

Innovation Activities and the Incentives for Vertical Acquisitions and Integration

Laurent Frésard, Gerard Hoberg and Gordon Phillips*

August 31, 2015

ABSTRACT

We examine the incentives for firms to vertically integrate through acquisitions and production. We develop a new firm-specific measure of vertical relatedness and integration using 10-K product text. We find that firms in high R&D industries are less likely to become targets in vertical acquisitions or to vertically integrate. These findings are consistent with the idea that firms with unrealized innovation avoid integration to maintain *ex ante* incentives to invest in intangible assets and to keep residual rights of control. In contrast, firms in high patenting industries with mature product markets are more likely to vertically integrate, consistent with control rights being obtained by firms to facilitate commercialization of already realized innovation.

*University of Maryland, University of Southern California, and University of Southern California and National Bureau of Economic Research, respectively. Frésard can be reached at lfresard@rhsmith.umd.edu, Hoberg can be reached at gerard.hoberg@marshall.usc.edu and Phillips can be reached at gordon.phillips@marshall.usc.edu. We thank Yun Ling for excellent research assistance. For helpful comments, we thank Kenneth Ahern, Jean-Noel Barrot, Thomas Bates, Giacinta Cestone, Robert Gibbons, Oliver Hart, Thomas Hellmann, Adrien Matray, Sébastien Michenaud, Steve Tadelis, Ivo Welch and seminar participants at the National Bureau of Economic Research Organizational Economics Meetings, 2015 International Society for New Institutional Economics Meetings, 2014 American Finance Association Meetings, Arizona State University, Carnegie Mellon, Dartmouth College, Humboldt University, IFN Stockholm, Harvard-MIT Organizational Economics joint seminar, Tsinghua University, University of Alberta, University of British Columbia, University of California Los Angeles Universidad de los Andes, University of Maryland, University of Washington, VU Amsterdam, and Wharton. All errors are the authors alone. Copyright ©2015 by Laurent Frésard, Gerard Hoberg and Gordon Phillips. All rights reserved.

The scope of firm boundaries and whether to organize transactions within the firm (integration) or by using external purchasing is of major interest in understanding why firms exist. Williamson (1971), Williamson (1979) and Klein, Crawford, and Alchian (1978) pioneered this area through their theory of transaction cost economics and ex post holdup given contractual incompleteness. Firms choose the organizational form that minimizes transaction costs and ex post holdup. Grossman and Hart (1986) in their property rights theory of the firm show that control rights are key to understanding firm boundaries and their influence on ex ante investment. They show that ex ante incentives for a firm to invest in relationship-specific assets are reduced under vertical integration for the firm that gives up its residual rights of control to the other contracting firm.¹

These theories particularly apply to innovation activities as technological investments are typically relationship-specific, not fully contractible and often unverifiable (Aghion and Tirole (1994) or Acemoglu (1996)).² In this paper, we argue and provide evidence that the costs and benefits of vertical integration, and hence firm boundaries, are related to the stage of development of innovation. In particular, we find that the distinction between unrealized innovation in the form of R&D and realized innovation characterized by legally enforceable patents is a key empirical determinant of firms' vertical organization and acquisitions.

We capture variations in firm boundaries based on new firm-specific measures of vertical relatedness that we construct using text-based analysis of firm 10-K product descriptions filed with the Securities and Exchange Commission (SEC), and vertically-linked product descriptions from the Bureau of Economic Analysis (BEA) Input-Output tables. This allows us to precisely determine vertical relatedness across firm pairs and identify which mergers and acquisitions are vertically related. This framework also allows us to develop a new measure of vertical inte-

¹Holmstrom and Milgrom (1991) and Holmstrom and Milgrom (1994) also emphasize the role of incentives in firm structure. Gibbons (2005) summarizes the large literature and highlights that the costs and benefits of vertical integration depend on transactions costs, rent seeking, contractual incompleteness, and the specificity of the assets involved in transactions.

²As highlighted by Acemoglu, Aghion, Griffith, and Zilibotti (2010) innovative activities are subjected to the types of problems featured by the transaction costs and property right theories.

gration at the firm-level based on whether firms use product vocabulary that spans vertically related markets. Because 10-Ks are updated annually, we relate firm vertical organization to the stage of innovation as these activities dynamically evolve.³ This more direct approach is not possible using measures of vertical integration based on static methods that rely on SIC or NAICS codes.

Using a sample of almost 7,000 publicly-traded firms over the 1996-2008 period, we find strong evidence that firms in R&D intensive industries are less likely to be acquired in vertical transactions. In contrast, firms in patent intensive industries are more likely to be targeted in vertical transactions. In addition, we show that vertical acquisitions tend to occur at times when target firms have accumulated more patents, in contrast to non-vertical acquisitions which are diametrically opposite.

The distinction between unrealized and realized innovation also matters for observed firm-level vertical integration. We find strong evidence that firms in R&D intensive industries are less likely to be vertically integrated whereas firms in high patenting industries are more likely to be vertically integrated. These findings are economically large: In our baseline specification, firm vertical integration decreases by 10% following a one-standard deviation increase in R&D intensity, and increases by 7% following a one-standard deviation increase in patenting intensity. These findings are consistent with firms with unrealized innovation avoiding integration to maintain *ex ante* incentives to invest in intangible assets and maintain residual rights of control as in Grossman and Hart (1986).

These results are consistent with a simple model based on incomplete contracting that links R&D and patent intensity to vertical acquisitions. The decision of two firms to integrate vertically depends on the stage of development of the specific asset exchanged in their relationship. The model predicts that when the asset is still in the form of R&D (unrealized innovation), firms optimally will not integrate.⁴

³Many studies in industrial organization take the single-industry approach. Earlier studies include Monteverde and Teece (1982) focusing on automobile manufacturing, Masten (1984) focusing on airplane manufacturing, and Joskow (1987) focusing on coal markets. More recent studies include Lerner and Merges (1998) focusing on biotechnology, Baker and Hubbard (2003) focusing on trucking, or Hortacsu and Syverson (2007) focusing on the cement industry.

⁴Acemoglu (1996) argues that technological investments are partner-specific, thus creating

Separation maintains ex ante incentives for the upstream firm to invest in product development, and to maintain residual rights of control. In contrast, when the asset is a patented realized innovation, the owner has more legally enforceable residual rights of control as in Grossman and Hart (1986). At this mature stage, incentives for further development are less important compared to incentives to commercialize the innovation. Hence, integration optimally allocates the residual rights of control to the downstream firm that will use the realized innovation.

To further assess our conclusions, we follow Bloom, Schankerman, and van Reenen (2013a) and exploit variation in R&D tax credits across U.S. states to construct an instrument for industries' R&D intensity. Confirming our baseline results, this instrumental variables framework indicates that an increase in industry R&D significantly lowers the likelihood that a firm will be acquired in a vertical acquisition. Also, higher R&D intensity leads to significantly less firm-level vertical integration.

The distinction between high R&D and patents is empirically nontrivial. High R&D does not necessarily lead to high patenting rates, and patent rates vary across industries. As reported by Cohen, Nelson, and Walsh (2000) in a survey of 1,478 R&D labs, high R&D may not lead to patents due to concerns about appropriability. Their survey points to the ability of others to work around patents using information conveyed by the patent application, causing managers not to patent in many industries. Consistent with property rights for realized innovation being important, we find that the link between patents and vertical integration is only present in industries where patents provide effective coverage.

Two recent examples of the effects we document for acquisitions are Microsoft's recent purchases of Skype and Nokia. Skype specialized in making VoIP phone and video calls over the Internet. After purchasing Skype, Microsoft integrated Skype into Windows and also into Windows phones. Regarding Nokia in 2013, one insider indicated that the deal between the two companies would help to bring the

relationship-specific assets that are difficult to contract on. Allen and Phillips (2000) and Kale and Shahrur (2007) show that R&D increases with interaction between alliance partners, consistent with the needed R&D incentives to develop relationship-specific assets.

“hardware closer to the operating system and achieve a tighter integration.” Buying firms to gain control of their realized innovations facilitates commercialization either through reduced ex post hold-up or increased commercialization incentives.⁵

An industry that exemplifies our findings regarding dynamic vertical integration is the network equipment industry, which includes Cisco, Broadcom, Citrix, Juniper, Novell, Sycamore, and Utstarcom. During our sample, and using our measures, we find that firms in this industry jointly experienced (A) levels of R&D that peaked and began to decline, (B) levels of patenting activity that rose four to five fold, and (C) levels of vertical integration that also rose four to five fold. The conversion of unrealized innovation into realized patented innovation reduced the incentives for relationship-specific investment, and increased the incentives to vertically integrate in order to transfer control rights to the party commercializing the patents.⁶

We also consider the role of supply chain stability and maturity. For example, assuming the benefits of integration derive from ongoing operations, an unstable supply chain can reduce the horizon during which benefits are realized. Because reorganization typically entails a high level of fixed costs, firms in unstable supply chains should thus be less willing to vertically integrate as the duration of gains may not be adequate to cover the high fixed costs. We find empirical support for the proposed positive link between maturity and vertical integration. In particular, both vertical acquisitions and vertical integration are positively related to maturity as captured by firm age, lower market-to-book ratio, and more tangible assets.

Our results highlight that firms’ vertical acquisitions are related to the stage of development of *intangible* assets. This motive is distinct from other motives for acquisitions including neoclassical theories, agency theories, and horizontal theories.⁷ Consistent with seminal theories of the firm, our results indicate that both

⁵See <http://www.businessinsider.com/why-microsoft-bought-skype-an-insider-explains-2011-5>.

⁶The 2014 IBISWORLD industry report on the Telecommunication Networking Equipment Manufacturing confirms the trend towards more integration in this market. Firms in this industry seek to offer “end-to-end” and “all-in-one” solutions.

⁷See Maksimovic and Phillips (2001), Jovanovic and Rousseau (2002), and Harford (2005) for neoclassical and q theories and Morck, Shleifer, and Vishny (1990) for an agency motivation for acquisitions, and Phillips and Zhdanov (2013) for a recent horizontal theory of acquisitions.

ex post effects (Klein, Crawford, and Alchian (1978) and Williamson (1979)) and *ex ante* incentives to assign residual rights of control (Grossman and Hart (1986)) are empirically important. *Ex ante* effects occur when innovation is unrealized and it is important to encourage further R&D. All else equal, firms are more likely to remain separate in R&D intensive industries. In contrast, *ex post* effects occur when innovation is realized, as integration minimizes the potential for hold up costs and *ex post* bargaining over the already patented innovation. As a result, firms are more likely to integrate when the innovation is realized and patented.

Our findings contribute to a large literature examining the determinants of vertical integration (see Lafontaine and Slade (2007) and Bresnahan and Levin (2012) for surveys), and to recent papers linking vertical integration to innovation and intangible assets (Aghion and Tirole (1994), Acemoglu, Aghion, Griffith, and Zilibotti (2010)⁸, and Atalay, Hortacsu, and Syverson (2014)). Our results are complementary to the findings of Atalay, Hortacsu, and Syverson (2014). Using comprehensive data on ownership structure, production, and shipment patterns, they report that upstream units ship only small shares of their outputs to own-firm downstream plants. They provide suggestive evidence in support of intangible capital being important in vertical organization.⁹ Our results complement theirs by providing direct evidence that the stage of development of intangible assets created through innovation are relevant determinants of vertical integration.

Our paper also adds to the literature on vertical mergers and acquisitions. Fan and Goyal (2006) examine stock market reactions to vertical deals where vertical integration is identified at the industry level through SIC codes and the Input-Output tables. Kedia, Ravid, and Pons (2011) show that vertical mergers create more value in imperfectly competitive markets. Ahern (2012) shows that division of stock-market gains in mergers is determined in part by customer or supplier bargain-

⁸These authors show that, in a sample of UK manufacturing firms, the intensity of backward integration is positively (negatively) related to the R&D intensity of the downstream (upstream) industry. Our approach is complementary as we focus on the stage of innovation, vertical acquisitions, and within-firm integration in a large sample of U.S. firms.

⁹Specifically, they show a relative decline in non-production workers in acquired establishments that are vertically related. They also show an increase in products that were made by the acquiring firm previously in the acquired firms' establishments.

ing power. Ahern and Harford (2013) examine how supply chain shocks translate into vertical merger waves. We show that the stage of development of innovation is important to predicting vertical acquisitions. Our results thus complement Bena and Li (2013) and Seru (2014), who examine the impact of acquisitions on ex post innovation rates, and Phillips and Zhdanov (2013) who examine how acquisitions affect the incentives for R&D investments.

Our last contribution is methodological. By linking vocabulary in firm business descriptions to vocabulary describing commodities in the Input-Output tables, we are able to identify vertical relatedness directly at the firm level. Existing measures based on NAICS or SIC not only fail to provide firm-level measures, but are further problematic because they are based on production processes and not the products themselves.¹⁰ Our new measures rely neither on the quality of the Compustat segment tapes, nor on the quality of the NAICS classification.¹¹ Our focus on vertical links economically extends the work of Hoberg and Phillips (2015), who examine horizontal links using 10-K text. We further extend this work methodologically by providing a general framework for combining firm textual descriptions with any third source of textual network data such as BEA data. This framework can be used to construct firm-by-firm relatedness networks using any existing networks with verbal descriptions even if such data sources do not provide firm links.

The remainder of this paper is organized as follows. Section II develops a simple model of vertical integration. Section III presents the data and develops our measures of vertical relatedness. Section IV examines the effect of innovation activities on vertical acquisitions, and Section V examines the impact on firm vertical integration. Section VI concludes.

¹⁰See <http://www.naics.com/info.htm>. The Census Department states “NAICS was developed to classify units according to their production function. NAICS results in industries that group units undertaking similar activities using similar resources but does not necessarily group all similar products or outputs.”

¹¹Hyland (1999) and Villalonga (2004) show that the Compustat segment database has serious reporting biases. Measures of vertical integration using SIC and NAICS rely heavily on this database.

II A Simple Model of Integration

To illustrate the contrasting effects of realized and unrealized innovation on firm integration decisions, we develop a simple dynamic incomplete contracting model of vertical acquisition using the framework introduced by Grossman and Hart (1986). The model is meant to illustrate the trade-offs of vertical integration versus separation over time.

Consider an upstream supplier and a downstream producer. At each time t , they cooperate to produce a product at a base price P_t^b . The sale price P_t that can be charged on consumers further depends on commercialization and product integration investments chosen by the downstream firm as well as R&D investments by the upstream firm that can result in new patentable features. In the spirit of Grossman and Hart (1986) and Aghion and Tirole (1994), we assume that both R&D and commercialization investments are relationship-specific, non-contractible and non-verifiable. At each period, firms can either operate as separate entities or can decide to integrate.¹² Integration implies the acquisition of a firm (or the patent) from a firm by the other firm. The party that sells its assets is called the target and it loses control rights over the assets sold, and thus makes no further relationship-specific investment.

For each t , the upstream supplier chooses an x_t amount of R&D effort with a cost $k_t = c(x_t) = Sx_t^g$. We assume x_t is the non-contractible portion of R&D effort. Thus, if the downstream producer acquires the upstream supplier, x_t will be equal to zero.¹³ The downstream producer chooses an amount y_t of commercialization investment that can also boost the price of the product with a cost $m_t = c(y_t) = Ry_t^h$. Commercialization investments can include for instance marketing the product, building

¹²Our model can be thought of as a model of one firm doing R&D which results in a patent. This patent can be used in the supplier's production process to improve what is sold to the downstream firm. Thus, integration can be viewed as either a bundled sale of all the assets of the target or the sale of a patent that can be separated from the target firm and used by the downstream firm to improve its product. This would come with some cost associated with using for the patent that varies with ownership of the patent or the bundled assets. We discuss these potential ex post costs more later.

