Product Market Synergies and Competition in Mergers and Acquisitions

Gerard Hoberg and Gordon Phillips*

September 7, 2009

ABSTRACT

We examine how product similarity and competition influences mergers and acquisitions and the ability of firms to exploit product market synergies through asset complementarities. Using novel text-based analysis of firm 10K product descriptions, we find three key results. (1) Firms are more likely to enter mergers with firms whose language describing their assets is similar. (2) Transactions in competitive product markets with similar acquirer and target firms experience increased stock returns and real longer-term gains including higher growth in their product descriptions. (3) These gains are higher when the target is less similar to the acquirer's closest rivals, and when firms have the potential for unique products. Our findings are consistent with firms merging and buying assets to exploit asset complementarities and to create new products to increase product differentiation.

*University of Maryland and University of Maryland and National Bureau of Economic Research respectively. Hoberg can be reached at ghoberg@rhsmith.umd.edu and Phillips can be reached at gphillips@rhsmith.umd.edu. We thank Michael Faulkender, Denis Gromb, Kathleen Hanley, Rich Matthews, Nagpurnanand Prabhala, David Robinson, Matthew Rhodes-Kropf, Paul Tetlock and seminar participants at Duke University, Hong Kong University of Science and Technology, Insead, Lausanne, National University of Singapore, NBER, Ohio State University, the Securities Exchange Commission, the University of North Carolina, the University of Texas, and the University of Maryland for helpful comments. Any remaining errors are the authors alone.

It has long been viewed that product market synergies and competition are key drivers of mergers. 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. Rhodes-Kropf and Robinson (2008) model similarity and asset complementarities as a motive for mergers, but do not present direct evidence of their importance.[1] This paper provides evidence of synergies and new product introductions. We examine whether product similarity increases the incentives to merge and whether the change in merging firms' profits is consistent with the existence of asset complementarities.[2] We examine asset complementarity using novel new measures of product similarity and new product introductions. We calculate these measures using text-based analysis of firm product description sections of firm 10-Ks obtained from web-crawling algorithms that read and process 49,408 10-K filings from the SEC Edgar website.

Mergers are a quick way to potentially increase product offerings if synergies arise from asset complementarities. Firms may also wish to merge with a firm whose skills or technologies are different enough from the acquirer's *rivals* to help the acquirer differentiate its product offerings from its rivals and improve profitability.[3] We explore this "similar but different" tension in examining the likelihood of mergers, who merges with whom, and ex post merger outcomes. Our product based measures calculated from firm 10-Ks capture the degree to which products are different from rivals - both within and across industries - while independently measuring the similarity between the acquirer and target. We examine if there are post-merger gains in profits and product differentiation and if gains are related to our measures of target and acquirer similarity. Although existing studies document that acquirer and target in related industries have positive ex post gains in cash flows, no existing studies examine how competition from non-merging rival firms can further shape merger decisions, and no existing studies have direct evidence of post merger new product creation.

We report three main findings. Our first main finding is that firms are more likely

[1]Many papers have presented evidence that related mergers have higher cash flows post-merger than unrelated ones (see Andrade, Mitchell, and Stafford (2001) and Betton, Eckbo, and Thorburn (2007) for two surveys on mergers). Potential explanations include increased efficiencies, decreased agency costs, increased product-market power and synergies.

[2]Chamberlain (1933) and Hotelling (1929) famously show that the notion of product differentiation is fundamental to profitability and theories of industrial organization.

[3]Gugler and Sieberg (2007) examine the potential for mergers to increase technological progress in the semiconductor industry through the diffusion of know-how and increasing R&D incentives.

to merge with specific firms that are more pairwise similar to themselves and are more likely to engage in M&A if they have products that are more broadly similar to all other firms in the economy. We also find that firms having more very close rivals are less likely to participate in transactions. We interpret this latter effect as a "competitive effect", as firms with very near rivals must compete for restructuring opportunities. These transaction incidence and specific pair selection effects are significant, economically and statistically, for large and small firms alike.

Our last two main findings relate to outcomes. Our second main finding is that long-term real outcomes are superior (higher profitability, sales, and evidence of new product introductions) when the target and the acquirer are more pairwise similar. Gains are also larger when there is evidence of a higher potential for unique products, as captured by a higher usage of patent and copyright words in the product description, suggesting that rivals will not be able to replicate new products. Our results broadly suggest that merging firms with similar assets use synergies from asset complementarities to introduce new products and improve cash flows.

Our third main finding is that merger and acquisition transactions experience better market reactions and long-term real outcomes 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. Our results are robust to including controls for horizontal and vertical relatedness used in the existing literature.

We differ from the existing literature in two major ways. First, the literature has not focused on how product differentiation and competition jointly affect acquisition choices, subsequent performance, and new product development. Second, we present an innovative new methodology based on text analysis that allows us to separately measure a firm's similarity to its rivals, the product location of a target relative to these rivals, and the product location of the target relative to the acquirer itself (pairwise similarity) for a very large sample of firms. Existing studies have relied on SIC code based variables[4]. While partially informative, this past approach cannot measure the degree to which firms are similar within and across industries, and how this changes in time series. Our tests would not be possible using similarity measures

[4]For example, 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.

based on traditional industry classifications such as SIC or NAICS codes.

In a companion paper (Hoberg and Phillips (2009)), we show that our new text-based measures have broad applicability beyond mergers. For example, as predicted by many theories of industrial organization, we find that firms with greater product differentiation (measured using our 10-K similarities) have higher profitability. As predicted by Panzar and Willig (1981), we also find a strong role for economies of scope. As predicted by Sutton (1991), we also find that firms that spend more on advertising and R&D experience increases in subsequent product differentiation. SIC or NAICS industries cannot measure the extent that industries change in their product mix and differentiation over time, thus our text-based measures provide additional information not contained in SIC or NAICS industries. Hoberg and Phillips (2009) also show that 10-K text-based industry classifications offer significantly more explanatory power than SIC or NAICS in explaining numerous firm-specific variables in a large panel data setting, including the extent that groups of firms co-move in the stock market.

Our research contributes to the literature on mergers, similarity, asset complementarity, and industrial organization. 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. Rhodes-Kropf and Robinson (2008) model asset complementarity and synergies as a motive for mergers. Our evidence is consistent with their model. We discuss the related literature in greater detail in the next section.

We also add to a new growing literature that uses text analysis in finance. Hanley and Hoberg (2008) address theories of IPO pricing by examining prospectus disclosures on the SEC Edgar website, and separate text into standard and informative content. Loughran and McDonald (2009) examine the SEC's Plain English rules on corporate finance outcomes and disclosure. Outside of corporate finance, other studies that use text based analysis to study the role of the media in stock price formation include Tetlock (2007), Macskassy, Saar-Tsechansky, and Tetlock (2008), 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. Determinants of the likelihood of restructuring transactions are presented in section IV. A discussion of ex-post outcomes upon announcement and in the long term is in section V. Section VI 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 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.[5] 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 introductions. We use text based methods to directly measure similarity and potential asset complementarity of merger partners rather than inferring it using market-to-book ratios. Perhaps more importantly, using new techniques, our paper can measure potential synergies through asset complementarities, how similar merging firms are from rivals, and how product descriptions grow post merger.

[5] 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)).

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 (although Kedia, Ravid, and Pons (2008) show that this reverses after 1996). While partially informative, SIC code based measures cannot 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. More recently, following Berry, Levinsohn, and Pakes (1997), the approach of the product differentiation literature has been to estimate demand and cost parameters in specific markets including the ready-to-eat cereal market (Nevo (2000)).

Our study has a different focus. We focus on the incentives for firms to merge to differentiate themselves and to exploit asset complementarities. Using text based measures, we examine how firm's are related to each other without relying on predefined industries. We can measure product similarity across arbitrary firms and in time series, allowing us to test for new product introductions more broadly, and across multiple industries. Our analysis conceptually is related to the empirical question raised by Berry and Waldfogel (2001), who study competition in the radio broadcasting market and show that the number of broadcasting formats increases post-merger. Our approach is also related to other recent studies including Mazzeo (2002) and Seim (2006), who show that product differentiation continues to be important nearly a century after the seminal work by Chamberlain (1933) and Hotelling (1929). Although these latter studies do not examine mergers, they also focus on the incentives for firms to differentiate themselves.

