NBER WORKING PAPER SERIES

# SCOPE, SCALE AND CONCENTRATION: THE 21ST CENTURY FIRM

Gerard Hoberg
Gordon M. Phillips

Working Paper 30672
http://www.nber.org/papers/w30672

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2022

We thank Chad Syverson (NBER discussant), Jon Garfinkel, Kai Li (WFA discussant), Michael Roberts, Rene Stulz, and seminar participants at Aarhus University, Australia National University, Georgia State University, Goethe University Frankfurt, Iowa State University, Ohio State University, Swiss Economic Institute, Tulane University, University of Amsterdam, University of Iowa and the University of Pennsylvania (Wharton), and also conference participants at the 2022 NBER, 2021 Western Finance Association conferences. All errors are the authors' alone. We have no outside sources of research support to disclose. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Scope, Scale and Concentration: The 21st Century Firm
Gerard Hoberg and Gordon M. Phillips
NBER Working Paper No. 30672
November 2022
JEL No. D20,D43,G30,O31,O34

**ABSTRACT**

We provide evidence that over the past 30 years, U.S. firms have expanded their scope of operations. Increases in scope and scale were achieved largely without increasing traditional operating segments. Scope expansion significantly increases valuation and is primarily realized through acquisitions and investment in R&D, but not through capital expenditures. We show that traditional concentration ratios do not capture this expansion of scope. Our findings point to a new type of firm that increases scope through related expansion, which is highly valued by the market.

Gerard Hoberg
University of Southern California
Marshall School of Business
BRI-308 (FBE Faculty)
Los Angeles, CA 90089
United States
hoberg@marshall.usc.edu

Gordon M. Phillips
Tuck School of Business
Dartmouth College
100 Tuck Hall
Hanover, NH 03755
and NBER
gordon.m.phillips@gmail.com

"Product diversification came from opportunities to use existing production, marketing, and research facilities ... Such expansion was based on organizational capabilities that had been developed by exploiting economies of scope." *p. 38 Chandler and Hikino (1994).*

# 1 Introduction

The interplay between scope, scale and competition has been the focus of numerous authors including both business historians and economists.[1] A principle focus of authors has been defining firms by the basket of products firms produce and the industries to which these products belong. Recent authors have also documented the rise in firm size, a rise in traditional industry concentration measures, and a drop in the number of U.S. listed firms.[2]

In this paper, we provide a new perspective on the 21st Century version of a multi-product firm and how it produces in different-but-related markets. It differs markedly from the concept of a diversified conglomerate with a multi-division organizational structure producing products across unrelated industries that was the focus of the early corporate finance literature. Instead of multiple distinct segments, firms might have flexible production and redeployable assets that allow them to pursue multi-sector production without the potential negative consequences of a complex multi-division organization.

We develop new firm-year measures of product market scope using the text describing the product markets in which firms operate. We document that the average firm's scope in related industries has increased steadily and dramatically (71%) during our sample period from 1989 to 2017. Moreover, firms have increased scope without increasing the number of operating segments they report during the period of our sample. Our results are consis-

[1]See Chandler and Hikino (1994), Hart and Moore (1990), Panzar and Willig (1977) and Williamson (1975).

[2]See Autor, Dorn, Katz, Patterson, and Van Reenen (2020), Doidge, Karolyi, and Stulz (2017), Grullon, Larkin, and Michaely (2019), Kwon, Ma, and Zimmermann (2021) and Philippon and Gutierrez (2017) for aggregate trends. Matvos, Seru, and Silva (2018) examine the link between scope expansion and episodes of financial market frictions.

tent with multi-product firms having synergies across related products and with unrelated diversification not being a major consideration for the increases in scope we document.

An important question is how do corporate investment policies react to incentives to increase scope? To shed light on this question, we examine how plausibly exogenous variation in the ex ante incentives to increase scope impact ex post investment decisions and measures of firm performance. We consider two sources of such variation. The first measures each firm's scope-expansion opportunity set using the diversity of related markets served by each firm's distant peers. When these peers have operations that are distributed across multiple well-defined markets, it suggests that the focal firm itself likely sees a larger opportunity set of potential scope-enhancing projects all else equal. Our second measure considers the cost-side and asset redeployability. We examine the asset portfolios of the firm's closer industry peers as compared to the firm's more distant industry peers (where distance is in product market space). If the closer industry peers' assets are easily redeployed to more distant peers' industries, it follows that the focal firm likely faces a low relative cost to expanding scope as its assets can be redeployed to more potential markets with lower adjustment costs.

Both variables measure a focal firm's incentives to increase scope, and neither is measured using any data about the focal firm itself. Instead, both are based on the characteristics of more distant industry peers, whose characteristics are not easily influenced by the focal firm and vice-a-versa. Our approach thus follows prior studies in the network econometrics literature that highlight the fact that endogenous effects are mitigated by focusing on distant peers and not the focal firm itself.[3] Although the literature indicates that this approach reduces the scope for endogeneity, we nevertheless interpret these results conservatively and view them as strong tests of our predicted mechanisms.

We find that firm scope is significantly related to more acquisitions, fewer divestitures, more spending on research and development and increased vertical integration. In contrast,

[3] See Bramoulle, Djebbari, and Fortin (2009) for theory and Cohen-Cole, Kirilenko, and Patacchini (2014) for a recent application in finance. These studies indicate that distant peers can produce exogenous variation that can be used as instruments.

we find no link to capital expenditures. These results are consistent with an ongoing process of asset redeployment across and within firms, which is reinforced by innovation that facilitates flexible and efficient redeployment of assets for multi-industry production. Importantly, this suggests that acquisitions and innovation facilitate scope increases, but our results are novel and are missed by existing studies because these scope increases are accomplished without increasing the number of formal operating segments in the Compustat database (the primary database used to research scope in the existing literature).

We examine firm outcomes and find that that increased scope is associated with higher valuations, and thus it is unlikely that increases in scope are due to bad governance or private benefit extraction. This evidence of higher value is important given the prior literature and possible empire building incentives managers would have to increase scope. We also find evidence of higher ex post sales growth and asset growth. However, we do not find any significant impact for profitability in the form of return on assets. These results suggest that scope expansion creates positive net present value and sales growth opportunities, and profit maximizing firms likely pick the most profitable industries to operate in first, and then expand into still-profitable but lower return on assets industries. Our valuation results - which show that firm market-to-book equity ratios increase with scope - are in contrast to the historical conglomerate literature (Berger and Ofek (1995) and Lang and Stulz (1994)) that finds that firm valuations decrease as reported Compustat segments increase.

We also examine how firm scope expansion is likely financed. Firms with higher ex ante incentives to increase scope issue more shares and pay lower dividends ex post. We find no significant link to debt financing. These results favoring equity are consistent with intangibles and the redeployment of existing assets playing an important role in scope expansion. In particular, this method of expansion through intangibles and better utilizing existing assets does not create much new collateral, which favors financing with equity.

