Product Integration and Merger Success

Gerard Hoberg and Gordon Phillips*

May 8, 2016

ABSTRACT

We examine the importance of merger integration risk to the outcomes of mergers and acquisitions using new product-based ex ante measures of integration risk at the firm and firm-pair level. Our ex ante measures are significantly associated with ex post statements by managers in their 10-K indicating difficulties with merger and acquisition integration and also employee retention issues. We find that firms performing mergers and acquisitions in high product integration risk markets experience lower ex post profitability, higher ex post expenses, and a higher propensity to divest assets. Upon announcement, acquirers experience lower announcement returns and targets experience significantly higher announcement returns when product integration risk is high. Examining long-term stock market returns, we find that the well-known anomaly that acquiring firms have lower longer-term stock returns primarily occurs in markets where product integration risk is high.

*University of Southern California, and Tuck School at Dartmouth and National Bureau of Economic Research, respectively. Hoberg can be reached at hoberg@marshall.usc.edu and Phillips can be reached at gordon.m.phillips@tuck.dartmouth.edu. We thank Christopher Ball for providing us with access to the metaHeuristica database. We also thank seminar participants at the University of Amsterdam, Hong Kong Polytechnic University, Stockholm School of Economics, Tilburg University, Tsinghua University and the University of Southern California for helpful comments. All errors are the authors alone.

I Introduction

Participants engaging in mergers frequently claim that merger integration problems are a major reason why many mergers do not succeed. A recent survey of more than 800 executives by McGee, Thomas, and Thomson (2015) cites different cultures and difficulty of integrating product lines as partially being responsible for worse ex post merger outcomes and a lower chance of achieving merger synergies. Ahern, Daminelli, and Fracassi (2015) examine international mergers and find that country-level cultural difference in trust and individualism lead to lower merger volumes and lower combined abnormal announcement returns. Yet, currently there is only limited evidence other than case studies[1] that problems with product and firm integration are important for merger outcomes at the deal level within countries. It is not just a lack of resources to implement merger integration that causes many mergers to fail. In fact, Harford (1999) shows that acquisitions by cash rich acquirers are often followed by declines in operating performance.

We define merger integration risk as the possibility that there will be value loss from attempting to coordinate activities and product line offerings to achieve synergies from previously separate organizations. Rhodes-Kropf and Robinson (2008) model asset complementarity and synergies as a motive for mergers but do not consider the problems and risks associated with achieving these synergies. Bena and Li (2014) show that innovation increases for targets and acquirers that have similar technological links from patents - evidence consistent with ex post innovation synergies. Hoberg and Phillips (2010) establish that similar targets and acquirers have higher ex post cash flows and more new product introductions. However, despite this evidence of ex post merger gains, we do not know what factors give rise to risks of potentially not fully achieving the synergies that managers frequently cite as the rationale for mergers and acquisitions.

We focus on measuring the ex ante difficulty of integrating product lines across organizations at the firm level for mergers within the U.S. We use text-based analysis of business descriptions in firm 10-Ks to measure this quantity, which measures the

[1] Arnold (1983) examines 5 cases studies of merger integration and Epstein (2004) examines the merger of J.P. Morgan and Chase Manhattan Bank.

extent to which merging firms will face challenges integrating their various product lines in the post merger firm. Although the concept of product integration risk might seem narrow relative to a more classic view of integration risk, which is often linked to employees leaving the firm due to difficult work environments and problems of integration of different firm cultures, for example, we propose that these issues are deeply linked. That is, the more there is ex ante difficulty of integrating merging firms' products, the more employees of the two firms will have to work together and thus the more important are employee retention and culture issues for these deals.

Our empirical results support this proposition. When our ex ante measures of product integration risk are high, we observe a higher ex post incidence of managers discussing both integration difficulties and employee retention. These findings are consistent with product integration risk translating to increased likelihood of unexpected drains on managerial time and in retaining employees.

Examining outcomes after mergers and acquisitions, we also find evidence that ex ante product integration risk is associated with lower operating income post-merger and higher ex post SG&A/sales, which specifically relates to the cost of managing the firm's employees and organizations. We also find evidence that mergers and acquisitions with higher ex ante product integration risk experience higher ex post asset divestitures. These findings illustrate the importance of product integration risk and its real impact on acquiring firms.

One example of managers discussing integration difficulties in their 10-K (in a different section than the product description section) is Integrated Health Services in 1997:

> *"IHS has recently completed several major acquisitions, ..., and is still in the process of integrating those acquired businesses. The IHS Board of Directors and senior management of IHS face a significant challenge in their efforts to integrate the acquired businesses, including First American, RoTech, CCA, the Coram Lithotripsy Division and the facilities and other businesses acquired from HealthSouth. The dedication of management resources to such integration may detract attention from the day-to-day business of IHS. The difficulties of integration may be increased by the necessity of coordinating geographically separated organizations, integrating personnel with disparate business backgrounds and combining different corporate cultures."*

In all, we find that over 19% of all firms in our sample make ex post statements like the one above in their 10-K. Such statements typically appear in sections of the 10-K other than the business description (for example in the MD&A or in the discussion of risk factors). We view such statements as an indicator of ex post integration difficulties, and the existence of such statements allows us to assess the validity of our ex ante measures of integration risk. We note that measuring integration risk ex ante is far more difficult than identifying cases of failure ex post. For example, it is perhaps not clear to managers themselves how risky a transaction truly is, and the post-merger firm is not observable ex ante, making it difficult to forecast the difficulties that might arise.

We measure ex ante product integration risk using individual words and the paragraph structure of the product market descriptions (in the business description section) of firm 10-Ks. We define a perfectly integrated word as one that is equally likely to appear in any paragraph in the given firm's 10-K business description. This atomistic word-level approach allows us to view any real or hypothetical firm as a collection of building blocks (words). A firm is thus in a market that requires extensive product integration if the words the firm uses in its business descriptions appear uniformly integrated across the paragraphs in this business summary. This approach allows us to compute levels of integration for individual firms, for hypothetical counterfactual firms, and even for hypothetical post-merger firms that do not yet exist. For example, we can compute integration levels for the target, the acquirer, and the part of the post merger firm that reflects newly anticipated product market synergies.

The intuition behind this approach can be seen if we consider the following generative process for business descriptions after a merger. Suppose that the instantaneous effect of merging two firms together (without any initial integration) can be characterized by simply appending the text of the target's business description to that of the acquirer. At this point, the text associated with both firms, while in the same document, is disjoint and unintegrated. As the firm proceeds to integrates, the product text from the two parts then becomes mixed. As a result, words from the target's vocabulary effectively move in the document into the paragraphs that previously just discussed the acquirer's products (and vice-a-versa). When this is

successfully achieved, the result is an integrated firm.

An example of an unintegrated firm is Harris Teeter, a firm operating in the grocery business. Unlike Apple, whose products share many features that were deliberately built into the products as the firm evolved, such is not typical in the grocery business, where goods are purchased from producers with little or no modification by Harris Teeter itself. As a result, its expected baseline level of integration in its business description is likely to be low. Such a firm faces less risk of integration failure because its products and lines of business are easier to separate.

Our first finding regarding outcomes is that proposed mergers and acquisitions are more likely to be withdrawn when the ex ante gap between expected integration and realized firm integration is high. Moreover, both sides of the gap calculation matter: deals are less likely to be withdrawn when ex ante realized firm integration is high, and are more likely to be withdrawn when expected integration of rival firms is high. This test supports the hypothesis that many deals are canceled when parties raise opposition to them. These results also support the conclusion that our measures indeed capture ex-ante integration risks.

For firms that do complete the announced deal, we observe lower ex post profits and higher selling and general administration (SG&A) expenses when the acquirer is ex ante less integrated and has a higher integration gap. These results are consistent with the acquiring firm having to spend additional resources and compensate employees to integrate the firms. We also document that ex ante product integration risk is associated with a higher rate of ex post divestiture of assets, consistent with difficulties in integrating firms with high ex ante product integration risk.

We find that acquirers have modestly negative announcement returns and targets have large positive announcement returns when expected product integration risk from potential product synergies is high. We find that integration risk relating specifically to synergies is most responsible for these announcement returns, and to subsequent negative real outcomes. This conclusion is based on using the integration properties word-by-word and by considering word-pair combinations that only exist in post-merger firms but not in pre-merger targets or acquirers. Our results are

consistent with targets receiving high announcement returns when integration risk is high to compensate agents affiliated with the target for the taking on the risk and providing the requisite effort to successfully integrate the firms.

Examining stock market longer-term outcomes, we show that ex post negative stock returns to acquiring firms can be explained by product integration risk and that the well-known anomaly of negative stock returns to acquiring firms only exists in the subsample of mergers and acquisitions where integration risk is high. These results are also robust to controlling for product similarity as measured by Hoberg and Phillips (2010), which captures potential synergies between merging firms. We conclude that product integration risk is distinct and separate from measures of product similarity.

Our paper adds to previous research on mergers which examines ex post outcomes after mergers. Healy, Palepu, and Ruback (1992) and Andrade, Mitchell, and Stafford (2001) document increases in industry-adjusted cash flows following mergers. Maksimovic and Phillips (2001) document increases in productivity after mergers that are related to demand shocks and acquirer skill. Rhodes-Kropf and Robinson (2008) model asset complementarity and synergies as a motive for mergers. Bena and Li (2014) and Hoberg and Phillips (2010) document evidence of synergies post merger, showing that there are increases in cash flows, new products and patents post merger that are related to ex ante similarity of acquirer and target.

However, these studies do not shed light on the difficulties of merger integration even for related firms. Our paper measures and captures merger integration risk that results from product integration. We directly show that product integration risk is related to merger success in a domestic context. This adds product integration risk to the list of international cultural integration risk factors that have been shown to impact mergers documented in Ahern, Daminelli, and Fracassi (2015).

The rest of this paper is organized as follows. Section II discusses our data and method for measure product integration risk. Section III provides tests which validate that our ex ante measure of product integration risk is correlated with ex post managerial discussions of problems with merger integration and employee retention.

Section IV provides our tests examining the relation between ex ante measures of product integration risk and M&A announcement returns. Section V examines ex post real outcomes and section VII examines ex post stock returns. Section VII concludes.

II Firm Integration and Transaction Integration Risk

A key objective of the methods used in our paper is to examine ex ante expected levels of integration failure risk for any candidate merger pair (even if the target and the acquirer have not yet merged). This presents two challenges. First, we do not observe the post-merger firm until later, and we have to rely on ex ante available information. Second, a post merger firm is more than the sum of its parts. Generally, a post merger firm has three parts: acquirer assets in place, target assets in place, and synergies and assets created from the business combination. Ideally, our measures of ex ante integration risk will be capable of assessing integration risks for each component. We predict that integration risks are more salient for the synergy component than for the assets in place. In particular, synergies likely draw strongly on the product market expertise of both firms and thus are more dependent on integration before they can be realized.

Our initial methodology is based on measuring the ex ante integration risk associated with each existing firm's assets in place. This can be computed for all public firms, even those not involved in a transaction. We then extend our methodology to compute the ex ante integration risk of firms involved in transactions. This approach can separately assess assets in place and potential synergies of the transacting firms. This flexibility is achieved by first defining the concept of integration at the atomistic word or word-pair level, and then by computing integration risk for any firm (or parts of a transacting firm) by averaging the integration of its atomistic parts (the words associated with each part's business description vocabulary). This general framework not only allows us to explore integration specifically for merger transactions as in the current research, but it also provides a foundation for computing integration

risk in other corporate settings. Examples of such future research might include divestitures, IPOs, new ventures, or even proposed early-stage business plans that can benefit from pre-implementation ex-ante measurable information on integration. In all cases, the integration properties of each such project can be computed by linking each project's product market text to the word-specific integration risk scores computed from the general population of public firm 10-Ks.

Before explaining the specific calculations used for measuring integration risk, we first discuss the conceptual foundation for the empirical measures. Our measures capture three different concepts: 1.) Firm realized integration, 2.) Firm expected integration and 3.) Transaction or synergy integration risk.

A The Integration Gap: Expected versus Actual Integration

Central to our analysis is the ability to measure a firm's level of integration success relative to a strong counterfactual or benchmark. A key issue is that, in some product markets such as agriculture, overall integration levels are low. In this setting, a firm that achieves an average level of realized integration relative to economy-wide averages can be viewed as quite successful. In contrast, in markets where integration levels are high, such as medical devices and services, a firm that achieves an average level of integration relative to economy-wide averages can be viewed as a laggard given expectations should be higher in such markets. This issue is particularly important when we assess longer-term integration success.

We assess each firm's integration success by comparing its realized integration to an appropriate counterfactual level of expected integration. We define a firm's "integration gap" as the difference between a firm's expected integration and its realized integration as follows (with the specific formulas and methods given in the next section):

$$\text{Integration Gap}_{i,t} = \text{Expected Integration}_{i,t} - \text{Actual Integration}_{i,t} \qquad (1)$$

A firm with a high integration gap has a realized level of integration that is low relative to its expected counterfactual level of expected integration. We might expect

that such firms are failing to fully integrate their acquired product offerings, and are thus more likely to experience negative outcomes when they acquire. In particular, firms with a larger integration gap might realize lower profits, higher administrative costs in the form of SG&A, higher rates of ex post divestiture, and lower ex post stock returns if they acquire other firms.