B 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 Similarity: *Firms are more likely to merge with other firms whose assets are highly similar or related to their own assets.*

Hypothesis H2a: Differentiation from Rivals: *Acquirers in competitive product markets should be more likely to choose targets that help them to increase product differentiation relative to their nearest ex-ante rivals.*

Hypothesis H2b: 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.*

Our second hypothesis H2b suggests that not only will competition reduce the likelihood that any given firm will merge, but it can also reduce the premium realized by a target when it does merge. Key to testing H1 and H2b is the degree of similarity. The effects of H2b are likely to be stronger 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. Our later tests on product introductions examine whether this is just similarity alone, or if complementary assets result in new products.

Our remaining hypotheses focus on long term outcomes.

Hypothesis H3: 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 3, are numerous as discussed earlier in section I.A. 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).

Hypothesis H4: Outcomes and Differentiation from Rivals: *Acquisitions that differentiate or have the potential of differentiating the acquirer from its rivals should experience positive outcomes post-acquisition.*

Hypothesis 4 follows from two ideas. First, recent (Mazzeo (2002) and Seim (2006)) and historical studies in industrial organization beginning with Hotelling (1929) suggest that firms that differentiate themselves 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.

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 profitability can be improved (Hypothesis 4), and highly similar targets are more likely to exist (Hypothesis 3). 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 from rivals".

Hypothesis H5: Product Uniqueness: *Gains from restructuring will be even larger if firms have unique assets that can be jointly used to create new products, as rivals would be less able to replicate the acquirer's strategy.*

We measure product uniqueness as the fraction of times firms mention patents and copyrights in their product description section of their 10-Ks. We then test whether ex post increases in product descriptions increase with these measures of uniqueness.

We illustrate these hypotheses using an example: the Symantec and Veritas (SV) 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.

[**Insert Figure 1 Here**]

The SV merger is consistent with Symantec choosing Veritas because it is similar enough to permit successful managerial integration and new product synergies (Hypothesis H3), 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 H4), 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 H5). 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 H4 and H5 are likely not possible, explaining why Veritas might be a more appropriate target.

II Data and Methodology

A Data Description

A 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).[6] We construct these text-based measures of product similarity (described later in this section) using firm product descriptions obtained directly from 10-K filings on the SEC Edgar website starting from the period electronic 10K records became available (1997 to 2005). 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 acquisitions. 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.[7] 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.[8]

We extract the product description section from each linked 10-K. This section

[6]We also considered whether variables based on the properties section of the 10-K might help to separate asset similarities and product similarities. After hand examining this section for over 100 documents, however, we concluded that the properties section almost exclusively describes real estate holdings and the location of the firm's corporate headquarters, rendering it not applicable for this purpose.

[7]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.

[8]The COMPUSTAT tax ID number is only available as a header variable and thus reflects the firm's most recent tax ID number. Because some firms maintain the same COMPUSTAT gvkey and CRSP permno even as their Tax ID number changes, these hand corrections ensure that our database has uniform coverage over all sample years. Hand corrections are 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 corrections are hand verified.

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. We exclude these firms.

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.

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 10-K product description. This method results in a real number in the interval (0,1) that captures the similarity of words used for each pair of firms. Details are discussed in Appendix 1 and also in Hoberg and Phillips (2009). Whereas the current paper's focus is on mergers, Hoberg and Phillips (2009) show that these new similarity measures are more broadly applicable in constructing industry classifications and measuring industry competition. The resulting industry classifications outperform traditional SIC and NAICs classification in explaining a broad array of financial variables in large panel data setting including operating margins, R&D and advertising intensity, and industry co-movement with the broader stock market. Hoberg and Phillips (2009) show that measures of competition based on these new industries (unlike those based on SIC and NAICS) have empirical properties that are broadly consistent with theories of industrial organization including theories of product differentiation (Chamberlain (1933)), economies of scope (Panzar and Willig (1981)), and endogenous barriers to entry (Sutton (1991)).

The main idea our method captures is that firms having descriptions with more words in common 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 the potential for unique products based on usage of word roots including "patent", "copyright", and "trademark".

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 roots "patent", "copyright", and "trademark". Because our focus is on whether a product market has or has the potential for new products (hypothesis H5), we then compute this variable's average over each firm's ten nearest rivals.[9]

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.

In independent analyses, we tested whether our similarity variables (constructed from the 10-K product descriptions) are valid measures of product market competitiveness and future product differentiation (low product similarity). We first examine

[9]Our inferences are unchanged if we only use the "patent" word root alone (not reported).

whether firm profitability increases with our measure of product differentiation (negatively with product similarity), and also whether firms with high advertising and R&D have higher future product differentiation. This "validation" is based on a large body of supporting theoretical and empirical literature, including the predictions of recent models that endogenize product choice (Mazzeo (2002) and Seim (2006)), and historical studies that first considered product differentiation (Chamberlain (1933) and Hotelling (1929)). In Appendix 2, we find strong evidence that our product similarity variable (the opposite of differentiation) is negatively related to profitability. In additional results available from the authors, we also find that firms with high advertising and R&D have significantly higher future product differentiation (lower similarity to rivals). These findings are both statistically significant (at the 1% level) and economically significant.

Product similarities are most intuitive when firms have only one segment. Our computer-based algorithms are not able to separate the text associated with each segment of conglomerate firms.[10] However, we believe that similarities measured relative to conglomerates are still informative regarding the competition faced 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.[11]

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.

---

[10] 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.

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

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 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.

Although SIC codes are informative in many applications, we present evidence that SIC codes only weakly measure 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.

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 equal to one if the given transaction is a merger and zero for an acquisition of assets. We only examine these two transactions.

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 3 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 we only consider 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 content. The list of similar words for each pair is also substantial, indicating that our identification of 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 roots “patent”, “copyright”, and “trademark” (patent words) is 0.227%.

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).[12] 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

[12] 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.

announcement return very close to zero (0.2%), and the average target experienced an announcement return of 6.9%. 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 horizons.

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 to the one hundred and ten nearest neighbors, respectively.[13] 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 2 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 2 Here]

[13] 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.

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 within these groups illustrates that SIC codes are too granular to capture all product heterogeneity. Also striking is the relatively high level of merger pair similarity observed for merger pairs residing in different two-digit SIC codes in the bottom most 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 chosen 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 very similar. We conclude that product similarity adds information not contained in SIC codes and measures similarity within and across industries.

IV Merger and Asset Acquisition Likelihood

In this section, we use choice models to test Hypotheses H1, H2a, and H2b, and we examine the link between transaction decisions and measures of product market competition and similarity. We first examine the decision to merge or not merge in section A. We report economic magnitudes in section B. We then extend our analysis in section C and examine the decision of which target is most suitable for a given acquirer. These tests allow us to more finely test Hypotheses H1 and H2a.

A Transaction Incidence

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 an acquirer of a restructuring transaction in a given year (Panel A) and a dummy equal to one if the given firm is a target (Panel B). In all specifications, we report $t-$statistics that account for clustering at the year and industry level.

[Insert Table VI Here]

Consistent with Hypothesis H1, 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.[14] Our later tests support the interpretation that our similarity measures capture asset complementarity, and not similarity associated with cost reductions. In particular, we present evidence consistent with new products being introduced when similar firms merge, and additional evidence that our similarity variables are not significantly related to cost reductions.

Second, consistent with Hypothesis H2b, 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 asset complementarity effect and the competitive effect are economically significant.

Table VI also shows that firms using more patent words in their product descriptions are more likely to restructure, both as targets and as acquirers. A likely explanation is that patents, copyrights and trademarks serve as measures of the existence of, and potential for, unique products. 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 new product creation via organic investment.

The table also shows that the SIC-based sales HHI variable is negatively related to

[14]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.

restructuring for acquirers, and weakly positively linked to restructuring for targets. 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 asset complementarity 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 predicts future restructuring for both SIC-based and product similarity based measures, confirming that both contain distinct information.

We also examine whether our results are explained by industry demand shocks. Although industry shocks likely cannot explain our similarity results due to the fact that firm similarities appear to be stable over time, whereas shocks are by definition unstable, we examine the role of shocks to ensure robustness. We consider a proxy for industry shocks based on industrial product shipment data from the Bureau of Labor and Statistics. We define an industry's lagged demand shock at the three-digit SIC level, as the logarithmic growth in its shipments from year t-2 to year t-1.[15] We consider two variations: own industry shock, and downstream industry shock (downstream industries are identified by the input/output tables previously discussed). Not surprisingly, both are less than ten percent correlated with our similarity variables. Hence, our results are virtually unchanged when we control for industry shocks. Although industry shocks do not explain our current findings, we can report that the results are independently interesting. We find that firms in industries with positive demand shocks are more likely to be acquirers, and less likely to be targets. When we limit attention to downstream shocks, the result for acquirers is robust, and the result for targets becomes insignificant.[16]

[15] We also include a dummy variable for industries where BLS shipment data is not available.