We note that these conclusions are drawn without using the traditional measure of scope: the number of Compustat segments. In particular, the average Compustat segment count of firms has not increased over time. A key issue with reported segment data is that it does not facilitate measurement of the "new conglomerate" we hypothesize, which emphasizes expansion in related markets. Our results favor the conclusion that this "new conglomerate" can span multiple product spaces while preserving a single segment organizational form. Additionally, the accounting standard, SFAS 131, which governs segment reporting, does not require that reported segments need to be based on the number of industries served. We thus further examine the comparative validity of our textual scope measures and the Compustat segments by relating both to the frequency of explicit statements in firm 10-Ks where firms indicate they serve multiple product markets. Our new measures are highly significant in predicting these statements, whereas the number of Compustat segments are only borderline significant and they become mostly insignificant when included in the same regressions as our new text-based scope measures.

We conclude with an exploratory analysis that we provide in the online appendix on whether increases in scope can provide new insights on why traditional measures of concentration are increasing over time. We use two methods to adjust traditional HHI industry concentration ratios to account for increases in scope. Using both methods, we find that that horizontal concentration measured using scope-adjusted HHIs is not increasing.

Our results suggest that increases in purely horizontal concentration are smaller when scope is taken into account, and thus changes in purely horizontal configurations may not explain why market power might be higher. Increases in scope, however, might motivate different concerns. Although scope increases reduce horizontal concentration as the same number of firms serve more markets, scope can increase anti-competitive conduct through product bundling, by increasing barriers to new entry, or it can induce kill zones in the market for innovation (see Kamepalli, Rajan, and Zingales (2020)).

Understanding the separate influences of horizontal competition and higher scope at the micro level, or within industries, should be a fruitful area for future research. In particular, regulatory interventions may differ for scope-induced versus horizontal market power.

## 2 Literature

Theories of economies of scale and scope were first developed by Panzar and Willig (1977) and Panzar and Willig (1981). Teece (1980) further develops relevant theory and suggests that a multi-product enterprise is particularly likely to emerge when economies of scope are based on a recurrent use of proprietary know-how. This theory therefore illustrates why our finding that R&D spending is increased when scope expansion incentives are high is consistent with theories that examine economies of scope. Henderson and Cockburn (1996) provide empirical support for a link between economies of scope and innovation investment in the pharmaceutical industry. Braguinsky, Ohyama, Okazaki, and Syverson (2020) provide further empirical support for these ideas in a novel historical setting: Japan's cotton spinning industry from 1893 to 1914. In particular, the authors show that technological capability was a major ingredient that fostered horizontal scope expansion and firm success in this early environment. Kwon, Ma, and Zimmermann (2021) focus on how scale economies over a long period of time have been achieved through the use of R&D and information technology. We also show the importance of R&D to increases in scope and show how scope needs to be measured before attributing size increases to scale.

More recent theory by Maksimovic and Phillips (2002) postulates an efficiency based view of multi-industry operations based on neoclassical profit optimization. In the model, a conglomerate discount can still emerge even when governance is aligned with shareholders and optimal policies are undertaken. However, if low cost scope expansion is possible through economies of scope with sharing a scarce resource such as innovation or managerial talent, this discount may result in a premium . Our results favor this perspective on a few dimensions. In particular, we find that scope expansion with higher R&D brings higher valuations and sales growth, consistent with rational expansion.

Our paper also has implications for the older literature documenting a diversification discount (Berger and Ofek (1995) and Lang and Stulz (1994)). Although recent studies call the discount into question (see Custodio (2012) and Hund, Monk, and Tice (2020) for example), our paper brings an entirely new perspective: modern multiple-industry firms share scarce resources and serve multiple industries while maintaining an efficient single segment organizational form with a high market valuation. Our study also contributes to the literature on acquisition motives and synergies (see Hoberg and Phillips (2010a), Rhodes-Kropf and Robinson (2008), Bena and Li (2014), and Fresard, Hoberg, and Phillips (2020)).

# 3 Data and Methods

## 3.1 Sample Selection and Panel Structure

Our sample begins with the universe of Compustat firm-years with available 10-K filings on the EDGAR system (later years) or scanned 10-Ks from the Dartmouth and Harvard libraries (earlier years of our sample). As the standard TNIC database of Hoberg and Phillips 2016 (HP2016) is based purely on EDGAR filings, its coverage begins in 1996. A contribution of the current study is thus that we back-extend the TNIC database to 1988 using 10-Ks from the aforementioned libraries. To remain in our sample, a firm must have a 10-K filing both in the current year of observation, and in the previous year. We exclude firms operating in financial industries and regulated utilities (SIC 6000 - 6999 and 4900 - 4949, respectively) and limit the sample to firm-years with sales and assets of at least $1 million. We are left with 100,525 firm-year observations from 1989 to 2017.

## 3.2 Novel Measures of Scope

We develop new measures of firm-year product market scope using the 10-K text-based framework of HP2016. These descriptions are updated as products evolve every year, and are required by Regulation S-K to accurately represent the products sold in each fiscal year.

Measuring scope requires scoring firm $i$ in year $t$ regarding how many industry vocabularies (industries are indexed $j \in J$) are discussed in its product description. Following HP2016, we define the product market space as the set of all product market words used in the TNIC database.[4] If there are $N$ words in the product space, then each firm is represented by a vector $V_{i,t}$, which is an $N$ dimensional vector containing a one for words that firm $i$ uses in year $t$ and a zero for words that it does not use. Following HP2016, we omit stop words, which are those appearing in more than 25% of all Item 1's in year $t$.

Analogously, we generate a vector representation for each industry $j \in J$. Unlike firm vectors, we lock in each industry's vocabulary vector and do not allow industry vectors to vary with time. As we discuss later, this ensures that we can measure changes in scope in its purest form, avoiding the potentially confounding matter of product market innovation and the creation of entirely new markets.[5] Studying the creation of new markets is separately interesting, but it is not the focus of the current study. Our focus on a fixed set industry vocabularies is also conservative, as it favors under-measuring the true increase in firm scope that has occurred due to the separate impact of new industries.

We consider two methods for identifying the spatial location of industries. The first draws industry vocabularies from the "Fixed Industry Classification" technology developed in HP2016, which uses a clustering algorithm based on single segment firms to identify product market clusters. This approach develops vocabularies for each industry, which are computed as the average of the firm vectors $V_{i,t}$ after normalization based on the firms assigned to each cluster. We specifically use the FIC-300 (which is based on 300 industries) from 1997, which is the base year used by HP2016. We denote the resulting industry vocabulary vectors by $D_{FIC,j}$, where $j \in 1, ..., 300$. The vocabularies are obtained directly from HP2016 without modification. Each vector is then normalized to sum to unity and resides in the same $N$-dimensional space as do the firm-vectors $V_{i,t}$.