Figure 1 provides four illustrative examples of firm realized integration levels over time (with the specifics of how we calculate these integration levels in the next section). For example, Apple's integration was quite low in the 1990s when its most noteworthy product offering was the Macintosh computer. However, as Apple launched a number of innovative new products in the 2000s, its integration soared to become one of the highest in the economy by the middle to the end of our sample. This suggests that Apple's new products are not only innovative, but also well-integrated as they share many common features, presumably relating to internet, software, casings, screen technology and other aspects. The figure also shows that Microsoft, in contrast, has experienced a more gradual ascent from a relatively disintegrated firm to a firm that has above median integration. The figure also illustrates two examples of firms with pervasively low levels of integration: Harris Teeter (grocery) and Black and Decker (hand tools). The products offered by these firms are likely more modular and are thus less integrated.

[Insert Figure 1 Here]

B Synergy Integration Risk

To compute synergy integration risk, we consider a framework that relies on the expected integration of word-pairs rather than integration levels of singleton words. The need to use word-pairs is best illustrated in Figure 2. The figure illustrates that every pair of words in the acquirer and target vocabularies falls into one of two categories. The shaded regions of the grid indicate word-pairs that already exist in the the union of the target and acquirer firms before the transaction, and that will thus continue to exist in the post-merger firm. These word-pairs represent assets in place. The unshaded word-pairs, however, do not exist in the pre-merger firms, but do exist in the combined post-merger firm. These unshaded word-pairs thus

represent the synergies of the merger. By dividing word-pairs into these groups, we can separately measure expected levels of integration for both assets in place and synergies.

[Insert Figure 2 Here]

Figure 2 also illustrates why it is necessary to use word-pairs to specifically identify the synergy (unshaded) word-pairs. For example, all singleton words exist in the pre-merger firms and in the post merger firms, and hence a researcher cannot separately identify synergy integration risk using singletons. A method based on word-pairs that we describe in the next section is thus necessary.

C Integration at the Word-Pair Level

Unlike most studies using linguistic analysis in business disciplines, we start with the assumption that paragraph boundaries are meaningful and we thus consider a document to be a collection of ideas delineated by paragraph boundaries. Empirically, we begin by representing each paragraph as a collection word-pairs. A word-pair is a permutation of any pair of words that exist in the given paragraph. For example, a paragraph containing 4 unique words (A, B, C, D) has 6 unique word-pairs (AB, AC, AD, BC, BD, CD). Note that although paragraph boundaries are important, the order of words within each paragraph is not important. Word-pairs defined this way also have frequencies attached to them, and in the previous example the frequency of each pair is one. If instead the 4 words in the paragraph were (A, B, B, C), then we have three unique word-pairs (AB, AC, BC) and they respectively have frequencies (2, 1, 2). We consider the atomistic unit of observation to be a word-pair instead of singleton words because (as we explain in the next section). This allows us to identify vocabularies associated with merger synergies of a candidate merger ex-ante, before we even observe the post-merger firm's 10-K.

Our goal is to compute an empirical measure of integration for each word-pair in the universe of all word-pairs. We analyze word-pairs from firm 10-K business descriptions filed with the SEC. In the business description section, firms discuss the products they offer as part of their operating business. We do not use all words

in this section as we exclude word-pairs from the analysis if they are (A) based on common words such as "the" and "and" or (B) are are too infrequent to be informative. Following Hoberg and Phillips (2015), we define common words as those used in more than 25% of all 10-K business descriptions in the given year. We define infrequent word-pairs at the document level as word-pairs that are used less than three times in a given 10-K. A word is in the overall universe of word-pairs if it is used at least three times in any firm's 10-K business description in the given year.

Once we code paragraphs based on their universe of observed word-pairs and their frequencies, our second step is to compute the degree to which each word-pair is integrated within the firm's 10-K overall. We define a word-pair k as integrated within a given firm i's 10-K business description if it appears in the text in a way that is uniformly mixed across paragraphs. Formally, we measure this as the similarity in distributions between word-pair k's appearance across paragraphs and the distribution of paragraph total word-pair counts (paragraph lengths). Intuitively, the empirical distribution of overall paragraph lengths is also the distribution that a randomly chosen word-pair will appear in each of the given paragraphs. A "fully integrated" word-pair is one that has an observed distributional similarity to the distribution of paragraph lengths of one (identically distributed), indicating it is perfectly mixed across the paragraphs in a proportional way, and thus it is fully integrated within the firm as a whole. A perfectly non-integrated word-pair would have a distributional proximity of zero. For example, the word-pair (internet, software) is likely integrated for firms like Microsoft and Apple, as these word-pairs are likely used in a more uniform way across their 10-K business descriptions. Visual examples of distributions of integrated and non-integrated word-pairs are depicted in Figure 3.

[Insert Figure 3 Here]

III Methodology and Data

A Methodology: Measuring Integration Using Words

We now discuss in more detail how we use individual words and word pairs to measure integration. Consider a firm i that has a business description with N_i paragraphs. Further let L_i denote the number of word-pairs in each paragraph. We then define the baseline aggregate-frequency distribution of paragraph word-pair lengths, which represents full integration, as the following N_i-vector $D_{i,full}$ (where $\mathbb{1}$ is a vector of ones):

$$D_{i,full} = \frac{L_i}{L_i \cdot \mathbb{1}}. \tag{2}$$

Let k denote a given word-pair and let $D_{i,k}$ denote the N_i-vector distribution of word-pair k's usage in the N_i paragraphs for firm i. For example, a word-pair that appears in just one paragraph would have a vector $D_{k,i}$ that is zero in all elements and one in the row corresponding to that paragraph. A firm that uses a word-pair twice in one paragraph and once in another would have a vector $D_{i,k}$ that contains all zeros, except one element would contain two-thirds and one element would contain one third.

We thus define word k's realized integration for firm i ($IW_{i,k}$) as the distributional proximity of word-pair k's usage to firm i's aggregate usage distribution of word-pair paragraph lengths:

$$IW_{i,k} = \frac{D_{i,k}}{||D_{i,k}||} \cdot \frac{D_{i,full}}{||D_{i,full}||} \tag{3}$$

We note that $IW_{i,k}$ can be computed fully from firm i's 10-K. We thus define this construct as a measure of "realized integration", as it is the observed level of integration for word-pair k in firm i's 10-K in the given year (note that all variables in this section have an implied t subscript for the given year, which we omit for parsimony).

In addition to realized integration levels, we also compute levels of benchmark "expected integration" for each word-pair k and firm i. This is done by simply computing the average of $IW_{j,k}$ across all firms j such that $j \neq i$ such that firm j uses word-pair k in its 10-K. Expected integration is thus a quantity that is also unique for each firm i and word-pair k, and we denote expected integration as $\overline{IW_{i,k}}$

and realized integration as $IW_{i,k}$. Expected integration indicates the extent to which word-pair k normally appears as an integrated word-pair across firms in the economy that use word-pair k. Therefore, it serves as a natural benchmark to which realized integration can be compared. For example, we propose that a given firm has an integration shortfall if the word-pairs it uses generally have low levels of realized integration and high levels of expected integration. This concept will be important when we later introduce firm-level measures.

We also note that the above calculations are general, and can be computed based on individual words, or based on word-pairs. For example, each individual word that exists in at least one paragraph in a 10-K thus also has a distribution $D_{i,k}$ as described above representing its appearance across paragraphs in firm i's 10-K. Similarly, the aggregate usage distribution $D_{i,full}$ can be computed based on the total word counts used in the N_i paragraphs. Importantly, we use word-pairs as the basis of our calculation of synergy integration risk as this concept relies on pairs of words that are in the combined target and acquirer, but not the pre-merger target and acquirer.

B Measuring Firm-level Integration

We now describe how we compute firm-level actual and expected integration levels for any firm in isolation, regardless of whether the given firm is experiencing or has experienced a transaction. Because there are no synergies to assess in this simpler environment, we base these calculations on singleton word distributions instead of on word-pair distributions as above, as this approach for stand alone firms is more parsimonious.

We first define I_i for firm i as a Q-vector where each element k contains each word's level of realized integration $IW_{i,k}$, which we defined in equation (3) (but now based on singleton words instead of word pairs). Q denotes the number of unique words in the sample of all firms in a given year. Firm realized integration is then computed by averaging the realized integration of the words the firm uses. Firm i's

realized integration is thus defined using the cosine similarity method as follows:

$$Actual\ Integration_i = \frac{V_i}{||V_i||} \cdot \frac{I_i}{||I_i||} \tag{4}$$

where V_i is a Q-vector that contains the number of times each word k is used by firm i in its overall business description section of its 10-K.

We next consider firm "expected integration", which is computed in a parallel fashion as realized integration, except that it is based on expected word-level integration ($\overline{IW_{i,k}}$) instead of realized word-level integration ($IW_{i,k}$). We thus define $\overline{I_i}$ for firm i as a Q-vector where each element k contains each word's level of expected integration $\overline{IW_{i,k}}$ (as defined in the previous section). Firm expected integration is thus the average expected integration of the words the firm uses as follows:

$$Expected\ Integration_i = \frac{V_i}{||V_i||} \cdot \frac{\overline{I_i}}{||\overline{I_i}||} \tag{5}$$

We emphasize that both realized and expected integration are not highly correlated with measures of similarity or competitiveness such as those used in Hoberg and Phillips (2015). This is by design, as the concept of integration has a different foundation than does competitiveness or the concept of across-firm relatedness. In particular, firm integration is a property of the paragraph structure and its distributional properties *within* a firm (measuring the degree to which words are mixed), and is not a property of how similar a firm's disclosure is to other firms.

From the expected and actual integration levels, we then can compute a firm's integration gap as:

$$Integration\ Gap_{i,t} = Expected\ Integration_{i,t} - Actual\ Integration_{i,t} \tag{6}$$

C Measuring Synergy Integration Risk

To measure synergy integration risk on actual or proposed merger transactions, we move to word pairs and consider the following triple product calculation:

$$PairInt_{i,j} = \frac{V_i}{||V_i||}' \cdot I_{pairs} \cdot \frac{V_j}{||V_j||}, \tag{7}$$

where V_i is a Q-vector that contains the number of times each word-pair k is used by a firm i overall in paragraphs in its 10-K, and Q denotes the number of unique word-pairs in the sample of all firms in a given year.

The matrix I_{pairs} is an NxN matrix of expected integration levels for each pair of words. Each data point in this matrix is computed using the expression in equation (3) above applied to word-pairs instead of singleton words. Because it is based on all word-pairs from both the target and acquirer (and the synergies), we can compute the variable $PairInt_{i,j}$ for a given acquirer and target pair (denoted by subscripts "a" and "t" respectively) to obtain the combined firm integration risk variable as follows:

$$Post\ Merger\ Integration\ Risk = PairInt_{a,t} \tag{8}$$

We can also identify the integration risk inherent to the synergies in isolation by subtracting out the integration risk from the acquirer and target assets in place.

$$Synergy\ Integration\ Risk = PairInt_{a,t} - Expected\ Integration_a - Expected\ Integration_t \tag{9}$$

Here, the expected integration levels of the stand-alone acquirer and target are firm-level variables, as described earlier. The synergy integration risk thus subtracts the expected levels of integration due to assets in place from the overall level of expected integration for the post-merger firm as a whole. Because this calculation does not require any transactional data, we also note that this calculation is general. In particular, we are not only able to compute combined firm and synergy integration risk for announced mergers, but also for transactions that are hypothetical or not-yet announced.

D Data

We use the Compustat sample of firm-years from 1996 to 2008. We then identify, extract, and parse machine readable 10-K annual firm business descriptions from the SEC Edgar database. We thus require that firms have machine readable filings of the following types on the SEC Edgar database: "10-K," "10-K405," "10-KSB,"

or "10-KSB40." These 10-Ks are merged with the Compustat database using using the central index key (CIK) mapping to gvkey provided in the WRDS SEC Analytics package. These minimum criteria leave us with a baseline panel database of 81,982 observations in our merged Compustat/Edgar universe. Following Hoberg and Phillips (2015), we only consider words that are nouns or proper nouns, and we only include words that appear in no more than 25% of all 10-Ks in the given year. We also drop any words that appear in less than three 10-Ks to reduce the size of our underlying data matrices and because these words are not highly informative about integration due to their scarcity.

We also consider metaHeuristica, which is also based on 10-K data and is merged to Compustat using the same procedure. Data from metaHeuristica is specifically used to identify managerial mentions of integration difficulties and employee retention issues in the 10-K, which we discuss more in the next section.

We identify merger and acquisition of asset transactions using SDC Platinum. We obtain 52,400 announced transactions where the acquirer is in our merged Compustat/Edgar universe and 13,380 announced transactions in which the target is in this universe. We use this full sample to examine stock returns and long-term real outcomes following acquisition transactions. We also identify a subsample of 5,981 of these transactions for which we have machine readable 10-Ks available for both the target and the acquirer. We use this subsample to examine announcement returns.