[16] Although we do not report the results to conserve space, we separately reproduce the tests in Table VI for small and large firms as well as separately by transaction type for mergers and acquisition of asset transactions. Both the asset complementarity effect and the competitive effect are especially robust for large firms, and results for acquirers are also robust for small firms. By transaction type, we find the asset complementarity effect remains highly positive and statistically significant for both mergers and acquisition of assets transactions. The competitive effect is stronger for acquisition of asset transactions than for mergers, given that the acquisition of assets are more common, as noted earlier.

B 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 magnitudes in terms of the dependent variable itself around its mean.

[**Insert Table VII Here**]

Table VII confirms that our findings regarding transaction incidence are economically relevant. The effect of changing similarity (10 nearest) from the 10%ile to the 90%ile changes acquirer incidence from 30.6% to 23.3%. 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 targets at 16.3% to 13.8%. The overall similarity variable is also large and moves from 24.6% to 29.4% for acquirer incidence. This similarity effect is considerably smaller for target incidence with a 1.8% spread. The patent words variable also has a rather substantial economic impact on both acquirers and targets, as it changes acquirer incidence from 24.0% to 30.0%, and target incidence from 12.3% to 17.8%. This variable's role is consistent with acquisitions being a necessity when rival firms need to acquire patent protected technology as organic investment is less feasible. Acquisitions should be more common when they are the only option.

C Merger Pair Selection

In this section, we examine which target firms are more likely to pair with which acquirers. Our focus is on more finely testing Hypotheses H1 and H2a. Hypothesis H1 suggests that a specific target and acquirer pair should be more likely if the firms are more pairwise similar. H2a explores the role of product market competition. In competitive product markets, where gains from differentiation should be largest, more appealing targets will be situated further from the acquirer's nearest rivals, as they allow ex-post improvements in previously hard-to-obtain pricing power.

We employ the nested logistic model because it offers the flexibility needed to account for the possibility that the decision to merge, and with whom, are likely considered jointly. This model uses a tree structure, and relaxes the independence of irrelevant alternatives assumption of non-nested models (see McFadden (1981) for details, and see Yasuda (1995) for a similar application used to examine bank relationships and underwriter choice). Figure 3 displays the tree underlying our application. We use a smaller sample of transactions for this test, as the necessary pairwise similarity calculations require that candidate target and acquiring firms both be publicly traded with machine readable 10-Ks. In this sample, 8.6% of all firms are acquirers, and 6.4% are targets (compared to 28.2% and 15.1% in our overall sample). This additional requirement does not change our main inferences.

Although a nested logistic model considering every possible merger pair would be ideal[17], we instead place candidate target firms into ten groups that maximize uncorrelated variation along two key dimensions needed to jointly test hypotheses H1 and H2a. Consider firm "A" considering an acquisition. The first dimension is the average pairwise similarity between A and each group of candidate targets, and we sort candidate target firms into five quintiles based on this variable (sorts are annual). The second dimension is each candidate target's similarity to firm A's nearest rivals, and we separate firms in each quintile into two bins: close or far from A's ten nearest rivals using spreading sorts (see Teoh, Welch, and Wong (1998), for example). We first sort firms in each quintile by their pairwise similarity to A, and consider sorted firms two-at-a-time (eg. the first two sorted firms are one sub-group, the third and

[17]Although our data is rich enough for this model, it would entail a nested tree with over 5,000 terminal nodes, and 5,000*5,000/2 observations per year, which is not feasible for us to fit. Such a structure would also result in a "rare events" problem.

fourth are one sub-group, etc.). Within each sub-group, the firm that is more similar to A's ten nearest rivals is placed in the close "C" bin, and the other firm in the far "F" bin. Each candidate target thus resides in one of ten possible bins: 1C, 1F, 2C, 2F, ..., 5C, and 5F.[18]

The nested logit permits us to account for specific target choice (terminal node variables), while accounting for the decision to merge or not merge (inner node variables). Our inner node variables include all included in our baseline merger logit in Table VI. Our terminal node variables vary across target groups and include each group's average pairwise similarity to firm A, and a dummy indicating whether the group is near or far from A's ten nearest rivals. We also interact these variables with product market competition variables (similarity between A and its ten closest rivals).

[Insert Table VIII Here]

Table VIII displays the coefficients, their t-statistics, and marginal effects for inner node variables (Panel A) and terminal node variables (Panel B). We consider two models: row (1) displays results for the decision tree from the acquirer's perspective, and row (2) displays analogous results for the decision tree from the target's perspective (whether to accept an acquisition and with which acquirer). The nested logit's results for the inner nodes in Panel A are similar to those in Table VI, which models only the decision to merge or not merge. Analysis of the terminal nodes in Panel B enables us to conduct more precise tests examining whether similarity affects the actual choice of merger and acquisition target.

Consistent with H1, Panel B shows that both acquirers and targets are overwhelmingly more likely to choose counterparts that are more pairwise similar to themselves. The similarity of group to firm coefficient is positive and very significant (T=+8.4), and marginal effect calculations show that a one standard deviation perturbation of pairwise similarity increases the average group's likelihood of being chosen from 10% to 26.1%. If we focus on the most important group (group Q1C, which is chosen most often), this group's likelihood of being chosen increases from 45.7% to 86.0% if its pairwise similarity to A is increased by one standard deviation.[19] This variable's

[18]Spreading sorts generate variation along both dimensions (pairwise similarity and proximity to A's rivals) that is uncorrelated. In unreported correlation calculations, we confirm that these variables are less than 1% correlated across the ten groups.

[19]More specifically, perturbing group Q1C changes the percentage probability of choosing groups

especially high economic significance provides very strong support for H1.

Table VIII also supports H2a, as the "Group close to rivals x Simm (10 nearest)" cross term in Panel B is negative and highly significant. This cross term is high when a group is close to A's ten nearest rivals (ie, groups 1C, 2C, 3C, 4C, and 5C), and when the acquirer resides in a competitive product market. We conclude that, holding constant a group's distance from the acquirer itself, acquirers in competitive markets favor firms that are more distant from the acquirer's nearest rivals. This effect is also economically relevant, as perturbing the average group by one standard deviation of this cross term decreases its likelihood of being chosen by 1.3%. The same perturbation decreases the most important group's selection probability by 5%.

Our results broadly show that firms are more likely to merge with, or acquire assets of, firms who have similar complementary assets. Firms are also less likely to favor firms that are similar to their rivals when they are in more competitive product markets.

V 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.

{Q1C, Q1F, Q2C, Q2F, ..., Q5C, Q5F} from {45.7, 39.5, 4.2, 3.7, 2.0, 1.7, 1.1, 0.9, 0.6, 0.5} to {86.0, 10.2, 1.1, 0.9, 0.5, 0.4, 0.3, 0.2,,0.2, 0.1}.

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.[20] Note that our 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 that are less related to the acquirer's existing rival firms. Rows 1, 3, 5, and 7 support this conclusion at the five percent level for longer horizons (five days or more), and at the ten percent level or better for shorter horizons of two days or less. This rather strong evidence supports the notion that the market rewards acquirers that buy assets permitting new product synergies.

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 H4 for all horizons, as the gain in product

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

differentiation is positive and significant at the 5% level or the 1% level in all rows. We also see some support for H3 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. We also find that the % neighbor patent words variable is positive in every row, and also significant for shorter horizons. Hence, we also find some support for H5.

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 firm (mergers, unlike asset acquisitions, involve all of the target's assets). The vertical similarity variable is negative and significant for short horizons (consistent with Kedia, Ravid, and Pons (2008)), but this result becomes insignificant for longer horizons.

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.[21]

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

[21] 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.

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.

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).[22]

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

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

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, as shown in the next table, also have increases in their product descriptions - consistent with them developing new products and increasing 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 in Panel A (sales is in the denominator in Panel B). The strong results for sales growth are especially consistent with the development of new products.

In Table X, Rows (3) and (4) in Panel A, and rows (11) and (12) of Panel C, show strong support for Hypothesis H4. In particular, the gain in product differentiation coefficient is positive and significant. H3 is also supported by the positive and significant target+acquirer pairwise similarity coefficient. We do not see significant coefficients in rows (7) and (8) in Panel B for these variables because Panels A and C suggest that results for 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.

C Product Descriptions

In this section, we consider the prediction that new product development will accompany positive real outcomes. New product development is especially likely when the target firm is more similar to the acquirer due to complementary assets (H3), and also likely when patents indicate the potential for unique products (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 in this section is smaller than that used in section B. This arises because COMPUSTAT data (needed for the dependent variables in Section B) is available for a slightly longer horizon than our collected 10-K data (needed in this section).