Our second method is based on NAICS industry definitions, as defined in the 2017 NAICS Manual, which is a 963 page document providing detailed verbal descriptions of each NAICS industry. We use the four-digit NAICS granularity, and group vocabulary for all industries

[4] This is the set of all words that appear in the union of all firm 10-K Item 1's.

[5] Although our baseline approach maximizes clarity, we note that our results are very similar if we instead re-define industries every year (we run this alternative specification using dynamically recomputed FIC-300 industries).

having the same four digits of their NAICS code into each NAICS' code's total vocabulary. This determines the vector for each 4-digit NAICS code in our $N$-dimensional space. In addition to applying the TNIC stop-word filter above, we also reviewed the common words in the NAICS manual and identified a list of additional stop words (see Appendix). The result is a vector $D_{NAICS,j}$ for each four digit NAICS industry, where $j \in 1, ..., 311$, as there are 311 four digit NAICS that we capture using this approach. Each element of this vector is populated with the number of times each word is used in the NAICS manual for the given industry, and each vector is normalized to sum to unity. These vectors reside in the same $N$-dimensional space as the firm-vectors $V_{i,t}$.

Next, we score each firm-year based on how much of each industry's vocabulary it uses. We avoid pairwise cosine calculations as that would overly penalize firms that have large product descriptions that cover many industries (because the fraction of each industry in the overall vector would be small). Instead, we compute the fraction of each industry's vocabulary that appears in each firm-year's product description as the following overlap ratio (we develop an analogous ratio for the NAICS-based vocabularies):

$$Q_{i,j,t,FIC} = \frac{\#\text{words overlapping in } D_{FIC,j} \text{ and } \ V_{i,t}}{\#\text{words in } D_{FIC,j}} \tag{1}$$

Finally, to compute how many FIC industries a given firm might operate in, we identify a fixed threshold $\bar{Q}_{FIC}$ above which we deem a firm having $Q_{i,j,t,FIC} > \bar{Q}_{FIC}$ to be operating in industry $j$ in the given year $t$ (we call this an "operating pair"). We hold $\bar{Q}_{FIC}$ fixed and do not allow it to vary with time, as otherwise we might create false inferences in our time series analysis. Following HP2016, we use our base year of 1997 to compute $\bar{Q}_{FIC}$ as the threshold such that 2% of all firm-industry combinations are operating-pairs in 1997. The 2% threshold is also used in HP2016 as it matches the granularity of three-digit SIC industries. Our results are robust to using 1% or 5%. Our main variable FIC-scope is then the number of industries the given firm likely operates in, i.e., the number of industries with a similarity above $\bar{Q_{FIC}}$ (we also winsorize scope variables at the 1/99% level):

$$FIC - Scope_{i,t} = \sum_{j=1,...300} Indicator\{Q_{i,j,t,FIC} > \bar{Q_{FIC}}\} \tag{2}$$

$$NAICS - Scope_{i,t} = \sum_{j=1,...300} Indicator\{Q_{i,j,t,NAICS} > \bar{Q_{NAICS}}\} \tag{3}$$

Finally, we consider a third measure of scope based on the Latent Dirichlet Allocation (LDA) topic model developed by Blei, Ng, and Jordan (2003). We summarize this measure in the Online Appendix, and we report a series of tables illustrating that all of our main results are fully robust.

## 3.3 Local Asset Redeployability and Outward Scope Expansion

We develop two ex-ante measurable shifters of firm scope, which allow us to examine the extent to which plausibly exogenous incentives for firms to increase scope predict ex-post corporate strategies for increasing scope and subsequent performance. This approach provides the foundation for strong tests of our predicted mechanisms. Our first instrument is based on asset redeployability, specifically each firm's ability to redeploy its assets in product markets that are nearby in the product space. Intuitively, a firm that can easily redeploy assets into spatially proximate product markets likely has strong incentives to increase scope because the cost of doing so is likely to be low. We focus on spatially local redeployability rather than broader measures of redeployability to increase the power of our instrument. This approach is motivated by the language theory of Crémer, Garicano, and Prat (2007) and tests in Hoberg and Phillips (2018) showing that firms expand scope by operating in groups of highly related industries and not distant industries.

Our approach to measuring local redeployability follows Kim and Kung (2017) (KK2017), who use the capital flows tables from the Bureau of Economic Analysis to compute measures of broad asset redeployability. This approach is also motivated by Boehm, Dhingra, and Morrow (2022) who show that input complementarities are important for product entry decisions in India. We refine the KK2017 methodology to focus on localized redeployability. In particular, the BEA tables indicate the extent to which each of the 123 BEA industries (which can be mapped to NAICS codes) utilizes a set of 180 assets. Intuitively, if two BEA industries utilize the 180 assets in similar proportions, we conclude that a firm operating in one faces a high degree of asset redeployability and a low cost of entry into the other. This is an incentive to expand scope in that direction. A lack of asset redeployability, conversely, indicates a barrier to scope expansion. To the extent that the asset allocation vectors across different industries are exogenous from a firm's perspective, it would follow that a firm facing high levels of asset redeployability for nearby industries faces an exogenously higher incentives to increase scope in the future.

Of course, as innovation can change the distribution of asset allocations within an industry, it follows that asset-allocation vectors are not fully exogenous. We thus take additional precautions to further improve identification. First, following KK2017, we use a single BEA table from 1997 for our entire sample and fix the asset allocation vectors in time. Second, we compute local redeployability for each firm using an approach that examines the product market around the firm, thus strictly avoiding using data from the focal firm itself. We do so by computing the redeployability of the firm's close peers (based on the TNIC3 classification) to expand into the industries covered by the focal firm's more distant peers (those in the focal firm's TNIC2 classification but not those in its TNIC3 classification).[6]

We compute local asset redeployability by first mapping each BEA industry to a four digit NAICS code (following KK2017), and representing the underlying assets used by each NAICS industry as a 180 element vector, which we denote as $A_j$ for a given NAICS-4 industry j. Each vector is obtained directly from the 1997 capital flows table, which has reported dollar amounts for 180 assets tracked by BEA for each industry. Next, for each focal firm in each year, we obtain two sets of peers. Close peers are those in the focal firm's

[6]TNIC2 industries are the text-based industry classification from HP2016 that is calibrated to be as granular as two-digit SIC industries. TNIC3 is finer and is as granular as SIC-3. Thus firms that are in TNIC2 but not TNIC3 are "distant peers" as they are still in nearby markets but they are not in the focal firm's current market.