We use the CRSP database for two purposes. First, we use the daily return tapes to compute the announcement returns for both targets and acquirers. Second, we use the monthly CRSP return tapes to construct a database of monthly stock returns that we use to test our predictions regarding the negative ex post acquirer stock return anomaly. After merging the monthly stock return database our with the standard Davis, Fama, and French (2000) and momentum controls, and our merged Compustat/Edgar universe, we are left with 626,839 monthly stock return observations from July 1997 to June of 2010.

IV Statistics and Validation

Table I displays the summary statistics for the key variables considered in our study. Panel A reports summary statistics for firm-level variables based on 10-K business descriptions and also for control variables. Although the mean values for realized and expected integration do not have a simple interpretation, the table shows that both variables have similar means and also comparable standard deviations. Hence it is not suprising that their difference, the integration gap, has a mean that is closer to zero and a distribution that spans both negative values (for firms whose realized integration is low relative to expectations) and also positive values.

Panel B of Table I reports the mean value of the dummy variables we compute based on verbal statements in the 10-K indicating integration difficulties surrounding mergers (integration challenges dummy) and employee retention issues surrounding mergers (employee retention dummy). We explain the construction of these variables in the next Section. Here we note that 19.4% of firms in our sample disclose direct statements indicating concerns about risks of failed merger integration, and 11.4% disclose statements indicating employee retention issues surrounding acquisition transactions. These results indicate that integration risk is salient for a large number of firms in our sample, as they discuss this issue directly in their 10-K. Finally, Panel C reports the summary statistics for our variables based on real outcomes, including profitability, SG&A expenses, and post merger rates of divestiture and acquisition.

[Insert Table I Here]

Table II displays the Pearson correlation coefficients. The table shows that, not surprisingly, realized and expected levels of integration are strongly positively correlated at 70.8%. This indicates that when firms operate in markets where high integration is the norm, they usually are able to generate a realized level of integration that is also quite high. However, there is also material differences in the information in these variables. For example, realized integration is lower for larger and older firms, and also for firms facing more competition in the form of total product similarity. In contrast, expected levels of integration do not strongly correlate with these variables.

We also consider the integration gap, which is the difference between expected and realized integration. A high value indicates that a firm's realized integration is low relative to its benchmark, which in turn should be an indicator of integration failure following a merger. In rows (5) and (6), we thus report correlations between our key variables and dummy variables indicating whether managers directly indicate challenges with merger integration in their 10-K (these variables are formally explained in the next section). We find that the integration gap, as we would predict, most strongly correlates with these variables. In particular, when a firm's level of integration is low relative to its benchmark, managers are more likely to report that the firm is facing difficulties in integrating its business lines following a merger. The results also suggest that the integration gap is most strongly with managerial statements about challenges regarding employee retention, a matter that is also fundamentally related to integration challenges.

[Insert Table II Here]

Table III displays sample industries based on the Fama-French 12 classification and average levels of realized and expected integration.[2] We report results both in the first year of our study (1997) and the last year (2008). The results suggest that for many of these broad industry classifications, that average realized integration is in a rather narrow band between 1.8 and 2.0. The health industry has materially higher average levels of integration at 2.07, and the finance industry has materially lower integration at 1.75. Comparing realized to expected integration, we observe similar patterns. Also, comparing 2008 in Panel B to 1997 in Panel A, we only observe modest shifts in the industry rankings. However, one notable difference in the panels is that average levels of integration have increased materially during our sample and are notably higher in 2008. This finding is consistent with firms building integration capacity over time, which likely adds to firm value, and which might also have product market benefits potentially in the form of better product offerings and also in the form of barriers to entry.

[Insert Table III Here]

[2]We thank Ken French for providing classification data on his website.

A Managerial Mentions of Integration Difficulties

We also use 10-K text to identify instances where managers explicitly indicate that they are facing difficulties with merger integration, and also instances where they are facing challenges with employee retention issues. We use these measures primarily for validation of our aforementioned measures of ex-ante integration risk based on business descriptions. We also use these managerial mention measures to further illustrate the importance of integration to managers. For example, we will show that 19.1% of firm 10-Ks contain a direct statement about integration challenges, and moreover, these statements are detailed and specific, and hence are not boilerplate.

To identify managerial mentions, we use the metaHeuristica software package and run queries on the entire 10-K - thus we use content in 10-Ks that is distinct from the firm's business description (which we use to construct our aforementioned measures of ex ante integration risk). The majority of managerial mentions relating to integration challenges are in the managerial discussion and analysis (MD&A) and risk factor sections of the 10-K. Our objective is to use the results of this query for validation, and in particular, to examine if our ex ante measures of integration risk based on product descriptions indeed predict ex post instances of managers explicitly complaining about integration difficulties. Strong evidence regarding this prediction would mitigate concerns that our ex-ante measures based on distributional mixture and product market vocabulary primitives are measuring something other than integration.

In order to identify firms that complain about integration difficulties, we run a metaHeuristica query requiring that one word from each of the following three buckets must all jointly appear in a paragraph. We use word buckets that contain an array of synonyms because there is a number of ways to express to a reader that the firm is experiencing integration difficulties. We identified the synonyms to use in these queries using the sentence tree views in metaHeuristica following Hoberg and Maksimovic (2015).

Integration Difficulty List 1: merger OR mergers OR merged OR acquisition OR acquisitions OR acquired

Integration Difficulty List 2: integration OR integrate OR integrating

Integration Difficulty List 3: challenge OR challenging OR difficulties OR difficulty OR inability OR failure OR unsuccessful OR substantial expense

If a given firm has a hit on this query, we define the "Integration Failure Dummy" to be one. We also compuate an "Integration Failure Intensity" variable as the total number of words in the paragraphs of firms that hit on this query.

We run a separate query also based on three word buckets to identify whether a firm is experiencing issues relating to employee retention. The buckets are as follows:

Employee Retention List 1: merger OR mergers OR merged OR acquisition OR acquisitions OR acquired

Employee Retention List 2: employee OR employees OR personnel

Employee Retention List 3: retention OR departure OR departures

If a given firm has a hit on this query, we define the "Employee Retention Dummy" to be one. We also compute an analogous variable "Employee Retention Intensity" based on word counts.

We identify all firms with these discussions of merger integration problems and employee retention issues and then create the resulting dummy and continuous intensity measures based on these mentions for each firm.

Table IV presents examples of the first ten paragraphs returned from metaHeuristica in 1997 that hit on our verbal query intended to measure managerial mentions of integration difficulties, where we query metaHeuristica using the word list searches discussed above. The identification of a relevant paragraph requires that at least one word from each of the three integration difficulty buckets discussing acquisitions and

integration problems appears in a paragraph. The examples clearly indicate specific mergers being discussed and integration problems with these mergers. We also note that these discussions appear ex post, after the acquisitions have taken place.

[Insert Table IV Here]

Table V shows similar examples where we use the text searches to identify employee retention issues discussed in the context of mergers. Quotes include statements like "Such merger-related costs, ..., include change in control payments and severance and retention bonuses for management and employees of the merged entity ..."

[Insert Table V Here]

We now regress these merger integration and employee discussion variables on our ex ante measures of merger integration risk. Table VI presents the results. We include control variables for size, age, overall textual similarity to rivals and Tobin's q. All regressions also include industry and year fixed effects with standard errors clustered by industry.

[Insert Table VI Here]

The results presented in Table VI show that firms are more likely to mention both integration problems and employee retention issues when expected integration is high. These results are based on how integrated individual firms are in comparison to the set of all acquiring firms. If the firm has a high level of ex ante realized integration, they mention integration failure problems ex post less often. However, these firms do identify employee retention issues as being important as shown by the positive and significant coefficient on firm integration when mentioning employee retention in Panel B. More importantly, our composite measure "Integration Gap" captures the differences between expected integration and actual firm integration. This measure is positively related to both the mentions of integration problems and also employee retention issues. Hence firms with a larger ex ante integration gap, indicating that their realized level of integration falls short of their expected level, experience more ex post managerial discussions of merger integration failure and employee retention issues.

Overall, these results provide a validation test that our measures of product integration based on words across paragraphs are picking up integration risk. Our integration measures are calculated using product description text and are based on products, whereas the measures of integration difficulties and employee retention issues are calculated based on firm discussions in the entire 10-K. In all, these tests strongly support the conclusion that our ex ante measures of integration risk do predict observed instances of integration failure being discussed directly in the 10-K, which is a key result motivating the use of our variables as valid measures of ex ante integration risk.

V Withdrawn Acquisitions

Before examining outcomes of mergers with high integration risk, we first examine if announced mergers are more likely to be canceled if realized integration is low and the gap between expected and realized integration is high. This test is based on the premise that many deals are canceled when parties raise opposition to them.

Table VII reports the results of regressions in which the dependent variable is a measure of withdrawn transactions. In Panel A, one observation is one firm in one year, and the dependent variable is the fraction of a given firm's announced transactions in the given year that were withdrawn. A firm-year observation is only included in the regression if the firm had at least one announced acquisition in the given year. In Panel B, we consider a larger panel database in which one observation is one announced transaction, and the dependent variable is a dummy that is equal to one if the transaction was withdrawn. The key independent variables are realized integration, expected integration, and the integration gap variables. We also include controls for size, age, TNIC total similarity and Tobins Q. All regressions include industry fixed effects and year fixed effects, and standard errors are clustered by industry.

[Insert Table VII Here]

Inspection of Table VII reveals that proposed mergers and acquisitions are more

likely to be withdrawn when the gap between expected integration and realized ex ante firm integration is high. These results hold both at the firm-year level in Panel A and at the deal level in Panel B. Both components of the ex ante integration gap are significant. Deals are less likely to be withdrawn when ex ante realized firm integration is high and are more likely to be withdrawn when expected integration of rival firms is high. In addition, when rivals and targets are similar, as measured by TNIC similarity, deals are less likely to be withdrawn. Highly valued acquirers are also less likely to withdraw deals. Overall the results support the conclusion that our measure of integration and integration gap captures ex ante information that firms and market participants are using to assess the potential success of acquisitions. When the risk of integration is high and firm integration is low, deals are more likely to be withdrawn.

VI Ex-Post Real Outcomes

We now examine the relationship between post-merger real outcomes and ex ante integration risk. We examine the ex post change in operating income and also the ex post change in operating costs (SG&A). Lastly, we examine if firms with high integration risk are more likely to divest assets ex post. Table VIII reports the results of OLS regressions in which the dependent variable is a measure of ex post operating income to assets and SG&A to sales. As our goal is to examine ex post outcomes for acquirers, we limit the sample to firms that were an acquirer in year t. We consider outcomes measured as changes for both a one-year horizon and a three-year horizon, where the horizon begins in year t of the merger and ends in year $t+1$ or $t+3$. We consider the following outcomes: ex post changes operating income scaled by assets and expenses captured by ex post changes in SG&A /sales. All regressions include industry and year fixed effects and all right-hand-side variables are standardized prior to running regressions for ease of interpretation.

[Insert Table VIII Here]

Inspection of the results in Table VIII reveal that operating income is lower for firms with high merger integration risk - especially when we consider the full panel

of acquisitions that include public and private targets in Panel B. We also find that operating expenses as captured by SG&A are higher when there is higher ex ante measured expected post-merger integration risk.

In particular, Panel B, rows 1 and 3, show that operating income is 1.0 to 1.3% lower for acquirers with a 1 standard deviation higher expected integration risk. We see equivalently sized negative effects on operating income when we look at the integration gap - which is the difference between the expected and realized integration, prior to the acquisition. The interpretation of the integration gap is very intuitive. When the ex ante difference between the expected integration and actual integration is high for the acquirer, it indicates that the firm's realized integration is below its potential expected integration. Hence such a firm is less likely to realize the full potential of its M&A activity, and we predict worse outcomes when such a firm does participate in acquisitions. The aforementioned results strongly support this conclusion.

Table IX examines whether post-merger divestitures are related to ex ante merger integration risk. We regress ex post divestitures on ex ante merger integration risk, both for the decomposed measures based on synergies and assets in place (available when both the acquirer and the target are public) in Panel A, and also for our standard measures of ex ante realized and expected integration risk (available for all acquirers in our sample) in Panel B. We also include controls for size, age, target fraction of acquirer, market to book and also previous text-based similarity measures shown to impact mergers. Panel A examines divestitures after purchase of publicly traded targets and Panel B examines divestitures after all acquisitions, both public and private.

[Insert Table IX Here]

Table IX reveals that divestitures in the year after the merger increase when there is higher ex ante merger integration risk - both combined and when separately examining acquirers and targets expected integration risk. Panel B examines firm-level expected integration, realized integration and also the integration gap (the difference between expected integration and realized integration). The results in in Panel B

show that divestitures increase with expected integration as well as with the integration gap. All of the integration variables are measured before the transaction, thus providing evidence that past difficulties or shortcomings in integration are associated with subsequent divestitures.

VII Stock Market Returns

Given we have documented outcomes differ on the real side, we turn to an examination of the impact of integration risk in the stock market. We examine whether merger integration risk and specifically the risk that synergies will experience integration difficulties, affect how the stock market is affected by merger integration risk. We examine both announcement returns and also longer term ex post stock market returns.

