

Internet Appendix for “Innovation Activities and the Incentives for Vertical Acquisitions and Integration”

(Not for publication)

Laurent Frésard, Gerard Hoberg, and Gordon Phillips

August 31, 2015

This appendix contains additional material that we mention in the paper, but do not report to preserve space. Section I lists words from the BEA commodity vocabulary that we exclude because they are used in a large number of commodities. Section II lists the phrase exclusions from firms’ 10-Ks that we apply to construct vertical links between firms. Section III provides validation tests for our text-based measure of firm-level vertical integration. Finally, we report in Section IV additional tests that assess the robustness of our main results.

I Excluded BEA words

Because they are used in the largest number of commodities, we exclude the following words from the BEA commodity vocabulary we use to compute vertical relatedness:

Accessories, accessory, air, airs, attachment, attachments, commercial, commercials, component, components, consumer, consumers, development, developments, equipment, exempt, expense, expenses, ga, gas, industrial, industrials, net, part, parts, processing, product, products, purchased, purchase, receipt, receipts, research, researches, sale, sales, service, services, system, systems, unit, units, work, works, tax, taxes, oil, repair, repairs, aids, aid, air, apparatuses, apparatus, applica-

tions, application, assemblies, assembly, attachments, attachment, automatic, auxiliary, bars, bar, bases, base, blocks, block, bodies, body, bulk, business, businesses, byproducts, byproduct, cares, care, centers, center, collections, collection, combinations, combination, commercials, commercial, completes, complete, components, component, consumers, consumer, consumption, contracts, contract, controls, control, covers, cover, customs, custom, customers, customer, cuts, cut, developments, development, directly, distributions, distribution, domestic, dries, dry, equipments, equipment, establishments, establishment, exempt, expenses, expense, facilities, facility, fees, fee, fields, field, finished, finish, finishings, finishing, gas, generals, general, greater, hands, hand, handling, high, hot, individuals, individual, industrials, industrial, industries, industry, installations, installation, lights, light, lines, line, maintenances, maintenance, managements, management, manmade, manufactured, manufacture, materials, material, naturals, natural, nets, net, offices, office, only, open, operated, operate, organizations, organization, others, other, pads, pad, paid, pay, parts, part, permanent, portable, powers, power, processing, products, product, productions, production, public, purchased, purchase, purposes, purpose, receipts, receipt, reclassified, reclassify, repairs, repair, researches, research, sales, sale, self, services, service, sets, set, shares, share, shipped, similar, singles, single, sizes, size, small, soft, specials, special, stocks, stock, storages, storage, supplies, supply, supports, support, surfaces, surface, systems, system, taxes, tax, taxable, technical, this, trades, trade, transfers, transfer, types, type, units, unit, used, without, work, works.

II 10-K Phrase Exclusions

Because we use 10-K text only to identify a firm's own-product market location (vertically related vocabulary is identified using BEA data), we exclude any part of a sentence that follows any of the following 81 phrases:

Buy, buys, sells its, are sold, buying, products for, for sale, for their, used in, used by, used as, used for, used with, used primarily, used mainly, used commonly,

primarily used, mainly used, commonly used, for use, uses, utilized, serve, serving, serves, sold to, sold primarily, sold mainly, sold commonly, designed for, supply of, supply for, supplier to, supplied to, service to, purchase, purchaser, purchasers, customer, customers, user, users, for application, equipment for, equipment to, equipment by, product for, product to, product by, solution for, solution to, solution by, component for, component to, component by, application for, application to, application by, system for, system to, system by, equipments for, equipment for, equipment to, equipments to, equipments by, products for, products to, products by, solutions for, solutions to, solutions by, components for, components to, components by, applications for, applications to, applications by, systems for, systems to, systems by.

III Validation of Text-Based Vertical Integration

We perform two analyses to provide additional external validation for the text-based measure of vertical integration VI we introduce and use in the paper. First, we compare our measure to a direct but binary measure of vertical integration based on whether firms explicitly mention that they are vertically integrated in their 10-Ks. Second, we compare our measure of vertical integration to (industry) measures of related-party trade (RPT).

A Detecting Explicit Integration

We identify whether a firm explicitly indicates that it is vertically integrated by searching for the terms ‘vertical integration’ and ‘vertically integrated’ in each firm’s 10-K. We exclude cases where a firm indicates it is not integrated or lacks integration. We then create a dummy variable VI_{search} that is equal to one when a firm explicitly states that it is vertically integrated in a given year, and zero otherwise. Because this measure is based on direct statements by firms and does not rely on firms’ product description or the input-output economic flows, it enables us to gauge the ability of our text-based measure to identify firms that mention being integrated,

and to compare it with the existing measure based on Compustat segments, that we label $VI_{segment}$ in the text.

Table IA.III.1 presents results from probit regressions estimating the probability that a firm explicitly indicates that it is vertically integrated ($VI_{search} = 1$) as a function of VI and $VI_{segment}$. To provide more meaningful economic comparisons, we standardize both independent variables so that they have unit standard deviation. The first column indicates that our text-based measure of vertical integration has a much higher propensity to detect explicitly stated vertical integration compared to the Compustat segment-based measure. The estimated coefficient on VI is roughly four times larger than that on $VI_{segment}$ (0.213 versus 0.061). The statistical significance is also much larger on VI . The superior performance of VI continues to hold when we include VI and $VI_{segment}$ separately (columns (2) and (3)). In these columns, we also observe that the explanatory power of VI (measured by pseudo R^2) is much larger than that of $VI_{segment}$. Columns (4) to (6) reveal that the differences are robust to the inclusion of year and industry fixed effects.