¹³The contractible portion of R&D effort need not be equal to zero. For simplicity, we focus on the non-contractible portion.

a new factory, and hiring sales people. We assume that both $g > 1$ and $h > 1$ so that costs are convex. The discount rate is r .

The base price P_t^b takes a value in the set $\{P_0, P_1, \dots, P_N\}$, with $P_s < P_{s+1}$ ($0 \leq s \leq N - 1$) and $P_{s+1} - P_s < P_s - P_{s-1}$ ($0 \leq s \leq N - 1$). A success in R&D at time t leads to new features and product enhancements. These product enhancements result in a legally enforceable patent, and boost the base price from P_s to P_{s+1} ($0 \leq s \leq N - 1$). Additional product features and extensions have a positive but decreasing effect on prices. We use X_t to denote the result of R&D investment which is realized and observed by both parties at the end of time period t , such that $X_t = 1$ corresponds to a success and $X_t = 0$ to a failure. The probability of success is determined by the R&D investments $p(X_t = 1) = x_t$.

For simplicity, we assume that the increase in price resulting from commercialization investment is deterministic, and it increases the base price P_t^b by an amount y_t if the firms are separate, and $\rho(y_t)$ if the firms are integrated. Both the level of price impact and the marginal product of commercialization investments are higher under integration, such that $\rho(y_t) > y_t$ and $\rho'(y_t) > 1$.¹⁴ The bargaining power of the upstream supplier is α (and the downstream producer $1 - \alpha$) in both the ex-ante acquisition negotiations that result in the integration of the two firms, and ex-post renegotiation for splitting total surplus when firms are separate.

The model's timing is summarized in Figure 1. At time t , given the outcome of the R&D investments by the upstream supplier in the last period X_{t-1} , we have:

1. The downstream producer decides whether to acquire the upstream supplier,

¹⁴This assumption can arise from the supplier not cooperating fully (withholding some information or selling related products to other firms) with the downstream firm if separate. We do not model the specific reason for the marginal product of commercialization expenditures being higher under integration. In the end, what is crucial is that the marginal product is higher for some types of expenditures if one firm has full control of the assets which can include a patent that is used in the production process. Clearly this is a crucial assumption but one that is likely to be satisfied for production when timely delivery of components are important and when the quality of the engineers or people involved in the production of the components cannot be perfectly observed. It would also be satisfied in situations when it is difficult to contract on all aspects of product quality as in the recent case of Boeing and other firms reintegrating with some of their suppliers given supply chain problems (See: <http://www.industryweek.com/companies-amp-executives/rebalancing-business-model>.)

and if so, negotiates with the supplier based on each party's bargaining power.

2. R&D investments x_t and commercialization investment y_t are decided by both parties as ex-ante investments.¹⁵
3. Renegotiation occurs if firms are separated.
4. By the end of the period, the success of R&D investments is realized, so that at the beginning of next period $t + 1$, both firms observe the value of X_t .

The realization of R&D and the grant of a patent is key to determining whether firms will integrate or remain separate. Since we make the assumption that X_t is realized at the end of each period, the final price charged on consumers is equal to $P_t = P_t^b(1 + y_t)$ under separation and $P_t = P_t^b(1 + \rho(y_t))$ under integration, with the base price $P_t^b = P_N$ if the last period base price is $P_{t-1}^b = P_N$ or $P_t^b = P_s + (P_{s+1} - P_s)X_{t-1}$ if the last period base price is $P_{t-1}^b = P_s < P_N$. Note that the base price is a contingent variable given the last-period R&D outcome X_{t-1} .

We model the decision of the producer to acquire the supplier and integrate (I) as a real option that, when exercised, is costly to reverse. We denote $I = 1$ as the situation where firms are integrated, and $I = 0$ when firms remain separate. In line with Grossman and Hart (1986), the integration decision is made to protect the two parties' investments in the relationship and to maximize total surplus. Firms thus do not integrate until the marginal benefit of staying separate decreases and is lower than that of integrating. Because product enhancements are cumulative, integration will also be positively related to firm maturity. We solve the model in Appendix 1 and discuss the central prediction below.

The central prediction of the model, shown as Proposition 1 in Appendix 1 is that R&D expenditures are higher when the firms are separate, while commercialization and product integration expenditures are higher when firms are integrated. We

¹⁵We could equivalently consider the case where the upstream firm buys the downstream firm. This would occur if the downstream firm does the R&D and the upstream firm customizes the product features before supplying the product. Note that this is not a crucial assumption. The model can thus be applied in either direction. We focus on the case of the downstream firm buying the upstream firm for simplicity, which is empirically the most frequent case as the previously cited Industry Week article notes.

show in Appendix 1 how the integration decision depends on the product price over time. Proposition 2 shows that when $P_t^b = P_N$ (i.e., the maximum price) both firms prefer to be integrated. This result arises because at that price the marginal effect of R&D on the price is zero.¹⁶ Formally, we show that $V(P_N) = V(P_N; I = 1) > V(P_N; I = 0)$ with V denoting the value function that measures total surplus. Moreover, we show as Proposition 3 in Appendix 1 that when P_t^b is below P_N , there exists a state s^* such that $V(P_s) = V(P_s; I = 1) \geq V(P_s; I = 0)$ for any $s \geq s^*$, and $V(P_s) = V(P_s; I = 1) < V(P_s; I = 0)$ for any $s < s^*$. The state s^* is the triggering state for integration.

This equilibrium is illustrated in Figure 2. Intuitively, separation is optimal when further incentives for R&D (x) benefit the overall relationship. In that case, separation maintains ex ante incentives for the upstream supplier to invest in R&D. Separation optimally allocates residual rights of control to the party whose incentives are more important (the upstream supplier). In contrast, when the asset is more fully developed and its features are protected by a patent (i.e. higher state s resulting from successful R&D), incentives for further R&D by the supplier (x) decline because of the decreasing marginal effect of R&D on the product price. At that time, the incentives for the downstream producer to spend on commercialization to further boost the product price (y) increases. Yet, without legal control rights on the asset (i.e. ownership of the patent), the producer faces hold up risk from the supplier. To encourage commercialization incentives, it is thus optimal for the overall relationship to allocate the residual rights of control to the downstream producer, whose incentives are more important. Hence, integration maximizes total surplus. The model thus delivers the following central prediction:

Central Prediction: *Firms are likely to remain separate when innovation is unrealized and R&D is important. Firms are more likely to be integrated when the innovation is realized and is protected by patents.*

¹⁶What is necessary is that marginal product of the non-contractible R&D declines over time such that the gain from R&D is less than the cost of not-integrating and getting the benefits of commercialization.

We test this proposition using new text-based measures of vertical relatedness, and by examining the distinct roles played by R&D and patenting intensity. However, we note that varying the assumptions about contractibility and how the marginal products of innovation and commercialization evolve will give different predictions. Hence the model is mainly provided to illustrate the economic forces that deliver this central prediction.

III Data and Methodology

We consider multiple data sources: 10-K business descriptions, Input-Output (IO) tables from the Bureau of Economic Analysis (BEA), COMPUSTAT, SDC Platinum for transactions, and data on announcement returns from CRSP.

A Data from 10-K Business Descriptions

We start with the Compustat sample of firm-years from 1996 to 2008 with sales of at least \$1 million and positive assets. We follow the same procedures as Hoberg and Phillips (2015) to identify, extract, and parse 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. Item 101 of Regulation S-K requires business descriptions to accurately report (and update each year) the significant products firms offer. We thus obtain 74,379 firm-years in the merged Compustat/Edgar universe.

B Data from the Input-Output Tables

We use both commodity text and numerical data from the Input-Output (IO) tables from the BEA, which account for the dollar flows between all producers and purchasers in the U.S. economy (including households, the government, and foreign buyers of U.S. exports). The tables are based on two primitives: ‘commodity’

outputs (any good or service) defined by the Commodity IO Code, and producing ‘industries’ defined by the Industry IO Code. In 2002, there were 424 distinct commodities and 426 industries in the “Make table”, which reports the dollar value of each commodity produced by a given industry. There are 431 commodities purchased by 439 industries or end users in the Use table in 2002, which reports the dollar value of each commodity purchased by each industry.¹⁷ We compute three data structures from the IO Tables: (1) Commodity-to-commodity (upstream to downstream) correspondence matrix (V), (2) Commodity-to-word correspondence matrix (CW), and (3) Commodity-to-‘exit’ (supply chain) correspondence matrix (E).

In addition to the numerical values in the BEA data, we use an often overlooked resource: the ‘Detailed Item Output’ table, which verbally describes each commodity and its sub-commodities. The BEA also provides the dollar value of each sub-commodity’s total production and a commodity’s total production is the sum of these sub-commodity figures.¹⁸ Each sub-commodity description uses between 1 to 25 distinct words (the average is 8) that summarizes the nature of the good or service provided.¹⁹ Table I contains an example of product text for the BEA ‘photographic and photocopying equipment’ commodity. We label the complete set of words associated with a commodity as ‘commodity words’.

[Insert Table I Here]

We follow the convention in Hoberg and Phillips (2015) and only consider nouns and proper nouns. We then apply four additional screens to ensure our identification of vertical links is conservative. First, because commodity vocabularies identify a stand-alone product market, we manually discard any expressions that indicate a

¹⁷An industry can produce more than one commodity: in 2002, the average (median) number of commodities produced per industry is 18 (13). Industry output is also concentrated as the average commodity concentration ratio is 0.78. Costs are reported in both purchaser and producer prices. We use producer prices. There are seven commodities in the Use table that are not in the Make table including for example compensation to employees. There are thirteen ‘industries’ in the Use table that are not in the Make table. These correspond to ‘end users’ and include personal consumption, exports and imports, and government expenditures.

¹⁸There are 5,459 sub-commodities and 427 commodities in 2002. The average number of sub-commodities per commodity is 12, the minimum is 1 and the maximum is 154.

¹⁹For instance, the commodity ‘Footwear Manufacturing’ (IO Commodity Code #316100) has 15 sub-commodities including those described as ‘rubber and plastics footwear’ and ‘house slippers’.

vertical relation such as ‘used in’, ‘made for’ or ‘sold to’. Second, we remove any expressions that indicate exceptions (e.g, we drop phrases beginning with ‘except’ or ‘excluding’). Third, we discard uninformative common words from commodity vocabularies.²⁰

Finally, we remove any words that do not frequently co-appear with the other words in the given commodity vocabulary. This further ensures that horizontal links or asset complementarities are not mislabeled as vertical links. We compute the fraction of times each word in a given IO commodity co-appears with other words in the same IO commodity when the given word appears in a 10-K business description (using all 10-Ks from 1997 to avoid any look ahead bias). We then discard words in the bottom tercile by this measure (the broad words). For example, if there are 21 words in an IO commodity description, we would discard 7 of the 21 words.²¹ We are left with 7,735 commodity words that identify vertically related product markets.

The ‘Detailed Item Output’ table also provides the economic importance of each commodity word. We compute the *relative* economic contribution of a given sub-commodity (ω) as the dollar value of its production relative to its commodity’s total production. Each word in a sub-commodity’s textual description is assigned the same ω . Because a word can appear in several sub-commodities, we sum its ω ’s within a commodity. A given commodity word is economically more important if this fraction is high. We define the commodity-word correspondence matrix (CW) as a three-column matrix containing: a commodity, a commodity word, and its economic importance.

Because the textual description in the Detailed Item Output table relates to commodities (and not industries), we focus on the intensity of vertical relatedness between pairs of commodities. We construct the sparse square matrix V based on the extent to which a given commodity is vertically linked (upstream or downstream)

²⁰There are 250 such words including accessories, air, attachment, commercial, component. See the Internet Appendix for a full list.

²¹This tercile-based approach is based on Hoberg and Phillips (2010), who also discard the most broad words.

to another commodity. From the Make Table, we create *SHARE*, an $I \times C$ matrix (Industry \times Commodity) that contains the percentage of commodity c produced by a given industry i . The *USE* matrix is a $C \times I$ matrix that records the dollar value of industry i 's purchase of commodity c as input. The *CFLOW* matrix is then given by $USE \times SHARE$, and is the $C \times C$ matrix of dollar flows from an upstream commodity c to a downstream commodity d . Similar to Fan and Goyal (2006), we define the *SUPP* matrix as *CFLOW* divided by the total production of the downstream commodity d . *SUPP* records the fraction of commodity c that is used as an input to produce commodity d . Similarly, the matrix *CUST* is given by *CFLOW* divided by the total production of the upstream commodities c , and it records the fraction of commodity c 's total production that is used to produce its commodity d . The *V* matrix is then defined as the average of *SUPP* and *CUST*. A larger element in *V* indicates a stronger vertical relationship between commodities c and d .²² Note that *V* is sparse (i.e., most commodities are not vertically related) and is non-symmetric as it features downstream ($V_{c,d}$) and upstream ($V_{d,c}$) directions.

Finally, we create an exit correspondence matrix *E* to account for production that flows out of the U.S. supply chain. To do so, we use the industries that are present in the Use table but *not* in the Make table ('final users'). *E* is a one-column matrix containing the fraction of each commodity that flows to these final users.

C Text-based Vertical Relatedness

We identify vertical relatedness between firms by jointly using the vocabulary in firm 10-Ks and the vocabulary defining the BEA IO commodities. We link each firm in our Compustat/Edgar universe to the IO commodities by computing the similarity between the given firm's business description and the textual description of each BEA commodity. Because vertical relatedness is observed from BEA at the IO commodity level (see description of the matrix *V* above), we can score every pair of firms i and j based on the extent to which they are upstream or downstream by

²²Alternatively, we consider in unreported tests the maximum between *SUPP* and *CUST*, and also *SUPP*, or *CUST* alone, to define vertical relatedness. Our results are robust.

(1) mapping i 's and j 's text to the subset of IO commodities it provides, and (2) determining i and j 's vertical relatedness using the relatedness matrix V .

When computing all textual similarities, we limit attention to words that appear in the Hoberg and Phillips (2015) post-processed universe. We also note that we only use text from 10-Ks to identify the product market each firm operates in (vertical links between vocabularies are then identified using BEA data as discussed above). Although uncommon, a firm will sometimes mention its customers or suppliers in its 10-K. For example, a coal manufacturer might mention in passing that its products are “sold to” the steel industry. To ensure that our firm-product market vectors are not contaminated by such vertical links, we remove any mentions of customers and suppliers using 81 phrases listed in the Internet Appendix.²³

Ultimately, we represent both firm vocabularies and the commodity vocabularies from BEA as vectors with length 60,507, which is the number of nouns and proper nouns appearing in 10-K business descriptions in Hoberg and Phillips (2015). Each element of these vectors corresponds to a single word. If a given firm or commodity does not use a given word, the corresponding element in its vector will be set to zero. By representing BEA commodities and firm vocabularies as vectors in the same space, we are able to assess firm and commodity relatedness using cosine similarities.