A Announcement Returns

We first examine stock market announcement returns. We regress stock market announcement returns on our measures of merger integration risk and synergy integration risk. We include both joint measures of synergy integration risk for the acquirer and target, and separate measures for the assets in place of the acquirer and target. We consider announcement returns measured just on day $t = 0$, and also a 3-day window, where all windows are centered around $t = 0$. Announcement returns are market-adjusted. We include control variables for size, age, the fraction of the acquirer the target represented, the firm market to book, and text variables based on the Hoberg and Phillips (2010) previous study of merger success. These text variables are are pairwise firm similarity, the expected gain in product differentiation and the firm's product description size.

The key independent variables of interest are the Synergy Integration Risk, and the Target and Acquirer Assets in Place (existing firm specific ex ante measures) Integration Risk. These measures are computed in a fashion analogous to the Expected Integration variable discussed in Table I with two primary exceptions. First, we compute integration risk for three firms: the acquirer, the target, and the hy-

pothetical post-merger firm (based on the combined firm vocabularies). This allows us to isolate synergies as discussed in Section II. Second, we use the distribution of word-pairs instead of single words for the expected synergy integration variable, as only this approach can isolate the combined firm's synergies that are not present in the individual firms. The resulting measures of integration risk are ex ante measurable and target specific parts of the post merger firm based on assets in place and likely synergies.

[Insert Table X Here]

Table X shows that the M&A announcement returns for acquirers and targets are affected by merger integration risk and synergy risk. Panel A shows that when there is expected synergy integration risk, acquirers have 3 day abnormal returns that are negative. More significantly, announcement returns to targets are strongly affected. Panel B shows when there is higher target expected integration risk, target announcement returns are lower. However, we also see that targets have higher abnormal announcement returns when expected total post-merger integration risk, and specifically synergy integration risk, is higher. These results are consistent with the target needing to be compensated in order to accept a situation where they will both lose control and accept a high risk of integration failure.

Consistent with this interpretation based on target compensation, we find in unreported tests that these results are driven by the subset of transactions in our sample that are mergers. In particular, we continue to find significant target premia for the 2,364 transactions in our sample that are mergers, and we do not find such premia for the 3,444 transactions that are acquisition of asset transactions. This distinction is relevant in this case because mergers more directly result in a loss of control for target managers.

Although we do not report combined target and acquirer announcement returns to conserve space, we note that combined firm announcement returns are not significantly related to our merger integration variables. This is not surprising given that integration risk is negatively associated with acquirer announcement returns, and positively associated with target announcement returns. We also note that the

pairwise TNIC similarity variable is positively associated with combined firm announcement returns despite its negative link to target announcement returns and insignificant link to acquirer announcement returns. These results replicate those in Hoberg and Phillips (2010), and they also illustrate that our measures of merger integration are distinct from firm pairwise similarities (which, as we pointed out earlier, is by construction given that we focus on within-firm integration using word-paragraph distributions, and the pairwise similarity variable focuses on vocabulary overlap across firms).

B Ex Post Long-run Stock Returns

In this section, we explore the extent to which ex-ante measures of integration are associated with the ex post stock returns of acquiring firms. This issue of the stock returns to acquiring firms is important and has been studied by Asquith (1983), Aggarwal, Jaffe, and Mandelker (1992), Fama (1998), Loughran and Vijh (1997) and Mitchell and Stafford (2000). These studies show that acquiring firms underperform in the years after an acquisition. Our study extends this work and we examine the extent to which acquiring firms with higher levels of integration risk experience lower stock returns than do acquirers with lower levels of integration risk. Evidence supporting this link can further explain why some acquiring firms underperform, as market participants might not have full information about the extent of integration risk and its potential adverse affect on acquiring firms.

C Asset Pricing Variables

We consider monthly excess stock returns as our dependent variable. Our primary independent variables of interest include ex ante realized integration, expected integration, and the integration gap. In particular, we consider interactions of these variables with an acquisition dummy. Our acquisition dummy is set to one when a firm has a completed acquisition as indicated by the SDC Platinum database. The dummy is set to one during the one year period starting six months after the acquisition date and is otherwise set to zero. The use of a six month lag is to maintain

consistency with our other variables, and also to reflect the fact that integration failure likely materializes after the firm has had ample time to attempt to properly integrate the acquired division. This allows us to examine if the well known anomaly that acquiring firms underperform can be explained by integration failure, and also allows us to more broadly examine the cross sectional role of merger integration failure in explaining monthly stock returns.

We also include controls for size, book to market and momentum. We construct size and book to market ratio variables following Davis, Fama, and French (2000) and Fama and French (1992). Market size is the natural log of the CRSP market cap. Following the lag convention in the literature, we use size variables from each June, and apply them to the monthly panel to use to predict returns in the following one year interval from July to June.

The book-to-market ratio is based on CRSP and Compustat variables. The numerator, the book value of equity, is based on the accounting variables from fiscal years ending in each calendar year (see Davis, Fama, and French (2000)) for details). We divide each book value of equity by the CRSP market value of equity prevailing at the end of December of the given calendar year. We then compute the log book to market ratio as the natural log of the book value of equity from Compustat divided by the CRSP market value of equity. Following standard lags used in the literature, this value is then applied to the monthly panel to predict returns for the one year window beginning in July of the following year until June one year later.

For each firm, we compute our momentum variable as the stock return during the eleven month period beginning in month $t-12$ relative to the given monthly observation to be predicted, and ending in month $t-2$. This lag structure that avoids month $t-1$ is intended to avoid contamination from microstructure effects, such as the well-known one-month reversal effect.

After requiring that adequate data exist to compute our integration variables and the aforementioned asset pricing control variables, and requiring valid return data in CRSP, our final sample has 626,877 observations.

D Fama MacBeth Regressions

Table XI displays the results of monthly Fama and MacBeth (1973) regressions in which the dependent variable is the monthly excess stock return. Row (1) shows our baseline model, where do not include any integration variables. Although the standard book to market and momentum variables are not significant in our sample, we critically note that the acquisition dummy is negative and significant at the 5% level.[3]

In row (2), we add a dummy that is set to one when a firm has a value for its expected integration that is in the highest tercile in the given year. We also include a critical cross term based on the acquisition dummy. The results in row (2) indicate that the acquirer dummy coefficient declines by an economically large 71% form 0.169 to 0.049 when the integration variable and its cross term are included. At the same time, the cross term is negative and significant at the 1% level with a coefficient of -0.335. These results indicate that acquiring firms do not underperform if expectations for their integration are low, and do underperform significantly when integration expectations are high. Because these tests are based on dummy variables, the coefficients can be interpreted. If we consider the acquisition dummy (without integration variables included) as in row (1), we find that acquiring firms underperform by 16.7 basis points per month (-2.0% annualized). However, the impact of integration expectations being high is 33.5 basis points per month (-4.0% per year).

An additional result from the table is that the expected integration dummy has a positive and (in some rows) weakly significant coefficient at the 10% level. Because we include the cross term with the acquirer dummy, the interpretation of this variable is that firms that are not acquirers experience marginally higher stock returns when their expected integration is higher. This result is consistent with integration potentially being valuable when it is not accompanied by mergers. In particular, mergers require that managers bring the acquired assets up to the level of integration enjoyed by the assets in place, a task that is very challenging and that entails risk. One po-

[3]The weak results for book to market and momentum likely relate to the relatively short nature of our sample. Also, it is well known that the momentum anomaly performed poorly during the time of the financial crisis, which is included in our study. As our main objective is to examine the impact of acquisitions, we note that these existing variables are included mainly as controls.

tential concern with the regressions in Table XI is potential multicollinearity. We rule out this possibility when we consider separate quintile regressions later in this section where cross terms are not necessary.

[Insert Table XI Here]

Row (3) of Table XI shows similar results for ex-ante realized integration. In particular, including the above median ex ante realized integration variable also reduces the magnitude of the acquisition dummy, and the cross term with ex ante realized integration is negative and only marginally significant at the 1% level. In particular, it is not as significant statistically or economically as the expected integration variable in row (2). Following our approach in earlier sections, we thus consider both together in row (4), and we find that only the expected integration variable's cross term is significant. This suggests that acquirers with high expected integration underperform, but those with high realized integration do not.

We consider above a dummy for above median levels of the integration gap in row (5). Although the results are marginally weaker for the gap than they are for the expected integration variable, we find that the integration gap cross term is negative and significant at the 5% level. Moreover, the acquisition dummy also falls below standard levels of statistical significance.

In rows (6) and (7), we repeat the key regressions in rows (4) and (5) with two additional variables. The first is the fraction of consideration paid in the acquisition that is in the form of stock. The second is a dummy that is one when the form of consideration is not available, as the consideration variables are frequently missing in SDC Platinum. This allows us to retain the full sample and we set the missing values for the first variable to zero as their impact is then absorbed by the dummy. The objective of these tests is to examine if our results are robust to the findings of Loughran and Vijh (1997), who find that longer term stock returns are strongly negative when acquisitions are done using stock. In our setting, the fraction stock variable is indeed negative, although its significance level just misses the 10% level. This is likely due to the fact that our sample is newer than those in existing studies, and our sample is also somewhat limited in time series. Nevertheless, the objective

in our case is to simply control for the fraction stock.

We find in rows (6) and (7) that our results are entirely robust to including controls for the form of consideration. Hence, our results are distinct from existing studies. We also note that, at least in row (6), that the acquisition dummy actually becomes positive (although not significant) when both the consideration and the integration variables are included. This suggests that acquisitions with low integration risk and that do not use stock do not underperform, whereas the mirror image transactions that have high integration risk and that use stock underperform in an economically material fashion.

To ensure that our cross term tests are not influenced by multicollinearity, and to further explore how the economic magnitude of the integration variables change as integration risk becomes more extreme, we next consider quintile subsamples in Table XII. In particular, we first sort firms in each month into quintiles based on their level of ex ante expected integration (Panel A), ex ante realized integration (Panel B) and ex ante integration gap (Panel C). For each quintile subsample, we then run Fama MacBeth regressions similar to those in Table XI but with a couple important changes. First, because we form subsamples based on the integration variables, we do not have to include them in the regression, and nor do we need to include a cross term. Instead we focus on the acquisition dummy in each regression and only include the standard size, book to market and momentum controls.

[Insert Table XII Here]

By examining the significance and the economic size of the acquisition dummy coefficient in each quintile, we can then explore how strong the acquirer underperformance anomaly is in each subsample. Our prediction is that the acquire dummy will be strongly negative in high integration risk quintiles, and negligible in low integration risk quintiles.

We first consider Panel A, where quintiles are formed based on ex ante expected firm integration. The table shows that the acquisition dummy coefficient is negative but insignificant in row (1) for the lowest quintile of expected integration. However, it is negative and highly significant with a t-statistic of -3.78 in the high expected

integration quintile. We also note that the economic magnitude of the high quintile coefficient is large at -0.504. This indicates that acquirers facing high levels of expected integration underperform by 50.4 basis points per month. This is an economically meaningful 6.05% per year.

We find similar results in Panel B for ex ante realized integration quintiles. In this case, the acquisition dummy is actually *positive* (although insignificant) for the first two quintiles indicating low integration. The acquisition coefficient is negative and significant at the 1% level for the two highest quintiles. Here the highest quintile coefficient of 0.449 indicates underperformance of 44.9 basis points per month (-5.39% per year).

Finally, Panel C displays results for the ex ante integration gap quintiles. Although results are weaker for this variable than they are in Panels A and B, we once again observe that the acquisition dummy is insignificant for low integration gap quintiles, and is negative and significant at the 5% level in the highest integration gap quintile. Here the coefficient indicates underperformance of 32 basis points per month (-3.84% per year).

Overall, our results indicate that stock returns are lower among firms that are acquirers when they face higher levels of ex ante integration measures. The two most likely explanations of this finding are that (A) the market does not ex ante fully predict the extent of integration failure among acquirers with high integration risk or (B) these findings might be related to a new systematic risk factor. Because integration risk is likely driven most by individual firm managers and their unique abilities to make integration work, we believe that (A) is most likely. In particular, the likely ingredients that separate firms regarding the ability to integrate are likely idiosyncratic. Assuming that integration risk and failure is also difficult to predict ex ante, the possibility that market participants are unable to have the full information about the extent of possible integration failure is quite plausible. On the other hand, given the nature of integration and its link to the managerial team, it would seem much less likely that our results can be explained by systematic risk. We also note that because we control for standard predictors of stock returns including the book to market ratio and momentum, that existing potential sources of systematic risk

also cannot explain our findings.

VIII Conclusions

We examine the importance of merger integration risk to merger outcomes - both for stock market and real outcomes. Our findings support the view that poor merger outcomes arise in part from the difficulty of integrating the product lines offered by the pre-merger firms and the intended synergies.

We focus on measuring the difficulty of integrating product lines across organizations at the firm level for acquisitions in the U.S. We use text-based analysis of business descriptions in firm 10-Ks to measure ex ante merger integration risk to capture the extent to which merging firms will face challenges integrating their product lines. The measures are general and are based on measuring integration at the atomistic level of individual words or word-pairs. Using our approach we can assess ex ante integration risk separately for assets in place and merger synergies. These integration risk components can even be computed before a candidate post merger firm is observed.