B Related-Party Trade

As an alternative way of to provide external validation, we relate our text-based measure of vertical integration to industry measures of related-party trade (RPT) provided by the U.S. Census Bureau.¹ The data measure the intensity of trade (both imports or exports) that occurs between related parties, where “related party trade” is defined as trade with an entity located outside the United States in which the importer (exporter) holds at least a 6% (10%) equity interest (as defined by the Census). The data thus captures the intensity of international transactions that occur within firms’ boundaries. Arguably, related party trade could capture both horizontal and vertical flows of goods. Yet, to the extent that our text-based measure of vertical integration builds on vertical relations between products described in firms’ 10Ks, any correlation between our measure and RPT should be related to international transactions that are vertical in nature (see Antras (2013), or Antras

¹<http://sasweb.ssd.census.gov/relatedparty/>

and Chor (2013) for instance).

The RPT data is available over the 2000-2010 period at the NAICS 6-digit level. We aggregate the data at the NAICS 4-digit and 5-digit levels to map it to our Compustat sample. For each industry, we compute the share of related-party imports to total imports to capture the propensity of firms to integrate foreign supplier activities (RPT(import)). Similarly, we compute the share of related-party exports in total exports to capture the propensity of firms to integrate foreign customers (RPT(export)). We also consider the average share between the import and export shares (RPT). We then aggregate VI and $VI_{segment}$ at the industry-level (NAICS 4-digit and 5-digit levels) using equally-weighted averages.

Table IA.III.2 presents the results of OLS regressions of industry-level VI (or $VI_{segment}$) on the three measures of related-party trade. Across all specifications, we observe a positive correlation between our text-based measure vertical integration and measures of RPT. Focusing on the average level of RPT in the first column, the correlation with VI is 0.580 at the NAICS 4-digit industry level, and 0.833 at the NAICS 5-digit industry level. Both coefficients are statistically significant at the 5% confidence level (t-statistics of 2.12 and 4.37 respectively). At both aggregation levels, our measure of vertical integration is also more strongly related to related-party import transactions compared to related-party export transactions (columns (2) and (3)). The coefficients on related-party import are 0.644 and 0.702, compared to 0.077 and 0.519 for related-party export. Moreover, columns (5) to (6) indicate that related-party trade is only weakly related to vertical integration when measured using Compustat segments as an alternative.

IV Additional Results

This section contains additional tables that are mentioned and described in the paper but were not reported there to preserve space. Specifically, this appendix includes:

- **Table IA.IV.1:** Probit and OLS regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration. We focus on own-firm independent variables instead of industry variables.
- **Table IA.IV.2:** Probit and OLS regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration. We focus on (one-year) lagged independent variables instead of contemporaneous variables.
- **Table IA.IV.3:** Probit and OLS regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration. We focus on sales-weighted industry average variables instead of equally-weighted average variables.
- **Table IA.IV.4:** Probit and OLS regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration. We only consider industry $R\&D/Sale$ and $\#Patents/assets$ as independent variables, and not the baseline full set of independent variables.
- **Table IA.IV.5:** Probit and OLS regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration. We include industry $R\&D/Sale$ and $\#Patents/assets$ individually and not together.
- **Table IA.IV.6:** Probit and OLS regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration. Vertical and non-vertical targets are identified using the NAICS-10% vertical network instead of our text-based 10% vertical network. The measure of vertical integration is based in Compustat segments ($VI(Segment)$) instead of our text-based measure.
- **Table IA.IV.7:** Probit regressions whose dependent variables are the probability of being an acquirer in a vertical or non-vertical transaction.

- **Table IA.IV.8:** OLS regressions of firm-year R&D/sales on the user cost of R&D capital. We use the estimation presented in column (3) to predict R&D for each firm-year, and use these predicted values to construct our instrument.
- **Table IA.IV.9:** Linear Probability Model (LPM) regressions whose dependent variables are the probability of being a target in a vertical or non-vertical transaction or vertical integration.
- **Table IA.IV.10:** Difference in average R&D/sales before and after vertical transactions for targets that continue to exist post-transactions, and for combined entities that aggregate R&D and sales across acquirers and targets for each vertical deal.
- **Table IA.IV.11:** List of the 30 most vertically integrated firms in 2008 based on our text-based measure of vertical integration (*VI*).
- **Table IA.IV.12:** OLS regressions whose dependent variable is the logarithm of vertical integration instead of its level.