TNIC-3 industry (excluding the focal firm itself). Distant peers are those in the focal firm's TNIC-2 industry but not in its TNIC-3 industry. There are no overlapping peers in these two sets. Next we compute the fraction of each set of peers in each NAICS-4 industry.[7] $F_{i,t,j,near}$ is the fraction of focal firm i's close peers that are in 4-digit NAICS industry $j$ in year $t$. $F_{i,t,j,distant}$ is analogously defined for distant peers. These two industry distributions reflect likely paths that firms would take if they outwardly expand scope into the nearest distinct markets. Our key variable, Outward-focused local asset rededployability, is then the weighted average asset-complementarity (cosine similarity of the asset vectors for the two industries in a pair j,k) summed over the distribution of industry pairs spanned by the two sets of peers:

$$LocalAssetRedep._{i,t} = \sum_{j,ks.t.j \neq k \in NAICS-4} F_{i,t,j,near} F_{i,t,k,distant} < \frac{A_j}{A_j \cdot 1} \cdot \frac{A_k}{A_k \cdot 1} > \tag{4}$$

We note that the above calculation does not depend on the focal firm itself and instead focuses on the industries served by peers that are more distant. This helps to reduce potential channels for violation of the exclusion requirement, as is reinforced by econometric theories of network identification (Bramoulle, Djebbari, and Fortin, 2009). Cohen-Cole, Kirilenko, and Patacchini (2014) is a related application in finance. Importantly, when local asset redeployability is high, it indicates that the focal firm likely has many ways to increase scope with relatively low adjustment costs. These strong ex-ante incentives are thus a shifter of the focal firm's ex-post scope-expansion strategy.

## 3.4 The Local Scope-Expansion Opportunity Set

Our second shifter of scope incentives is based on the size of the outward-scope expansion opportunity set as seen from the focal firm's perspective. As was the case for our above

[7] We use NAICS and not TNIC industries for this part of the calculation because BEA tables are linked to NAICS.

cost-of-entry shifter, we construct our second instrument by focusing on more distant peers and avoid using the characteristics of the focal firm itself. We thus identify distant peers as rivals that are in the focal firm's TNIC-2 industry but are not in the focal firm's TNIC-3 industry, and we compute the distribution of NAICS-4 industries served by these distant peers as ($F_{i,t,j,distant}$). To compute the local scope expansion opportunity set (our second shifter), we compute one minus the concentration ratio (HHI) based on this distribution:

$$LocalScopeExpansionOpp.Set_{i,t} = 1 - \sum_{j \in NAICS-4} F^2_{i,t,j,distant} \tag{5}$$

When this variable is high, it indicates that nearby peers serve a wide array of closely related product markets. From the focal firm's perspective, this indicates a high quality opportunity set for scope-expansion. Because this measure is only a function of the firm's distant peers, as noted above, it is plausibly exogenous from the perspective of the focal firm's policies. Focal firms facing a higher value of this variable are likely to increase scope given the wider array of growth opportunities available.

## 3.5 R&D, Investment and Acquisitions

We examine four investment policies: R&D/assets, CAPX/assets, acquisitions, and divestitures. The R&D (XRD) and CAPX variables are from Compustat. We scale each by beginning of period total assets (AT). When R&D is missing, we assume it to be zero.[8] We obtain acquirer and target data using both full-firm and partial-firm asset acquisition data from SDC Platinum. SDC Acquirer is an indicator equal to one if the given firm acquires assets from any seller (public or private) in the given year and is zero otherwise. SDC Target is an indicator equal to one if the given firm sells any assets to any buyer (public or private) in the given year and is zero otherwise.

We also consider four other outcome variables including sales growth and asset growth, which are the log of the ratio of current sales to past-year sales and current assets to past-

[8] If we exclude firms with missing R&D, we obtain similar results.

year assets, respectively. We compute firm valuation ratios as the firm market value (market equity plus book assets minus book equity) scaled by total assets, and we compute profitability as operating income before depreciation scaled by total assets. Finally, we consider four financing policies including equity issuance, debt issuance, equity repurchases, and dividends, with all four being scaled by assets. All accounting ratios are winsorized in each year at the 1%/99% level. A complete variable description is in the Appendix.

### 3.6 Summary Statistics and Correlations

Table 1 displays summary statistics for our 1989 to 2017 panel of 100,525 firm-year observations. The average value of our key FIC-scope and NAICS-scope variables are 6.9 and 6.3, respectively. This suggests that, using 2% granularity, the average firm in our sample is operating in markets that are related to roughly six well-defined FIC or NAICS-4 industries, respectively. This is larger than the average number of Compustat Operating segments, which is just 1.4 in our sample. We measure scope in this relatively broad way to ensure there is adequate power to compare firms in the cross section, and because operating segments likely understate the true girth of the product portfolios offered by public firms in the United States. Notwithstanding that, we also note that our results are robust if we measure scope more narrowly using a 1% threshold or more broadly using a 5% threshold.

We also note that our accounting variables have values that are similar to those in other studies. The average firm in our sample spends roughly 5.5% of its assets each on R&D and CAPX, and 29% of our sample firm-years are involved in an acquisition. The average firm's valuation ratio (market to book) is roughly 1.76, and the average firm spends roughly 1.6% of its assets on repurchases and 0.8% on dividend payments.

Table 2 displays Pearson correlation coefficients. Our two scope variables (FIC-scope and NAICS-scope) are about 20% correlated with the number of Compustat operating segments. This is significant and positive as expected, but it is far from unity, illustrating why evaluating scope using segment counts is likely inadequate. Both measures are also roughly

27% correlated with firm size, illustrating that larger firms serve a wider array of product markets. This finding also illustrates why controlling for size is important. For example, both scope measures are also positively correlated with both the acquisition dummy and the target dummy. However, both dummies are even more positively correlated with firm size, as it is well known that larger firms are more active in restructuring. Given these facts, it is not surprising that when we run formal regression analysis, we find that scope is associated with more acquisitions but less divestitures (targets), which conforms to the intuition that firms with high incentives to increase scope are indeed net acquirers once size is held fixed.

# 4 Relation to Traditional Scope Variables

In this section, we explore and validate the properties of our scope variables and compare them to the Compustat segments database and to firm size. We first illustrate the nature of our scope measures via examples. Table 3 displays the markets that Disney operates in for three representative years: 1990, 2003, and 2017. The results are based on the LDA version of our scope measures as this approach additionally provides intuitive labels. Our LDA model is based on 300 topics, and we fix the topic structure to 1997 for consistency with our baseline variables. The table shows that Disney operates in six potential sub-markets in 1990 including studio productions, sports, cable programming, motion pictures, publishing, and music. This increases to 10 sub-markets in 2003 and 11 sub-markets in 2017 illustrating Disney's increase in scope. Online Appendix Table IA1 shows a similar report for Pfizer, whose scope has remained stable over time.