Our next step is to compute the ‘firm to IO commodity correspondence matrix’ B . This matrix has dimension $M \times C$, where C is the number of IO commodities, and M is the number of firms. An entry $B_{m,c}$ (row m , column c) is the cosine similarity of the text in the given IO commodity c , and the text in firm m 's business description. In this cosine similarity calculation, commodity word vector weights are assigned based on the words' economic importance from the CW matrix (see above), and firm word vectors are equally-weighted following Hoberg and Phillips (2015). We use cosine similarity because it controls for document length and is well-established in computational linguistics (see Sebastiani (2002)). The cosine similarity is the normalized dot product (see Hoberg and Phillips (2015)) of the word-distribution vectors of the two vocabularies being compared. The result is

²³Although we feel this step is important, our results are robust if we exclude this step.

bounded in $[0,1]$, and a value close to one indicates that firm i 's product market vocabulary is a close match to IO commodity c 's vocabulary. The matrix B thus indicates which IO commodity a given firm's products is most similar to.

We then measure the extent to which firm i is upstream relative to firm j :

$$UP_{ij} = [B \cdot V \cdot B']_{i,j}. \quad (1)$$

The triple product $(B \cdot V \cdot B')$ is an $M \times M$ matrix of unadjusted upstream-to-downstream links between all firms i to firms j . Note that direction is important, and this matrix is not symmetric. Upstream relatedness of i to j is thus the i 'th row and j 'th column of this matrix. Firm-pairs receiving the highest scores for vertical relatedness are those having vocabulary that maps most strongly to IO commodities that are vertically related according to the matrix V (constructed only using BEA relatedness data), and those having vocabularies that overlap non-trivially with the vocabularies that are present in the IO commodity dictionary according to the matrix B . Thus, firm i is located upstream from firm j when i 's business description is strongly associated with commodities that are used to produce other commodities whose description resembles firm j 's product description.

Downstream relatedness is simply the mirror image of upstream relatedness, $DOWN_{ij} = UP_{ji}$. By repeating this procedure for every year in our sample (1996-2008), the matrices UP and $DOWN$ provide a time-varying network of vertical links among individual firms.

D NAICS-based Vertical Relatedness

Given we are proposing a new way to compute vertical relatedness between firms, we compare the properties of our text-based vertical network to those of the NAICS-based measure used in previous research, which we describe now. One critical difference is that the NAICS-based vertical network is computed using the BEA industry space, and not the BEA commodity space. This is by necessity because the links to NAICS are at the level of BEA industries. Avoiding the need to link to BEA industries is one advantage of the textual vertical network. More generally, the com-

pounding of imperfections in BEA industries and NAICS industries may result in horizontal contaminations, especially when firms are in markets that do not cleanly map to NAICS industries. In particular, the Census Department states “NAICS was developed to classify units according to their production function. NAICS results in industries that group units undertaking similar activities using similar resources but does not necessarily group all similar products or outputs.”

To compute the NAICS-based network, we use methods that parallel those discussed above for the BEA commodity space (matrix V), but we focus on the BEA industry space and construct an analogous matrix Z . We first compute the BEA industry matrix $IFLOW$ as $SHARE \times USE$, which is the dollar flow from industry i to industry j . We then obtain $ISUPP$ and $ICUST$ by dividing $IFLOW$ by the total production of industry j and i respectively (using parallel notation as was used to describe the construction of V). The matrix Z is simply the average between $ICUST$ and $ISUPP$.

Following common practice in the literature (see for example Fan and Goyal (2006)), we map IO industries to NAICS industries and use two numerical thresholds to identify meaningful relatedness using the NAICS-based vertical network: 1% and 5%. A given industry i is deemed to be upstream (downstream) relative to industry j when the flow of goods Z_{ij} (Z_{ji}) is larger than this threshold. We find that the 1% and 5% flow thresholds generates NAICS-based vertical relatedness networks that have granularity of 1.37% and 9.48% (9.48% granularity means that 9.48% of randomly chosen industry pairs are vertically related in this network), respectively. For simplicity, we thus label the two resulting vertical networks as ‘NAICS-1%’ and ‘NAICS-10%’, respectively.

To ensure our textual networks are comparable, we choose two analogous textual granularity levels: 10% and 1%. These two text-based vertical networks define firm pairs as vertically related when they are among the top 10% and top 1% most vertically related firm-pairs using the textual scores. We label these networks as ‘Vertical Text-10%’ and ‘Vertical Text-1%’. Note that the textual networks generate a set of vertically related peers that is customized to each firm’s unique product offerings.

These firm level links provide considerably more information than is possible using broad industry links such as those based on NAICS.

E Vertical Network Statistics

We compare the properties of five key relatedness networks: Vertical Text-10%, Vertical Text-1%, NAICS-10%, NAICS-1%, and the TNIC-3 network developed by Hoberg and Phillips (2015). The first four networks are intended to capture vertical relatedness, and the TNIC-3 network is calibrated to be as granular as are three-digit SIC industries, and is intended to capture horizontal relatedness.

[Insert Table II Here]

The first row of Panel A in Table II presents the level of granularity of each network. The NAICS-10% and NAICS-1% networks have granularity levels of 9.48% and 1.37% respectively. The ‘Vertical Text-10%’ and ‘Vertical Text-1%’ networks have 10% granularity and 1% granularity, respectively. Finally, the TNIC-3 network has a granularity of 2.33%. The Vertical Text-1% network and the NAICS-1% network are thus twice as fine as SIC-3.

Reassuringly, the second to fourth rows in Panel A show that the four vertical networks exhibit low overlap with the horizontal TNIC-3, SIC and NAICS networks. Hence, none of the vertical networks are severely contaminated by known horizontal links. Despite this, the fifth and sixth rows illustrate that the vertical networks are quite different. Only 10.48% of firm-pairs in the NAICS-10% network are also present in the Vertical Text-10% network. Similarly, only 1.21% of firm-pairs are in both the Vertical Text-1% and NAICS-1% networks.

The eighth row reports the fraction of firm-pairs that include at least one financial firm (SIC code ranging between 6000 and 6999). The presence of financial firm pairs is quite low in the text-based vertical networks, at 9.20% and 1.80% of linked pairs, respectively. In contrast, financial firms account for a surprisingly large 48.44% and 34.31% of firm-pairs in the NAICS-based vertical networks. These results illustrate that treatment of financials is a first-order dimension upon which these networks

disagree. When we discard financial firms, the overlap between our text-based network and the NAICS-based network roughly doubles (e.g. 19.90% of non-financial firm-pairs in the NAICS-10% network are also in the Vertical Text-10% network). As theories of vertical relatedness and integration often focus on non-financial concepts such as relationship-specific investment and ownership of assets, these results support the use of the text-based network as being more relevant.

Although we do not report full details here to conserve space, we conduct two validation tests in the Internet Appendix to this paper. The goal is to compare the ability of the text-based and NAICS-based vertical networks to identify actual instances of vertical relatedness from orthogonal data sources. In the first test, we search all firm 10-Ks to identify direct verbal statements indicating the firm is vertically integrated. We find that the text-based network is roughly four times stronger in predicting these direct 10-K statements than is the NAICS-based network (Table IA.III.1). In a second validation test, we examine related party trade data from the U.S. Census Bureau, and examine which network better predicts vertical integration through offshore activities. Once again, we find strong evidence that the text-based network better predicts vertical integration (Table IA.III.2). Overall, both tests strongly support the conclusion that the text-based network is substantially more informative about vertical firm-to-firm linkages than is the NAICS based network.

IV Vertical Acquisitions

We now use our text-based vertical network to examine innovation activities and vertical organization. We start by studying vertical acquisitions, as these transactions represent a direct way firms can alter their boundaries and modify their degree of integration. To test our main hypothesis and theoretical predictions in the literature (e.g., Grossman and Hart (1986)), we concentrate on targets (the sellers of assets) as they are the party that loses control rights due to the transaction, and for which the trade-off between ex ante investment incentives and ex post hold-up should be important. We thus examine how R&D and patenting intensity are related to the

likelihood of being a target in vertical and non-vertical transactions.

Comparing vertical to non-vertical transactions is important, as the hypothesized issues of ex-ante incentives, contracting frictions, and potential ex-post hold up are most salient for vertically related firms and our theoretical predictions do not extend to other transactions such as horizontally related acquisitions.²⁴

A Transactions Sample

Our sample of transactions is from the Securities Data Corporation SDC Platinum database. We consider all announced and completed U.S. transactions with announcement dates between January 1, 1996 and December 31, 2008 that are coded as a merger, an acquisition of majority interest, or an acquisition of assets. As we are interested in situations where the ownership of assets changes hands, we only consider acquisitions that give acquirers majority stakes. To be able to distinguish between vertical and non-vertical transactions, we further require that both the acquirer and the target have available Compustat and 10-K data.

Table III displays summary statistics of our transactions sample. Following the convention in the literature, we limit attention to publicly traded acquirers and targets, and we exclude transactions that involve financial firms and utilities (SIC codes between 6000 and 6999 and between 4000 and 4999). Panel A shows that the sample consists of 3,460 transactions. Panel A further reports how many of these transactions are classified as vertical using the Vertical Text-10% and the NAICS-10% networks.

[Insert Table III Here]

Given that the Vertical Text-10% and NAICS-10% networks are designed to

²⁴Theories based on horizontal patent races for example would predict that firms engaged in R&D have higher incentives to merge to internalize high R&D competition. Phillips and Zhdanov (2013) provide additional theory regarding innovation and horizontally aligned firms, and also predict that R&D should be positively related to transaction likelihood. The predictions of both horizontal theories are thus opposite to our predictions for vertical acquisitions, making horizontal acquisitions a natural benchmark for comparing vertical acquisitions. Any evidence illustrating that vertical transactions are indeed different would support our hypotheses even in the context of the broader literature.

have similar granularity levels, it is perhaps surprising that the networks disagree sharply regarding the fraction of transactions that are vertically related. For our primary sample excluding financials, we observe that 39% are vertically related using the Vertical Text-10% network. Using the NAICS-10% network, we observe that just 13% are vertically related. For any network with a granularity of 10%, if transactions are random, we expect to see 10% of transactions belonging to this network. The fact that we find 39% is strong evidence that many transactions occur between vertically related parties. The results also suggest that the accumulated noise associated with NAICS greatly reduces the ability to identify vertically related transactions. We also note that with both networks, vertical deals are almost evenly split between upstream and downstream transactions.²⁵

Panel B of Table III displays the average abnormal announcement returns (in percent) of combined acquirers and targets in vertical and non-vertical transactions. We present these results mainly to compare with previous research (based on either SIC or NAICS codes). Confirming existing evidence, the combined returns across all transactions are positive and range from 0.49% to 0.94%. Notably, when vertical transactions are identified using our text-based measure, the combined returns are larger for vertical relative to non-vertical transactions. This supports the idea that vertical deals are value-creating on average as in Fan and Goyal (2006). Yet, this conclusion is not significant when vertical relatedness is identified using the NAICS-based network.

B Profile of Targets in Vertical Transactions

Table IV presents the R&D and patenting profile of targets in vertical and non-vertical deals. We focus on all transactions and we use our text-based network (10%) to identify vertical deals. We consider both industry- (i.e. TNIC-3) and

²⁵We also find that transactions classified as vertical are followed by an increase in our firm-level measure of vertical integration (*VI*). Using the Vertical Text-10% network, acquirers in vertical transactions experience an increase of 6% in *VI* from one year prior to one year after the acquisition. In contrast, acquirers in non-vertical transactions experience a decrease of 0.70% in *VI*. When we use the NAICS-10% network to identify vertical mergers, vertical acquirers see a negligible increase of 0.30% in *VI*.

firm-level measures of R&D and patenting activity. We measure R&D intensity as R&D divided by sales, and patenting intensity as the number of patents divided by assets. We describe all variables used in the paper and display summary statistics in Appendix 2. In Panel A, we observe a large difference between targets in vertical and non-vertical deals. When compared to firms that never participate in any acquisitions over the sample period (labeled as non-merging firms), vertical targets exhibit lower levels of R&D and hold more patents. In contrast, targets in non-vertical deals are more R&D intensive with lower patenting intensity.

[Insert Table IV Here]

To formally test these differences, we account for the fact that targets in vertical and non-vertical deals can differ on dimensions other than their R&D and patent profile. In Panels B and C, each actual target (vertical and non-vertical) is directly compared to a matched target with similar characteristics. For every actual transaction, we select matched targets from the subset of firms that did not participate in any transaction over the three years that precede the actual transaction. Matched targets are the nearest neighbors from a propensity score estimation. In panel B, we obtain matched targets based on industry (defined using the Fixed Industry Classification (FIC) of Hoberg and Phillips (2015))) and size. In panel C, we obtain matched targets based on FIC industries, size, age, market-to-book, PPE/Assets, the fraction of BEA End Users, and the number of segments.

The results in Table IV are consistent with our main hypothesis. High R&D firms remain separate to preserve ex ante incentives to create new innovation consistent with Grossman and Hart (1986). In contrast, target firms have higher patenting activity consistent with acquirers buying high patent firms to reduce ex post hold-up when firms are attempting to commercialize the innovation.

[Insert Figure 3 Here]

These patterns are confirmed in Figure 3 when we look at the average patenting and R&D intensity of target firms *prior* to their acquisition. Strikingly, vertical acquisitions tend to occur after targets experience a period of increased patenting activity (either measured with $\log(1+\#Patents)$ or $\#Patents/assets$). The realiza-

tion of successful innovation (i.e. the grant of patents) marks a time of increased firm maturity, and may indicate the end of the innovation cycle. As the marginal product of additional R&D investment declines, we observe increased integration at this time. The mirror image appears true for non-vertical acquisitions, which tend to cluster after a period of lower patenting activity. Although the dynamics are less clear-cut, Figure 3 confirms that there are large differences in R&D intensity between the firms that are acquired in vertical and non-vertical deals. Non-vertical targets have much higher R&D intensity than vertical targets.

C Multivariate Analysis

We complement the above univariate tests by estimating probit regressions to examine how R&D and patenting intensity affect the likelihood of becoming a target. The dependent variable is an indicator variable indicating whether a given firm is a target in a vertical or a non-vertical transaction, as noted in the column headers, in a given year. We consider our text-based network when identifying which transactions are vertically related (Vertical Text-10%). Our sample covers the period 1996-2008 and excludes regulated utilities and financial firms. We further require observations to have non-missing values for each variable we use in the estimations. We have 45,198 firm-year observations corresponding to 6,924 distinct firms.

For the explanatory variables of interest (in particular R&D and patent intensity), we consider equally-weighted averages across TNIC-3 industries instead of own-firm variables. We note, however, that our conclusion is qualitatively unchanged if we use own-firm R&D and patent variables instead of industry-level variables (see Internet Appendix IA.IV.1). This choice is driven by two considerations. First, focusing on industry lessens endogeneity concerns for both vertical and non-vertical transactions (see Acemoglu, Aghion, Griffith, and Zilibotti (2010)). Indeed, while a firm directly chooses its own degree of vertical integration, it has little choice regarding its industry's level of R&D or patenting activities. Second, the theoretical incentives to vertically integrate should be driven mostly by the characteristics of product markets, which is best captured using industry variables. For instance, as

in Acemoglu, Johnson, and Mitton (2009), the incentives to invest in intangibles are primarily determined by the specific product being exchanged between firms.