Table IA.III.1: VI Detection

Dep. Variable:	Prob($VI_{search} = 1$)					
	(1)	(2)	(3)	(4)	(5)	(6)
VI	0.213 ^a (0.007)	0.224 ^a (0.007)		0.123 ^a (0.010)	0.130 ^a (0.010)	
$VI_{segment}$	0.061 ^a (0.007)		0.096 ^a (0.006)	0.052 ^a (0.008)		0.062 ^a (0.008)
Year FE	No	No	No	Yes	Yes	Yes
Industry FE	No	No	No	Yes	Yes	Yes
#.Obs.	45,198	45,198	45,198	45,198	45,198	45,198
Pseudo R^2	0.037	0.034	0.007	0.127	0.125	0.122

Note: This table reports Probit estimations where the dependent variable is VI_{search} , a dummy that equals one if a firm mentions being vertically integrated in its annual 10-K report, and zero otherwise. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.III.2: VI and Related-Party Trade

Dep. Variable:	<i>VI</i>			<i>VI_{segment}</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: NAICS 4-digit industries						
RPT	0.580 ^b			0.361		
	(0.272)			(0.273)		
RPT(import)		0.644 ^a			0.706 ^a	
		(0.208)			(0.208)	
RPT(export)			0.077			-0.554 ^c
			(0.295)			(0.294)
Obs.	636	636	636	636	636	636
Pseudo R ²	0.005	0.013	0.001	0.001	0.016	0.004
Panel B: NAICS 5-digit industries						
RPT	0.833 ^a			-0.510 ^a		
	(0.190)			(0.191)		
RPT(import)		0.702 ^a			-0.122	
		(0.146)			(0.147)	
RPT(export)			0.519 ^a			-0.876 ^a
			(0.198)			(0.197)
#.Obs.	1,122	1,122	1,122	1,122	1,122	1,122
Pseudo R ²	0.015	0.019	0.005	0.005	0.001	0.016

Note: Columns (1) to (3) report OLS estimations where the dependent variable is our new text-based measure of vertical integration *VI*. Columns (4) to (6) report OLS estimations where the dependent variable is a measure of vertical integration based on Compustat segments *VI_{segment}*. In Panel A, all variables are aggregated at the NAICS 4-digit industry level (averages). In Panel B, all variables are aggregated at the NAICS 5-digit industry level (averages). The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.1: Vertical Acquisitions and Integration: Own-Firm Variables

Dep. Variables:	Prob(Target) (Probit)				VI (OLS)			
	Vert.	Non-Vert.	Vert.	Non-Vert.	Baseline		Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
R&D/sales	0.017 (0.019)	0.091 ^a (0.013)	0.035 (0.021)	0.113 ^a (0.013)	-0.027 ^a (0.005)	-0.004 (0.004)	-0.025 ^a (0.005)	-0.007 (0.004)
#Patents/assets	0.101 ^a (0.013)	-0.027 ^b (0.013)	0.118 ^a (0.015)	0.021 (0.017)	0.017 ^a (0.004)	0.006 (0.004)	0.020 ^a (0.006)	0.003 (0.005)
R&D/sales × #Patents/assets			-0.013 (0.009)	-0.028 ^a (0.007)			-0.002 (0.002)	0.003 (0.002)
PPE/assets	-0.028 ^c (0.015)	-0.077 ^a (0.018)	-0.025 ^c (0.015)	-0.072 ^a (0.018)	0.030 ^a (0.010)	0.037 ^a (0.009)	0.031 ^a (0.010)	0.037 ^a (0.009)
HHI	-0.028 ^b (0.015)	-0.078 ^a (0.014)	-0.028 ^c (0.015)	-0.077 ^a (0.014)	-0.105 ^a (0.006)	-0.056 ^a (0.005)	-0.105 ^a (0.006)	-0.056 ^a (0.005)
End User	-0.166 ^a (0.015)	0.095 ^a (0.012)	-0.165 ^a (0.015)	0.098 ^a (0.012)	-0.241 ^a (0.008)	-0.142 ^a (0.007)	-0.241 ^a (0.008)	-0.142 ^a (0.007)
#Segment (NAICS)	0.088 ^a (0.010)	-0.011 (0.011)	0.089 ^a (0.010)	-0.022 (0.011)	0.132 ^a (0.006)	0.041 ^a (0.006)	0.132 ^a (0.006)	0.041 ^a (0.006)
log(Assets)	0.285 ^a (0.016)	0.195 ^a (0.013)	0.285 ^a (0.016)	0.196 ^a (0.013)	0.049 ^a (0.005)	0.127 ^a (0.011)	0.049 ^a (0.005)	0.127 ^a (0.011)
log(Age)	0.086 ^a (0.015)	-0.025 ^b (0.011)	0.085 ^a (0.015)	-0.026 ^b (0.011)	0.024 ^a (0.005)	0.012 (0.010)	0.024 ^a (0.005)	0.012 (0.010)
MB	-0.145 ^a (0.021)	-0.012 (0.012)	-0.146 ^a (0.021)	-0.014 (0.012)	-0.019 ^a (0.003)	0.005 ^b (0.003)	-0.019 ^a (0.003)	0.005 ^b (0.003)
Industry Fixed Effects	No	No	No	No	Yes	No	No	Yes
Firm Fixed Effects	No	No	No	No	No	Yes	No	Yes
#obs.	45,198	45,198	45,198	45,198	45,198	45,198	45,198	45,198
Pseudo R^2 / Adj. R^2	0.113	0.037	0.114	0.038	0.525	0.855	0.525	0.855

Note: Columns (1) to (4) report probit estimations where the dependent variable 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. Columns (5) to (8) report OLS estimation where the dependent variable is vertical integration VI . All estimations include year fixed effects. All variables are defined in Appendix 2 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.2: Vertical Acquisitions and Integration: Lagged Variables