## 4.1 Compustat Segments

The existing literature uses the Compustat segment tapes when exploring scope and takes the perspective that Conglomerate firms have high product scope, and they are diversified as they operate in unrelated product markets. We take the perspective that the segment tapes are problematic for measuring scope, not only due to mismeasurement (see Villalonga 2004), but also because we expect modern firms to increase scope without increasing the number of rigid operating segments. In particular, we expect modern firms to use innovation to increase product scope through more flexible production and by redeploying existing assets. A consequence would be increasing scope but Compustat segments would remain stable.

We begin our analysis by computing summary statistics for subsamples of firms with different numbers of Compustat segments. Table 4 displays the results and shows that moving from one segment to two increases FIC-scope and NAICS-scope by roughly one unit. For example, FIC-scope increases from 6.42 to 7.53. The table also shows that adding segments beyond two also roughly adds one more unit to our scope measures. These results conform to intuition, and suggest that each additional operating segment adds one additional product market to our measures of scope. However, these expected positive relationships greatly under-state the true variation in our measures of scope, which have standard deviations ranging from five to seven, and thus most of the variation in our measures cannot be explained by the number of Compustat segments a given firm operates in.

## 4.2 Scope Trends

Figure 1 plots the average number of Compustat segments over time in the upper panel, and the average FIC-Scope and NAICS-Scope in the lower panel. The number of segments initially declines from roughly 1.5 in 1989 to 1.3 by 1997. This conforms to the intuition that older conglomerates, which might have been formed in the 1970s and 1980s, were gradually disbanding over time. However, from 1997 to 1999, we observe a major structural break and the number of segments suddenly increases to more than 1.5. This jump can be explained by SFAS 131, in which FASB changed segment disclosure requirements for fiscal years ending after December 15, 1997. This rule required that managers must report segments based on how managers internally evaluate operating performance. Prior to this rule, segment reporting was instead based on an industry approach. Yet we note that the rule change was precipitated by concerns by market participants that segments were being under-reported, perhaps for strategic reasons. We refer readers to Song (2020) for a more detailed summary of these important events.

The events leading up to SFAS 131, and the rule change itself, suggest that any trend for Compustat segments in Figure 1 should be interpreted with caution. For example, the alleged practice of under-reporting prior to the rule change calls the declining trend from 1989 to 1997 into question. The flat trend after 1997 is also questionable because post rule change segment counts are based on how managers internally evaluate performance and not how many product markets the firm actually operates in. A major result we report later is that firms increasing scope choose to operate in industries that are closely related, not industries that are diversified. Intuitively, many managers might prefer to internally assess performance of such related industry product lines together. If so, the number of Compustat segments would understate the increase in scope, as it is likely to only capture instances where firms operate in unrelated product markets where separate internal evaluation is likely.

The lower panel of Figure 1 displays the average FIC-scope and NAICS-scope over our sample. The coverage dating back to 1989 was made possible by our backward extension of the TNIC database to the late 1980s. The figure illustrates that scope was increasing during our sample, with the most rapid rate of increase between 1997 and 2013. The average scope of firms in our sample increased by a full 50%. Notably, the upper panel shows that the number of Compustat segments did not change during this period. These results are consistent with the view that product market scope did increase, but the increases were mainly driven by firms serving multiple industries that are related (for example selling computers and cell phones) rather than diversified and unrelated (such as selling oil and cat food). Put together, our results are consistent with managers evaluating these highly related industry markets together, resulting in relatively few Compustat segments despite the high product counts.

A final note on scope trends is that the increased scale and scope we report can also help to explain why the length of 10-Ks has also increased over time. The upper panel of Figure

2 plots the average number of words in the 10-K Item 1 over time, which increased in the first half of our sample steadily until 2003. After this year, the size of the Item 1 has become relatively stable. This suggests that document size is related to scope, but also different, as Figure 1 shows that scope has been increasing throughout our sample, even after 2003.

To further understand the evolution of Item 1 over time, we note that there are 3 sources of variation that might drive its length: (1) increased scope across product markets, (2) increased product variety within markets, and (3) increased boilerplate content. Regarding boilerplate content, our removal of stop words, or any word that appears in more than 25% of all filings in a given year, should reduce the impact of boilerplate content. Regarding increased scope, as noted above, its time series is different from the trend in document size overall, which suggests that a shift in product variety might also have occurred in parts of our sample. The lower panel of Figure 2 supports this intuition, and reports the average time trend of the number of 10-K Item 1 words per product market the firm likely operates in. This is computed as the number of words in the given firm's Item 1 divided by FIC-scope (number of likely product markets). The figure suggests that within-market product variety has likely increased some during the earlier part of our sample. These findings motivate future research on this topic, although our focus is on across-market scope.

## 4.3 Scale and Scope

We now examine the role of increasing firm scale and its relation to firm scope. Figure 3 plots average firm size over time (based on book assets) both in nominal terms and in inflation-adjusted terms, and shows that average firm size has increased substantially over time. Using the conservative inflation adjusted metric, firm size has roughly tripled during our sample. This increase likely reflects the increases in firm scope we document above, but also increases in firm scale. We thus assess scale and scope properties using sorts.

Table 5 reports average scale and scope statistics for size quintiles. Quintiles are formed by sorting on Compustat assets separately in each year. We report these statistics separately for the full sample and for firms that report just one Compustat operating segment.

The table confirms that all measures of scope sort strongly with firm size. Even for Compustat segments, the smallest quintile firms have an average of 1.22 segments, which grows to 1.94 for the largest firms. Regarding FIC-scope, the interquartile range is from 5.65 product markets to 8.62 product markets. The range is larger for NAICS-scope at 4.02 to 9.17 markets. Yet the growth in scope by any measure across these quintiles pales in comparison to the range of firm size itself. Small quintile firms have about $23 million in assets, whereas the largest quintile firms average $11.7 billion. We conclude that some of the increase in firm size is likely related to a corresponding increases in firm scope. However, the sheer magnitude of the scale increase suggests that other drivers also matter. For example, U.S. firms have achieved not only economies of scope, but also economies of scale.

We conclude with a note about single segment firms. The rightmost columns in Table 5 report the same statistics for the subset of single segment firms. We find that the large variation in firm scope with firm size is not much diminished. This reinforces our conclusion that Compustat segments are not a reliable source of information about firm scope.

# 5 Scope but not Diversification

In this section, we examine if the high levels of scope we find are related to companies spanning distant and highly diversified product markets, or more proximate related product markets. We first compute the average product market distance between every permutation of pairs of FIC-300 industries. For a given pair of industries in a given year, this is computed as the average TNIC pairwise similarity (see HP2016) between all of the firms in the first industry relative to those in the second industry in the pair. We thus observe which industry pairs are proximate and distant.

