	Acquirer Industry HHI (SIC-3)	Target Relative Size	Merger Dummy	Merger x Relative Size	Log Total $ Size	Initial Prod. Desc. Size	R^2	Obs
							Panel A: Ex post growth in product description								
(1)	1 Year	0.610			0.074	-0.024	0.122	0.021	-0.051	-0.011	0.108	0.015	-0.275	0.118	3,898
		(3.56)			(1.19)	(-1.24)	(2.69)	(0.15)	(-1.43)	(-0.34)	(1.69)	(2.86)	(-11.72)		
(2)	2 Year	0.727			0.074	-0.021	0.119	0.056	-0.037	-0.013	0.094	0.018	-0.361	0.155	3,898
		(4.22)			(1.13)	(-0.88)	(2.17)	(0.30)	(-0.94)	(-0.38)	(1.16)	(2.98)	(-15.83)		
(3)	3 Year	0.716			0.209	-0.002	0.162	0.056	-0.027	-0.014	-0.009	0.014	-0.396	0.177	3,898
		(3.58)			(2.97)	(-0.10)	(2.85)	(0.27)	(-0.67)	(-0.40)	(-0.10)	(2.59)	(-17.01)		
(4)	1 Year		0.213	0.891	0.068	-0.034	0.125	0.076	-0.047	-0.026	0.100	0.019	-0.273	0.120	3,898
			(1.35)	(4.68)	(1.10)	(-1.77)	(2.72)	(0.55)	(-1.32)	(-0.83)	(1.59)	(3.47)	(-11.83)		
(5)	2 Year		0.263	0.943	0.064	-0.031	0.123	0.124	-0.031	-0.029	0.085	0.022	-0.357	0.156	3,898
			(1.66)	(4.54)	(0.98)	(-1.31)	(2.28)	(0.69)	(-0.80)	(-0.80)	(1.05)	(3.49)	(-15.84)		
(6)	3 Year		0.181	0.812	0.195	-0.011	0.166	0.133	-0.020	-0.027	-0.019	0.018	-0.389	0.176	3,898
			(1.06)	(3.69)	(2.77)	(-0.45)	(2.91)	(0.67)	(-0.50)	(-0.74)	(-0.20)	(3.10)	(-16.90)		

Table XII: Economic Magnitudes of Returns and Real Outcomes

The table displays economic magnitudes associated with various findings reported earlier in this study. All magnitudes are predicted values, and all magnitudes are conditional and thus account for the effects of industry, year and all control variables. For each dependent variable being considered (noted in the panel headers and the description column), we first set all control variables to their mean values and compute the model's predicted value. The result of this calculation is the value displayed in the "mean" column in each category. For each independent variable whose economic magnitude we are measuring (product similarity 10 nearest, product similarity overall, and neighbor patent worlds), which is noted in the column headers, we also compute the model's predicted value assuming the given independent variable is expected to be in the 10th and 90th percentile of its distribution, while still holding all control variables fixed at their mean. Product similarities are measures that lie in the interval (0,1) based on the degree to which two firms use the same words in their 10-K product descriptions (see Appendix 1). We consider similarity based on a firm's ten closest rivals, and similarity based on all firms in the universe excluding these ten firms. A higher similarity implies that the firm has a product description more closely related to those of other firms. The % neighbor patent words variable is the average over a firm's ten nearest neighbors of the percentage of words in the 10-K product description having the same word root as the word patent. The sample is from 1997 to 2005.

			Product Similarity (10 Nearest)			***Neighbor Patent Words***	
Row	Description	10 %ile	Mean	90 %ile	10 %ile	Mean	90 %ile
			Panel A: Announcement Returns (Based on models in Table IX)				
1	Combined Firm Ann Returns (t=0)	0.3%	0.5%	0.6%	0.4%	0.5%	0.6%
2	Combined Firm Ann Returns (t=-10 to t=0)	2.2%	2.6%	3.0%	2.5%	2.6%	2.7%
			Panel B: Profitability and Sales Growth (Based on models in Table X)				
3	Δ OI/Assets: 1 Year (A)	-0.9%	-0.5%	-0.2%	-0.7%	-0.5%	-0.4%
4	Δ OI/Assets: 3 Year (A)	-2.3%	-1.6%	-0.9%	-1.6%	-1.6%	-1.6%
5	Δ OI/Sales: 1 Year (A)	-0.6%	-0.5%	-0.3%	-1.0%	-0.5%	0.1%
8	Δ OI/Sales: 3 Year (A)	-2.4%	-2.0%	-1.5%	-2.4%	-2.0%	-1.5%
7	Sales Growth: 1 Year (A)	11.9%	15.9%	19.9%	15.8%	15.9%	15.9%
8	Sales Growth: 3 Year (A)	20.0%	27.0%	33.9%	27.3%	27.0%	26.6%
			Panel C: Growth in Product Descriptions (Based on models in Table XI)				
9	Prod Desc Growth: 1 Year (A)	-2.4%	3.3%	9.1%	1.7%	3.3%	5.0%
10	Prod Desc Growth: 3 Year (A)	-3.4%	3.3%	10.1%	-1.3%	3.3%	7.9%