Validating our approach, we find that when ex-ante merger integration risk is high, that the post-transaction incidence of managers discussing integration difficulties increases. These discussions are specific and often refer to issues such as drains on managerial time, drains on other corporate resources, or specific failures in integration. These findings are consistent with ex ante product integration risks predicting an increased likelihood of such ex post unexpected drains on managerial time and also in retaining employees.

We document the impact of ex ante integration risk throughout the merger process and on ex post outcomes. We find that when ex ante merger integration risk is high, proposed deals are more likely to be withdrawn consistent with market participants recognizing that some deals have higher integration costs. For deals that are finalized and are not withdrawn, we find that ex ante merger integration risk is associated with lower ex post operating income and higher ex post SG&A/sales, which specifically relates to the cost of managing the firm's employees and organi-

zations. We also find evidence that divestitures are higher when there is higher ex ante product integration risk. These findings illustrate the importance of product integration risk and its real impact on acquiring firms. Because our results indicate that integration risk poses a greater challenge for synergies than for assets in place, they also highlight the elevated role that synergies play in determining successful instances of merger integration.

Examining the impact in the stock market, we find that ex ante product integration risk is associated with lower stock market announcement returns and lower ex post monthly stock returns for the acquirer, and higher announcement returns for the target. The former is consistent with the market only learning the negative consequences of high ex ante integration risk over time. These results further suggest that the longer term underperformance of acquirers can be explained at least in part by integration failure. Although more research is needed to fully understand the latter effect, we note that it is consistent with agents associated with the target demanding a higher premium to compensate them for accepting a transaction that entails integration risk, especially as the acquirer management and its employees will take over control of the combined firm.

References

Aggarwal, Anup, Jeffrey Jaffe, and Gershon Mandelker, 1992, The post-merger performance of acquiring firms: a reexamination of an anomlay, *Journal of Finance* 47, 1605–1621.

Ahern, Kenneth, Daniele Daminelli, and Cesare Fracassi, 2015, Lost in translation? the effect of cultural values on mergers around the world, *Journal of Financial Economics* 117, 165–189.

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

Arnold, JD, 1983, Saving corporate marriages: Five cases, *Mergers and Acquisitions* 23, 55–58.

Asquith, Paul, 1983, Merger bids, uncertainty and stockholder returns, *Journal of Financial Economics* 11, 51–83.

Bena, Jan, and Kai Li, 2014, Corporate innovations and mergers and acquisitions, *Journal of Finance* 69, 19231960.

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

Epstein, Marc, 2004, The drivers of success in post-merger integration, *Organizational dynamics* 33, 174–189.

Fama, Eugene, and Kenneth French, 1992, The cross section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene, and J. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Fama, Eugene F., 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.

Harford, Jarrad, 1999, Corporate cash reserves and acquisitions, *Journal of Finance* 54, 1969–1997.

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

Hoberg, Gerard, and Vojislav Maksimovic, 2015, Redefining financial constraints: a text-based analysis, *Review of Financial Studies* 28, 1312–1352.

Hoberg, Gerard, and Gordon Phillips, 2010, Product market synergies in mergers and acquisitions: A text based analysis, *Review of Financial Studies* 23, 3773–3811.

——— , 2015, Text-based network industry classifications and endogenous product differentiation, Forthcoming Journal of Political Economy.

Loughran, Timothy, and Anand Vijh, 1997, Do long-term shareholders benefit from corporate acquisitions?, *Journal of Finance* 52, 1765–1790.

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.

McGee, Tom, Trevear Thomas, and Russell Thomson, 2015, Integration report 2015: Putting the pieces together, available at http://www2.deloitte.com/content/dam/Deloitte/us/Documents/mergers-acqisitions/us-ma-integration-report-030415.PDF.

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

Rhodes-Kropf, Matthew, and David Robinson, 2008, The market for mergers and the boundaries of the firm, *Journal of Finance* 63, 1169–1211.

Table I: Summary Statistics

Notes: Summary statistics are reported for our sample from 1996 to 2008. Realized integration is the extent to which a firm's individual words appear within its own paragraphs in a distribution close to a uniform distribution. Expected integration is the extent to which a firm uses vocabulary that generally appears in a uniform distribution across paragraphs in all firms that use the given word in the economy in the given year. The integration gap is expected minus realized integration. TNIC total similarity is the summed TNIC similarity of firms in the given firm's TNIC industry. The integration challenges dummy is one if the firm's 10-K has a paragraph where the firm mentions integration in the context of a discussion about acquirers and along side vocabulary that indicates difficulty. The employee retention dummy is a dummy that is one if the firm mentions employee retention issues in a paragraph that also discusses acquisitions. The profitability and expense variables are based on Compustat data. The change in target (acquirer) rate is the natural logarithm of one plus the number of asset sales (purchases) in year t divided by one plus the number of asset sales (purchases) in year $t-1$.

Variable	Mean	Std. Dev.	Minimum	Median	Maximum
Panel A: Integration Variables and Firm Characteristics					
Realized Integration	1.820	0.315	0.908	1.797	5.727
Expected Integration	1.737	0.457	0.383	1.706	3.938
Integration Gap	-0.083	0.323	-3.150	-0.094	1.837
TNIC Total Similarity	10.156	19.907	1.000	2.577	132.939
Log Assets	5.701	2.254	-6.908	5.664	14.598
Log Age	2.459	0.819	0.000	2.398	4.078
Panel B: Managerial Mentions of Integration Difficulties					
Integration Challenges Dummy	0.194	0.396	0.000	0.000	1.000
Employee Retention Dummy	0.114	0.317	0.000	0.000	1.000
Panel C: ex post Outcome Variables					
Δ OI/Assets	-0.012	0.153	-1.286	-0.001	0.961
Δ SG&A/Sales	-0.021	0.410	-8.918	0.000	3.466
Δ Target Rate	0.040	0.391	-2.833	0.000	2.565
Δ Acquirer Rate	-0.011	0.529	-3.526	0.000	3.807

Table II: Pearson Correlation Coefficients

Notes: Pearson Correlation Coefficients are reported for our sample from 1996 to 2008. Please see Table I for variable descriptions.

Row	Variable	Realized Integration	Expected Integration	Integration Gap	TNIC Total Similarity	Log Assets	Log Age	Integration Diffic. Dummy
				Correlation Coefficients				
(1)	Expected Integration	0.708						
(1)	Integration Gap	0.024	0.723					
(2)	TNIC Total Similarity	-0.120	0.021	0.147				
(3)	Log Assets	-0.123	0.008	0.131	0.164			
(4)	Log Age	-0.032	-0.104	-0.116	-0.196	0.344		
(5)	Integration Challenges Dummy	0.020	0.067	0.075	-0.090	0.049	-0.065	
(6)	Employee Retention Dummy	-0.002	0.038	0.055	-0.001	0.091	-0.002	0.124

Table III: Integration Across Industries

The table displays the average realized and expected integration for the Fama-French-12 industries in 1997 (Panel A) and 2008 (Panel B). Realized integration is the extent to which a firm's individual words appear within its own paragraphs in a distribution close to a uniform distribution. Expected integration is the extent to which a firm uses vocabulary that generally appears in a uniform distribution across paragraphs in all firms that use the given word in the economy in the given year.

Row	FF12 Industry	Realized Integration	Expected Integration	# Obs.
		Panel A: 1997 Industries		
1	Hlth	2.065	2.101	759
2	NoDur	1.956	1.826	395
3	Chems	1.922	1.744	143
4	Telcm	1.912	1.989	211
5	BusEqSv	1.905	1.825	1366
6	Utils	1.903	1.882	170
7	Manuf	1.898	1.739	725
8	Other	1.882	1.746	1007
9	Durbl	1.859	1.706	183
10	Enrgy	1.857	1.821	248
11	Shops	1.821	1.647	765
12	Money	1.751	1.670	1359
		Panel B: 2008 Industries		
1	Hlth	2.384	2.433	561
2	Chems	2.160	2.018	101
3	Enrgy	2.108	2.114	218
4	Telcm	2.106	2.107	121
5	BusEqSv	2.095	1.981	803
6	NoDur	2.085	1.938	194
7	Manuf	2.071	1.884	379
8	Utils	2.039	2.058	116
9	Other	1.999	1.840	558
10	Shops	1.946	1.758	382
11	Durbl	1.930	1.840	101
12	Money	1.826	1.754	1053

Table IV: Sample Managerial Statements of Integration Risk

The table displays the first ten paragraphs returned from metaHeuristica in 1997 that hit on our verbal query intended to measure managerial measures of integration risk. The query was run using metaHeuristica and requires that one word from each of three buckets must appear in a paragraph. The first bucket is acquisition words: {merger, mergers, merged, acquisition, acquisitions, acquired}. The second bucket is integration words: {integration, integrate, integrating}. The third bucket is an indication of difficulty: {challenge, challenging, difficulties, difficulty, inability, failure, unsuccessful, substantial expense}. The results from this query are then used to compute the integration challenges dummy and the integration challenges intensity variables.

Row	Sample Paragraph
1	**[Integrated Health Services]** IHS has recently completed several major acquisitions, including the acquisitions of First American, RoTech, CCA and the Coram Lithotripsy Division and the Facility Acquisition, and is still in the process of integrating those acquired businesses. The IHS Board of Directors and senior management of IHS face a significant challenge in their efforts to integrate the acquired businesses, including First American, RoTech, CCA, the Coram Lithotripsy Division and the facilities and other businesses acquired from HEALTHSOUTH. The dedication of management resources to such integration may detract attention from the day-to-day business of IHS. The difficulties of integration may be increased by the necessity of coordinating geographically separated organizations, integrating personnel with disparate business backgrounds and combining different corporate cultures.
2	**[Siebel Systems]** The Company has acquired in the past, and may acquire in the future, other products or businesses which are complementary to the Company's business. The integration of products and personnel as a result of any such acquisitions has and will continue to divert the Company's management and other resources. There can be no assurance that difficulties will not arise in integrating such operations, products, personnel or businesses. The failure to successfully integrate such products or operations could have a material adverse effect on the Company's business, financial condition and results of operations.
3	**[Cable Design Technologies]** Although the Company has been successful in integrating previous acquisitions, no assurance can be given that it will continue to be successful in integrating future acquisitions. The integration and consolidation of acquired businesses will require substantial management, financial and other resources and may pose risks with respect to production, customer service and market share. While the Company believes that it has sufficient financial and management resources to accomplish such integration, there can be no assurance in this regard or that the Company will not experience difficulties with customers, personnel or others. In addition, although the Company believes that its acquisitions will enhance the competitive position and business prospects of the Company, there can be no assurance that such benefits will be realized or that any combination will be successful.
4	**[Star Telecommunications]** Additionally, on November 19, 1997, the Company entered into an agreement to acquire UDN. The acquisition of UDN is subject to approval of UDN's stockholders and to various regulatory approvals, and the Company may not complete this acquisition. These acquisitions have placed significant demands on the Company's financial and management resources, as the process for integrating acquired operations presents a significant challenge to the Company's management and may lead to unanticipated costs or a diversion of management's attention from day-to-day operations.
5	**[Sun Healthcare Group]** The integration of the operations of Retirement Care and Contour, to the extent consummated, will require the dedication of management resources which will detract attention from Sun's day-to-day business. The difficulties of integration may be increased by the necessity of coordinating geographically- separated organizations, integrating personnel with disparate business backgrounds and combining different corporate cultures. As part of the RCA and Contour Mergers, Sun is expected to seek to reduce expenses by eliminating duplicative or unnecessary personnel, corporate functions and other expenses.
6	**[Sunquest Information Systems]** management has limited experience in identifying appropriate acquisitions and in integrating products, technologies and businesses into its operations. The evaluation, negotiation and integration of any such acquisition may divert the time, attention and resources of the Company, particularly its management. There can be no assurance that the Company will be able to integrate successfully any acquired products, technologies or businesses into its operations, including its pharmacy systems.
7	**[Waterlink Inc]** Waterlink has grown by completing ten acquisitions consisting of seventeen operating companies. The success of the Company will depend, in part, on the Company's ability to integrate the operations of these businesses and other companies it acquires, including centralizing certain functions to achieve cost savings and developing programs and processes that will promote cooperation and the sharing of opportunities and resources among its businesses. A number of the businesses offer different services, utilize different capabilities and technologies, target different markets and customer segments and utilize different methods of distribution and sales representatives. While the Company believes that there are substantial opportunities in integrating the businesses, these differences increase the difficulty in successfully completing such integration.

Table V: Sample Managerial Statements of Employee Retention Issues

The table displays the first ten paragraphs returned from metaHeuristica in 1997 that hit on our verbal query intended to measure managerial mentions of employee retention issues. The query was run using metaHeuristica and requires that one word from each of three buckets must appear in a paragraph. The first bucket is acquisition words: {merger, mergers, merged, acquisition, acquisitions, acquired}. The second bucket is employee words: {employee, employees, personnel }. The third bucket is an indication of retention or departures: {retention, departure, departures}. The results from this query are then used to compute the employee retention dummy and the employee retention intensity variables.