Next, we consider the firm-to-industry mapping created when computing the FIC-score. As noted in Section 3, this calculation first requires us to identify the set of FIC-300 industries that each firm likely operates in. We use these firm configurations across industries to create a database of observed "operating pairs". A firm that maps to industry $i$ and industry $j$ is thus an observation of the operating pair $ij$. A firm that maps to three industries {i,j,k} is an observation of three operating pairs: {ij,ik,jk}. A firm that maps to just one industry does not have any observed operating pairs. We then tabulate the number of operating pairs for each pair of industries and obtain the distribution of operating pairs for each year.

In each year, we sort industry pairs into deciles based on the TNIC similarity of the pair. Table 6 then reports the fraction of all observed operating pairs that are in each decile. We report this distribution for all firms, and separately for single segment and multi-segment firms. The table shows that firms overwhelmingly operate in industry pairs that are close together in the product space. Almost 40% of all operating pairs are in the highest decile of TNIC industry pairwise similarity. An additional 13% to 15% are in the next decile. These results indicate that modern multi-industry firms are not the diversified conglomerates portrayed in the early literature. Rather, these firms operate in highly related industries with value-adding synergies (we present evidence of higher valuations later).

We next examine the risk properties of our scope measures. Our prediction is that firms having operating segments that are far in the product space should be less risky due to the diversifying effects of unrelated markets. In order to examine whether different levels of scope reduce or increase risk, we separate out scope into terciles from "close scope" to"far scope". This is done by computing the average pairwise distance of the FIC-300 industry pairs each firm operates in (using the same operating pair database discussed above) and sorting firms into terciles in each year.

Table 7 reports the results of regressions where measures of firm risk is the dependent variable. We also include controls for size, age, and year fixed effects. We additionally include firm fixed effects in Panels C and D. The first dependent variable Market Volatility is the standard deviation of the firm's daily stock returns in year $t$, and Cashflow Volatility is the standard deviation of a firm's quarterly operating income scaled by assets, computed over the 8 quarters of year $t$ and $t + 1$.

Panels A and C of Table 7 show that firms with close-scope operations are significantly more risky for stock volatility. These results are robust with or without firm fixed effects. Panel B shows that cash flow volatility is broadly higher for firms with high scope, with the coefficient being largest for near-scope firms. Panel D shows that the relation between scope and cash flow volatility generally becomes insignificant when we additionally include firm fixed effects. These insignificant to positive links to risk broadly support our conclusion that diversification is unlikely to be a major motive for scope increase. Because scope expansions focus on related markets, and because scope increase requires innovative investment (which is risky), a high scope strategy tends to associate with higher (not lower) risk.

## 5.1 Validation of Scope Measures

The results presented support the conclusion that text-based measures of scope have many advantages over using Compustat segments. In this section, we provide a formal validation test that uses direct firm statements to evaluate candidate measures of scope.

We consider four queries of firm 10-Ks to identify direct statements indicating that a firm offers products with a high degree of scope. These queries are based on three lists:

**List A:** product lines, product categories

**List B:** product lines, product categories, service lines, service categories

**List C:** breadth, broad, broader, wide, multiple, numerous, diverse, categories, divisions

We use the metaHeuristica software to compute four variables of interest. "Product Breadth" is the number of paragraphs in each firm's 10-K that mentions a phrase in List A, scaled by the total number of paragraphs in the firm's 10-K. "Product/Svc Breath" is analogously defined based on List B. "Product Breadth Detail" is the number of paragraphs that contain a phrase in List A and also a word from List C. "Product/Svc Breadth Detail" is analogously defined using List B and List C. Intuitively, when these scores are higher, it is likely that the firm offers a high scope array of products and services. We also note that the latter two variables are quite stringent and are based on proximity searches. This anchor-phrase approach is more rigorous than basic word counts as it would be less informative if words from List C appeared in different paragraphs than those where Lists A and B appear.

To validate our measures of scope and compare them to Compustat segments, we regress the four above variables on candidate scope measures. We additionally include controls for size, age, market to book, and the TNIC HHI, and we also include firm and year fixed effects. For validation, a measure of scope should have a strong positive and significant coefficient.

Table 8 displays the results. The first four rows only include the controls as a baseline, and illustrate for example that size (log assets) is not surprisingly related to our direct statements of scope, and the $t$-statistic is between 3.5 to 4.0. Rows (5) to (8) add the number of Compustat segments to the regression. This variable is positive and significant for the first two variables ($t$-statistic of roughly 2.7) but is only significant at the 10% level for the more stringent variables. Rows (9) to (12) additionally add FIC-scope to the regression. We find that FIC-scope is much more positive and significant than both firm size and Compustat segments. Its $t$-statistic is roughly 8.5 for the broad validation measures and 6.5 for the strict validation measures. Additionally, including FIC scope reduces the significance of the Compustat segment variable by roughly one third. Rows (13) to (16) reproduce this test for the NAICS-scope variable, and we find similar but slightly weaker validation ($t$-statistics range from 5.3 to 6.2). We conclude that FIC-scope has the strongest validation, NAICS-scope also performs well, and Compustat segment counts are noisy.

# 6 Scope Incentives and Corporate Finance Policies

We now explore how firms seeking to increase scope modify their corporate finance policies. This question touches upon many issues of high importance for understanding corporate finance in general, and also issues of relevance to regulators. For example, is the increase in scope we report related to the high level of acquisition activity reported in the popular press over the past couple decades? Additionally, is innovation investment in R&D associated with increases in scope, or is scope achieved instead through acquisitions and capital expenditures? We examine these potential mechanisms in this section using plausibly exogenous shifters of the incentives firms have to increase scope. We also assess the link between scope and firm performance, and we explore how increases in scope are financed.

## 6.1 First-Stage Analysis

We first examine the relation between our measures of scope incentives and our novel measures of scope. This analysis constitutes the first stage that we will ultimately use in our two-stage least squares analysis of the impact of scope incentives on ex-post investment, performance and financing in the next section. We use these instruments to assess the relevance of plausibly exogenous incentives to increase scope on various corporate finance outcomes. This is thus a test of mechanism relevance more than a test of pure causality.

Our first ex-ante measure of scope incentives we label *Sectoral Redeployment Potential*, which we explain in detail in Section 3.3. This variable measures the extent to which the assets owned by a focal firm's close peers can be easily redeployed for use in the product markets covered by the focal firm's more distant peers. When this variable is high, it indicates that the focal firm likely has the ability to increase its scope outward at low cost, as its assets are likely redeployable to assist in production in these nearby product markets. Our second measure we label *Sectoral Opportunity Set Potential*, which is based on the supply of scope-expansion opportunities rather than the cost of executing them. This measure is one minus the concentration ratio of the distribution of industries spanned by the focal firm's more distant peers. When this quantity is high, it indicates that these peers span many related product markets. As a consequence, the focal firm likely sees a favorable distribution of industries to which it can increase its scope (a "thick" opportunity set). Importantly, both measures are computed without using the characteristics of the focal firm itself, and are weighted heavily on the more distant peers. The use of distant peers, as explained earlier, is supported by econometric research as being more plausibly exogenous due to the the second-degree (rather than first-degree) network linkages of these peers.