Figure 1:

The large dashed circles give a visual depiction of Symantec's and Veritas's ten closest rival firms determined using our measure of product similarity described in Section II. Symantec and Veritas are both within each other's circle of ten nearest rivals. Each firm has a header beginning with the letter "S" or "V" followed by a number. This identifies which firm's circle of ten nearest rivals the given firm exists in, and also how close the given firm is to either Symantec or Veritas. For example, McAfee has a code "S7" and is thus Symantec's seventh closest rival. Veritas is Symantec's 18th closest rival, and thus is an example of a firm that is similar to Symantec, but a firm that also might offer Symantec added product differentiation relative to its very closest rivals. For each firm, we also report its primary three-digit SIC code in parentheses.

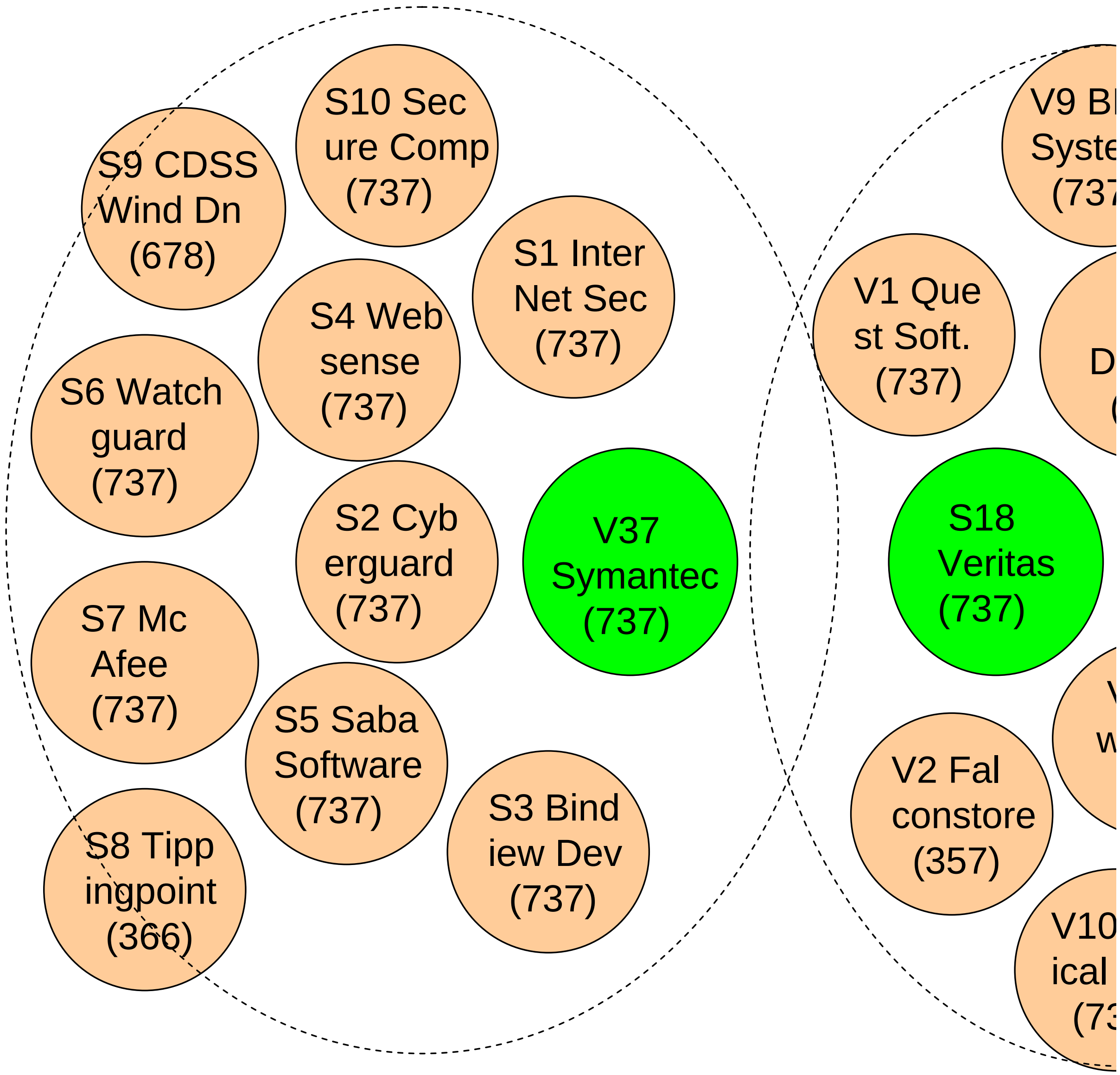

Figure 2:

The large dashed circles give a visual depiction of Disney's and Pixar's ten closest rival firms determined using our measure of product similarity described in Section II. Disney and Pixar are both within each other's circle of ten nearest rivals. The one firm in the middle (News Corp) is common to both Disney and Pixar's circle of ten nearest rivals. Each firm has a header beginning with the letter "D" or "P" followed by a number. This identifies which firm's circle of ten nearest rivals the given firm exists in, and also how close the given firm is to either Disney or Pixar. For example, Six Flags has a code "D4" and is thus Disney's fourth closest rival. Pixar is Disney's ninth closest rival, and thus is an example of a firm that is similar to Disney, but a firm that also might offer Disney added product differentiation relative to its very closest rivals. For each firm, we also report its primary three-digit SIC code in parentheses.