Dep. Variables:	Prob(Target) (probit)				VI (OLS)			
	Vert.	Non-Vert.	Vert.	Non-Vert.	Baseline		Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	-0.113 ^a (0.014)	0.184 ^a (0.016)	-0.097 ^b (0.038)	0.252 ^a (0.023)	-0.097 ^a (0.008)	-0.018 ^a (0.006)	-0.084 ^a (0.009)	-0.020 ^a (0.007)
Ind.(#Patents/assets)	0.167 ^a (0.015)	-0.091 ^a (0.018)	0.175 ^a (0.020)	-0.020 (0.022)	0.067 ^a (0.008)	0.011 ^b (0.005)	0.077 ^a (0.010)	0.008 (0.007)
Ind.(R&D/sales × #Patents/assets)			-0.009 (0.018)	-0.048 ^a (0.012)			-0.010 ^b (0.005)	0.002 (0.003)
Ind.(PPE/assets)	0.004 (0.016)	-0.071 ^a (0.023)	0.006 (0.017)	-0.049 ^b (0.023)	0.027 ^c (0.014)	0.021 ^b (0.010)	0.028 ^b (0.014)	0.021 ^b (0.010)
HHI	-0.015 (0.018)	-0.083 ^a (0.015)	-0.015 (0.018)	-0.075 ^a (0.015)	-0.089 ^a (0.007)	-0.023 ^a (0.005)	-0.088 ^a (0.007)	-0.023 ^a (0.005)
End User	-0.149 ^a (0.015)	0.083 ^a (0.014)	-0.149 ^a (0.015)	0.091 ^a (0.014)	-0.223 ^a (0.008)	-0.088 ^a (0.006)	-0.222 ^a (0.008)	-0.088 ^a (0.006)
#Segment (NAICS)	0.104 ^a (0.011)	0.017 (0.012)	0.104 ^a (0.011)	0.017 (0.012)	0.129 ^a (0.006)	0.032 ^a (0.006)	0.129 ^a (0.006)	0.032 ^a (0.006)
log(Assets)	0.284 ^a (0.017)	0.199 ^a (0.014)	0.284 ^a (0.017)	0.201 ^a (0.014)	0.049 ^a (0.005)	0.102 ^a (0.013)	0.049 ^a (0.005)	0.103 ^a (0.013)
log(Age)	0.063 ^a (0.017)	-0.038 ^a (0.013)	0.063 ^a (0.017)	-0.036 ^a (0.013)	0.023 ^a (0.006)	0.012 (0.014)	0.023 ^a (0.006)	0.012 (0.014)
MB	-0.079 ^a (0.023)	-0.011 (0.012)	-0.079 ^a (0.023)	-0.013 (0.0212)	-0.016 ^a (0.003)	0.004 (0.003)	-0.016 ^a (0.003)	0.003 (0.003)
Industry Fixed Effects	No	No	No	No	Yes	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	Yes	No	Yes
#Obs.	37,820	37,820	37,820	37,820	37,820	37,820	37,820	37,820
Pseudo R^2 / Adj. R^2	0.116	0.052	0.116	0.053	0.521	0.855	0.521	0.855

Note: Columns (1) to (4) report probit estimations where the dependent variable 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. Columns (5) to (8) report OLS estimation where the dependent variable is vertical integration VI . All estimations include year fixed effects. All variables are defined in Appendix 2 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.3: Vertical Acquisitions and Integration: Sales-Weighted Variables

Dep. Variables:	Prob(Target) (probit)				VI (OLS)			
	Vert.	Non-Vert.	Vert.	Non-Vert.	Baseline		Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	-0.056 ^b (0.023)	0.148 ^a (0.015)	-0.041 (0.027)	0.184 ^a (0.018)	-0.057 ^a (0.006)	-0.016 ^a (0.005)	-0.031 ^a (0.008)	-0.006 (0.005)
Ind.(#Patents/assets)	0.129 ^a (0.014)	-0.074 ^a (0.016)	0.149 ^a (0.019)	-0.004 (0.023)	0.060 ^a (0.007)	0.013 ^a (0.004)	0.093 ^a (0.011)	0.025 ^a (0.007)
Ind.(R&D/sales × #Patents/assets)			-0.013 (0.011)	-0.041 ^a (0.010)			-0.021 ^a (0.005)	-0.008 ^b (0.003)
Ind.(PPE/assets)	-0.002 (0.016)	-0.088 ^a (0.018)	-0.001 (0.016)	-0.075 ^a (0.018)	0.011 (0.009)	0.006 (0.007)	0.014 (0.009)	0.006 (0.007)
HHI	-0.027 (0.018)	-0.078 ^a (0.016)	-0.071 (0.018)	-0.077 ^a (0.016)	-0.106 ^a (0.007)	-0.056 ^a (0.005)	-0.104 ^a (0.007)	-0.056 ^a (0.005)
End User	-0.153 ^a (0.014)	0.099 ^a (0.013)	-0.152 ^a (0.014)	0.102 ^a (0.013)	-0.241 ^a (0.008)	-0.142 ^a (0.007)	-0.240 ^a (0.008)	-0.142 ^a (0.007)
#Segment (NAICS)	0.091 ^a (0.010)	-0.008 (0.012)	0.091 ^a (0.010)	-0.008 (0.012)	0.131 ^a (0.006)	0.041 ^a (0.006)	0.131 ^a (0.006)	0.041 ^a (0.006)
log(Assets)	0.277 ^a (0.016)	0.203 ^a (0.013)	0.277 ^a (0.016)	0.205 ^a (0.012)	0.050 ^a (0.005)	0.122 ^a (0.011)	0.050 ^a (0.005)	0.122 ^a (0.011)
log(Age)	0.083 ^a (0.015)	-0.021 ^c (0.011)	0.083 ^a (0.015)	-0.023 ^b (0.011)	0.023 ^a (0.005)	0.014 (0.010)	0.022 ^a (0.005)	0.014 (0.010)
MB	-0.118 ^a (0.020)	-0.028 ^b (0.012)	-0.120 ^a (0.020)	-0.032 ^b (0.013)	-0.019 ^a (0.003)	0.004 ^c (0.003)	-0.019 ^a (0.003)	0.005 ^c (0.003)
Industry Fixed Effects	No	No	No	No	Yes	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	Yes	No	Yes
#Obs.	45,198	45,198	45,198	45,198	45,198	45,198	45,198	45,198
Pseudo R^2 / Adj. R^2	0.113	0.043	0.113	0.044	0.525	0.855	0.526	0.855