Row	Sample Paragraph
1	**[Tellabs Inc]** The Company has a number of employee retention programs under which certain employees, primarily as a result of the Company's acquisitions, are entitled to a specific number of shares of the Company's stock over a two-year vesting period.
2	**[Marvel Entertainment Group]** The Company has been in bankruptcy since December 27, 1996. There is a general uncertainty amongst the Company's employees regarding the outlook of the Company. The Company believes its relationship with its employees is satisfactory, however, it is not known if a merger or sale of the Company under a plan of reorganization would negatively affect employee retention.
3	**[Rational Software Corp]** The ability of the Company to attract and retain the highly trained technical personnel that are integral to its direct sales and product development teams may limit the rate at which the Company can develop products and generate sales. Competition for qualified personnel in the software industry is intense, and there can be no assurance that the Company will be successful in attracting and retaining such personnel. Merger activities, such as the proposed acquisition of Pure Atria, may have a destabilizing effect on employee retention at all levels within the Company. Departures of existing personnel, particularly in key technical, sales, marketing or management positions, can be disruptive and can result in departures of other existing personnel, which in turn could have a material adverse effect upon the Company's business, operating results and financial condition.
4	**[Peoples Bancorp]** Expenses for human resources also increased through the acquisitions and corresponding expansion of the Company's services and geographic area. For the year ended December 31, 1997, salaries and benefits expense increased $844,000 (or 11.2%) to $8,358,000 compared to 1996. The acquisitions increased the number of employees due to the retention of many customer service associates. At December 31, 1997, the Company had 314 full-time equivalent employees, up from 304 full-time equivalent employees at year-end 1996. The Company had 261 full-time equivalent employees at March 31, 1996, before the combined impact of recent acquisition activity. Management expects salaries and employee benefits to increase in 1998 due to the pending West Virginia Banking Center Acquisition and normal merit increases. Management will continue to strive to find new ways of increasing efficiency and leveraging its resources while concentrating on maximizing customer service.
5	**[Whitney Holding Corp]** The Company and its merger candidates incur various non-recurring costs to complete merger transactions and to consolidate operations subsequent to a merger. Such merger-related costs, which are expensed for business combinations accounted for as poolings-of-interests, include change in control payments and severance or retention bonuses for management and employees of the merged entity, investment banker fees, fees for various professional services, including legal, audit and system conversion consulting services, and losses on the disposition of obsolete facilities and equipment and the cancellation of contracts. Total merger-related expenses will vary with each transaction.
6	**[Sinclair Broadcast Group]** Except as otherwise provided in this Section 3.5 or in any employment, severance or retention agreements of any Transferred Employees, all Transferred Employees shall be at-will employees, and Time Broker may terminate their employment or change their terms of employment at will. No employee (or beneficiary of any employee) of Seller may sue to enforce the terms of this Agreement, including specifically this Section 3.5, and no employee or beneficiary shall be treated as a third party beneficiary of this Agreement. Except to the extent provided for herein, Time Broker may cover the Transferred Employees.
7	**[Ensearch Corp]** Mr. Hunter, Mr. Pinkerton and certain other key employees of ENSERCH have entered into retention bonus arrangements, effective as of August 1997, pursuant to which ENSERCH will pay the employee a bonus equal to a percentage of the employee's current annual salary (typically 50% and 100%, respectively) upon the attainment of six and eighteen months of employment. Mr. Biegler was paid a retention bonus of $900,000 by ENSERCH for services up until the consummation of the Merger in August 1997.

Table VI: Post-Merger Integration and Employee retention

The table reports the results of a linear probability model in which the dependent variable is either the ex post (year $t+1$) integration challenges dummy/intensity or the employee retention dummy/intensity as noted in the second column. As our goal is to examine ex post outcomes for acquirers, we limit the sample to firms that were an acquirer in year t. The integration challenges dummy is one if the firm's 10-K has a paragraph where the firm mentions integration in the context of a discussion about acquirers and along side vocabulary that indicates difficulty. The employee retention dummy is a dummy that is one if the firm mentions employee retention issues in a paragraph that also discusses acquisitions. The corresponding intensity variable for both dummies measures the number of paragraphs that contain this kind of content. The key independent variables are realized integration, expected integration, and the integration gap variables. Results are similar if we instead use logistic regressions. All regressions include industry fixed effects and year fixed effects, and standard errors are clustered by industry.

Row	Dependent Variable	Expected Integration	Firm Integration	Integration Gap	Log Assets	Log Age	TNIC Total Similarity	Tobins Q	Obs.
				Panel A: Ex Post Integration Failure					
(1)	Integration Failure Dummy	0.065	-0.032		0.010	-0.067	-0.000	0.007	20,905
		(6.57)	(-2.67)		(5.53)	(-11.05)	(-0.56)	(4.06)	
(2)	Integration Failure Intensity	0.433	-0.172		0.080	-0.472	-0.001	0.050	20,905
		(6.31)	(-2.13)		(6.49)	(-11.50)	(-0.50)	(3.89)	
(3)	Integration Failure Dummy			0.065	0.010	-0.068	-0.000	0.007	20,905
				(6.55)	(5.36)	(-11.20)	(-0.39)	(4.08)	
(4)	Integration Failure Intensity			0.434	0.080	-0.479	-0.001	0.050	20,905
				(6.31)	(6.27)	(-11.69)	(-0.32)	(3.91)	
				Panel B: Ex Post Employee Retention					
(5)	Employee Retention Dummy	0.013	0.017		0.017	-0.018	0.001	0.001	20,905
		(1.95)	(1.98)		(6.78)	(-4.87)	(2.50)	(2.08)	
(6)	Employee Retention Intensity	0.087	0.141		0.122	-0.125	0.004	0.009	20,905
		(1.83)	(2.22)		(6.88)	(-4.61)	(2.60)	(2.02)	
(7)	Employee Retention Dummy			0.013	0.017	-0.019	0.001	0.002	20,905
				(1.95)	(6.81)	(-4.98)	(2.75)	(2.18)	
(8)	Employee Retention Intensity			0.087	0.121	-0.131	0.004	0.009	20,905
				(1.84)	(6.93)	(-4.72)	(2.89)	(2.18)	

Table VII: Withdrawn Transactions and Ex Ante Integration Risk

The table reports the results of regressions in which the dependent variable is a measure of withdrawn transactions. In Panel A, one observation is one firm in one year, and the dependent variable is the fraction of a given firm's announced mergers or acquisitions in the given year that were withdrawn. A firm-year observation is only included in the regression if the firm had at least one announced acquisition in the given year. In Panel B, we consider a larger panel database in which one observation is one announced transaction, and the dependent variable is a dummy that is equal to one if the transaction was withdrawn. The key independent variables are realized integration, expected integration, and the integration gap variables. We also include controls for size, age, TNIC total similarity and Tobins Q. All regressions include industry fixed effects and year fixed effects, and standard errors are clustered by industry.

Row	Dependent Variable	Expected Integration	Firm Integration	Integration Gap	Log Assets	Log Age	TNIC Total Similarity	Tobins Q	Obs.
				Panel A: Firm-year regressions					
(1)	% Withdrawn	0.004	-0.005		0.000	0.001	-0.000	-0.000	21,068
		(2.21)	(-2.86)		(1.13)	(1.11)	(-2.65)	(-3.03)	
(2)	% Withdrawn			0.004	0.000	0.001	-0.000	-0.000	21,068
				(2.22)	(1.13)	(1.15)	(-2.74)	(-3.07)	
				Panel B: Deal-level regressions					
(3)	Withdrawn Dummy	0.003	-0.004		0.000	0.001	-0.000	-0.000	49,594
		(2.38)	(-2.33)		(0.34)	(1.44)	(-2.10)	(-1.85)	
(4)	Withdrawn Dummy			0.003	0.000	0.001	-0.000	-0.000	49,594
				(2.33)	(0.34)	(1.49)	(-2.20)	(-1.89)	

Table VIII: Post-Merger Real Outcomes and Ex ante Integration Risk

The table reports the results of OLS regressions in which the dependent variable is a measure of ex post real outcomes as noted in the second column. As our goal is to examine ex post outcomes for acquirers, we limit the sample to firms that were an acquirer in year t. We consider outcomes measured as changes for both a one-year horizon and a three year horizon, where the horizon begins in year t of the merger and ends in year $t+1$ or $t+3$. We consider the following outcomes: ex post changes operating income scaled by assets and ex post changes in SG&A/sales. All regressions include industry and year fixed effects, RHS variables are standardized prior to running regressions, and standard errors are clustered by industry.

Panel A: Acquisitions with Public Acquirer and Target

Row	Dep. Var.	Expected Post-Merger Integration Risk	Expected Synergy Integration Risk	Acquirer Expected Integration Risk	Target Expected Integration Risk	Log Assets	Log Age	Target Fraction of Acquirer	Pairwise TNIC Simil. Score	Expected Gain in Product Diff.	Market to Book	Document Size	Lagged Depend. Variable	Obs.
(1)	Yr 1 Δ OI/Assets	-0.005				0.007	0.003	-0.001	0.003	0.002	-0.006	-0.002	-0.264	5,466
		(-2.00)				(2.98)	(1.50)	(-0.23)	(1.50)	(1.68)	(-1.96)	(-1.42)	(-5.23)	
(2)	Yr 1 Δ OI/Assets		-0.003	-0.003	-0.002	0.007	0.003	-0.001	0.003	0.002	-0.006	-0.001	-0.266	5,466
			(-2.35)	(-1.67)	(-1.51)	(3.03)	(1.63)	(-0.22)	(1.59)	(1.68)	(-2.00)	(-0.73)	(-5.24)	
(3)	Yr 3 Δ OI/Assets	-0.003				0.012	0.004	-0.002	0.005	0.002	-0.004	-0.004	-0.397	5,466
		(-1.15)				(3.32)	(1.44)	(-0.41)	(2.24)	(0.95)	(-0.82)	(-1.74)	(-6.69)	
(4)	Yr 3 Δ OI/Assets		-0.004	-0.001	-0.001	0.012	0.004	-0.002	0.005	0.002	-0.004	-0.003	-0.398	5,466
			(-1.58)	(-0.55)	(-0.64)	(3.33)	(1.49)	(-0.41)	(2.38)	(0.93)	(-0.81)	(-1.23)	(-6.70)	
(5)	Yr 1 Δ SG&A/Sales	0.006				-0.007	0.001	-0.005	0.002	0.001	0.010	-0.003	-0.142	5,466
		(2.74)				(-2.62)	(0.64)	(-2.32)	(1.19)	(0.86)	(4.15)	(-2.39)	(-8.18)	
(6)	Yr 1 Δ SG&A/Sales		0.003	0.003	0.001	-0.007	0.001	-0.005	0.002	0.001	0.010	-0.004	-0.142	5,466
			(2.46)	(1.61)	(1.17)	(-2.62)	(0.46)	(-2.34)	(1.54)	(0.71)	(4.17)	(-2.34)	(-8.14)	
(7)	Yr 3 Δ SG&A/Sales	0.006				-0.013	0.005	-0.002	0.002	0.002	0.009	-0.005	-0.274	5,466
		(2.20)				(-3.34)	(2.10)	(-0.73)	(1.21)	(0.93)	(2.28)	(-1.92)	(-11.66)	
(8)	Yr 3 Δ SG&A/Sales		0.004	0.005	0.001	-0.013	0.005	-0.002	0.003	0.002	0.009	-0.006	-0.275	5,466
			(1.78)	(1.32)	(0.87)	(-3.35)	(1.94)	(-0.75)	(1.41)	(0.86)	(2.31)	(-2.07)	(-11.65)	

Panel B: All Acquirers

Row	Dependent Variable	Expected Integ-ration	Realized Integ-ration	Integ-ration Gap	Log Assets	Log Age	TNIC Total Simil.	Tobins Q	Lagged Dep. Var.	Obs.
(1)	Yr 1 Δ OI/Assets	-0.010	0.005		0.005	0.004	-0.000	-0.001	-0.177	21,454
		(-4.57)	(1.75)		(9.12)	(4.20)	(-1.01)	(-0.66)	(-11.89)	
(2)	Yr 1 Δ OI/Assets			-0.010	0.005	0.005	-0.000	-0.001	-0.177	21,454
				(-4.56)	(9.09)	(4.35)	(-1.13)	(-0.68)	(-11.87)	
(3)	Yr 3 Δ OI/Assets	-0.013	0.003		0.008	0.002	-0.000	-0.004	-0.315	21,454
		(-3.57)	(0.72)		(10.43)	(1.50)	(-1.62)	(-3.28)	(-15.43)	
(4)	Yr 3 Δ OI/Assets			-0.013	0.008	0.003	-0.000	-0.004	-0.314	21,454
				(-3.55)	(10.40)	(1.66)	(-1.76)	(-3.31)	(-15.40)	
(5)	Yr 1 Δ SG&A/Sales	0.005	0.001		-0.003	0.000	0.000	0.004	-0.152	21,454
		(2.17)	(0.22)		(-6.12)	(0.03)	(2.75)	(4.78)	(-19.50)	
(6)	Yr 1 Δ SG&A/Sales			0.005	-0.003	-0.000	0.000	0.004	-0.151	21,454
				(2.17)	(-6.11)	(-0.11)	(2.85)	(4.80)	(-19.49)	
(7)	Yr 3 Δ SG&A/Sales	0.004	0.005		-0.007	0.003	0.001	0.006	-0.275	21,454
		(1.32)	(1.13)		(-8.33)	(1.62)	(5.08)	(6.06)	(-25.06)	
(8)	Yr 3 Δ SG&A/Sales			0.004	-0.007	0.002	0.001	0.007	-0.275	21,454
				(1.32)	(-8.30)	(1.47)	(5.19)	(6.07)	(-25.03)	

Table IX: Post-Merger Restructuring and Ex ante Integration Risk

The table reports the results of OLS regressions in which the dependent variable is a measure of ex post divestiture or acquiring activity as noted in the second column. As our goal is to examine ex post outcomes for acquirers, we limit the sample to firms that were an acquirer in year t. We consider outcomes measured as changes for both a one-year horizon and a two year horizon, where the horizon begins in year $t+1$ after the merger and ends in year $t+2$ or $t+3$. The use of a forward window avoids having the calculation load on the year of the merger itself, and reflects our objective of examining longer-term outcomes. We consider the following restructuring variables: ex post increases in the incidence of the firm being a target and divesting (Panel A) or being an acquirer (Panel B). Where $N_{div,t}$ is the number of divestiture transactions a given firm has in year t, we compute the one-year increase in divestitures using the following logarithmic formula: $log[\frac{1+N_{div,t+1}}{1+N_{div,t}}]$. The two year growth is computed in an analogous fashion using year $t+2$ instead of year $t+1$. This form computes growth in a relative way while avoiding the overweighting of outliers. The number of acquisitions is computed in an analogous way. The key independent variables are the expected integration, and realized integration. We also consider the integration gap, which is expected integration minus realized integration. All regressions include industry fixed effects and year fixed effects, and standard errors are clustered by industry.