Table XI presents the results of these tests. We find support for hypotheses H3 and 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) patents indicate the potential for unique products. Evidence supporting H3 is particularly strong, as pairwise similarity is significant at the 1% level. For the three year horizon, the % neighbor patent words variable is positive and significant at the 1% or 5% level. This evidence is consistent with unique products requiring more time to develop.

Table XI also shows that vertical mergers, as described in Fan and Goyal (2006), independent of our similarity measures, experience ex-post growth in the size of product descriptions. This evidence linking vertical mergers to the introduction of new products supports the conclusion of Fan and Goyal (2006) that vertical mergers create value. However, this result loses significance for longer horizons.

The table also shows that a same-industry SIC dummy variable is insignificant and thus 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 broadly measure the degree of product similarity. The table also documents that product descriptions have a tendency to mean revert over time, a feature that is likely due to writing style and a preference for brevity. We control for this feature 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 using 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 variables. We also reproduce these tests after excluding repeat acquiring firms. 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 as in section IV.

Table XII displays the results. 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 0.4%, 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 event day 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% of the total standard deviation, and this variable impacts the announcement return by more than the mean itself. The relatively small nominal size of these returns is consistent with our data containing partial acquisitions, and with target firms being smaller than acquiring firms.

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 -2.2% to -1.0% (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 12.0% to 19.8% (one year) and 20.3% to 33.6% (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 -2.2% to 8.9% (results similar for one and three year horizons). This spread is economically large. Overall, our most economically significant findings relate to transaction incidence, sales growth, and product description growth.

E Robustness

In this section, we examine alternative theories and summarize additional 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). We 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 information not contained in these vertical relationships.

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 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 complementarities. 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. 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 are due to this hypothesis for two reasons. First, the word lists driving our similarity measure (see Table II) fit a product market interpretation, and not corporate culture. 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.

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 and are thus not driven by repeat acquiring firms.

VI 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 similarity, product market competition, and new product development have advantages over SIC code measures as they can capture similarity within and across product markets, and changes in similarity over time.

We find that firms choose merger partners that are "similar to self but different from rivals". The goal of choosing a target firm that is similar to the acquirer itself is consistent with asset complementarity. When localized product market competition is high, then firms are also more likely to choose targets that are more dissimilar to their rivals. This is consistent with firms merging to increase their differentiation versus their rival firms. We also find that firms with a higher potential for unique products, as captured by patents, copyrights and R&D, are also more likely to enter into restructuring transactions. This finding is consistent with merging firms exploiting unique products and the importance of protecting product-market differentiation that might result from asset complementarities.

Examining post-merger outcomes, we find that value creation upon announcement, long-term profitability, sales growth, and most interestingly, increases in ex post product descriptions are higher when acquirers purchase targets that (1) have high pairwise similarity to the acquirer's own products, and (2) increase the acquirer's product differentiation relative to its nearest rivals especially in competitive product markets. These gains are larger when there are unique products and patents increasing the potential for new product introductions. Our findings suggest these economically relevant gains are associated with new product introductions.

Overall, our results are consistent with firms merging to use asset complementarities to create value through sales growth and new product introductions. More broadly, our results suggest that firms facing high ex-ante competition can actively improve profitability via strategic restructuring transactions that increase ex-post product differentiation through new product synergies.

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 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 one if the word associated with the given element is in the given firm's product description. To focus on products, we restrict the words in this vector to less commonly used words. Very common words include articles, conjunctions, personal pronouns, abbreviations, and legal jargon, for example. 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 how similar the products of firms i and j are, we take the dot product of their normalized vectors, which is then "product similarity".

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

We define product differentiation as one minus 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. The differentiation between two products is zero if they are the same, and one if they are entirely different.

Appendix 2: 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)			

Appendix 3 (Table II Cont)

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 3 Continued

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.

Berry, Steven, James Levinsohn, and Ariel Pakes, 1997, Automobile prices in market equilibrium, *Econometrica* 63, 841–890.

Berry, Steven, and Joel Waldfogel, 2001, Do mergers increase product variety? evidence from radio broadcasting, *Quarterly Journal of Economics* 116, 1009–25.

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, *Journal of Finance* 55, 389–406.

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

Gugler, Klaus, and Ralph Sieberg, 2007, Market power versus efficiency effects of mergers and research joint ventures: evidence from the semiconductor industry, *Review of Economics and Statistics* 89, 645–659.

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, forthcoming, Journal of Finance.

——— , 2009, New dynamic product based industry classifications and firm performance, 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, *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, *Journal of Finance* 47, 107–138.

Kedia, Simi, Abraham Ravid, and Vicente Pons, 2008, Vertical mergers and the market valuation of the benefits of vertical integration, Rutgers Business School Working Paper.

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

Loughran, Tim, and Bill McDonald, 2009, Plain english, Working paper, University of Notre Dame.

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 C. Tetlock, 2008, More than words: Quantifying language to measure firms' fundamentals, *Journal of Finance* 63, 1437–1467.

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

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

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

Mazzeo, Michael, 2002, An empirical model of firm entry with endogenous product choices, *Rand Journal of Economics* 33, 221–42.

McFadden, Daniel, 1981, *Econometric models of probabilistic choice, in Charles Manski and Daniel McFadden, eds.: Structural Analysis of Discrete Data with Econometric Applications* (MIT Press: Cambridge, Mass).

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.

Panzar, J., and R. Willig, 1981, Economies of scope, *American Economic Review* 71, 268–272.

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

Seim, Katja, 2006, An empirical model of firm entry with endogenous product choices, *Rand Journal of Economics* 37, 619–40.

Sutton, John, 1991, *Sunk Costs and Market Structure* (MIT Press: Cambridge, Mass).

Teoh, Siew-Hong, Ivo Welch, and T.J. Wong, 1998, Earnings management and the long-run market performance of initial public offerings, *Journal of Finance* 53, 1935–1974.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

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

Yasuda, Ayako, 1995, Do bank relationships affect the firm's underwriter choice in the corporate-bond underwriting market, *Journal of Finance* 60, 1259–92.

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.