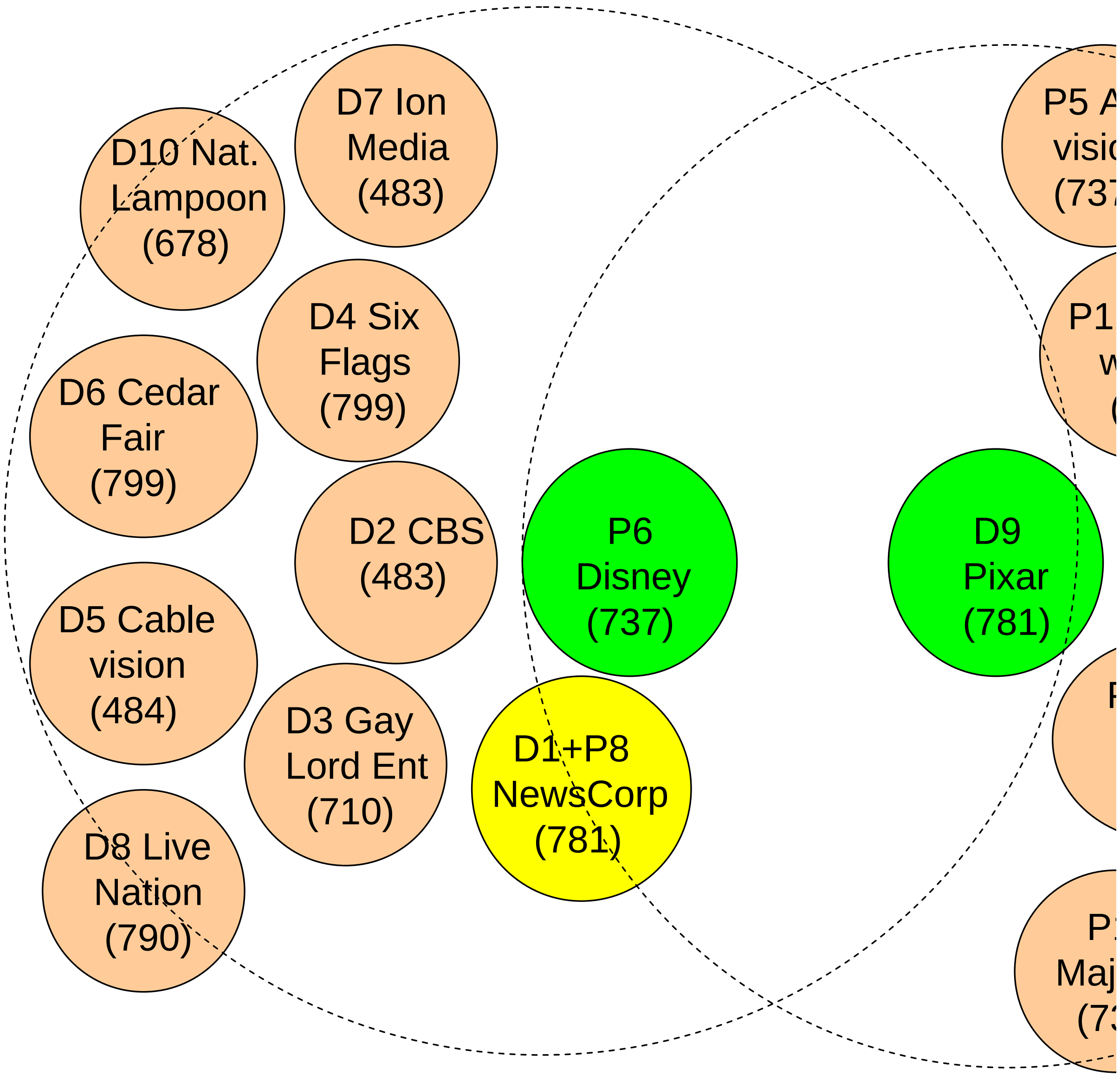

Figure 3:

Distribution of product similarity for random firm pairings and merger pairings. Each plot is an empirical density function, and total probability mass sums to one. The lower axis reflects similarities between zero and 100 (similarities are displayed as percentages for convenience). We truncate displayed results at 50%. The small number of outliers with values higher than 50% are represented by the probability mass assigned to the last bin. The random firm pairings group is based on the subsample of firms that merged, but the differences are taken with respect to a randomly chosen pair of firms in this subsample (results nearly identical in set of firms that did not merge). The lower four plots are based on actual merger pairs.