[Insert Table V Here]

Table V displays the results of these probit regressions.²⁶ We first focus on columns (1) to (4). Consistent with our predictions, we find strong differences between the types of firms targeted in vertical and non-vertical deals. In particular, column (1) indicates that even after we control for other factors, firms in high R&D industries are less likely to be a target in a vertical transaction. In contrast, column (2) shows that firms in these same R&D intensive industries are in fact more likely to be targets in non-vertical transactions. This different finding for horizontal acquisitions is consistent with theories of patent races where horizontally related firms engaged in R&D have an incentive to merge to internalize high R&D competition. It is also consistent with recent research by Phillips and Zhdanov (2013) who model and provide evidence that small target firms conduct more R&D when they have a high probability of selling out to larger horizontally related firms.

We next focus on the level of patenting activity. Consistent with our hypothesis that ex post successful innovation indicates maturity and lowers the returns from separate R&D investment, we find that vertical targets are more likely to be in high patenting industries. The opposite is true for non-vertical acquisitions, where firms in high patenting industries are less likely to be acquired.

Table V provides strong and robust evidence that R&D and patenting activities have opposite effects on the vertical transactions. Yet, as R&D and patenting activities are positively related, some industries with high levels of R&D also have high levels of patenting.²⁷ To account for this link, we introduce an interaction term between R&D and patenting activities and report the results in columns (3) and (4). The coefficient on this interaction is only significant for non-vertical transactions, which further confirms that these transactions are different. The positive coefficient

²⁶Note that we cluster standard errors at the industry (using the FIC data from Hoberg and Phillips (2015)) and year level. Also, we obtain similar results if we use linear probability models instead on non-linear probit models. The results are presented in the Internet Appendix (IA.IV.9).

²⁷In our sample, R&D and patenting activity are not perfectly correlated. This correlation is 0.33 across firms, and 0.58 across industries.

for industry patenting activity also remains robust for vertical acquisitions.²⁸

Table V also supports our hypothesis that maturity is an important positive determinant of vertical transactions. For instance, column (1) indicates that firms with lower market-to-book ratios and older firms are more likely to be targets of vertical deals. In contrast, targets in non-vertical deals are more likely to be young and are in less capital intensive industries. These findings are consistent with the following interpretation of the U-shaped relationship between firm maturity and restructuring activity noted in Arikian and Stulz (2011). Younger firms engage in non-vertical transactions likely to capitalize on asset complementarities, whereas more mature firms increase their acquisition activity as their focus turns to vertical acquisitions.

Although our main focus is on transaction targets, we also examine the link between R&D and patenting intensity and the likelihood of becoming a vertical or non-vertical acquirer. We report the results in the Internet Appendix for brevity (Table IA.IV.7). Mirroring the target results, we find that firms in R&D-intensive industries are less likely to engage in vertical acquisitions but more likely to engage in non-vertical purchases. In contrast, patenting intensity positively predicts the incidence of vertical acquisitions, and negatively predicts the incidence of non-vertical acquisitions. These additional results lend further support for our prediction that firms avoid vertical integration when R&D intensity is high, but embrace it when innovation is realized and translates into legally-enforceable patents.

D State R&D Tax Credit as Instrument

Our results so far reveal significant associations between the propensity to be purchased by vertically-related acquirers and the R&D and patenting intensity of the

²⁸In additional tests that we present in the Internet Appendix for brevity, we show that the results hold when we use lagged values of the independent variables (Table IA.IV.2), when we use sales-weighted industry measures instead of equally-weighted measures (Table IA.IV.3), when we focus solely on industry R&D and patent intensity and exclude the additional control variables (Table IA.IV.4), and when we include industry R&D and patent intensity separately (Table IA.IV.5). We also consider the NAICS-based measure of vertical relatedness (NAICS-10%) and report the results in the Internet Appendix (Table IA.IV.6) and find the results are much weaker.

targets’ industries. These associations however do not necessarily correspond to the causal effect of the stage of innovation on firms’ boundaries. For example, unobserved variables related to industry innovation activities might affect firms’ integration decisions. The presence of such omitted variables could bias our estimates and threaten our interpretation.

Fully addressing this concern would require instruments that separately affect industry R&D and patenting intensity, without influencing a given firm’s propensity to be acquired through other channels. Although such an ideal instrument is not available, we follow Bloom, Schankerman, and van Reenen (2013a) and use tax-induced changes to the user cost of R&D to construct an instrument for industry R&D intensity. State R&D tax credits offer firms credits against state income tax liability based on the amount of qualified research done within the state.²⁹ In practice, different states have different levels of R&D tax credits, and hence the user cost of R&D is dependent on firm locations and time. As discussed in Bloom, Schankerman, and van Reenen (2013a), the existing literature suggests that the introduction and level of R&D tax credits is quite random.

We use the state-by-year tax-induced user cost of R&D capital ($\rho_{s,t}$) to capture exogenous variation in industry R&D intensity. The user cost of R&D capital in state s and year t is given by the Hall-Jorgenson formula:

$$\rho_{s,t} = \frac{1 - (k_{s,t} + k_t^f) - (\tau_{s,t} + \tau_t^f)}{1 - (\tau_{s,t} + \tau_t^f)} [r_t + \delta] \quad (2)$$

where $k_{s,t}$ and k_t^f are the state and federal R&D tax credit rates, $\tau_{s,t}$ and τ_t^f are the state and federal corporation income tax rates, r_t is the real interest rate, and δ is the depreciation rate of R&D capital. The data are from Wilson (2009) and cover the period 1996-2006. Following the methodology of Bloom, Schankerman, and van Reenen (2013a), we implement the instrumental variable approach by first

²⁹As detailed in Wilson (2009), state and federal tax credits are based on the amount of qualified research within the state or country. States generally follow the Federal Internal Revenue Code (IRC) definition of qualified research: the wages, material expenses, and rental costs of certain property and equipment incurred in performing research “undertaken to discover information” that is “technological in nature” for a new or improved business purpose. Because we do not know the state location of each firm’s R&D spending, we assume that all R&D activities are performed in the firm headquarter’s state.

projecting the firm-year endogenous variable ($R\&D/sales$) on the instrument (ρ) as well as firm and year fixed effects. We show the results of this estimation in the Internet Appendix (Table IA.IV.8). From this, we calculate the value of $R\&D/sales$ predicted by the tax credit for each firm and year. Next, we average firms' tax-credit predicted $R\&D/sales$ across each (TNIC) industry-year to create our instrumental variable, which we label $Ind.(PredictedR\&D/sales)$. We then use this variable to instrument industry R&D intensity in a second stage equation (correcting the standard errors appropriately). Bloom, Schankerman, and van Reenen (2013b) provide a detailed explanation of this two-step approach, and discuss the exogeneity of R&D tax credit policy changes.

Our instrument is designed to isolate variation in industry R&D that is purely driven by tax rules and firm locations. We argue that this variation is plausibly unrelated to any specific firm's organizational choice. In particular, we posit that differences in the average user cost of R&D across industries and time – which measure the combined heterogeneity of firm locations and state tax rules – are related to firm integration decisions only through their impact on industry R&D.³⁰

Columns (5) to (7) of Table V report results from the instrumental variable estimations.³¹ Column (5) presents the first-stage results. We observe a positive and highly significant coefficient on the tax-induced industry predicted R&D, indicating that industries where more firms benefit from tax credits are indeed more R&D intensive. The second-stage estimates are reported in columns (6) and (7). In column (6), we continue to observe that industry R&D is negatively related to the intensity of vertical deals, and industry patenting intensity is positively related to vertical deals. These results confirm that the R&D intensity of industries significantly decreases the likelihood that a firm will be targeted in a vertical transaction.

³⁰ We recognize that individual firm locations are not random, as some firms may have incentives to move operations across states to reap larger R&D tax credits. Such moves can generate variation in industry R&D intensity. Yet, to invalidate our instrumental variable strategy, one would have to argue that mass relocations of firms in a given industry to exploit tax credits would have a direct effect on the propensity of firms to be purchased by a vertically-related acquirer from a different industry. Although we cannot formally rule out this explanation, we find it implausible.

³¹ Due to the binary nature of the dependent variable, we estimate instrumental variables using probit regressions using Maximum Likelihood. We obtain similar results if we use Linear Probability Models (see Table IA.IV.9 of the Internet Appendix).

Remarkably, the opposite result is again observed for non-vertical transactions where the IV-coefficient on industry R&D is positive. This is consistent with our central hypotheses being uniquely tied to vertical transactions, where the necessary assumptions of ex-ante incentives, contracting frictions, and potential ex-post hold up are most salient.

E Patents versus Secrecy to Protect Innovation

Next, we examine if our results are different in industries where patents provide effective legal protection versus industries where patents are ineffective and easy to work around. Indeed, high R&D does not necessarily lead to high patenting rates because some inventions are better protected with secrecy than with patents. For example, many firms maintain trade secrets, and in some cases, require employees to sign agreements protecting them. Based on the Carnegie Mellon Survey (CMS) on industrial R&D in the manufacturing sector, Cohen, Nelson, and Walsh (2000) report large differences across industries in the use of patents to protect inventions. In many cases, inventors indeed rely on secrecy to limit rivals from using information in the patent application. We use data from the CMS study to separate manufacturing industries into ‘Low’ and ‘High’ secrecy sub-samples, and estimate the effect of R&D and patent intensity on vertical acquisitions across both sub-samples.³² Our prediction is that our patent results should be strongest in low secrecy industries, as patents are less relevant in the high secrecy industries.

[Insert Table VI Here]

Table VI present the results. Consistent with the importance of property rights, we observe that patenting intensity is positively and significantly associated with vertical transactions in low secrecy industries where patents are effective at protecting innovation. Vertical acquisitions are largely unrelated to patenting intensity in high secrecy industries where legal protection is weaker. This suggests that the

³²The survey was performed with 1,478 R&D units or laboratories in 1994 and covers 34 manufacturing industries defined at the SIC 3-digit level. We measure the importance of secrecy using the average of two variables capturing the importance of secrecy for product and process innovation. We then assign firms into the low and high secrecy groups based on the sample median.

allocation of residual rights of control to the party whose commercialization incentives are large is less effective in these industries consistent with our predictions. Columns (3) and (4) present the results using the instrumental variable approach that are similar to the previous columns.

F Post-Acquisition R&D Intensity

Our tests so far concentrate on participation in vertical and non-vertical transactions. Another implication of our hypothesis is that vertical transactions should be followed by lower innovation incentives for the party that loses control rights. Testing this prediction is difficult because the majority of targets cease to exist as separate entities after the transaction, making it impossible to individually observe their R&D expenditures.

[Insert Figure 4 Here]

To provide indirect evidence on the evolution of R&D intensity post-acquisition, we perform two additional analyses. First, we track the evolution of R&D intensity around vertical transactions for the subset of targets that continue to exist for at least one year post-acquisition.³³ Second, we track R&D intensity for our full sample by assessing the “combined” target plus acquirer’s aggregate R&D for each deal and year. Consistent with the idea that changes of control rights can influence innovation incentives, Figure 4 provides suggestive evidence that R&D intensity declines by about 10% in response to vertical acquisitions in both samples.³⁴

V Vertical Integration within Firms

In this section, we use our text-based approach to develop a new firm-level measure of vertical integration. We then examine how firm innovation activities are linked to vertical organization both within industries and firms.

³³This represents 35% of all vertical targets in our sample.

³⁴Table IA.IV.10 shows that the decrease in R&D intensity is statistically significant.

A Text-based Integration

Using the notation from Section III, we define the extent to which a given firm is ‘vertically integrated’ by looking at a firm’s vertical relatedness to itself ($UP_{i,i}$ or interchangeably $DOWN_{i,i}$) as follows:

$$VI_i = [B \cdot V \cdot B']_{i,i}. \quad (3)$$

With this measure, a firm is more vertically integrated when its 10-K business description contains words that are vertically related (upstream or downstream) to other words in its own business description. This occurs when a firm offers products or services at different stages of a supply chain.³⁵ We also characterize whether a firm supplies products or services that are related to commodities that exit the U.S. supply chain using the exit correspondence matrix E defined in Section III as End Users $_i = [B \cdot E]_i$. We compute analogous measures of the extent to which firms sell to retail, government, or export by replacing E with the fraction sold to each sub-group.

In the Internet Appendix, we provide evidence that further validates our measurement by showing that our text-based measure is highly correlated with explicit mentions by firms about their vertical organization, and to the intensity of transactions that take place within the firm boundaries (i.e., intra-firm trade) using industry level data on related-party trade from the Census (see for instance Nunn and Treffer (2013) or Antras and Chor (2013)).

Statistics for the text-based variables are reported in Appendix 2. The average and median value of vertical integration (VI) are 0.012 and 0.008 respectively, and the maximum is 0.116. Although the nominal magnitudes do not have a direct interpretation, we note that there is a fair amount of right skewness as we would expect. In particular, most firms are not vertically integrated, but a smaller fraction do feature business descriptions that contain many words that are strongly vertically related. Hence, the firms situated toward the right tail are likely the set of firms

³⁵Unfortunately, data limitations prevent us from determining the economic importance of each product from firms’ product descriptions. Hence, while VI is a novel measure that uniquely captures firm-level vertical integration, it cannot account for each product-by-product importance.

that are vertically integrated. Figure 5 displays the evolution of vertical integration over time, and we note a trend away from integration especially in the late 1990s.

[Insert Figure 5 Here]

Table VII displays averages across quartiles of vertical integration. Consistent with the predictions of the model in Section II and the results on vertical acquisitions, integrated firms spend less on R&D than non-integrated firms. The average R&D/sales is roughly four times larger in low integration quartiles than in the high integration quartiles (9.8% versus 2.7%). In contrast, vertically integrated firms appear to own larger portfolios of patents. The (log) number of patents is two times larger in the high integration quartile (0.835 versus 0.420). These univariate results illustrate this section’s key findings.

[Insert Table VII Here]

The distinction between R&D and patents is well exemplified by the networking equipment industry, which includes Cisco, Broadcom, Citrix, Juniper, Novell, Sycamore, and Utstarcom. Figure 6 shows that these firms became four to five fold more vertically integrated over our sample. They also experienced (A) levels of R&D that peaked in 2002 and then began to sharply decline, and (B) levels of patenting activity that increased four to five fold starting in 2001. These dynamics are broadly consistent with the idea that the conversion of unrealized innovation into realized patented innovation increased the incentives to vertically integrate as the importance of ownership and control rights shifts from the smaller innovative firms to the larger commercializing firms.

[Insert Figure 6 Here]

Table VII also reveals that vertically integrated firms are further from the end of the supply chain and sell less output to retail and the government. Consistent with the fact that a large fraction of U.S. trade takes place intra-firm, vertically integrated firms are also more focused on exports (e.g. Zeile (1997) or Antras (2003)). Vertically integrated firms also operate in more concentrated markets (based on the TNIC HHI from Hoberg and Phillips (2015)). Consistent with Atalay, Hortacsu, and Syverson (2014), they are larger, older, and are more capital-intensive firms

with lower growth opportunities (MB). Vertically integrated firms also have more business segments, as identified from the (NAICS-based) Compustat segment tapes. We also report an alternative measure of vertical integration based on NAICS industries and the Compustat segment tapes: $VI_{Segment}$, as used in Acemoglu, Johnson, and Mitton (2009). $VI_{Segment}$ is larger when the segments a given firm operates in share stronger vertical relations. Although this measure can only be obtained for multi-segment firms (less than 33% of observations in our sample), it is significantly positively 20% correlated with our text-based measure.³⁶

B R&D, Patents, and Vertical Integration

To further examine the effect of innovation activities on firm’s degree of vertical integration, we estimate panel data regressions where the dependent variable is our measure of vertical integration (VI). We focus on within-industry and within-firm specifications and include year fixed effects in all specifications to isolate the apparent trend towards dis-integration in our sample.