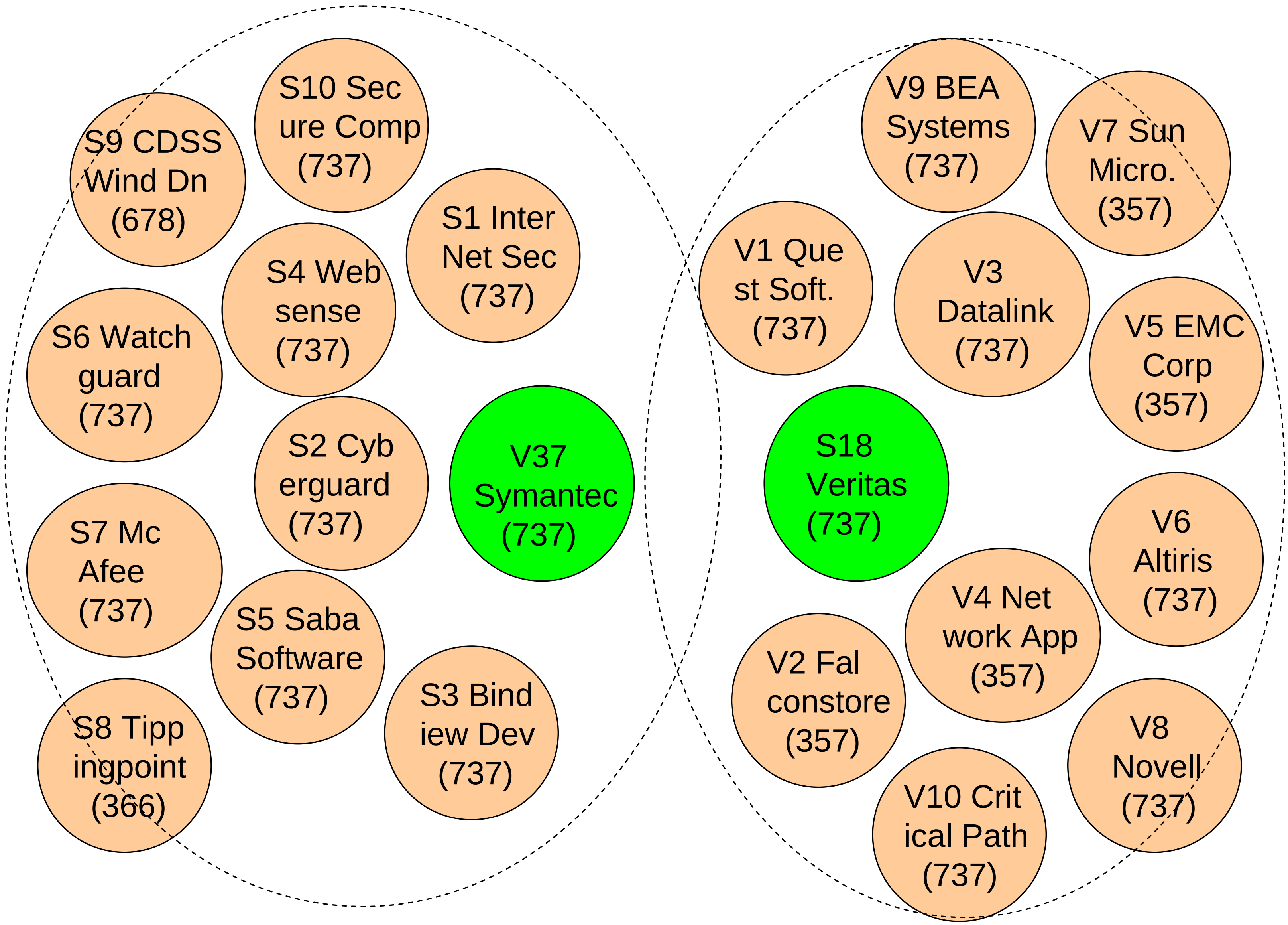

Figure 2:

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.

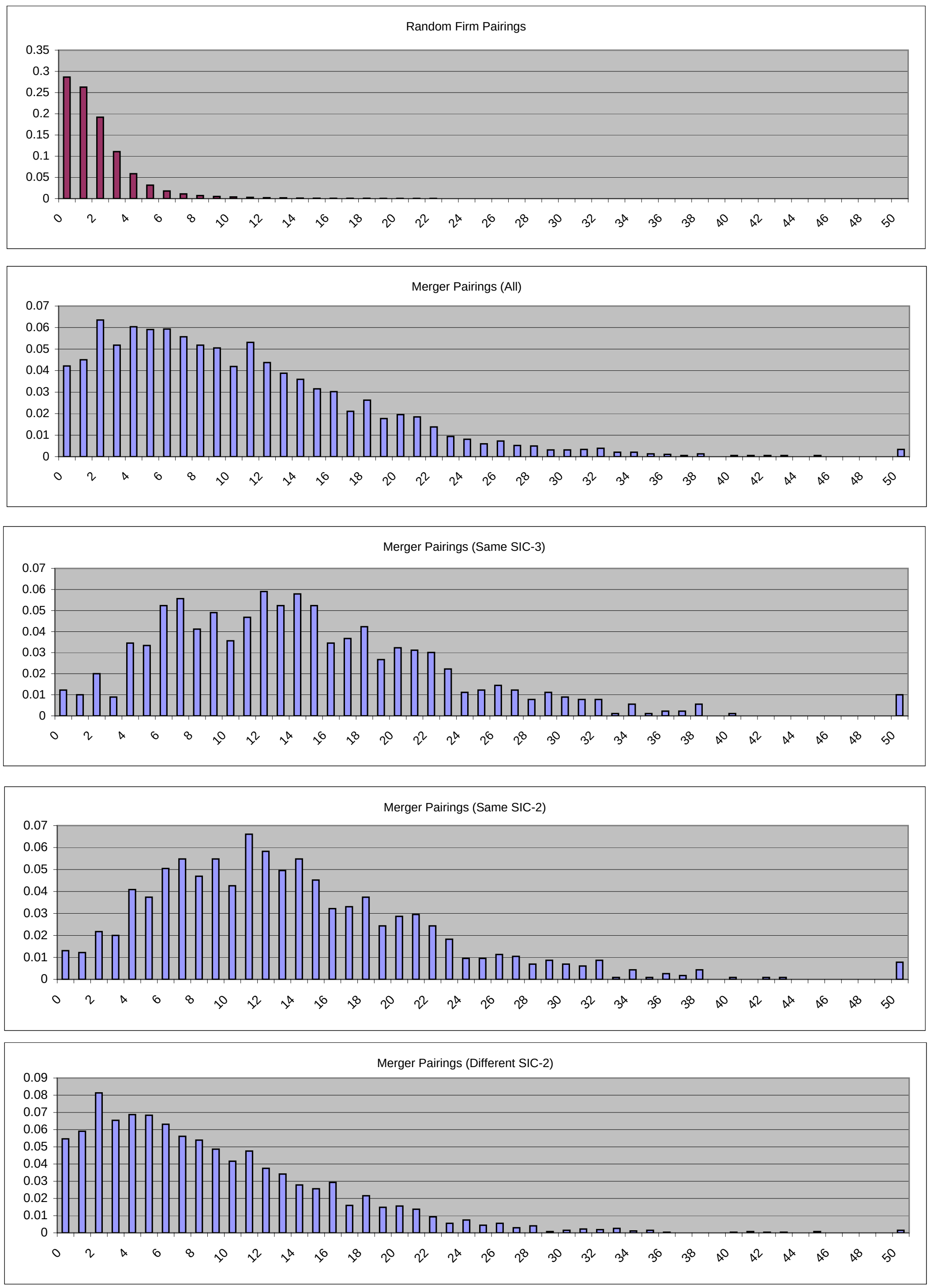

Figure 3:

Tree diagram for the nested logistic model used in Table VIII. The first branch of the nested logit models the decision of a firm “A” to merge or not merge in a given year. If a firm decides not to merge, the tree ends. If it decides to merge, we then model A’s choice among groups of candidate target firms designed to maximize the dispersion in candidate targets across two dimensions. The first is the pairwise similarity between A and a candidate target, and firms are first sorted into quintiles 1 to 5 along this dimension (sorts are annual). Firms in each quintile are then separated into bins that are close or far from A’s ten nearest rivals using spreading sorts. This is done by sorting firms in each quintile by pairwise similarity and then taking groups of two firms at a time (eg. the first two sorted firms are one group, the third and fourth are one group, etc.). Within each group of two firms, the firm with the highest average similarity between itself and A’s ten nearest rivals is placed into bin “C” (close). The other firm is placed into bin “F” (far). As a result, each candidate target firm resides in one of ten possible bins, each having a number of firms that differs by no more than one: Group 1C, 1F, 2C, 2F, ..., 5C, and 5F. The spreading sort ensures that pairwise similarity and the similarity relative to A’s ten rivals are uncorrelated. The choice among these ten groups makes up the second branch of the nested logit tree.