Panel A: Acquisitions and Subsequent Divestures with Public Targets

Row	Dep. Var.	Ex Ante Merger Integration Risk	Expected Synergy Integration Risk	Acquirer Expected Integration Risk	Target Expected Integration Risk	Log Assets	Log Age	Target Fraction of Acquirer	Pairwise TNIC Simil. Score	Expected Gain in Product Diff.	Market to Book	Document Size	Lagged Depend. Variable	Obs.
(1)	Yr 1 Δ Divestitures	0.019				0.086	0.039	0.006	-0.003	0.003	-0.002	-0.021	-0.600	5,288
		(2.87)				(13.45)	(3.01)	(1.87)	(-2.64)	(1.50)	(-1.35)	(-1.06)	(-16.01)	
(2)	Yr 1 Δ Divestitures		0.002	0.095	0.046	0.088	0.035	0.006	-0.002	0.002	-0.002	-0.031	-0.604	5,288
			(0.48)	(2.69)	(2.73)	(13.58)	(2.67)	(1.75)	(-1.76)	(1.23)	(-0.84)	(-1.40)	(-16.77)	
(3)	Yr 3 Δ Divestitures	0.003				0.086	0.004	0.008	-0.003	0.001	0.005	0.018	-0.692	4,441
		(0.39)				(10.36)	(0.24)	(1.72)	(-2.24)	(0.65)	(1.74)	(0.92)	(-14.33)	
(4)	Yr 3 Δ Divestitures		-0.006	0.097	0.020	0.087	-0.000	0.008	-0.002	0.001	0.005	-0.003	-0.697	4,441
			(-0.93)	(1.70)	(0.93)	(10.57)	(-0.01)	(1.67)	(-1.92)	(0.59)	(2.01)	(-0.13)	(-15.24)	

Panel B: All Acquirers and all Divestitures

Row	Dependent Variable	Expected Integration	Ex ante actual Integration	Integration Gap	Log Assets	Log Age	TNIC Total Similarity	Tobins Q	Lagged Dep. Var.	Obs.
(1)	Yr 1 Δ Divestitures	0.037	-0.043		0.057	0.032	-0.001	-0.003	-0.714	20,807
		(2.95)	(-2.54)		(20.22)	(7.32)	(-3.11)	(-2.75)	(-38.09)	
(2)	Yr 3 Δ Divestitures	0.064	-0.059		0.055	0.021	-0.002	0.001	-0.787	16,412
		(3.86)	(-2.84)		(15.82)	(3.68)	(-3.16)	(0.71)	(-42.16)	
(3)	Yr 1 Δ Divestitures			0.037	0.057	0.032	-0.001	-0.003	-0.714	20,807
				(2.95)	(20.20)	(7.36)	(-3.15)	(-2.77)	(-38.16)	
(4)	Yr 3 Δ Divestitures			0.064	0.055	0.021	-0.002	0.001	-0.787	16,412
				(3.86)	(15.83)	(3.67)	(-3.17)	(0.72)	(-42.27)	

Table X: Ex-ante Merger Integration Risk and Announcement Returns

The table reports the results of OLS regressions in which the dependent variable is either the announcement return of the acquirer (Panel A), the target (Panel B) or the combined acquirer and target only (Panel C). We consider announcement returns measured just on day $t = 0$, and also a 3-day, a 5-day, and a 7-day window, where all windows are centered around $t = 0$. The key independent variables of interest are the Synergy Integration Risk, and the Target and Acquirer Assets in Place Integration Risk. These is computed in a fashion analogous to the Expected Integration variable discussed in Table I with three primary exceptions. First, we compute integration risk for three firms: the acquirer, the target, and the hypothetical synergies of the post-merger firm (based on the combined firm vocabularies). Second, we use the distribution of word-pairs instead of single words, as only this approach can isolate the combined firm's synergies that are not present in the individual firms. This calculation is discussed in detail in Section II.B. The resulting measures of integration risk are ex ante measurable and target specific parts of the post merger firm based on assets in place and likely synergies. The remaining variables are discussed in Table I. All regressions include industry and year fixed effects, RHS variables are standardized prior to running regressions, and standard errors are clustered by industry.

Row	Group	Expected Post-Merger Integration Risk	Expected Synergy Integration Risk	Acquirer Expected Integration Risk	Target Expected Integration Risk	Log Assets	Log Age	Target Fraction of Acquirer	Pairwise TNIC Simil. Score	Expected Gain in Product Diff.	Market to Book	Document Size	Obs.
						Panel A: Acquirer Firm Announcement Returns							
(1)	Acquirer	-0.000				-0.004	0.001	-0.002	0.001	0.003	-0.001	0.000	5,808
	1 day	(-0.18)				(-3.62)	(1.45)	(-1.02)	(1.23)	(6.07)	(-1.10)	(0.05)	
(2)	Acquirer	-0.001				-0.008	0.000	-0.004	-0.000	0.002	-0.000	0.002	5,808
	3 days	(-1.08)				(-5.36)	(0.28)	(-1.92)	(-0.27)	(2.03)	(-0.47)	(1.72)	
(3)	Acquirer		-0.001	-0.001	0.001	-0.004	0.001	-0.002	0.001	0.002	-0.001	0.001	5,808
	1 day		(-1.29)	(-0.93)	(1.12)	(-3.63)	(1.49)	(-1.03)	(1.62)	(6.10)	(-1.09)	(1.03)	
(4)	Acquirer		-0.004	-0.001	0.002	-0.008	0.000	-0.004	0.000	0.001	-0.000	0.003	5,808
	3 days		(-2.00)	(-0.58)	(2.78)	(-5.40)	(0.26)	(-1.93)	(0.01)	(1.64)	(-0.42)	(2.13)	
						Panel B: Target Firm Announcement Returns							
(5)	Target	0.005				-0.041	0.001	0.001	-0.005	-0.004	-0.015	-0.005	5,808
	1 day	(3.07)				(-7.03)	(0.30)	(0.83)	(-3.23)	(-1.77)	(-6.50)	(-1.45)	
(6)	Target	0.009				-0.060	0.001	0.004	-0.008	-0.007	-0.017	-0.009	5,808
	3 days	(3.86)				(-8.56)	(0.22)	(2.24)	(-2.97)	(-2.54)	(-4.81)	(-2.51)	
(7)	Target		0.013	0.001	-0.003	-0.040	0.000	0.001	-0.006	-0.003	-0.015	-0.006	5,808
	1 day		(4.02)	(0.23)	(-2.09)	(-6.83)	(0.19)	(0.88)	(-3.46)	(-1.51)	(-6.45)	(-1.21)	
(8)	Target		0.022	0.003	-0.003	-0.059	0.000	0.004	-0.009	-0.007	-0.017	-0.010	5,808
	3 days		(4.32)	(0.62)	(-1.50)	(-8.39)	(0.07)	(2.33)	(-3.12)	(-2.31)	(-4.81)	(-2.03)	

Table XI: Fama MacBeth Monthly Return Regressions

The table displays Fama-MacBeth regressions form July 1997 to June 2010 in which the dependent variable is the firm's monthly excess stock return. The acquirer dummy is one if the firm was an acquirer in a merger or an acquisition of assets transaction in the previous one-year period (based on effective date and lagged 6 months for consistency with other variables). The key integration variables are from the past fiscal year, lagged using the minimum 6 month lag required in Davis, Fama, and French (2000). For each of the three integration variables we examine, we use a dummy variable indicating whether the given value is in the high tercile in the given year. Realized integration is the extent to which a firm's individual words appear within its own paragraphs in a distribution close to a uniform distribution. Expected integration is the extent to which a firm uses vocabulary that generally appears in a uniform distribution across paragraphs in all firms that use the given word in the economy in the given year. The integration gap is expected integration minus realized integration. We also consider cross terms based on the acquirer dummy and each integration variable. Finally, we include controls for the log book to market ratio, the log of firm market capitalization and the past one year stock return, where these variables are measured following Davis, Fama, and French (2000). All variables are ex ante measurable and quantities from any given fiscal year follow the lag structure of Davis, Fama, and French (2000). For example, any variable from a fiscal year ending in calendar year t will not be used to predict returns until July of year $t+1$. We discard penny stock firms from our sample if they have a stock price of one dollar or less.

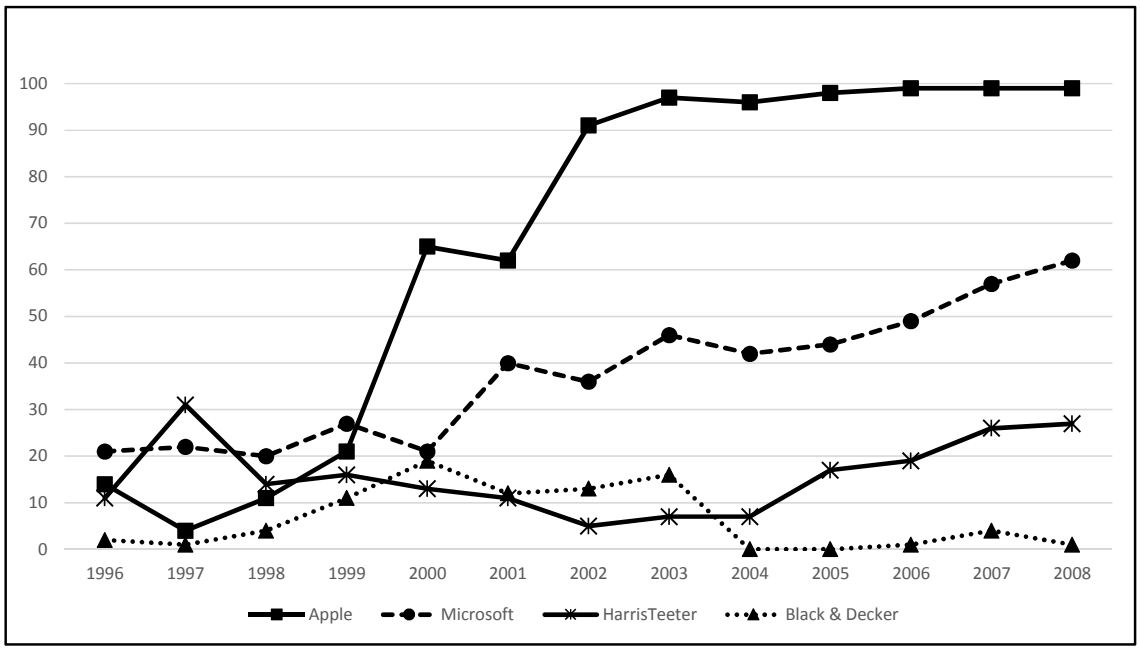
Row	Acquirer Dummy	High Ex ante Expected Integration	Acquirer Dum x Hi Expected Integration	High Ex ante Realized Integration	Acquirer Dum x Hi Realized Integration	High Ex ante Integration Gap	Acquirer Dum x Hi Integration Gap	Fraction Stock	Missing Fraction Stock	Log B/M Ratio	Log Size	Past Year Stock Return	Periods / Obs.
(1)	-0.167									0.142	-0.159	-0.014	144
	(-2.37)									(0.88)	(-0.83)	(-0.04)	626,877
(2)	-0.049	0.352	-0.335							0.149	-0.166	-0.012	144
	(-0.66)	(1.63)	(-2.76)							(0.95)	(-0.86)	(-0.03)	626,877
(3)	-0.088			0.251	-0.223					0.149	-0.156	-0.011	144
	(-1.11)			(1.35)	(-1.93)					(0.96)	(-0.82)	(-0.03)	626,877
(4)	-0.040	0.302	-0.297	0.090	-0.063					0.150	-0.163	-0.011	144
	(-0.50)	(1.72)	(-2.40)	(0.74)	(-0.55)					(0.96)	(-0.85)	(-0.03)	626,877
(5)	-0.103					0.176	-0.189			0.141	-0.163	-0.014	144
	(-1.46)					(1.63)	(-2.29)			(0.87)	(-0.84)	(-0.04)	626,877
(6)	0.034	0.302	-0.293	0.090	-0.063			-0.005	-0.071	0.149	-0.165	-0.011	144
	(0.31)	(1.72)	(-2.37)	(0.74)	(-0.55)			(-1.54)	(-0.78)	(0.96)	(-0.86)	(-0.03)	626,877
(7)	-0.029					0.176	-0.182	-0.005	-0.074	0.140	-0.165	-0.014	144
	(-0.31)					(1.64)	(-2.22)	(-1.49)	(-0.82)	(0.87)	(-0.86)	(-0.04)	626,877