[Insert Table VIII Here]

Table VIII presents results that confirm the univariate evidence: Firms operating in industries with high levels of R&D are less vertically integrated. This result obtains both within industries (when we include industry fixed effects in column (1)) and within firms (with firm fixed effects in column (2)). This latter result is important as it indicates that firms modify their degree of vertical integration over time as industry R&D varies. Economically, the negative link between R&D and integration is substantial: A one standard deviation increase in R&D intensity is associated with a 10% decrease in our text-based measure of integration in the within-industry specification, and with a 1.7% decrease in the within-firm specification where much variation is absorbed. The lower magnitude for the within-firm specification reflects

³⁶In the Internet Appendix, we further illustrate our text-based measure of vertical integration by displaying the 30 most vertically integrated firms in 2008 (Table IA.IV.11). A close look at these firms suggests a high degree of actual vertical relatedness among product offerings. Moreover, although they are highly integrated, these firms rank rather low on existing non-text measures of integration based on Compustat segments.

the high degree of persistence of VI at the firm-level.³⁷ Statistically, both findings are significant at the 1% level.

In sharp contrast, the coefficients on $\#Patents/assets$ are positive and significant. All else equal, firms operating in high patenting industries are more likely to be vertically integrated. The economic magnitude is also large: Integration increases by 6.8% (1.9%) following a one standard deviation increase in patenting intensity in the within-industry (within-firm) specification. This is further consistent with the ex post realization of successful innovation alleviating the ex ante need to incentivize relation-specific investment. Firms with successful innovation are more likely to increase integration to reduce the threat of ex post holdup.

Results in columns (3) and (4) indicate that the coefficient on the interaction between R&D and patenting is negative, while patenting remains positive and R&D remains negative. The results confirm that unrealized and realized innovation have opposite effects on firms' propensity to vertically integrate and suggest that the R&D effect dominates. The results with industry fixed effects are stronger but the results with firm fixed effects remain significant.³⁸

Firms are also more likely to be integrated when they are more mature. In particular, integration is positively related to capital intensity and size in all specifications. It is also positively related to firm age and negatively related to market-to-book in within-industry specifications. The link to maturity is likely related to the irreversibility of integration, and firms will be more willing to commit to integration when product markets are more stable, and they are less likely to need to dis-integrate later due to changes in the product market. This issue of irreversibility also relates to the high fixed costs of integration, and integration is likely only profitable if gains are expected to remain stable over a suitably long horizon to amortize the fixed costs. Firms that are closer to the end of the supply chain (*End Users*) are also less likely to integrate, and firms with an observed conglomerate structure ($\#Segments$) are more likely to be vertically integrated.

³⁷The coefficient of autocorrelation for VI is 0.931.

³⁸The Internet Appendix presents additional tests showing the robustness of our results (Tables IA.IV.1-IA.IV.12).

C Instrumental Variables and Sub-Samples

As we did in Section IV for vertical acquisitions, we re-examine vertical integration using an instrumental variables framework based on State R&D tax credits. This test is important because the same omitted variables that might threaten our identification for acquisitions might also matter in the current setting. As discussed in Section IV, tax-induced changes to the user cost of R&D allow us to construct a powerful instrument for industry R&D intensity that is plausibly exogenous.

We report the instrumental variables tests using average industry tax-credit predicted $R\&D/sales$ as an instrument for industry R&D intensity in the last four columns of Table VIII. Columns (5) and (6) display the results from the estimation with industry fixed effects, which confirm our baseline results. We continue to observe a negative and significant coefficient on instrumented industry R&D, which implies that an exogenous increase in industry R&D leads to lower levels of vertical integration at the firm level. The coefficient for patent intensity remains virtually identical in the instrumented estimation. Columns (7) and (8) report results with firm fixed effects. Remarkably, we continue to find a significant negative coefficient on instrumented industry R&D intensity even when we control for unobserved heterogeneity across firms.

[Insert Table IX Here]

Table IX highlights that the effect of patenting intensity on firm integration is only present in the low-secrecy subsample, where patents provide effective protection for realized innovation. Firm boundaries appear unrelated to patents in industries that rely on secrecy.

VI Conclusions

Our paper examines vertical acquisitions and changes to firm-specific vertical integration. We consider how the distinction between incentives to invest in R&D, and the potential for ex post holdup influence vertical transactions and integration. We also consider maturity, as vertical integration is more likely when the supply chain

is mature enough to support long-term gains from operational synergies.

We measure vertical relatedness using computational linguistics analysis of firm product descriptions and how they relate to product vocabularies from the BEA Input-Output tables. The result is a dynamic network of vertical relatedness between publicly-traded firms. We thus observe the extent to which acquisitions are vertical transactions and develop a new firm-level measure of vertical integration.

We show that unrealized innovation through R&D and realized innovation through patents affect the propensity to vertically integrate. Firms in high R&D industries are less likely to vertically integrate through own-production and vertical acquisitions. These results are robust to using state-level tax credits as an instrument for R&D. These findings are consistent with firms remaining separate to maintain ex ante incentives to invest in intangible capital and to maintain residual rights of control, as in the property rights theory of Grossman, Hart and Moore.

In contrast, firms in high patenting industries with high realized innovation are more likely to vertically integrate. In these industries, owners have more legally enforceable residual rights of control. They are more likely to integrate via acquisitions as giving control to commercializing firms should mitigate ex-post holdup. These results reconcile some of the tension between the ex post hold-up literature of Klein, Crawford, and Alchian (1978) and Williamson (1979), and the ex ante incentives of assigning residual rights of control as in Grossman and Hart (1986)).

Appendix 1: Model Proofs

In this appendix, we first show how the integration decision can be viewed as a real option. We then present the proofs of the three propositions from the text, along with a lemma needed for the proofs.

Optimal Timing of Integration As A Real Option

First note that after integration $I_t^* = 1$, integration stays forever ($I_{t+\tau}^* = 1$ for any $\tau > 0$). Here, the R&D investment by the supplier $x_{t+\tau}^* = 0$ for any $\tau \geq 0$ since it is non-contractible. For the producer's investment in integration, we have $y_t^* = \operatorname{argmax}_{y_t} [P_s(1 + \rho(y_t)) - Ry_t^h]$ in each period if the base price in integration has been improved to P_s . Denote the maximized value and the perpetuity value by:

$$\begin{aligned} v(P_s; I = 1) &= P_s(1 + \rho(y_t^*)) - Ry_t^{*h} \\ V(P_s; I = 1) &= \max_{y_t} [P_s(1 + \rho(y_t)) - Ry_t^h] + \frac{V(P_s; I = 1)}{1 + r} \\ &= v(P_s; I = 1) \left[1 + \frac{1}{1 + r} + \frac{1}{(1 + r)^2} + \dots \right] = \frac{1 + r}{r} v(P_s; I = 1) \end{aligned}$$

The optimal y_t^* thus depends on P_s in the following way

$$P_s \rho'(y_t^*) = Rhy_t^{*h-1}$$

A second observation is that the only state variable for the value function is the base price P_t^b , which is assumed to be equal to P_s ($s < N$) at time t . Therefore, we can define continuation value of separation and the value function $V(P_s)$ recursively as follows for $s < N$

$$V(P_s; I = 0) = \max_{\{x_t, y_t\}} \underbrace{[P_s(1 + y_t) - Ry_t^h - Sx_t^g]}_{\text{time } t \text{ profit}} + \underbrace{\frac{1}{1 + r} [x_t V(P_{s+1}) + (1 - x_t) V(P_s)]}_{\text{expected future profit}}$$

For $s = N$ any additional R&D expenditures thus cannot increase the base price

Table A1:

$P_t^b = P_s \ (s < N)$	$V(P_s)$	x^*	y^*
Integration $I = 1$	$V(P_s; I = 1) > V(P_s; I = 0)$	0	$P_s \rho'(y^*) = Rhy^{*h-1}$
Separation $I = 0$	$V(P_s; I = 1) < V(P_s; I = 0)$	$\frac{V(P_{s+1}) - V(P_s)}{1+r} = Sgx^{*g-1}$	$P_s = Rhy^{*h-1}$

anymore so:

$$V(P_N; I = 0) = \max_{y_t} P_N(1 + y_t) - Ry_t^h + \frac{V(P_N)}{1+r}$$

The optimal y_t^* and x_t^* also depend on P_s :

$$\begin{aligned} P_s &= Rhy_t^{*h-1} \\ \frac{V(P_{s+1}) - V(P_s)}{1+r} &= Sgx_t^{*g-1} \end{aligned}$$

The value function is thus:

$$V(P_s) = \max\{V(P_s; I = 1), V(P_s; I = 0)\}$$

The optimal decisions in each state can be summarized in Table A1 (above). We now prove the propositions that we gave earlier in the paper.

Proposition 1

R&D expenditures are higher in separation, while commercialization and product integration expenditures are higher in integration.

Proof:

In integration we have $x^* = 0$. In separation we must have $x^* > 0$, otherwise assuming $x^* = 0$, by definition $V(P_s) = V(P_s; I = 0) = \max_y P_s(1 + y) - Ry^h + \frac{V(P_s)}{1+r}$. So we solve that $V(P_s) = \frac{1+r}{r} [\max_y P_s(1 + y) - Ry^h] < \frac{1+r}{r} v(P_s; I = 1) = V(P_s; I = 1)$, which gives a contradiction. So as long as separation is chosen, $x^* > 0$, which from the FOC, we can derive that $V(P_{s+1}) > V(P_s) = V(P_s; I = 0)$ (if separation is chosen when the base price last period is P_s).

Proposition 2

If $P_t^b = P_N$, then both firms prefer to integrate so $V(P_N) = V(P_N; I = 1) > V(P_N; I = 0)$.

Proof:

Assuming separation is chosen, then $V(P_N) = V(P_N; I = 0) = \max_y [P_s(1 + y) - Ry^h] + \frac{V(P_N)}{1+r}$. So we can solve that $V(P_N) = V(P_N; I = 0) = \frac{1+r}{r} \max_y [P_N(1 + y) - Ry^h] < \frac{1+r}{r} \max_y [P_N(1 + \rho(y)) - Ry^h] = V(P_N; I = 1)$, which is a contradiction. Therefore, we must have $V(P_N) = V(P_N; I = 1) > V(P_N; I = 0)$.

Lemma 1

Value function $V(P_s)$ is increasing in P_s .

Proof:

First note that the value of integration $V(P_s; I = 1)$ is always increasing in P_s . By the Envelope theorem, we have that $\frac{\partial V}{\partial P_s}(P_s; I = 1) = \frac{1+r}{r}(1 + \rho(y^*)) > 0$. Now just analyze by cases, if separation is chosen, given base price P_s , by the proof in Proposition 1 we know that $V(P_{s+1}) > V(P_s)$; otherwise integration is chosen, then $V(P_s) = V(P_s; I = 1) < V(P_{s+1}; I = 1) \leq V(P_{s+1})$. So in both cases, value function is increasing in base price. Also, we could see directly that $V(P_s; I = 0) < V(P_{s+1}; I = 0)$ since $V(P_s; I = 0)$ is increasing in P_s , $V(P_s)$, and $V(P_{s+1})$.

Solution of $V(P_s)$ by Backward Induction

Integration is a real option, and the base price is the only state variable. The series of value functions $\{V(P_0), V(P_1), \dots, V(P_s)\}$ is solved by backward induction.

- $P_b = P_N$: we know that $V(P_N) = V(P_N; I = 1) = \frac{1+r}{r}v(P_N; I = 1)$ which can be solved directly
- $P_b = P_{N-1}$: note that the value of integration is pre-determined as $V(P_{N-1}; I =$

$1) = \frac{1+x}{r}v(P_{N-1}; I = 1)$, so if $V(P_{N-1}) = V(P_{N-1}; I = 0)$ then it must be true that $V(P_{N-1}; I = 0)$ is the solution solving the following equation on M and M must be greater than $V(P_{N-1})$

$$M = \max_{\{x,y\}} [P_{N-1}(1+y) - Ry^h] + \left[\frac{1}{1+r} (xM + (1-x)V(P_N)) - Sx^g \right]$$

- $P_b = P_s$: by now $V(P_{s+1})$ is known. Again, solve the following equation on M , then $V(P_s) = \max\{V(P_s; I = 1), M\}$

$$M = \max_{\{x,y\}} [P_s(1+y) - Ry^h] + \left[\frac{1}{1+r} (xM + (1-x)V(P_{s+1})) - Sx^g \right]$$

The above is a valid solution as long as the integration decision is monotonic in s , in other words, there is a triggering state s^* such that separation is chosen whenever $s < s^*$ and integration is chosen when $s \geq s^*$

Assumption 2

The increase in price P_s decreases with each successive innovation such that the series of value functions solved using the above method satisfies this condition: $V(P_{s+2}) - V(P_{s+1}) < V(P_{s+1}) - V(P_s)$.

The following Proposition then claims that there exists a triggering state s^* , so the series of value functions solved using the backward induction is the true solution.

Note that under this assumption, Lemma 1 holds, and the marginal benefit of R&D expenditures which equals $\frac{V(P_{s+1}) - V(P_s)}{1+r}$ in separation is decreasing in base price, and the optimal level of R&D expenditures in separation is also decreasing. Also note that even though the function $V(P; I = 1)$ is convex in P , we could make the increment in base price so small such that conditions in Assumption 3 hold.

Proposition 3

There exists a state s^* such that $V(P_s) = V(P_s; I = 1) \geq V(P_s; I = 0)$ for any $s \geq s^*$, and $V(P_s) = V(P_s; I = 1) < V(P_s; I = 0)$ for any $s < s^*$. The state s^* would then be the triggering state for integration.

Proof:

We only need to prove that there does not exist a state s such that integration is chosen with base price P_s while separation is chosen with base price P_{s+1} .

In state s , we have

$$V(P_s; I = 1) = \max_y [P_s(1 + \rho(y)) - Ry^h] + \frac{V(P_s; I = 1)}{1 + r}$$

$$V(P_s; I = 0) = \max_y [P_s(1 + y) - Ry^h] + \frac{V(P_s)}{1 + r} + \max_x \left[\frac{V(P_{s+1}) - V(P_s)}{1 + r} x - Sx^g \right]$$

Integration is chosen in state s meaning $V(P_s) = V(P_s; I = 1)$ and

$$\underbrace{\max_y [P_s(1 + \rho(y)) - Ry^h] - \max_y [P_s(1 + y) - Ry^h]}_{\text{Increments in TS by commercialization expenditures in integration if the integration as a real option is exercised right now}} >$$

Increments in TS by commercialization expenditures in integration
if the integration as a real option is exercised right now

$$\underbrace{\max_x \left[\frac{V(P_{s+1}) - V(P_s)}{1 + r} x - Sx^g \right]}_{\text{Increments in TS by R\&D exp. Continuation value in separation}}$$

First note that $\max_x \left[\frac{V(P_{s+1}) - V(P_s)}{1 + r} x - Sx^g \right]$ is always non-negative so we must have $\max_y [P_s(1 + \rho(y)) - Ry^h] - \max_y [P_s(1 + y) - Ry^h] > 0$.