Note: Columns (1) to (4) report probit estimations where the dependent variable 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. Columns (5) to (8) report OLS estimation where the dependent variable is vertical integration VI . All estimations include year fixed effects. All variables are defined in Appendix 2 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.4: Vertical Acquisitions and Integration: Limited Models

Dep. Variables:	Prob(Target) (probit)				VI (OLS)			
	Vert.	Non-Vert.	Vert.	Non-Vert.	Baseline		Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	-0.260 ^a (0.023)	0.186 ^a (0.013)	-0.243 ^a (0.032)	0.252 ^a (0.018)	-0.100 ^a (0.007)	-0.001 (0.005)	-0.067 ^a (0.008)	0.013 ^b (0.006)
Ind.(#Patents/assets)	0.135 ^a (0.013)	-0.103 ^a (0.0016)	0.146 ^a (0.015)	-0.032 (0.019)	0.073 ^a (0.008)	0.016 ^a (0.005)	0.103 ^a (0.010)	0.029 ^a (0.007)
Ind.(R&D/sales × #Patents/assets)			-0.013 (0.014)	-0.053 ^a (0.010)			-0.026 ^a (0.005)	-0.012 ^a (0.003)
Industry Fixed Effects	No	No	No	No	Yes	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	Yes	No	Yes
#Obs.	45,198	45,198	45,198	45,198	45,198	45,198	45,198	45,198
Pseudo R^2 / Adj. R^2	0.019	0.018	0.019	0.020	0.473	0.849	0.474	0.849

Note: Columns (1) to (4) report probit estimations where the dependent variable 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. Columns (5) to (8) report OLS estimation where the dependent variable is vertical integration VI . All estimations include year fixed effects. All variables are defined in Appendix 2 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.5: Vertical Acquisitions and Integration: R&D and Patents Individually

Dep. Variables:	Prob(Target) (probit)				VI (OLS)			
	Vert.	Vert.	Non-Vert.	Non-Vert.	Baseline			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	-0.023 (0.019)		0.128 ^a (0.014)		-0.072 ^a (0.007)		-0.013 ^b (0.005)	
Ind.(#Patents/assets)		0.109 ^a (0.013)		0.011 (0.013)		0.043 ^a (0.007)		0.017 ^a (0.005)
Ind.(PPE/assets)	-0.024 ^c (0.015)	0.019 (0.014)	-0.069 ^a (0.022)	-0.120 ^a (0.022)	0.021 (0.013)	0.038 ^a (0.013)	0.015 (0.009)	0.015 ^c (0.009)
HHI	-0.038 ^b (0.016)	-0.016 (0.015)	-0.072 ^a (0.014)	-0.105 ^a (0.014)	-0.108 ^a (0.007)	-0.097 ^a (0.007)	-0.055 ^a (0.005)	-0.054 ^a (0.005)
End User	-0.168 ^a (0.014)	-0.158 ^a (0.015)	0.095 ^a (0.012)	0.090 ^a (0.012)	-0.242 ^a (0.008)	-0.240 ^a (0.008)	-0.142 ^a (0.007)	-0.142 ^a (0.007)
#Segment (NAICS)	0.087 ^a (0.010)	0.093 ^a (0.010)	-0.005 (0.011)	-0.013 (0.011)	0.130 ^a (0.006)	0.132 ^a (0.006)	0.041 ^a (0.006)	0.041 ^a (0.006)
log(Assets)	0.264 ^a (0.016)	0.278 ^a (0.016)	0.198 ^a (0.013)	0.191 ^a (0.013)	0.048 ^a (0.005)	0.052 ^a (0.005)	0.122 ^a (0.011)	0.123 ^a (0.011)
log(Age)	0.085 ^a (0.015)	0.086 ^a (0.015)	-0.022 ^c (0.011)	-0.034 ^a (0.011)	0.024 ^a (0.005)	0.025 ^a (0.005)	0.014 (0.010)	0.014 (0.010)
MB	-0.110 ^a (0.020)	-0.135 ^a (0.021)	-0.032 ^b (0.013)	-0.003 (0.012)	-0.016 ^a (0.003)	-0.020 ^a (0.003)	0.004 ^c (0.003)	0.005 ^c (0.003)
Industry Fixed Effects	No	No	No	No	Yes	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	Yes	No	Yes
#Obs.	45,198	45,198	45,198	45,198	45,198	45,198	45,198	45,198
Pseudo R^2 / Adj. R^2	0.107	0.112	0.043	0.035	0.525	0.525	0.855	0.855