In our first stage analysis, we regress our scope measures on both scope incentive variables and include all control variables included in our two stage models, and also firm and year fixed effects in both stages. The results are in Table 9. Row (1) shows that both scope incentive variables are positively related to FIC-scope. *Sectoral Redeployment Potential* is positive with a $t$-statistic of 4.3, and *Sectoral Opportunity Set Potential* has a positive $t$-statistic of 11.9. Results are similar for NAICS scope. For comparison, we run the same regression with the # of Compustat segments as the dependent variable and we find the results are much weaker. The first scope incentive variable is not significant, and the second is moderately significant with a $t$-statistic of 2.9. These results indicate that our scope incentive variables are powerful shifters of FIC-scope and NAICS-scope, but not the number of segments. We thus only consider FIC-scope and NAICS-scope in our second stage models.

## 6.2 Corporate Finance Policies and Scope Expansion

We now consider the second stage regressions where we assess the impact of ex-ante scope increase incentives on ex-post investments, performance, and financing policies. We start with investments and consider ex post acquisitions, divestitures (target of acquisition), R&D, and CAPX. In particular, we consider two-stage least squares regressions where we instrument either FIC-scope or NAICS-scope using our two instruments based on ex-ante scope incentives. We also control for size, age, and firm and year fixed effects. We also consider a specification that also controls for ex ante valuation (market to book) and the TNIC HHI.

The results for investment policies are displayed in Table 10. In Panel A, we find that firms with high ex-ante incentives to increase scope are more likely to do an acquisition ($t-statistic$ of 4.7) and are less likely to divest ($t-statistic$ of -2.3). Economic magnitudes are also quite large. If we vary the predicted scope from the 25th percentile to the 75th percentile, the change in the probability of the firm acquiring another firm increases by 7.88 percentage points, which is 27% of the mean and 17.4% of the standard deviation of the probability of doing an acquisition. For the probability of divesting, the economic significance is smaller but still meaningful with the predicted probability declining by 2.65 percentage points, which is 20% of the mean and 8% of the standard deviation.

These results indicate that acquisitions are a natural way to increase scope, and avoiding divestitures is also necessary to avoid losing any previous gains in scope. These results also suggest that increases in product scope are an important acquisition motive, and that scope increases might help to explain why acquisitions became so prevalent over the past two decades. These results can also inform regulatory debates given the more controversial link between acquisitions and market power suggested in the popular press.

A second finding is that firms with high incentives to increase scope also increase R&D expenditures, but not capital expenditures. The increase in R&D is significant with a *t*-statistic of 3.9. Economically, if we vary the predicted scope from the 25th percentile to the 75th percentile, firms are predicted to increase their R&D/Assets ratio by .87 percentage points, which is 15.5% of the mean and 6.8% of the standard deviation of R&D / Assets.

This finding indicates that innovation likely facilitates scope expansion as do acquisitions. For example, when assets can be redeployed across product markets, innovation spending can be synergistic as it can serve to improve productive efficiency and flexibility. In turn, this model likely facilitates the creation of multi-industry firms that do not need multiple operating segments. For example, firms might use R&D to develop more universal and flexible production sites. Hence the benefits of increased scope might be feasible in the 21st century without having to accept the dark side of negative governance externalities. Indeed our earlier results suggest that increases in scope were achieved during our sample with almost no change in the average number of Compustat operating segments.

Finally, we examine the extent to which incentives to increase scope also shift vertical integration. We measure vertical integration following Fresard (2010) and find a significant positive relationship. This suggests that vertical integration is one way firms increase scope synergistically. Economically, if we vary predicted scope from the 25th percentile to the 75th percentile, firms are predicted to increase vertical integration by .80 percentage points, which is 70.2% of the mean and 72.6% of the standard deviation.

Table 11 documents novel supplemental results based on running the tests in Table 10 separately within each major industry sector, as defined using the Fama-French-5 classification. Panel A shows that firms in the technology sector primarily increase scope by investing in R&D and not in acquisitions. In contrast, firms in manufacturing sectors do the opposite and primarily invest in acquisitions but not in R&D. These results are intuitive given the focus on intangible assets in the tech industry and the focus on tangible assets in manufacturing. The other three sectors (consumer, health and miscellaneous) are consistent with both investment channels (R&D and net acquisitions) being relevant.

In Online Appendix Table IA2, we run the tests in Table 10 using a one-stage model where we regress ex post investments on the two ex ante scope-incentive measures directly. We find that increased R&D and reduced asset sales are significantly and positively related to the asset redeployability incentive, whereas increased acquisitions are most related to the opportunity set incentive. These results suggest that innovation spending indeed might be used to redeploy flexible assets to new localized product markets, thus also leading to a lower likelihood of asset sales. In contrast, acquisitions are more likely in less redeployable markets having numerous opportunities (redeployment is less feasible in these markets).

## 6.3 Ex-post Outcomes and Scope

Table 12 reports the results of analogous regressions for ex-post performance metrics. We consider ex-post valuations (market to book), sales growth, asset growth and return on assets. All right-hand-side variables are lagged one period and we use a two-stage least squares model with scope instrumented by our two ex ante scope incentive variables.

Table 12 shows that firms with high scope expansion incentives experience higher ex-post valuations, higher sales growth and higher asset growth.[9] However, we do not observe a significant coefficient for profitability measured as return on assets (ROA). Economically, if we vary the predicted scope from the 25th percentile to the 75th percentile, firms' valuation is predicted to increase by .24, which is 13.5% of the mean valuation and 11% of the standard deviation. For sales (asset) growth, respectively, these economic magnitudes are 9.33 (14.1) percentage points, which is 23.3% (39.7%) of the standard deviation of sales (asset) growth.

Overall, the higher valuations suggest that scope expansion is a positive net present value investment and that investors expect higher profits in the future. We also note that the non-result for ROA is consistent with the view that firms add new industries to their portfolio in an assortive way,[10] and focus on the most profitable markets first. This interpretation is consistent with all four of our findings on performance. Our valuation results - which show that firm market-to-book equity ratios increase with our scope measures - are in contrast to the historical conglomerate literature (Berger and Ofek (1995) and Lang and Stulz (1994)) that finds that firm market-to-book valuation measures go down with an increased number of reported Compustat segments. In Online Appendix Table IA3, we run these outcome tests in a one-stage model as before. The results suggest that value creation is strongest for scope expansions that are realized through the asset redeployability channel, but sales growth is highest for the opportunity set channel. These results are intuitive given our investment results, as they suggest that the return on investment is higher for organic investments such as R&D using existing redeployable assets. Consistent with the literature, valuation gains to acquisitions are lower by comparison, but acquisitions instead generate higher and more immediate sales growth (as acquisitions, unlike organic growth, typically entail the purchasing of an existing market presence in addition to the assets).