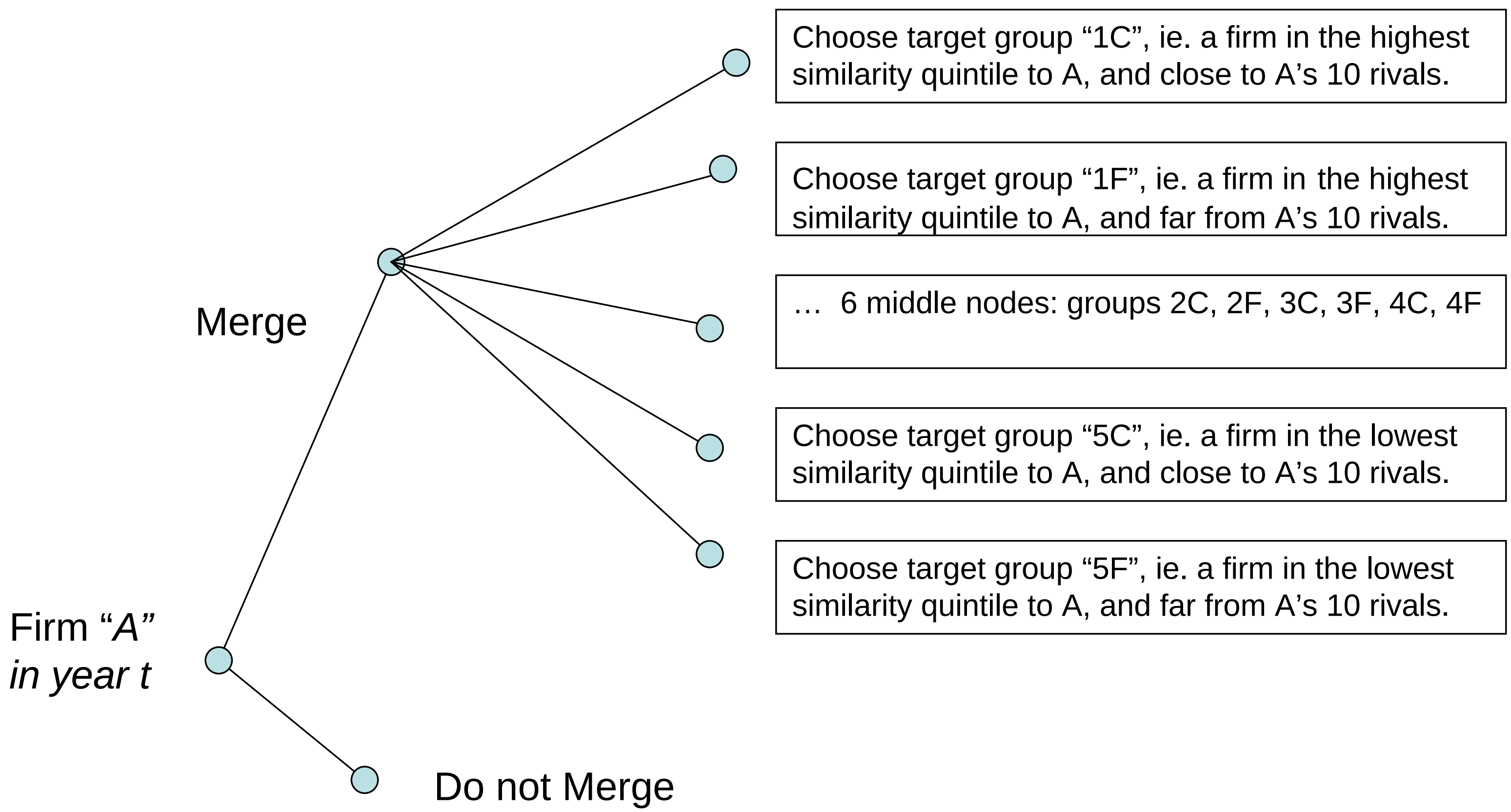

Table I: Merging firms in 2005 with highest percentile similarity but different two-digit SIC codes

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).

Acquirer	Target	Acquirer SIC-3	Target SIC-3
Allis-Chalmers Energy Inc	RPC Inc-Technical Services	SIC3=735, Equipment Rental + Leasing	SIC3=138, Drilling Oil + Gas Wells
Americredit Financial Services	Bay View Acceptance Corp	SIC3=619, Finance Services	SIC3=602, National Commercial Banks
Atlas Pipeline Partners LP	Energy Transfer Partners	SIC3=492, Natural Gas Transmission	SIC3=131, Crude Petroleum + Natural Gas
Belo Corp	WUPL-TV, New Orleans	SIC3=271, Newspapers: Publishing + Printing	SIC3=483, Radio Broadcasting Stations
Buckeye GP LLC	Atlas Oil Co-Refined Petroleum	SIC3=461, Pipe Lines	SIC3=590, Retail-Miscellaneous Retail
CareerStaff Unlimited Inc	ProCare One Nurses LLC	SIC3=805, Services-Nursing + Personal Care	SIC3=874, Services-Management Services
ChevronTexaco Corp	Unocal Corp	SIC3=291, Petroleum Refining	SIC3=131, Crude Petroleum + Natural Gas
Correctional Properties Trust	Geo Group Inc-Lawton	SIC3=679, Miscellaneous Investing	SIC3=874, Services-Management Services
Eagle Hosp Prop Trust Inc	Hilton Glendale	SIC3=679, Miscellaneous Investing	SIC3=701, Hotels + Motels
EMCORE Corp	JDS Uniphase Corp	SIC3=355, Special Industry Machinery	SIC3=366, Telephone + Telegraph Apparatus
Enterprise Products Partners	El Paso Corp-Natural Gas	SIC3=131, Crude Petroleum + Natural Gas	SIC3=492, Natural Gas Transmission
Express Scripts Inc	Priority Healthcare Corp	SIC3=641, Insurance Agents, Brokers + Service	SIC3=512, Wholesale-Drugs + Sundries
First Advantage Corp	Credit Information Group	SIC3=738, Miscellaneous Business Services	SIC3=636, Title Insurance
General Dynamics Corp	Anteon International Corp	SIC3=373, Ship + Boat Building + Repairing	SIC3=737, Services-Computer Programming
H&R Block Inc	American Express Tax & Bus	SIC3=720, Services-Personal Services	SIC3=619, Finance Services
Hampshire Group Ltd	Kellwood Co-David Brooks Bus	SIC3=225, Knitting Mills	SIC3=233, Women's and Juniors Outerwear
Hewlett-Packard Co	Peregrine Systems Inc	SIC3=357, Computer + of fice Equipment	SIC3=737, Services-Computer Programming
Highland Hospitality Corp	Hilton Boston Back Bay Hotel	SIC3=679, Miscellaneous Investing	SIC3=701, Hotels + Motels
IRIS International Inc	Quidel Corp-Urinalysis Bus	SIC3=382, Laboratory Apparatus + Furniture	SIC3=283, Medicinal Chemicals + Botanicals
Journal Communications Inc	Emmis Comm-Radio Stations	SIC3=271, Newspapers: Publishing Printing	SIC3=483, Radio Broadcasting Stations
Kinder Morgan Energy Partners	Terra Nitrogen Co-Blytheville	SIC3=492, Natural Gas Transmission	SIC3=287, Agricultural Chemicals
LaSalle Hotel Properties	Hilton San Diego Resort	SIC3=679, Miscellaneous Investing	SIC3=701, Hotels + Motels
Leggett & Platt Inc	Foamex International-Rubber	SIC3=251, Household Furniture	SIC3=308, Miscellaneous Plastics Products
Lionbridge Technologies Inc	Bowne Global Solutions	SIC3=737, Services-Computer Programming	SIC3=275, Commercial Printing
MarkWest Energy Partners LP	Javelina Gas Processing	SIC3=131, Crude Petroleum + Natural Gas	SIC3=492, Natural Gas Transmission
McDATA Corp	Computer Network Technology	SIC3=366, Telephone + Telegraph Apparatus	SIC3=357, Computer + office Equipment
Nektar Therapeutics	AeroGen Inc	SIC3=283, Medicinal Chemicals + Botanicals	SIC3=384, Surgical + Medical Instruments
NRG Energy Inc	West Coast Power LLC	SIC3=491, Electric Services	SIC3=131, Crude Petroleum + Natural Gas
Pacific Energy Partners LP	Valero LP- Terminal & Pipeline	SIC3=461, Pipe Lines (No Natural Gas)	SIC3=291, Petroleum Refining
Polo Ralph Lauren Corp	Ralph Lauren Footwear Co Inc	SIC3=232, Men's + Boys Work Clothg	SIC3=302, Rubber + Plastics Footwear
Renaissance Learning Inc	AlphaSmart Inc	SIC3=737, Services-Computer Programming	SIC3=357, Computer + office Equipment
RH Donnelley Corp	Dex Media Inc	SIC3=731, Services-Advertising	SIC3=274, Miscellaneous Publishing
Stonemor Partners LP	Service Corp Intl-Cemeteries	SIC3=650, Real Estate	SIC3=720, Services-Personal Services
Sunstone Hotel Investors Inc	Renaissance Hotels-Hotel	SIC3=679, Miscellaneous Investing	SIC3=701, Hotels + Motels
Sunstone Hotel Investors Inc	Renaissance Washington DC	SIC3=679, Miscellaneous Investing	SIC3=701, Hotels + Motels
Titan Tire Corp	Goodyear Tire & Rubber-North	SIC3=331, Steel Works, Blast Furnaces + Mills	SIC3=301, Tires + Inner Tubes

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.

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 words patent, copyright, and trademark. Announcement returns are net of the CRSP value weighted index and are measured from the day preceding the announcement to the day of the announcement itself. 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 the 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	0.227	0.200	0.000	0.206	0.994	47,394
Panel B: Transaction Level Variables						
Target Ann. Return (-1,0)	0.069	0.191	-0.861	0.012	3.784	5,149
Acquirer Ann. Return (-1,0)	0.002	0.060	-0.619	0.000	0.828	5,274
Combined Firm Ann. Return (-1,0)	0.005	0.042	-0.335	0.002	0.755	4,937
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 words patent, copyright, and trademark.