Note: Columns (1) to (4) report probit estimations where the dependent variable 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. Columns (5) to (8) report OLS estimation where the dependent variable is vertical integration VI . All estimations include year fixed effects. All variables are defined in Appendix 2 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.6: Vertical Acquisitions and Integration: NAICS-based Measure

Dep. Variables:	Prob(Target) (probit)				VI (OLS)			
	Vert.	Non-Vert.	Vert.	Non-Vert.	Baseline		Interaction	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ind.(R&D/sales)	0.012 (0.026)	0.110 ^a (0.014)	0.083 ^b (0.036)	0.202 ^a (0.019)	-0.141 ^a (0.031)	0.025 (0.017)	0.016 (0.032)	0.016 (0.018)
Ind.(#Patents/assets)	0.040 ^b (0.021)	-0.004 (0.013)	0.096 ^a (0.027)	0.074 ^a (0.016)	0.297 ^a (0.029)	0.006 (0.015)	0.438 ^a (0.038)	-0.002 (0.018)
Ind.(R&D/sales × #Patents/assets)			-0.051 ^b (0.020)	-0.065 ^a (0.010)			-0.124 ^a (0.017)	0.008 (0.008)
Ind.(PPE/assets)	-0.029 (0.027)	-0.052 ^a (0.014)	-0.014 (0.028)	-0.028 ^b (0.014)	-0.287 ^a (0.048)	-0.050 ^b (0.024)	-0.272 ^a (0.047)	-0.050 ^b (0.024)
HHI	-0.059 ^b (0.023)	-0.053 ^a (0.012)	-0.056 ^b (0.023)	-0.046 ^a (0.011)	-0.090 ^a (0.022)	-0.014 (0.015)	-0.079 ^a (0.022)	-0.015 (0.015)
End User	-0.014 (0.019)	-0.001 (0.010)	-0.009 (0.019)	0.005 (0.010)	-0.085 ^a (0.023)	-0.026 (0.018)	-0.083 ^a (0.023)	-0.027 (0.018)
#Segment (NAICS)	0.036 ^b (0.014)	0.057 ^a (0.009)	0.035 ^b (0.014)	0.057 ^a (0.009)	1.018 ^a (0.055)	0.948 ^a (0.050)	1.018 ^a (0.055)	0.948 ^a (0.050)
log(Assets)	0.141 ^a (0.023)	0.256 ^a (0.011)	0.142 ^a (0.023)	0.259 ^a (0.011)	0.104 ^a (0.024)	-0.086 ^b (0.037)	0.105 ^a (0.024)	-0.085 ^b (0.037)
log(Age)	0.078 ^a (0.021)	0.019 ^c (0.010)	0.080 ^a (0.022)	0.021 ^b (0.010)	0.090 ^a (0.019)	-0.041 ^c (0.024)	0.087 ^a (0.019)	-0.040 ^c (0.024)
MB	-0.034 (0.025)	-0.054 ^a (0.012)	-0.035 (0.025)	-0.056 ^a (0.012)	0.012 (0.016)	0.015 (0.010)	0.011 (0.016)	0.015 (0.010)
Industry Fixed Effects	No	No	No	No	Yes	No	Yes	No
Firm Fixed Effects	No	No	No	No	No	Yes	No	Yes
#Obs.	45,198	45,198	45,198	45,198	45,198	45,198	45,198	45,198
Pseudo R^2 / Adj. R^2	0.032	0.053	0.034	0.056	0.258	0.828	0.258	0.828

Note: Columns (1) to (4) report probit estimations where the dependent variable 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. Columns (5) to (8) report OLS estimation where the dependent variable is vertical integration VI . All estimations include year fixed effects. All variables are defined in Appendix 2 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.7: The Determinants of Vertical Acquirers

Dep. Variable: Specification: Deal type:	Prob(Acquirer)						
	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)	-0.154 ^a (0.029)	0.129 ^a (0.018)	-0.143 ^a (0.037)	0.220 ^a (0.024)		-0.151 ^a (0.030)	0.188 ^a (0.026)
Ind.(#Patents/asse)	0.194 ^a (0.014)	-0.059 ^a (0.020)	0.200 ^a (0.020)	0.032 (0.023)	-0.049 ^a (0.009)	0.192 ^a (0.015)	-0.096 ^a (0.024)
Ind.(R&D/sales) × Ind.(#Patents/assets)			-0.006 (0.017)	-0.067 ^a (0.011)			
Ind.(PPE/assets)	0.021 (0.016)	-0.064 ^a (0.024)	0.022 (0.017)	-0.037 (0.024)	0.039 ^a (0.007)	0.031 ^c (0.0167)	-0.048 ^c (0.026)
HHI	-0.031 ^b (0.018)	-0.038 ^a (0.014)	-0.031 ^b (0.018)	-0.029 ^b (0.014)	0.017 ^a (0.004)	-0.025 (0.018)	-0.033 ^b (0.015)
End User	-0.185 ^a (0.014)	0.076 ^a (0.012)	-0.185 ^a (0.014)	0.085 ^a (0.012)	0.008 ^a (0.002)	-0.190 ^a (0.015)	0.078 ^a (0.013)
#Segment (NAICS)	0.104 ^a (0.011)	-0.022 ^b (0.011)	0.104 ^a (0.010)	-0.022 ^b (0.011)	0.011 ^a (0.001)	0.100 ^a (0.011)	-0.021 ^c (0.011)
log(Assets)	0.342 ^a (0.015)	0.387 ^a (0.013)	0.342 ^a (0.014)	0.389 ^a (0.013)	0.027 ^a (0.003)	0.351 ^a (0.016)	0.386 ^a (0.013)
log(Age)	-0.002 (0.015)	-0.040 ^a (0.012)	-0.002 (0.015)	-0.037 ^a (0.012)	-0.020 ^a (0.003)	-0.002 (0.015)	-0.040 ^a (0.013)
MB	-0.030 ^c (0.020)	0.069 ⁴ (0.011)	-0.030 ^c (0.017)	0.068 ^a (0.010)	-0.005 (0.004)	-0.025 (0.018)	0.059 ^a (0.011)
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	0.131	0.094	0.131	0.096	N/A	N/A	N/A