The difference $\max_y [P_s(1 + \rho(y)) - Ry^h] - \max_y [P_s(1 + y) - Ry^h]$ is a function of P_s and the derivative with respect to P_s is $\rho(y^1) - y^0$ (by the Envelope Theorem), with y^1 and y^0 the optimum under integration and separation. Note that from the FOC we have $P_s = Rh(y^0)^{h-1}$ and $P_s = Rh(y^1)^{h-1}$, since $\rho'(y^1) > 1$ and $h > 1$ we must have $y^1 > y^0$ (commercialization expenditures are larger in integration) and thus $\rho(y^1) > \rho(y^0) > y^0$. So the difference is increasing in P_s so that

$$\max_y [P_{s+1}(1 + \rho(y)) - Ry^h] - \max_y [P_{s+1}(1 + y) - Ry^h] >$$

$$\max_y [P_s(1 + \rho(y)) - Ry^h] - \max_y [P_s(1 + y) - Ry^h]$$

The net benefit of R&D expenditures $\max_x [\frac{V(P_{s+1})-V(P_s)}{1+r}x - Sx^g]$, however, is decreasing in P_s because the increments in the value function $V(P_{s+1}) - V(P_s)$, by assumption, are decreasing in P_s , so must have

$$\max_x [\frac{V(P_{s+2}) - V(P_{s+1})}{1+r}x - Sx^g] < \max_x [\frac{V(P_{s+1}) - V(P_s)}{1+r}x - Sx^g]$$

Combining the three inequalities, we have that

$$\max_y [P_{s+1}(1 + \rho(y)) - Ry^h] - \max_y [P_{s+1}(1 + y) - Ry^h] > \max_x [\frac{V(P_{s+2}) - V(P_{s+1})}{1+r}x - Sx^g]$$

So the exercise value (exercise the option of integration) is greater than the continuation value (in separation) in state $s + 1$. From this it is easy to see that $V(P_{s+1}; I = 1) > V(P_{s+1}; I = 0)$ since otherwise $V(P_{s+1}; I = 0)$ would be equal to $\frac{1+r}{r} \left\{ \max_y [P_{s+1}(1 + y) - Ry^h] + \max_x [\frac{V(P_{s+2})-V(P_{s+1})}{1+r}x - Sx^g] \right\}$, which is less than $V(P_{s+1}; I = 1)$ which is equal to $\frac{1+r}{r} \max_y [P_{s+1}(1 + \rho(y)) - Ry^h]$.

Therefore, if in state s , integration is chosen, then in state $s + 1$, integration will be chosen too. By induction, all states after s will be under integration. Given the fact that the two firms start as separated, and in the final state N they must choose integration, there must exist a triggering state s^* such that integration is chosen in states $s \geq s^*$ and separation is chosen in states $s < s^*$. In other words, s^* would be the state in which the real option of integration is exercised in equilibrium.

Note that some states of the world could not be reached in equilibrium. For example, for any $s > s^*$, the base price P_s would never appear in equilibrium since the two firms have integrated at state s^* . So the total surplus $V(P_{s^*})$ would be the highest one reached in equilibrium, which is also the final value in integration.

Appendix 2: Variable Descriptions

In this appendix, we describe the variables used in this study and report summary statistics. We report Compustat items in parenthesis when applicable. All ratios are winsorized at the 1% level in each tail.

- *VI* measures the degree to which a firm offers products and services that are vertically related based on our new text-based approach to measure vertical relatedness (as defined in Section V.A)).
- *Retail* measures the degree to which a firm-year's products flow to retail customers (as defined in Section V.A)).
- *Government* measures the degree to which a firm-year's products flow to the government (as defined in Section V.A)).
- *Export* measures the degree to which a firm-year's products are exported (as defined in Section V.A)).
- *End users* is the sum of *Retail*, *Government*, and *Export*. It measures degree to which a firm-year's products exit the supply chain.
- *HHI* measures the degree of concentration (of sales) within TNIC-3 industries. We compute HHI as the TNIC HHI in Hoberg and Phillips (2015) using the text-based TNIC-3 horizontal industry network.
- *Industry R&D/sales* is equal to research & development expenses (XRD) scaled by the level of sales (SALE). This variable is set to zero when R&D is missing. We average this firm level measure over all firms in a given industry.
- *Industry #Patents/assets* is the number of (granted) patents a firm owns scaled by the level of assets (AT). We average this firm level measure over all firms in a given industry. Patents data are obtained from the US Patent and Trademark Office data (the NBER Patent data archive). Patent data are available until 2006. We assume patents for 2007 and 2008 remain at the level of 2006 in our tests.³⁹
- *PPE/assets* is equal to the level of property, plant and equipment (PPENT) divided by total assets (AT).
- *Log(assets)* is the natural logarithm of the firm assets (AT).
- *Log(age)* is the natural logarithm of one plus the firm age. Age is computed as the current year minus the firm's founding date. When we cannot identify a firm's founding date, we use its listing vintage (based on the first year the firm appears in the Compustat database).

³⁹All the results in the paper hold if we instead restrict our sample to the period 1996-2006 to have complete information on patents.

- *#Segments* is the number of operating segments observed for the given firm in the Compustat segment database. We measure operating segments based on the NAICS classification.
- *MB* is the firm's market-to-book ratio. It is computed as total assets (TA) minus common equity (CEQ) plus the market value of equity ((CSHO×PRCC.F)) divided by total assets.
- *VI_{segment}* measures firm-level vertical integration based on Compustat Segments. It is computed as the average vertical relatedness across a firm's distinct NAICS segments. Vertical relatedness is based on the matrix *Z* (defined in Section III.D) that relies on the 2002 BEA Input-Output table.
- *User cost of R&D capital* (ρ) is the user cost of R&D capital in state *s* and year *t* is given by the Hall-Jorgenson formula: $\rho_{s,t} = \frac{1-(k_{s,t}+k_t^f)-(\tau_{s,t}+\tau_t^f)}{1-(\tau_{s,t}+\tau_t^f)} [r_t + \delta]$, where $k_{s,t}$ and k_t^f are the state and federal R&D tax credit rates, $\tau_{s,t}$ and τ_t^f are the state and federal corporation income tax rates, r_t is the real interest rate, and δ is the depreciation rate of R&D capital. The data are from Wilson (2009) and cover the period 1996-2006.

Table A2: Summary Statistics

Variable	Mean	St. Dev	Min	p25	p50	p75	Max	#Obs.
<i>Panel A: Data from Text Analysis</i>								
VI	0.012	0.011	0.000	0.004	0.008	0.016	0.116	45,198
Retail	0.352	0.081	0.001	0.297	0.347	0.401	0.859	45,198
Government	0.031	0.021	0.000	0.018	0.026	0.038	0.971	45,198
Export	0.086	0.025	0.001	0.069	0.085	0.102	0.301	45,198
End users	0.470	0.076	0.086	0.421	0.470	0.516	0.973	45,198
<i>Panel B: Data from Existing Literature</i>								
HHI	0.259	0.225	0.014	0.103	0.177	0.334	1.000	45,198
R&D/sales	0.059	0.121	0.000	0.000	0.000	0.062	0.669	45,198
#Patents/assets	0.007	0.020	0.000	0.000	0.000	0.002	0.112	45,198
log(1+#Patents)	0.600	1.176	0.000	0.000	0.000	0.693	8.195	45,198
PPE/assets	0.263	0.221	0.000	0.089	0.195	0.374	0.888	45,198
log(assets)	5.684	1.784	2.529	4.330	5.505	6.839	10.881	45,198
log(age)	2.953	1.060	0.000	2.303	2.996	3.664	5.037	45,198
#Segments	1.550	0.999	1.000	1.000	1.000	2.000	12.000	45,198
MB	1.980	1.444	0.624	1.110	1.497	2.247	8.351	45,198
<i>VI_{segment}</i>	0.013	0.037	0.000	0.000	0.000	0.006	0.639	45,198
User cost of R&D capital	1.168	0.041	1.027	1.136	1.181	1.201	1.237	39,751

Note: This table displays summary statistics for all the variables used in the analysis.

References

- Acemoglu, Daron, 1996, A microfoundation for social increasing returns in human capital accumulation, *Quarterly Journal of Economics* 111, 779–804.
- , Phillippe Aghion, Rachel Griffith, and Fabrizio Zilibotti, 2010, Vertical integration and technology: Theory and evidence, *Journal of the European Economic Association* pp. 989–1033.
- Acemoglu, Daron, Simon Johnson, and Todd Mitton, 2009, Determinants of vertical integration: Financial development and contracting costs, *Journal of Finance* 64, 1251–1290.
- Aghion, Philippe, and Jean Tirole, 1994, The management of innovation, *Quarterly Journal of Economics* pp. 1185–1209.
- Ahern, Kenneth, 2012, Bargaining power and industry dependence in mergers, *JFE* 103, 530–550.
- , and Jarrad Harford, 2013, The importance of industry links in merger waves, *Journal of Finance (forthcoming)* University of Michigan and University of Washington Working Paper.
- Allen, Jeffrey W., and Gordon M. Phillips, 2000, Corporate equity ownership, strategic alliances, and product market relationships, *Journal of Finance* 55, 2791–2815.
- Antras, Pol, 2003, Firms, contracts, and firm structure, *Quarterly Journal of Economics* pp. 1375–1418.
- , and Davin Chor, 2013, Organizing the global value chain, *Econometrica* pp. 2127–2204.
- Arikan, Asli, and Rene M. Stulz, 2011, Corporate acquisitions, diversification, and the firm’s life-cycle, Ohio State University Working Paper.
- Atalay, Enghin, Ali Hortacsu, and Chad Syverson, 2014, Vertical integration and input flows, *forthcoming American Economic Review* pp. 1120–1148.
- Baker, Geroge, and Thomas Hubbard, 2003, Make versus buy in trucking: Asset ownership, job design, and information, *American Economic Review* pp. 551–572.
- Bena, Jan, and Kai Li, 2013, Corporate innovations and mergers and acquisitions, *Journal of Finance (forthcoming)*.
- Bloom, Nicholas, Mark Schankerman, and John van Reenen, 2013a, Identifying technology spillovers and product market rivalry, *Econometrica* pp. 1347–1393.
- , 2013b, Supplement to “identifying technology spillovers and product market rivalry”, *Econometrica Supplementary Material* pp. 1347–1393.
- Bresnahan, Timothy, and Jonathan Levin, 2012, Vertical integration and market structure, Working Paper.
- Cohen, Wesley M., Richard R. Nelson, and John P. Walsh, 2000, Protecting their intellectual assets: Appropriability conditions and why us manufacturing firms patent (or not), *National Bureau of Economic Research* 7552, 1–50.
- Fan, Joseph, and Vidhan Goyal, 2006, On the patterns and wealth effects of vertical mergers, *Journal of Business* 79, 877–902.
- Gibbons, Robert, 2005, Four formal(izable) theories of the firm?, *Journal of Economic Behavior and Organization* 58, 200–245.
- Grossman, Sanford J., and Oliver D. Hart, 1986, The cost and benefits of ownership: A theory of vertical and lateral integration, *Journal of Political Economy* 94, 691–719.
- Harford, Jarrad, 2005, What drives merger waves?, *Journal of Financial Economics* 77, 529–560.
- 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, conditionally accepted *Journal of Political Economy*.
- Holmstrom, Bengt, and Paul Milgrom, 1991, Multi-task principal-agent problems: Incentive contracts, asset ownership and job design, *Journal of Law, Economics and Organization* 7, 24–52.
- , 1994, The firm as an incentive system, *American Economic Review* 84, 972–991.
- Hortacsu, Ali, and Chad Syverson, 2007, Cementing relationships: Vertical integration, foreclosure, productivity and prices, *Journal of Political Economy* pp. 250–301.
- Hyland, David, 1999, Why firms diversify: An empirical investigation, Working Paper.
- Joskow, Paul L., 1987, Contract duration and relationship specific investments, *American Economic Review* 77, 168–175.
- Jovanovic, Boyan, and P Rouseau, 2002, The q-theory of mergers, *American Economic Review* 92, 198–204.
- Kale, Jayant R., and Husayn K. Shahrur, 2007, Corporate capital structure and the characteristics of suppliers and customers, *Journal of Financial Economics* 83, 321–365.
- Kedia, Simi, Abraham Ravid, and Vicente Pons, 2011, When do vertical mergers create value?, *Financial Management* 40, 845–877.
- Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian, 1978, Vertical integration, appropriable rents, and the competitive contracting process, *Journal of Law and Economics* 21, 297–326.
- Lafontaine, Francine, and Margaret Slade, 2007, Vertical integration and firm boundaries: The evidence, *Journal of Economic Literature* 45, 629–685.
- Lerner, Josh, and Robert Merges, 1998, The control of technology alliances: An empirical analysis of the biotechnology industry, *Journal of Industrial Economics* pp. 125–156.
- 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.
- Masten, Scott E., 1984, The organization of production: Evidence from the aerospace industry, *Journal of Law and Economics* 27, 403–417.
- Monteverde, Kirk, and David J. Teece, 1982, Supplier switching costs and vertical integration in automobile industry, *Bell Journal of Economics* 13, 206–213.
- Morck, Randall, Andrei Shleifer, and Robert Vishny, 1990, Do managerial motives drive bad acquisitions, *Journal of Finance* 45, 31–48.
- Nunn, Nathan, and Daniel Treffer, 2013, Incomplete contracts and the boundaries of the multinational firms, *Journal of Economic Behavior and Organization* pp. 330–344.
- Phillips, Gordon M., and Alexei Zhdanov, 2013, R&d and the incentives from merger and acquisition activity, *Review of Financial Studies* 34-78, 189–238.
- Sebastiani, Fabrizio, 2002, Machine learning in automated text categorization, *ACMCS* 34, 1–47.
- Seru, Amit, 2014, Firm boundaries matter: Evidence from conglomerates and r&d activity, *Journal of Financial Economics* 111, 381–405.
- Villalonga, Belen, 2004, Does diversification cause the diversification discount, *Financial Management* 33, 5–27.
- Williamson, Olivier E., 1971, The vertical integration of production: Market failure consideration, *American Economic Review* 61, 112–123.
- , 1979, Transaction-cost economics: The governance of contractual relations, *Journal of Law and Economics* 22, 233–261.

Wilson, Daniel, 2009, Beggar thy neighbor? the in-state, out-of-state, and aggregate effects of r&d tax credits, *Review of Economics and Statistics* pp. 431–436.

Zeile, William, 1997, U.s. intrafirm trade in goods, *Survey of Current Business* pp. 23–38.

Table I: BEA vocabulary example: Photographic and Photocopying Equipment

Description of Commodity Sub-Category	Value of Production (\$Mil.)
Still cameras (hand-type cameras, process cameras for photoengraving and photolithography, and other still cameras)	266.1
Projectors	72.4
Still picture commercial-type processing equipment for film	40.5
All other still picture equipment, parts, attachments, and accessories	266.5
Photocopying equipment, including diffusion transfer, dye transfer, electrostatic, light and heat sensitive types, etc.	592.4
Microfilming, blueprinting, and white-printing equipment	20.7
Motion picture equipment (all sizes 8mm and greater)	149.0
Projection screens (for motion picture and/or still projection)	204.9
Motion picture processing equipment	23.0

Note: This table provides an example of the BEA commodity ‘photographic and photocopying equipment’ (IO Commodity Code #333315). The table displays its sub-commodities and their associated product text, along with the value of production for each sub-commodity.