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	-0.086	-0.145	-0.232	-0.435	-0.296	-0.016

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 words patent, copyright, and trademark. 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	Downstream Demand Shock	Obs
					Panel A: Acquirer Likelihood					
(1)	Acquirer?	2.142	-2.830		6.461		2.031	10.303		47,394
		(8.07)	(-7.06)		(20.41)		(5.65)	(34.77)		
(2)	Acquirer?	0.949			7.359		2.873	9.910		47,394
		(4.36)			(22.52)		(7.57)	(32.73)		
(3)	Acquirer?		-1.558		6.674		2.146	10.278		47,394
			(-4.97)		(20.72)		(5.90)	(35.09)		
(4)	Acquirer?			-1.428		5.431	1.487	10.544		47,394
				(-3.12)		(16.10)	(3.22)	(35.79)		
(5)	Acquirer?	1.857	-2.825	-1.776	5.047	3.030	2.320	10.481	1.451	47,394
		(7.66)	(-7.40)	(-4.78)	(15.15)	(10.70)	(6.31)	(34.54)	(2.74)	
					Panel B: Target Likelihood					
(6)	Target?	0.700	-1.012		3.754		2.015	8.346		47,394
		(3.59)	(-3.37)		(15.09)		(6.71)	(45.14)		
(7)	Target?	0.285			4.048		2.312	8.229		47,394
		(1.56)			(17.10)		(7.82)	(45.35)		
(8)	Target?		-0.585		3.824		2.037	8.335		47,394
			(-2.20)		(15.23)		(6.67)	(45.30)		
(9)	Target?			0.433		2.452	1.450	8.572		47,394
				(1.95)		(12.91)	(5.21)	(44.78)		
(10)	Target?	0.710	-0.948	0.314	3.179	1.101	2.115	8.275	0.225	47,394
		(3.58)	(-3.13)	(1.51)	(12.05)	(5.63)	(6.86)	(44.02)	(0.86)	

Table VII: 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 words), 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 words patent, copyright, and trademark. 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	16.3%	15.1%	13.8%	14.2%	15.1%	16.0%	12.3%	15.1%	17.8%
2	All Firms: Acquirer Incidence	30.6%	27.0%	23.3%	24.6%	27.0%	29.4%	24.0%	27.0%	30.0%
	Panel B: Large Firms									
3	Big Firms: Target Incidence	22.5%	21.3%	20.0%	19.4%	21.3%	23.1%	16.1%	21.3%	26.4%
4	Big Firms: Acquirer Incidence	41.2%	37.1%	33.0%	34.5%	37.1%	39.7%	31.5%	37.1%	42.8%
	Panel C: Small Firms									
5	Small Firms: Target Incidence	9.3%	8.9%	8.5%	8.9%	8.9%	8.9%	8.8%	8.9%	9.0%
6	Small Firms: Acquirer Incidence	22.4%	19.2%	16.0%	16.2%	19.2%	22.2%	18.2%	19.2%	20.2%
	Panel D: Mergers Only									
7	Mergers Only: Target Incidence	3.9%	4.3%	4.6%	3.9%	4.3%	4.6%	4.7%	4.3%	3.8%
8	Mergers Only: Acquirer Incidence	11.5%	10.4%	9.4%	9.3%	10.4%	11.6%	9.3%	10.4%	11.5%
	Panel E: Acquisition of Assets Only									
9	Acq Assets Only: Target Incidence	12.8%	10.8%	8.8%	10.1%	10.8%	11.5%	7.7%	10.8%	13.9%
10	Acq Assets Only: Acquirer Incidence	20.8%	17.7%	14.7%	16.4%	17.7%	19.1%	15.9%	17.7%	19.6%

Table VIII: Nested Logit and Target Selection

The table displays coefficients and marginal effects of nested logistic regressions where the dependent variable is a dummy indicating whether the given firm chooses among eleven options: not to merge at all, or merge with one of ten possible groups of firms in a given year. We run the model from the Acquirer's perspective (Row 1 both panels) and from the Target's perspective (Row 2 both panels). The nested logit has a tree structure in which the first node of the tree is the decision to merge or not to merge. If a firm does not merge, the tree ends. If the firm decides to merge, we consider ten groups of firms to merge with. The following discussion is based on acquiring firms in Row 1, but Row 2 is analogously constructed. In Panel A, we report coefficients and marginal effects for variables associated with the merge or do not merge decision (variables that that differ at the firm level, but not across target choices). In Panel B, we report coefficients and marginal effects associated with the terminal branch (variables that differ for each group of candidate target firms). For each candidate acquiring firm (henceforth CA), we group available targets into one of ten groups in each year. We first sort all candidate target firms into quintiles based on their similarity to the CA. Then we further divide each of the five groups into two groups using spreading sorts that generate dispersion in how close firms are to the CA's ten nearest neighbors. The independent variables are as follows. The Similarity of Group to Firm is the average similarity between the CA and the firms in each group. The Group Close to Rivals dummy is one if the spreading sort places the group as being close to CA's rivals, and zero if the group is far from CA's rivals. Product similarity (10 nearest) is the average similarity between the CA and its ten nearest neighbors (our measure of product market competition). We consider two cross terms involving this product similarity variable and the two above mentioned group variables. The remaining variables are properties of the CA. These include the fraction of nearest neighbors who were involved in restructuring transactions in the past year, and 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 words patent, copyright, and trademark. The sample is from 1997 to 2005. Robust t-statistics are adjusted for heteroskedasticity. Marginal effects are displayed below t-statistics and are displayed as percentages. Marginal effects for merge/dont merge measure the change in the predicted probability of merging when the given variable is increased by one standard deviation, with all other variables held at the mean. Marginal effects for substitute target measure (all) the change in the predicted conditional probability of merging with a specific group (conditional on having decided to merge) when the given group has the given independent variable increased by one standard deviation (averaged over the ten groups). We also report this specification for the top group (group 1C, which is the closest group to the acquirer, ie the group with the highest ex-ante predicted probability of merging).

Panel A: Inner Branch: Merge or do Not Merge

Row	Decision Variable	Product Similarity (All-10)	Product Similarity (10 Near)	Industry HHI (SIC-3)	Last Year 10 Nearest % Rest.	Last Year SIC-3 % Rest.	Last Year % Neighbor Patent Wds	Last Year Log Total Assets	Obs
(1)	Acquirer's Coeff.	–10.0177	0.5892	–2.9460	2.2456	0.7254	1.2542	0.3749	46,770
	t-stat	(–0.070)	(0.400)	(–1.700)	(1.990)	(4.870)	(6.450)	(16.970)	
	marginal effect	–0.3%	0.2%	0.4%	1.6%	0.6%	1.5%	6.3%	
(2)	Target's Coeff.	8.3992	0.4716	–0.0717	2.0832	0.5662	0.9920	0.3648	46,888
	t-stat	(0.050)	(0.230)	(–0.030)	(1.550)	(3.300)	(3.870)	(13.310)	
	marginal effect	0.2%	0.1%	0.0%	1.1%	0.3%	0.9%	4.6%	

Panel B: Terminal Branch, Who to Merge With

Row	Decision Variable	Similarity of Group to Firm	Simil. Grp. to Firm x Simm (10 near)	Group Close to Rivals	Group Close to Rivals x Simm (10 near)	Obs
(1)	Acquirer's Coeff.	1.1944	–1.1460	0.4834	–2.1314	467,700
	t-stat	(8.400)	(–1.350)	(4.200)	(–3.470)	
	marginal effect (substitute target, all)	16.1%	–3.1%	1.6%	–1.3%	
	marginal effect (substitute target, top group)	40.2%	–11.3%	6.0%	–5.0%	
(2)	Target's Coeff.	1.1113	–1.0768	0.5583	–2.5356	468,880
	t-stat	(6.880)	(–1.050)	(4.320)	(–3.780)	
	marginal effect (substitute acquirer, all)	15.9%	–3.0%	1.9%	–1.6%	
	marginal effect (substitute acquirer, top group)	39.1%	–10.6%	7.0%	–5.9%	

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 words patent, copyright, and trademark. 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.