Note: The dependent variable in the probit models is a dummy indicating whether the given firm is an acquirer 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 ?. Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Table IA.IV.8: Predicting R&D using R&D tax credit

Dep. Variable:	R&D/sales		
	(1)	(2)	(3)
User cost of R&D capital	-0.706 ^a (0.016)	-0.115 ^a (0.020)	-0.060 ^a (0.021)
Firm FE	No	Yes	Yes
Year FE	No	No	Yes
#Obs.	39,751	39,751	39,751
Adj. R ²	0.057	0.862	0.863

Note: This table reports OLS estimations where the dependent variable is firm-year $R\&D/sales$, and the independent variable is the user cost of R&D capital. Standard errors are clustered by 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 IA.IV.9: The Determinants of Vertical Target Acquisitions: Linear Probability Models

Dep. Variable: Specification: Deal type:	Prob(Target)						
	LPM				IV LPM		
	Vertical (1)	Non-Vertical (2)	Vertical (3)	Non-Vertical (4)	1st Stage (5)	Vertical (6)	Non-Vertical (7)
Ind.(R&D/sales)	-0.004 ^a (0.001)	0.017 ^a (0.001)	-0.001 (0.001)	0.026 ^a (0.002)		-0.002 ^a (0.001)	0.021 ^a (0.002)
Ind.(#Patents/asse)	0.006 ^a (0.001)	-0.006 ^a (0.001)	0.008 ^a (0.001)	0.000 (0.001)	-0.049 ^a (0.009)	0.005 ^a (0.001)	-0.009 ^a (0.001)
Ind.(R&D/sales) × Ind.(#Patents/assets)			-0.001 ^a (0.001)	-0.006 ^a (0.001)			
Ind.(PPE/assets)	-0.000 (0.000)	-0.005 ^a (0.001)	0.000 (0.000)	-0.003 ^b (0.001)	0.039 ^a (0.007)	0.000 (0.000)	-0.004 ^a (0.001)
HHI	-0.001 ^b (0.001)	-0.003 ^a (0.001)	-0.001 ^c (0.001)	-0.003 ^a (0.001)	0.017 ^a (0.004)	-0.001 (0.000)	-0.003 ^a (0.001)
End User	-0.007 ^a (0.001)	0.007 ^a (0.001)	-0.007 ^a (0.001)	0.007 ^a (0.001)	0.008 ^a (0.002)	-0.007 ^a (0.001)	0.007 ^a (0.001)
#Segment (NAICS)	0.010 ^a (0.001)	-0.000 (0.001)	0.010 ^a (0.001)	-0.000 (0.001)	0.011 ^a (0.001)	0.010 ^a (0.001)	-0.000 (0.000)
log(Assets)	0.016 ^a (0.001)	0.016 ^a (0.001)	0.016 ^a (0.001)	0.017 ^a (0.001)	0.027 ^a (0.003)	0.016 ^a (0.001)	0.018 ^a (0.001)
log(Age)	0.005 ^a (0.001)	-0.000 (0.000)	0.005 ^a (0.001)	-0.000 (0.000)	-0.020 ^a (0.003)	0.005 ^a (0.001)	0.000 (0.000)
MB	-0.002 ^a (0.001)	-0.002 ^b (0.001)	-0.002 ^a (0.001)	-0.003 ^a (0.001)	-0.005 (0.004)	-0.002 ^a (0.001)	-0.003 ^a (0.001)
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	0.029	0.015	0.029	0.016	0.939	0.29	0.15

Note: The dependent variable in the LPM 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 instrumental variable estimations where we use tax-induced industry predicted R&D/sales (using exogenous variation in the user cost of R&D capital) as instrument for industry R&D intensity (Ind.(R&D/sales)). All estimations include year fixed effects. All independent variables are defined in Appendix 2. The independent variables are standardized for convenience. Standard errors are clustered by 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 IA.IV.10: Change in R&D intensity around vertical acquisitions

Variable:	R&D/sales			
	Pre (1)	Post (2)	Post-Pre (3)	P-value (4)
Target firms (existing post)	0.0467 (1,473)	0.0365 (493)	-0.0101 ^a	0.001
Combined entity	0.0380 (4,140)	0.0343 (3,324)	-0.0036 ^a	0.005

Note: This table displays the average differences in *R&D/sales* before (pre) and after (post) a vertical acquisition. The first row comprises all vertical targets that continue to exist for at least one year after being acquired. The second row comprises “combined” entities that aggregate R&D and sales of acquirers and targets. Vertical transactions are identified using the Vertical Text-10% network. Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.