Table II: Vertical Network Summary Statistics

Network:	Vert. Text-10%	Vert. Text-1%	NAICS-10%	NAICS-1%	TNIC-3
Granularity	10%	1%	9.48%	1.37%	2.33%
% of pairs in TNIC-3	1.33%	2.39%	2.67%	2.89%	100%
% of pairs in the same SIC	0.74%	1.03%	0.35%	0.20%	38.10%
% of pairs in the same NAICS	0.59%	0.66%	0.30%	0.18%	38.11%
% of pairs in the same SIC or NAICS	0.81%	1.09%	0.35%	0.20%	41.24%
% of pairs in Vert. Text-10%	100%	100%	10.48%	13.18%	6.15%
% of pairs in Vert. Text-1%	10%	100%	1.18%	1.21%	1.09%
% of pairs that include a financial firm	9.20%	1.80%	48.44%	34.31%	58.72%
% of (no fin.) pairs in Vert. Text-10%	100%	100%	19.90%	19.29%	11.63%
% of (no fin.) pairs in Vert. Text-1%	10%	100%	2.14%	1.71%	2.44%

Note: This table displays various characteristics for five networks: Vertical Text-10% and Vertical Text-1% vertical networks, NAICS-10% and NAICS-1% vertical networks, and the TNIC-3 horizontal network.

Table III: Mergers and Acquisitions - Sample Description

Measure:	All	Text-Based		NAICS-based	
Deal type:		Vertical	Non-Vertical	Vertical	Non-Vertical
<i>Panel A: Sample Description</i>					
# Transactions	3,460	1,368	2,092	460	3,000
% Vertical (Non-Vertical)		39.54%	60.46%	13.29%	86.71%
# Upstream		687		199	
# Downstream		681		261	
<i>Panel B: Combined Acquirers and Targets Returns</i>					
CAR(0)	0.49%	0.65%	0.38% ^a	0.46%	0.49%
CAR(-1,1)	0.86%	0.97%	0.79%	0.55%	0.91%
# Transactions	3,256	1,301	1,995	427	2,829

Note: Panel A displays statistics for vertical and non-vertical transactions (non-financial firms only). A transaction is vertical if the acquirer and target are pairs in the Vertical Text-10% network or the NAICS-10% network. Panel B displays the average cumulated abnormal announcement returns (CARs) of combined acquirers and targets. We include the superscript ^a when the difference in CARs between vertical and non-vertical transactions is significant at the 5% level.

Table IV: Vertical Transactions - Deal-level Analysis

Variable:	Ind.(R&D/ sales)	R&D/ sales	Ind.(#Patents/ Assets)	#Patents/ Assets
<i>Panel A: Whole Sample</i>				
(i) Vert. Targets	0.0555	0.0424	0.0076	0.0091
(ii) Non-Vert. Targets	0.1262	0.0813	0.0077	0.0069
(iii) Non-Merging Firms	0.0905	0.0622	0.0075	0.0082
<i>t</i> -statistic [(i)-(ii)]	(-16.67) ^a	(-9.32) ^a	(-0.30)	(3.36) ^a
<i>t</i> -statistic [(i)-(iii)]	(-9.20) ^a	(-5.34) ^a	(0.27)	(1.45)
<i>t</i> -statistic [(ii)-(iii)]	(11.03) ^a	(6.07) ^a	(0.74)	(-2.42) ^b
<i>Panel B: Matched Targets I</i>				
(i) Vert. Targets	0.0555	0.0424	0.0076	0.0091
(ii) Matched Vert. Targets	0.0953	0.0592	0.0072	0.0065
<i>t</i> -statistic [(i)-(ii)]	(-9.64) ^a	(-4.01) ^a	(0.94)	(3.44) ^a
(i) Non-Vert. Targets	0.1262	0.0813	0.0077	0.0069
(ii) Matched Non-Vert. Targets	0.1073	0.0696	0.0079	0.0072
<i>t</i> -statistic [(i)-(ii)]	(4.15) ^a	(2.66) ^a	(-0.62)	(-0.52)
<i>Panel C: Matched Targets II</i>				
(i) Vert. Targets	0.0555	0.0424	0.0076	0.0091
(ii) Matched Vert. Targets	0.0802	0.0477	0.0061	0.0048
<i>t</i> -statistic [(i)-(ii)]	(-6.63) ^a	(-1.36)	(3.81) ^a	(6.39) ^a
(i) Non-Vert. Targets	0.1262	0.0813	0.0077	0.0069
(ii) Matched Non-Vert. Targets	0.0931	0.0584	0.0072	0.0062
<i>t</i> -statistic [(i)-(ii)]	(7.67) ^a	(5.62) ^a	(1.62)	(1.21)

Note: Transactions are defined as vertical when the acquirer and target are in pairs in the Vertical Text-10% network. In Panel A, we compare targets of vertical and non-vertical deals, and non-merging firms. In Panel B, each target is compared to a “matched” non-merging target using a propensity score model based on industry, size, and year. In Panel C, the propensity score is based on industry, size, age, Market-to-Book, PPE/Assets, End Users, # of NAICS Segments, and year. We report *t*-statistics corresponding to tests of mean differences. Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Table V: The Determinants of Vertical Target Acquisitions

Dep. Variable: Specification: Deal type:	Prob(Target)						
	Probit				IV Probit		
	Vertical (1)	Non-Vertical (2)	Vertical (3)	Non-Vertical (4)	1st Stage (5)	Vertical (6)	Non-Vertical (7)
Ind.(R&D/sales)	-0.132 ^a (0.023)	0.168 ^a (0.016)	-0.097 ^a (0.036)	0.257 ^a (0.021)		-0.104 ^a (0.027)	0.232 ^a (0.024)
Ind.(#Patents/asse)	0.154 ^a (0.014)	-0.080 ^a (0.017)	0.174 ^a (0.016)	0.009 (0.021)	-0.049 ^a (0.009)	0.132 ^a (0.016)	-0.114 ^a (0.022)
Ind.(R&D/sales) × Ind.(#Patents/assets)			-0.024 (0.015)	-0.065 ^a (0.012)			
Ind.(PPE/assets)	-0.007 (0.015)	-0.075 ^a (0.022)	-0.003 (0.015)	-0.048 ^b (0.022)	0.039 ^a (0.007)	0.001 (0.016)	-0.060 ^a (0.023)
HHI	-0.037 ^b (0.016)	-0.075 ^a (0.014)	-0.037 ^b (0.016)	-0.066 ^a (0.014)	0.017 ^a (0.004)	-0.032 ^b (0.016)	-0.062 ^a (0.015)
End User	-0.154 ^a (0.015)	0.086 ^a (0.012)	-0.153 ^a (0.015)	0.096 ^a (0.012)	0.008 ^a (0.002)	-0.153 ^a (0.015)	0.086 ^a (0.013)
#Segment (NAICS)	0.089 ^a (0.010)	-0.007 (0.025)	0.089 ^a (0.010)	-0.007 (0.011)	0.011 ^a (0.001)	0.087 ^a (0.011)	-0.008 (0.012)
log(Assets)	0.279 ^a (0.016)	0.193 ^a (0.013)	0.279 ^a (0.016)	0.195 ^a (0.013)	0.027 ^a (0.003)	0.278 ^a (0.017)	0.204 ^a (0.014)
log(Age)	0.074 ^a (0.015)	-0.014 (0.011)	0.074 ^a (0.015)	-0.012 (0.011)	-0.020 ^a (0.003)	0.078 ^a (0.016)	-0.006 (0.012)
MB	-0.110 ^a (0.020)	-0.032 ^b (0.013)	-0.111 ^a (0.020)	-0.033 ^b (0.012)	-0.005 (0.004)	-0.106 ^a (0.020)	-0.041 ^a (0.013)
Ind.(Predicted R&D/sales)					1.004 ^a (0.018)		
#obs.	45,198	45,198	45,198	45,198	40,017	40,017	40,017
Pseudo R ²	0.115	0.045	0.116	0.048	N/A	N/A	N/A

Note: The dependent variable in the probit models is a dummy indicating whether the given firm is a target in a vertical or non-vertical transaction in a given year. Vertical transactions are identified using the Vertical Text-10% network. The first four columns compare vertical and non-vertical transactions for the full sample. The last three columns report the results of IV probit estimations where we use tax-induced industry predicted R&D/sales (using exogenous variation in the user cost of R&D capital) as an instrument for industry R&D intensity (Ind.(R&D/sales)). All independent variables are defined in Appendix 2. The independent variables are standardized for convenience. All estimations also include year and FIC industry fixed effects. Standard errors are clustered by FIC industry and year and are reported in parentheses. FIC industries are the transitive version of TNIC industries from Hoberg and Phillips (2015). Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Table VI: Vertical Target Acquisitions: High Secrecy vs. Low Secrecy

Dep. Variable: Sub-Sample: Specification:	Prob(Vertical Target)			
	Secrecy			
	Probit		IV Probit	
	Low	High	Low	High
	(1)	(2)	(3)	(4)
Ind.(R&D/sales)	-0.317 ^a (0.039)	0.027 (0.043)	-0.309 ^a (0.042)	0.053 (0.052)
Ind.(#Patents/assets)	0.137 ^a (0.024)	0.030 (0.030)	0.125 ^a (0.027)	-0.008 (0.037)
Ind.(PPE/assets)	-0.029 (0.037)	-0.042 (0.046)	-0.028 (0.039)	-0.015 (0.046)
HHI	-0.080 ^b (0.032)	-0.070 ^a (0.028)	-0.092 ^a (0.035)	-0.077 ^a (0.029)
End User	-0.092 ^a (0.026)	-0.097 ^a (0.030)	-0.094 ^a (0.028)	-0.091 ^a (0.032)
#Segment (NAICS)	0.056 ^b (0.021)	0.073 ^a (0.024)	0.055 ^b (0.022)	0.077 ^a (0.025)
log(Assets)	0.285 ^a (0.028)	0.294 ^a (0.031)	0.290 ^a (0.030)	0.271 ^a (0.031)
log(Age)	0.069 ^b (0.028)	0.050 ^c (0.029)	0.058 ^c (0.030)	0.058 ^c (0.031)
MB	-0.106 ^b (0.045)	-0.177 ^a (0.038)	-0.105 ^b (0.046)	-0.161 ^a (0.038)
#obs.	9,409	9,333	8,342	8,138
Pseudo R^2	0.136	0.094	N/A	N/A

Note: The dependent variable in the probit models is a dummy indicating whether the given firm is a target in a vertical transaction in a given year. Vertical transactions are identified using the Vertical Text-10% network. The ‘Low’ and ‘High’ groups in the column headers correspond to industries where the importance of secrecy (as opposed to patents) for protecting innovation is below and respectively above the sample median as defined in the text. This sample is limited to 34 manufacturing industries. The first two columns report results from probit estimations. The last two columns report the second-stage results of IV probit estimations where we use tax-induced industry predicted R&D/sales (using exogenous variation in the user cost of R&D capital) as an instrument for industry R&D intensity (Ind.(R&D/sales)). All independent variables are defined in Appendix 2. In all columns, the independent variables are standardized for convenience. All estimations also include year and FIC industry fixed effects. Standard errors are clustered by FIC industry and year and are reported in parentheses. FIC industries are the transitive version of TNIC industries from Hoberg and Phillips (2015). Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Table VII: Averages by Quartiles of VI

Variable	Quartile 1 (Low VI)	Quartile 2	Quartile 3	Quartile 4 (High VI)
<i>Panel A: Data from Text Analysis</i>				
VI	0.002	0.006	0.011	0.028
Retail	0.396	0.366	0.346	0.301
Government	0.039	0.031	0.027	0.025
Export	0.079	0.084	0.087	0.095
End users	0.515	0.482	0.462	0.422
<i>Panel B: Data from Existing Literature</i>				
R&D/sales	0.098	0.062	0.047	0.027
#Patents/assets	0.006	0.008	0.008	0.007
log(1+#Patents)	0.420	0.504	0.640	0.835
PPE/assets	0.199	0.247	0.285	0.320
HHI	0.231	0.258	0.272	0.274
log(assets)	5.355	5.485	5.771	6.123
log(age)	2.736	2.788	2.970	3.318
#Segments	1.318	1.422	1.572	1.890
MB	2.355	2.080	1.880	1.606
VI _{segment}	0.006	0.009	0.013	0.024

Note: This table displays averages by (annually sorted) quartiles based on text-based vertical integration (VI). The sample includes 45,198 observations. All variables are defined in Appendix 2.

Table VIII: The Determinants of Vertical Integration

Dep. Variable: Specification:	(Text-based) VI							
	OLS				Instrumental Variables			
	Baseline		Interaction		1st Stage	2nd Stage	1st Stage	2nd Stage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	-0.100 ^a (0.007)	-0.017 ^a (0.005)	-0.081 ^a (0.008)	-0.012 ^c (-0.006)		-0.126 ^a (0.009)		-0.016 ^b (0.007)
Ind.(#Patents/assets)	0.068 ^a (0.007)	0.019 ^a (0.004)	0.084 ^a (0.009)	0.025 ^a (0.006)	-0.024 ^a (0.005)	0.075 ^a (0.008)	-0.019 ^a (0.006)	0.015 ^a (0.004)
Ind.(R&D/sales) × Ind.(#Patents/assets)			-0.014 ^a (0.004)	-0.005 ^c (0.002)				
Ind.(PPE/assets)	0.022 ^c (0.012)	0.014 (0.009)	0.023 ^c (0.012)	0.014 (0.009)	0.027 ^a (0.004)	0.016 (0.014)	0.012 ^a (0.004)	0.009 (0.009)
HHI	-0.107 ^a (0.006)	-0.055 ^a (0.004)	-0.106 ^a (-0.006)	-0.055 ^a (0.004) ^a	0.006 ^a (0.002)	-0.111 ^a (0.007)	-0.000 (0.002)	-0.056 ^a (0.004)
End User	-0.240 ^a (0.007)	-0.142 ^a (0.006)	-0.240 ^a (0.007)	-0.142 ^a (0.006)	-0.004 ^a (0.001)	-0.238 ^a (0.008)	-0.005 ^b (0.002)	-0.136 ^a (0.006)
#Segment (NAICS)	0.131 ^a (0.005)	0.041 ^a (0.006)	0.131 ^a (0.005)	0.041 ^a (0.006)	0.003 ^a (0.001)	0.129 ^a (0.006)	0.003 (0.002)	0.036 ^a (0.005)
log(Assets)	0.051 ^a (0.004)	0.124 ^a (0.010)	0.051 ^a (0.004)	0.124 ^a (0.011)	0.014 ^a (0.001)	0.051 ^a (0.005)	0.026 ^a (0.007)	0.123 ^a (0.010)
log(Age)	0.021 ^a (0.004)	0.014 (0.010)	0.020 ^a (0.004)	0.014 (0.010)	-0.012 ^a (0.001)	0.020 ^a (0.005)	-0.020 ^a (0.005)	0.018 ^c (0.009)
MB	-0.016 ^a (0.003)	0.005 ^c (0.002)	-0.016 ^a (0.003)	0.005 ^b (0.002)	-0.005 ^b (0.002)	-0.013 ^a (0.003)	-0.007 ^c (0.004)	0.007 ^a (0.002)
Ind.(Predicted R&D/sales)					0.976 ^a (0.011)		1.032 ^a (0.016)	
Industry Fixed Effects	Yes	No	Yes	No	Yes	Yes	No	No
Firm Fixed Effects	No	Yes	No	Yes	No	No	Yes	Yes
#obs.	45,198	45,198	45,198	45,198	40,017	40,017	40,017	40,017
Adj. R^2	0.526	0.855	0.527	0.855	0.774	0.704	0.937	0.905