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	0.018	-0.013			0.008	-0.001	-0.006	0.016	0.000	-0.003	0.021	-0.002	0.023	4,937
	(2.47)	(-1.82)			(2.02)	(-0.40)	(-2.03)	(1.52)	(0.12)	(-1.48)	(3.18)	(-6.52)		
(2) t=0 only			0.018	0.009	0.009	-0.000	-0.006	0.014	0.000	-0.003	0.021	-0.002	0.023	4,937
			(2.55)	(0.72)	(2.30)	(-0.28)	(-1.98)	(1.37)	(0.09)	(-1.51)	(3.20)	(-6.49)		
(3) t=-1 to t=0	0.016	-0.015			0.008	0.001	-0.007	0.015	0.001	-0.002	0.023	-0.003	0.023	4,937
	(1.81)	(-1.78)			(1.83)	(0.88)	(-1.85)	(1.29)	(0.52)	(-1.05)	(3.08)	(-6.58)		
(4) t=-1 to t=0			0.021	0.004	0.009	0.002	-0.006	0.013	0.001	-0.002	0.023	-0.003	0.024	4,937
			(2.65)	(0.28)	(2.12)	(1.00)	(-1.80)	(1.11)	(0.47)	(-1.02)	(3.10)	(-6.69)		
(5) t=-5 to t=0	0.033	-0.039			0.007	0.002	0.002	0.006	0.003	0.003	0.022	-0.004	0.025	4,937
	(2.40)	(-3.28)			(1.31)	(0.82)	(0.48)	(0.43)	(0.91)	(1.13)	(2.38)	(-6.76)		
(6) t=-5 to t=0			0.028	0.027	0.010	0.002	0.002	0.002	0.002	0.003	0.023	-0.004	0.024	4,937
			(2.14)	(1.43)	(1.95)	(0.75)	(0.52)	(0.15)	(0.84)	(0.94)	(2.39)	(-6.54)		
(7) t=-10 to t=0	0.042	-0.035			0.005	0.003	0.005	-0.000	-0.001	0.006	0.028	-0.004	0.019	4,937
	(2.47)	(-2.47)			(0.84)	(1.30)	(0.79)	(-0.01)	(-0.13)	(1.87)	(2.81)	(-5.01)		
(8) t=-10 to t=0			0.035	0.049	0.008	0.003	0.005	-0.003	-0.001	0.005	0.028	-0.003	0.019	4,937
			(2.10)	(2.30)	(1.37)	(1.13)	(0.80)	(-0.19)	(-0.17)	(1.63)	(2.81)	(-4.70)		

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 words patent, copyright, and trademark. 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	0.034			0.003	-0.004	-0.025	0.033	0.005	0.002	-0.004	0.000	0.007	4,451
		(1.37)			(0.33)	(-1.37)	(-2.17)	(1.09)	(1.14)	(0.51)	(-0.27)	(0.44)		
(2)	3 Year	0.065			-0.007	-0.009	-0.039	0.073	0.017	0.000	0.004	-0.002	0.019	4,451
		(2.24)			(-0.54)	(-2.32)	(-3.78)	(1.72)	(2.76)	(0.05)	(0.27)	(-1.33)		
(3)	1 Year		0.033	0.054	0.004	-0.005	-0.025	0.034	0.005	0.001	-0.004	0.001	0.008	4,451
			(1.81)	(2.29)	(0.40)	(-1.47)	(-2.16)	(1.14)	(1.12)	(0.27)	(-0.30)	(0.81)		
(4)	3 Year		0.048	0.081	-0.007	-0.010	-0.039	0.076	0.017	-0.001	0.004	-0.001	0.020	4,451
			(2.08)	(2.63)	(-0.55)	(-2.40)	(-3.81)	(1.79)	(2.77)	(-0.19)	(0.23)	(-0.94)		
						Panel B: Operating Income/sales								
(5)	1 Year	0.022			0.019	-0.006	-0.013	0.059	0.013	-0.001	-0.006	-0.003	0.008	4,451
		(0.46)			(1.55)	(-1.23)	(-1.11)	(1.56)	(1.88)	(-0.23)	(-0.36)	(-2.37)		
(6)	3 Year	0.046			0.013	-0.012	-0.042	0.074	0.026	-0.002	0.008	-0.006	0.017	4,451
		(0.88)			(0.62)	(-2.03)	(-2.92)	(1.36)	(2.66)	(-0.27)	(0.32)	(-3.12)		
(7)	1 Year		0.019	-0.001	0.018	-0.005	-0.013	0.060	0.013	-0.001	-0.006	-0.004	0.008	4,451
			(0.58)	(-0.03)	(1.41)	(-1.09)	(-1.09)	(1.65)	(1.91)	(-0.14)	(-0.37)	(-2.43)		
(8)	3 Year		0.012	0.030	0.011	-0.012	-0.042	0.078	0.026	-0.003	0.008	-0.006	0.016	4,451
			(0.30)	(0.59)	(0.53)	(-1.94)	(-2.93)	(1.42)	(2.69)	(-0.31)	(0.31)	(-3.10)		
						Panel C: Log Sales Growth								
(9)	1 Year	0.402			-0.014	0.008	-0.028	-0.193	0.068	0.014	0.048	-0.015	0.030	4,451
		(5.79)			(-0.37)	(0.69)	(-1.19)	(-2.11)	(3.78)	(1.02)	(1.17)	(-4.25)		
(10)	3 Year	0.684			-0.070	0.010	-0.022	-0.075	0.070	-0.005	-0.024	-0.014	0.019	4,451
		(4.97)			(-0.83)	(0.57)	(-0.41)	(-0.50)	(2.34)	(-0.20)	(-0.32)	(-2.63)		
(11)	1 Year		0.268	0.288	-0.026	0.008	-0.027	-0.173	0.070	0.013	0.044	-0.014	0.028	4,451
			(4.25)	(3.72)	(-0.65)	(0.72)	(-1.14)	(-1.92)	(3.85)	(0.94)	(1.09)	(-3.80)		
(12)	3 Year		0.452	0.463	-0.091	0.011	-0.020	-0.040	0.073	-0.006	-0.030	-0.012	0.016	4,451
			(3.86)	(3.37)	(-1.08)	(0.62)	(-0.38)	(-0.26)	(2.41)	(-0.26)	(-0.40)	(-2.24)		

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 words patent, copyright, and trademark. 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	0.594			-0.000	-0.019	0.111	-0.032	-0.050	-0.009	0.104	0.015	-0.274	0.117	3,898
		(3.47)			(-0.00)	(-0.99)	(2.06)	(-0.21)	(-1.39)	(-0.30)	(1.64)	(2.83)	(-11.72)		
(2)	2 Year	0.720			0.018	-0.017	0.095	0.025	-0.036	-0.012	0.092	0.018	-0.359	0.155	3,898
		(4.14)			(0.30)	(-0.71)	(2.10)	(0.13)	(-0.90)	(-0.35)	(1.14)	(2.96)	(-15.82)		
(3)	3 Year	0.721			0.129	-0.001	0.029	0.050	-0.024	-0.013	-0.012	0.015	-0.394	0.174	3,898
		(3.57)			(1.96)	(-0.06)	(0.64)	(0.24)	(-0.60)	(-0.34)	(-0.12)	(2.65)	(-16.87)		
(4)	1 Year		0.202	0.873	-0.009	-0.029	0.109	0.018	-0.046	-0.025	0.096	0.019	-0.272	0.119	3,898
			(1.28)	(4.62)	(-0.16)	(-1.51)	(2.02)	(0.12)	(-1.29)	(-0.78)	(1.54)	(3.43)	(-11.81)		
(5)	2 Year		0.256	0.930	0.004	-0.027	0.093	0.085	-0.030	-0.027	0.082	0.022	-0.356	0.156	3,898
			(1.61)	(4.47)	(0.06)	(-1.13)	(2.09)	(0.45)	(-0.77)	(-0.77)	(1.02)	(3.47)	(-15.83)		
(6)	3 Year		0.169	0.805	0.109	-0.010	0.026	0.119	-0.018	-0.025	-0.022	0.018	-0.386	0.174	3,898
			(0.99)	(3.63)	(1.67)	(-0.41)	(0.60)	(0.58)	(-0.43)	(-0.68)	(-0.23)	(3.14)	(-16.73)		

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 words), 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 words patent, copyright, and trademark. 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%	0.7%	0.3%	0.5%	0.7%
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%	-0.6%	-0.5%	-0.5%
4	Δ OI/Assets: 3 Year (A)	-2.2%	-1.6%	-1.0%	-1.4%	-1.6%	-1.8%
5	Δ OI/Sales: 1 Year (A)	-0.7%	-0.5%	-0.2%	-0.9%	-0.5%	0.0%
8	Δ OI/Sales: 3 Year (A)	-2.4%	-2.0%	-1.5%	-2.3%	-2.0%	-1.7%
7	Sales Growth: 1 Year (A)	12.0%	15.9%	19.8%	16.2%	15.9%	15.5%
8	Sales Growth: 3 Year (A)	20.3%	27.0%	33.6%	28.7%	27.0%	25.2%
	Panel C: Growth in Product Descriptions (Based on models in Table XI)						
9	Prod Desc Growth: 1 Year (A)	-2.2%	3.3%	8.9%	3.3%	3.3%	3.3%
10	Prod Desc Growth: 3 Year (A)	-3.4%	3.3%	10.1%	0.1%	3.3%	6.6%