Table IA.IV.11: Examples of Vertically Integrated firms: Top 30 in 2008

Company	Rank	#Segments	VI	Perc.(VI)	Perc.(VI(<i>Segment</i>))
HANDY & HARMAN LTD	1	5	0.091	1	0.969
PARKER-HANNIFIN CORP	2	2	0.079	0.999	0.000
EATON CORP	3	5	0.076	0.999	0.966
EMERSON ELECTRIC CO	4	6	0.074	0.999	0.991
FRANKLIN ELECTRIC CO INC	5	1	0.073	0.998	0.717
COMMERCIAL VEHICLE GROUP INC	6	1	0.069	0.998	0.000
ROCKWOOD HOLDINGS INC	7	5	0.069	0.997	0.959
SCHNITZER STEEL INDS -CL A	8	3	0.064	0.997	0.000
LEGGETT & PLATT INC	9	3	0.062	0.997	0.710
DOVER CORP	10	4	0.058	0.996	0.641
SIFCO INDUSTRIES	11	2	0.055	0.996	0.994
MYERS INDUSTRIES INC	12	1	0.053	0.996	0.000
AMPCO-PITTSBURGH CORP	13	2	0.053	0.995	0.681
SONOCO PRODUCTS CO	14	3	0.052	0.995	0.000
LKQ CORP	15	1	0.052	0.995	0.000
P & F INDUSTRIES -CL A	16	2	0.052	0.994	0.760
BERKSHIRE HATHAWAY	17	9	0.051	0.994	0.000
PRECISION CASTPARTS CORP	18	2	0.051	0.993	0.790
MATTHEWS INTL CORP -CL A	19	6	0.051	0.993	0.884
RELIANCE STEEL & ALUMINUM CO	20	1	0.050	0.993	0.000
CARLISLE COS INC	21	6	0.050	0.992	0.962
UNVL STAINLESS & ALLOY PRODS	22	1	0.050	0.992	0.000
AMERICAN AXLE & MFG HOLDINGS	23	1	0.049	0.992	0.000
ENCORE WIRE CORP	24	1	0.049	0.991	0.000
HAWK CORP	25	1	0.049	0.991	0.000
KANSAS CITY SOUTHERN	26	1	0.049	0.991	0.000
AMERICAN ELECTRIC TECH INC	27	3	0.049	0.990	0.885
DREW INDUSTRIES INC	28	1	0.049	0.990	0.000
CHINA PRECISION STEEL INC	29	1	0.048	0.989	0.000
COLEMAN CABLE INC	30	1	0.048	0.989	0.000

Note: The table displays the 30 most vertically integrated firms in 2008 based on our text-based measure of vertical integration (*VI*). The table also presents the number of Compustat segments, the *VI* score, the firm's percentile *VI* ranking, and the firm's percentile *VI*(*Segment*) ranking.

Table IA.IV.12: The Determinants of Vertical Integration: Log(VI)

Dep. Variable:	(Text-based) log(VI)			
	Baseline		Interaction	
	(1)	(2)	(3)	(4)
Ind.(R&D/sales)	-0.042 ^a (0.003)	-0.003 (0.002)	-0.036 ^a (0.004)	-0.001 (0.003)
Ind.(#Patents/assets)	0.031 ^a (0.003)	0.008 ^a (0.002)	0.036 ^a (0.003)	0.010 ^a (0.003)
Ind.(R&D/sales × #Patents/assets)			-0.004 ^b (0.002)	-0.002 (0.001)
Ind.(PPE/assets)	0.014 ^a (0.005)	0.007 ^c (0.004)	0.015 ^a (0.005)	0.007 ^c (0.004)
HHI	-0.032 ^a (0.002)	-0.018 ^a (0.002)	-0.032 ^a (0.002)	-0.018 ^a (0.002)
End User	-0.110 ^a (0.003)	-0.070 ^a (0.003)	-0.110 ^a (0.003)	-0.070 ^a (0.003)
#Segment (NAICS)	0.048 ^a (0.002)	0.018 ^a (0.002)	0.048 ^a (0.002)	0.018 ^a (0.002)
log(Assets)	0.020 ^a (0.002)	0.050 ^a (0.004)	0.020 ^a (0.002)	0.050 ^a (0.004)
log(Age)	0.006 ^a (0.002)	0.003 (0.004)	0.005 ^a (0.002)	0.003 (0.004)
MB	-0.007 ^a (0.001)	0.003 ^a (0.001)	-0.007 ^a (0.001)	0.003 ^a (0.001)
Industry Fixed Effects	Yes	No	No	No
Firm Fixed Effects	No	Yes	Yes	Yes
#obs.	45,198	45,198	45,198	45,198
Adj. R ²	0.579	0.867	0.579	0.867

Note: The dependent variable is vertical integration measured by the logarithm of VI . All estimations include year fixed effects. All variables are defined in Appendix 4 of the paper. The independent variables are standardized for convenience. Standard errors are clustered by industry and year and are reported in parentheses. Symbols ^a, ^b, and ^c indicate statistical significance at the 1%, 5%, and 10% confidence levels.