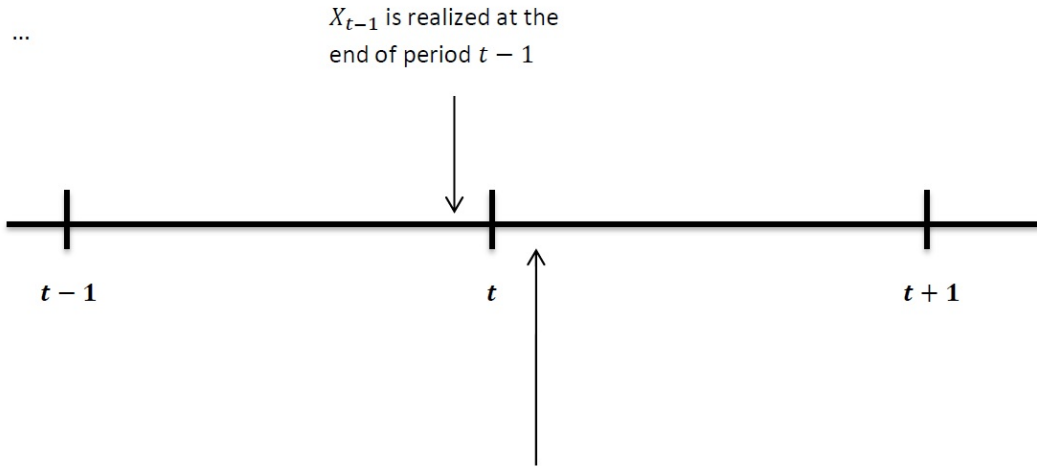
Note: The dependent variable is vertical integration VI . The first four columns are based on OLS regressions with industry or firm fixed effects as noted. The last four columns report results of instrumental variables estimations with industry or firm fixed effects as noted, where we use tax-induced industry predicted R&D/sales (using exogenous variation in the user cost of R&D capital) as an instrument for industry R&D intensity (Ind.(R&D/sales)). All estimations also include year fixed effects. Industry fixed effects are based on FIC industries (the transitive version of TNIC industries from Hoberg and Phillips (2015)). All independent variables are defined in Appendix 2. The independent variables are standardized for convenience. Standard errors are clustered by FIC industry and year and are reported in parentheses. Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Table IX: The Determinants of Vertical Integration: High vs. Low Secrecy Industries

Dep. Variable: Sub-Sample: Specification:	(Text-based) VI							
	Secrecy							
	OLS				Instrumental Variables			
	Low	High	Low	High	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	-0.141 ^a (0.015)	-0.104 ^a (0.016)	-0.036 ^a (0.013)	-0.004 (0.013)	-0.174 ^a (0.018)	-0.144 ^a (0.020)	-0.042 ⁴ (0.016)	-0.013 (0.018)
Ind.(#Patents/assets)	0.042 ^a (0.014)	0.002 (0.010)	0.023 ^a (0.009)	0.013 (0.008)	0.055 ^a (0.015)	0.007 (0.010)	0.020 ^b (0.008)	0.010 (0.007)
Ind.(PPE/assets)	-0.027 (0.029)	0.077 ^a (0.023)	0.017 (0.024)	-0.041 ^c (0.023)	-0.048 (0.030)	0.006 ^a (0.024)	0.022 (0.023)	-0.056 ^b (0.022)
HHI	-0.173 ^a (0.016)	-0.165 ^a (0.013)	-0.069 ^a (0.011)	-0.089 ^a (0.012)	-0.183 ^a (0.017)	-0.173 ^a (0.014)	-0.080 ^a (0.010)	-0.090 ^a (0.011)
End User	-0.25 ^a (0.017)	-0.311 ^a (0.014)	-0.227 ^a (0.019)	-0.179 ^a (0.018)	-0.252 ^a (0.018)	-0.298 ^a (0.015)	-0.215 ^a (0.018)	-0.174 ^a (0.019)
#Segment (NAICS)	0.102 ^a (0.014)	0.155 ^a (0.014)	0.051 ^a (0.016)	0.068 ^a (0.017)	0.104 ^a (0.014)	0.158 ^a (0.014)	0.037 ^b (0.015)	0.064 ^a (0.017)
log(Assets)	0.036 ^a (0.011)	0.057 ^a (0.010)	0.114 ^a (0.025)	0.238 ^a (0.032)	0.036 ^a (0.012)	0.056 ^a (0.010)	0.123 ^a (0.023)	0.251 ^a (0.032)
log(Age)	0.002 (0.014)	-0.012 (0.012)	-0.008 (0.032)	0.012 (0.030)	0.003 (0.014)	-0.013 (0.012)	-0.012 (0.032)	0.003 (0.029)
MB	-0.001 (0.007)	-0.046 ^a (0.008)	0.001 (0.006)	0.004 (0.006)	0.002 (0.007)	-0.041 ^a (0.007)	0.006 (0.005)	0.007 (0.006)
Industry Fixed Effects	Yes	Yes	No	No	Yes	Yes	No	No
Firm Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes
#obs.	9,409	9,333	9,409	9,333	8,342	8,139	8,342	8,139
Adj. R^2	0.570	0.509	0.857	0.815	0.749	0.779	0.916	0.917

Note: The dependent variable is vertical integration VI . The first four columns are based on OLS regressions with industry or firm fixed effects as noted. The last four columns report second-stage results of instrumental variables estimations with industry or firm fixed effects as noted, where we use tax-induced industry predicted R&D/sales (using exogenous variation in the user cost of R&D capital) as an instrument for industry R&D intensity (Ind.(R&D/sales)). The ‘Low’ and ‘High’ groups in all columns correspond to industries where the importance of secrecy (as opposed to patents) for protecting innovation is below and respectively above the sample median as defined in the text. This sample is limited to 34 manufacturing industries. All estimations also include year fixed effects. Industry fixed effects are based on FIC industries (the transitive version of TNIC industries from Hoberg and Phillips (2015)). All independent variables are defined in Appendix 2. The independent variables are standardized for convenience. Standard errors are clustered by FIC industry and year and are reported in parentheses. Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Figure 1:



At the beginning of period t

Actions:

- (1) Producer decides I_t given X_{t-1}
- (2) Choose x_t and y_t given I_t
- (3) Renegotiation

Prices:

$$P_t = \begin{cases} P_t^b(1 + y_t) & \text{if } I_t = 0 \\ P_t^b(1 + \rho(y_t)) & \text{if } I_t = 1 \end{cases}$$

$$P_t^b = \begin{cases} P_s + (P_{s+1} - P_s)X_{t-1} & \text{if } P_{t-1}^b = P_s \text{ with } 0 \leq s < N \\ P_N & \text{if } P_{t-1}^b = P_N \end{cases}$$

Payoffs:

- (1) $TS_t = P_t - Sx_t^g - Ry_t^h$
- (2) The split of the sales depends on α
- (3) Supplier's profit is αTS_t , and producer's profit is $(1 - \alpha)TS_t$

Figure 2:

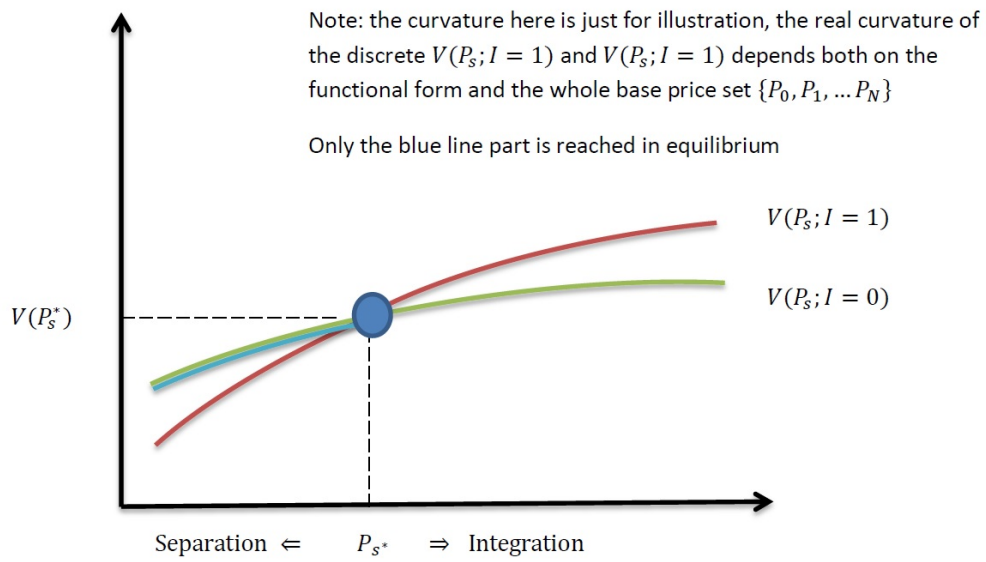


Figure 3: R&D and Patents prior to acquisitions. The figure shows the average R&D (lower panel) and patenting activity (upper panel) of firms that are targets in vertical and non-vertical acquisitions prior to the acquisition. Solid lines represent vertical transactions identified using the Vertical Text-10% network. Dashed lines represent non-vertical transactions.

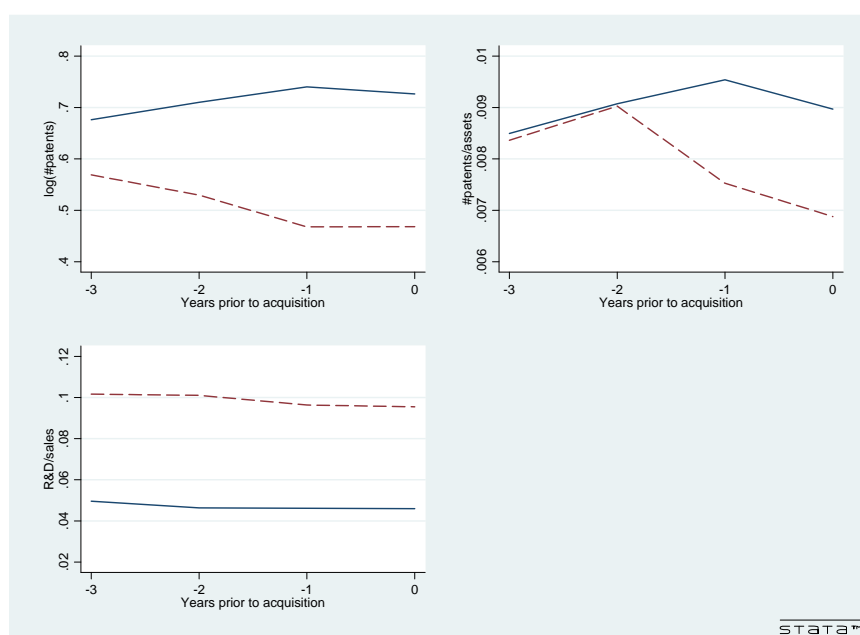


Figure 4: R&D intensity around vertical acquisitions. The figure shows the average R&D/sales around the years that surround vertical transactions. The solid line displays all vertical targets that continue to exist for at least one year after being acquired. The dashed line displays “combined” entities that aggregate R&D and sales of acquirers and targets. Vertical transactions are identified using the Vertical Text-10% network.

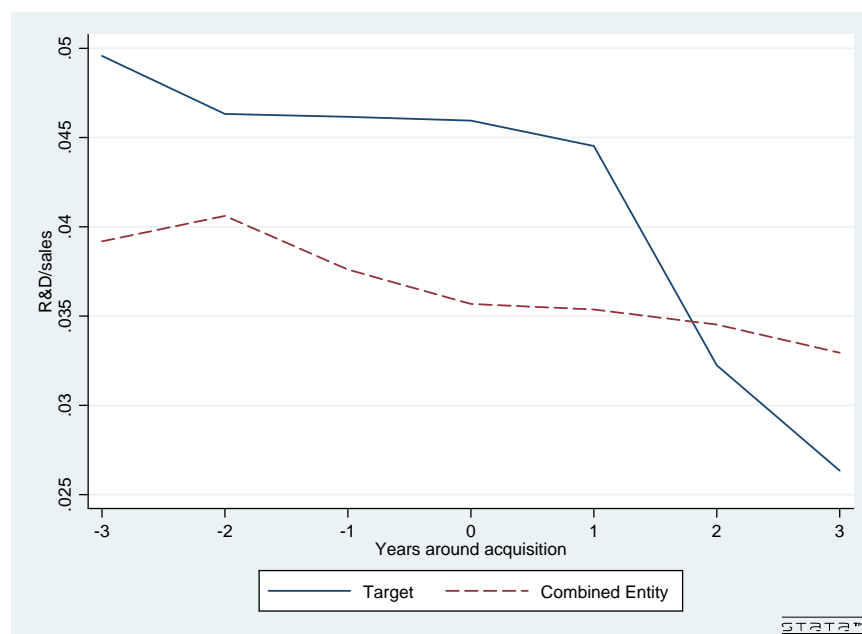


Figure 5: Evolution of sample-wide average (text-based) Vertical Integration over time. Vertical integration (VI) is defined in Section V.A. The solid blue line is the annual equal-weighted average VI . The dashed red line is the corresponding sales-weighted average.

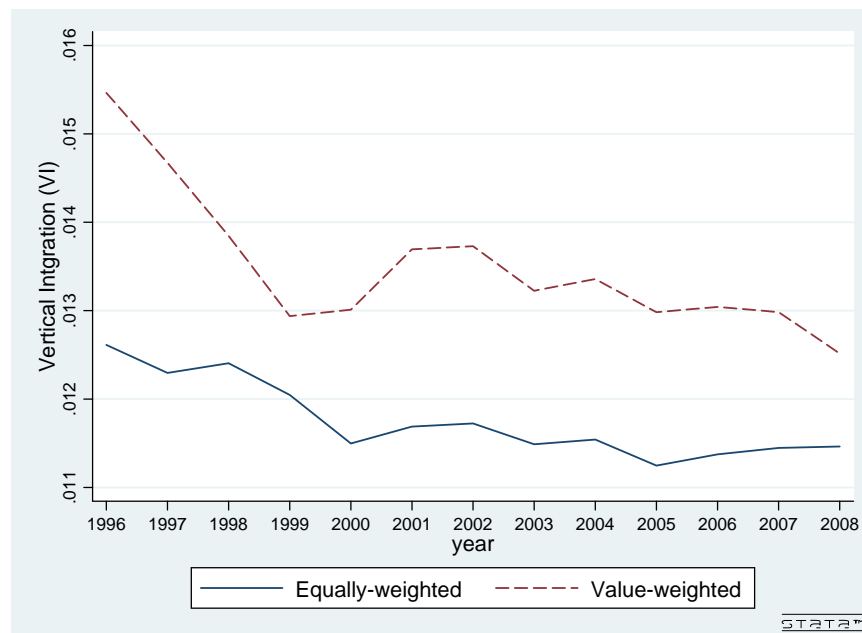


Figure 6: An Example: the Network Equipment Industry. The figure plots the evolution of text-based vertical integration (VI), patenting activity ($\log(\#\text{patents})$ and $\#\text{patents}/\text{assets}$) and R&D activity ($\text{R\&D}/\text{sales}$) for seven representative firms in the network equipment industry: Cisco, Broadcom, Citrix, Juniper, Novell, Sycamore, and Utstarcom.