Table XII: Fama MacBeth Monthly Return Regressions (by Integration Quintiles)

The table displays Fama-MacBeth regressions form July 1997 to June 2010 in which the dependent variable is the firm's monthly excess stock return. We present results for quintiles based on sorting firms based on their expected integration (Panel A), realized integration (Panel B), and the integration gap (Panel C). The acquirer dummy is one if the firm was an acquirer in a merger or an acquisition of assets transaction in the previous one-year period (based on effective date and lagged 6 months for consistency with other variables). The key integration variables are from the past fiscal year, lagged using the minimum 6 month lag required in Davis, Fama, and French (2000). Realized integration is the extent to which a firm's individual words appear within its own paragraphs in a distribution close to a uniform distribution. Expected integration is the extent to which a firm uses vocabulary that generally appears in a uniform distribution across paragraphs in all firms that use the given word in the economy in the given year. The integration gap is expected integration minus realized integration. Finally, we include controls for the log book to market ratio, the log of firm market capitalization and the past one year stock return, where these variables are measured following Davis, Fama, and French (2000). All variables are ex ante measurable and quantities from any given fiscal year follow the lag structure of Davis, Fama, and French (2000). For example, any variable from a fiscal year ending in calendar year t will not be used to predict returns until July of year $t+1$. We discard penny stock firms from our sample if they have a stock price of one dollar or less.

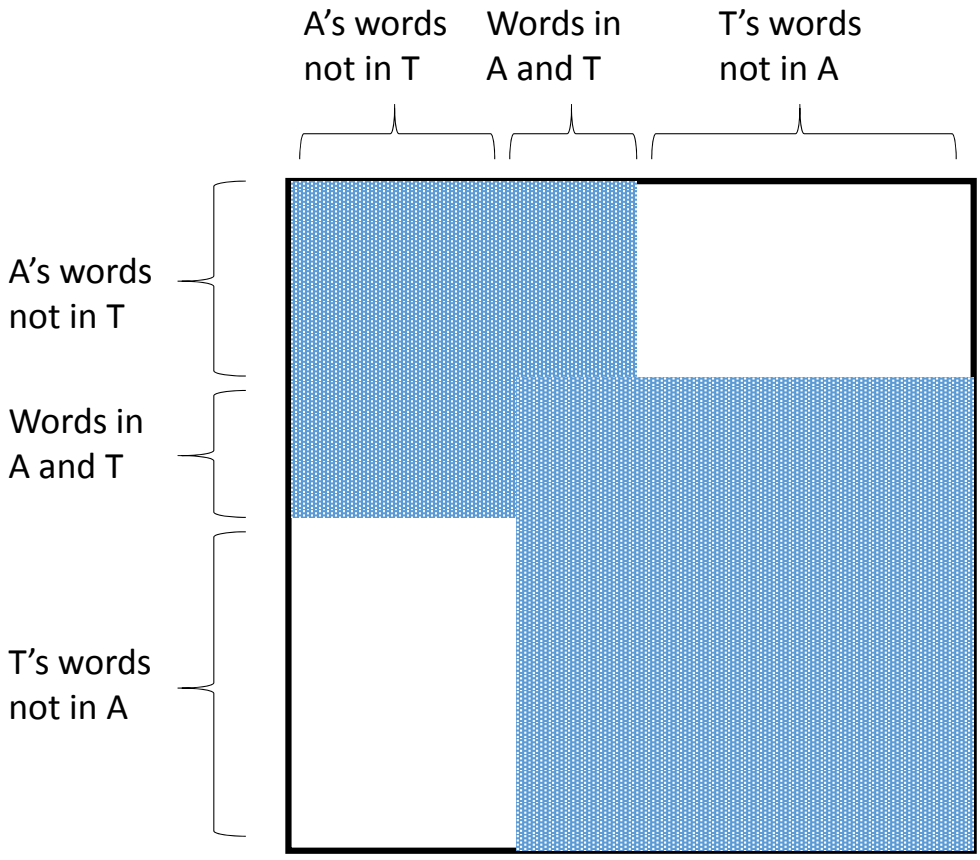
Row	Quintile	Acquirer Dummy	Log B/M Ratio	Log Size	Past Year Stock Return	Periods / Obs.
				Panel A: Expected Integration Quintiles		
(1)	Quintile 1	-0.097	0.101	-0.210	-0.159	144
		(-0.89)	(0.84)	(-1.23)	(-0.35)	126,513
(2)	Quintile 2	0.084	0.127	-0.165	0.211	144
		(0.78)	(0.78)	(-0.88)	(0.54)	126,609
(3)	Quintile 3	-0.141	0.110	-0.139	-0.026	144
		(-1.26)	(0.63)	(-0.75)	(-0.06)	125,951
(4)	Quintile 4	-0.192	0.263	-0.127	-0.079	144
		(-1.70)	(1.58)	(-0.60)	(-0.20)	124,792
(5)	Quintile 5	-0.504	0.164	-0.186	-0.104	144
		(-3.78)	(0.84)	(-0.72)	(-0.24)	123,012
				Panel B: Realized Integration Quintiles		
(6)	Quintile 1	0.008	0.152	-0.147	0.067	144
		(0.07)	(1.07)	(-0.91)	(0.14)	126,617
(7)	Quintile 2	0.010	0.046	-0.101	0.076	144
		(0.09)	(0.29)	(-0.54)	(0.18)	126,600
(8)	Quintile 3	-0.112	0.142	-0.262	0.010	144
		(-0.96)	(0.93)	(-1.34)	(0.02)	125,061
(9)	Quintile 4	-0.326	0.203	-0.128	-0.012	144
		(-3.05)	(1.24)	(-0.62)	(-0.03)	124,673
(10)	Quintile 5	-0.449	0.207	-0.141	-0.194	144
		(-3.51)	(1.02)	(-0.60)	(-0.49)	123,926
				Panel C: Integration Gap Quintiles		
(11)	Quintile 1	-0.093	0.150	-0.166	-0.217	144
		(-0.94)	(1.37)	(-0.95)	(-0.54)	126,026
(12)	Quintile 2	-0.097	0.165	-0.182	-0.102	144
		(-0.86)	(0.88)	(-0.90)	(-0.23)	126,319
(13)	Quintile 3	-0.187	0.130	-0.117	0.069	144
		(-1.48)	(0.81)	(-0.57)	(0.19)	125,244
(14)	Quintile 4	-0.188	0.158	-0.163	0.039	144
		(-1.53)	(0.83)	(-0.78)	(0.09)	125,087
(15)	Quintile 5	-0.320	0.108	-0.185	0.092	144
		(-2.82)	(0.53)	(-0.86)	(0.21)	124,201

Figure 1:



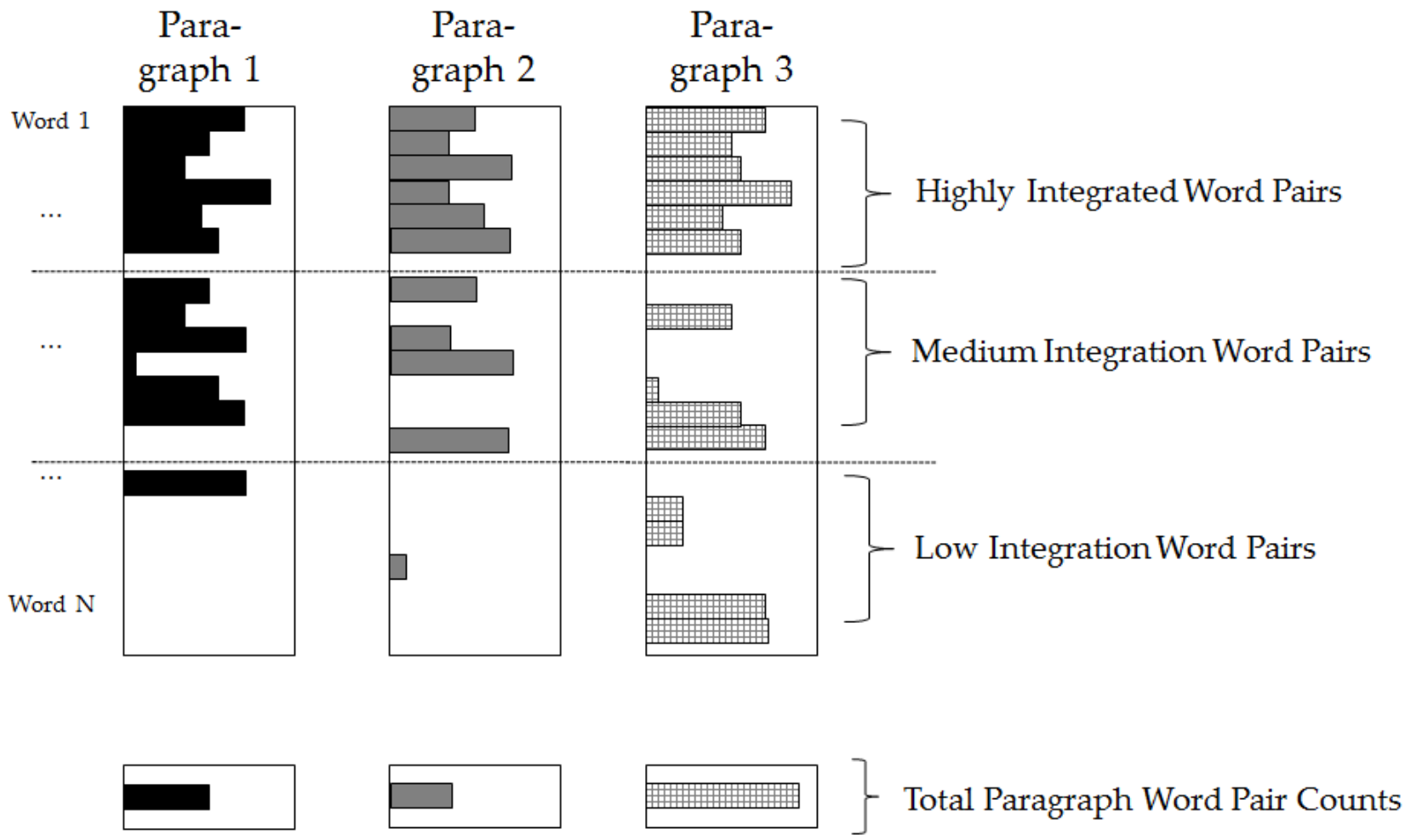
Notes: The Figure displays the realized integration over time for four sample firms of interest: Apple, Microsoft, Harris Teeter, and Black and Decker. The figure illustrates Apple's dramatic rise in integration compared to Microsoft's more gradual ascent. Conversely, the figure shows that Harris Teeter and Black and Decker are examples of firms with longer-term persistent levels of lower integration.

Figure 2:



Notes: The Figure illustrates the role of using word-pairs instead of individual words when computing integration risk for a specific transaction (for example when assessing announcement returns). The figure shows that all words in the union of the acquirer and target fall into three baskets: those just used by A, those just used by T, and those used by both. By examining pairs, the figure shows that the word-pairs in dark shading are those that appeared in either the acquirer, the target, or both prior to the hypothetical transaction. However, the word-pairs in white did not exist in either firm prior to the transaction. These word-pairs are thus associated with the merger synergies, and assessing the expected integration level of these word-pairs is thus a measure of how much integration risk exists for a firm trying to realize these synergies. We note that a similar quantity cannot be computed using single words rather than word-pairs as all single words used in the diagram do exist in the 10-K of either the acquirer or the target firm.

Figure 3:



Notes: The Figure visually illustrates examples of highly integrated and highly non-integrated word-pairs based on their distribution across paragraphs. Word-pairs that are thoroughly mixed in all paragraphs are integrated into the firm's product offerings fully. Word-pairs that appear only in one, or a small number, of paragraphs are not integrated. At the bottom of the figure, we depict the distribution of total word-pair counts across paragraphs, which motives our measure of word-pair level integration based on distributional proximity of a word-pair's distribution to this aggregate length distribution.