We next explore the importance of scope expansion to venture capital funded private firm

[9]Our results for valuation are robust if we use the Peters and Taylor (2017) measure of Q (which accounts for intangibles) as an alternative to market to book assets.

[10]The non-result for ROA is robust to alternative specifications such as (A) adding R&D back into operating income as it is not clear that R&D is an expense, (B) truncating ROA at zero when it is negative, or (C) scaling profitability by sales instead of assets.

entry and product market fluidity (the rate of change in product portfolios of existing public firms that operate in a focal firm's product markets) over time. We use two variables from Hoberg, Phillips, and Prabhala (2014) to measure both quantities. The first variable, *VC funding similarity* is the cosine similarity of the given focal firm's 10-K product description to the average vocabulary used by all startups in the given year (where the startup vocabulary is obtained from Venture Expert business descriptions of all startups receiving their first round of financing in the given year). The second variable *Product Market Fluidity* is the average market-wide change in the use of the given firm's 10-K product description vocabulary by all other firms. A high value indicates a large amount of product innovation by competing firms in the focal firm's product markets. We report these results in Table 13.

The results are shown in Table 13, which shows that high venture capital related entry occurs into the focal firm's markets when the firm has scope increases, and fluidity also increases consistent with existing public firms also turning over their product portfolios more actively when scope incentives are strong. Economically, the predicted impact on the firm is large. If we vary the predicted scope from the 25th percentile to the 75th percentile, venture capital financed entry is predicted to increase by 40.9% of the mean and 79.1% of the standard deviation of venture capital financed entry in our sample. For product market fluidity these magnitudes are 56% of the mean and 79% of the standard deviation.

These results are particularly compelling when interpreted alongside our earlier finding that high scope firms also do more acquisitions (see Table 10). Overall, these innovation and acquisition results are consistent with Phillips and Zhdanov (2013), who show that existing firms have strong incentives to purchase related start ups as their scope increases. This interpretation suggests that already-public firms tend to outsource scope-enhancing R&D and technology development to smaller startups (thus generating the higher levels of VC funding we observe) and later buy them (thus explaining in part the increased acquisitions).

Table 14 reports the results of analogous regressions for ex-post financing policies. The question of interest is how do firms likely finance scope expansions? Our results suggest that equity is more commonly used than debt. Increased equity in the capital structure appears to accrue both through the issuance of new shares and through lower overall dividend payments. If we vary the predicted scope from the 25th percentile to the 75th percentile, equity issuance is predicted to increase by 2.72 percentage points which is 53.6% of the mean and 17.9% of the standard deviation in our sample. Dividends are predicted to decrease by 45.2% of the mean and 13.8% of the standard deviation. These results are consistent with innovation and asset redeployment being used to facilitate scope expansion, as neither creates a significant amount of new fixed collateral that is traditionally associated with debt financing. The one stage results in Online Appendix Table IA4 shows that our results for the financing variables are rather evenly spread across the two scope incentive variables.

Although the goal of our tests is not to establish full exclusion, but rather to construct strong tests of mechanism consistency, we nevertheless report the results of statistical tests that summarize the quality of exogenous identification in Tables 10 to 14. The first finding is that our IV models produce a Kleibergen and Paap (2006) r-k statistic that is significant at the one percent level, indicating that our instruments have excellent power. Regarding the Sargan-Hansen (see Sargan (1958) and Hansen (1988)) test of overidentification, the majority of our dependent variables are not susceptible to overidentification as the Hansen J-test is not significant at the 5% level. These dependent variables include the acquisition dummy, the target dummy, vertical integration, sales growth, asset growth, equity issuance and dividends. The four variables that have a significant J-test, and hence where the exclusion requirement is more in doubt, are RD/assets, Tobins Q, VC funding similarity and product market fluidity. Overall the significant results for instrument power and the overall consistent results for overidentification support our conclusion that strong incentives to increase scope likely explain a wide array of important corporate finance policies.

As discussed in Section 3, we also consider an alternative measure of scope based on the Latent Dirichlet Allocation topic model. As we note in Online Appendix Tables IA5 to IA7, our results are fully robust to this alternative measure of scope.

## 6.4 Extent of Scope Expansions and the Role of Scale

We next examine if our findings regarding scope-induced corporate policies and performance differ as firms become more spread out regarding the markets they serve, and whether these results differ based on firm scale. These issues are important as it is natural to ask whether the gains to scope become larger or smaller as firms advance in scale and scope. For example, if the gains to adding more scope are largest for firms that already achieved high scale and scope, it would suggest that the trend toward increased scope and fewer publicly traded firms is likely to continue. Alternatively, if the gains are largest for smaller firms, then this trend might focus more on smaller firms, and less on already-expanded firms.

As in our earlier analysis on risk profiles, we first split our sample into three terciles based on the type of scope firms have achieved ex ante. Firms with “near” scope are those with segments that are similar to each other in terms of product market spatial distance (these firms have a narrow market presence and operate in highly related markets). We also consider terciles for “mid” scope and “far” scope (firms that have achieved a high level of spatial coverage across spatially more distant U.S. markets). These subsamples are constructed in each year by sorting firms based on past-year scope-spread (the average spatial distance of the segments each firm covers) resulting in have near, mid and far scope subsamples. Our instrumented variable of interest, as in the prior subsection, is our measure of scope (FIC-Scope). For parsimony, we only report this variable’s coefficient for each regression. Thus, in our table below, each coefficient we report is from a separate regression based on the three tercile subsamples as indicated in the column headers. We report these regressions across the key dependent variables we considered in the prior subsection, and the dependent variable for each model is indicated in the first column. All regressions include firm and year fixed effects as well as controls for size and age.

Table 15 shows that results are strongest for firms with near-scope. They do more acquisitions, R&D and they have higher valuations when scope-incentives are strong. These results support the predicted value gains to close-by scope as predicted by Panzar and Willig

(1977). The strong results for innovation for these firms is further consistent with Teece (1980) and Henderson and Cockburn (1996), as firms especially invest in innovation when they are expanding across closer markets, where the gains associated with these theories are likely to be largest. The table also shows that there are essentially no valuation gains to increasing scope when firms are operating across product markets that are more spatially distant. This result suggests that there are diminishing gains to scope when firms increase scope too far. This finding is consistent with potential inefficiencies with excessive scope, as suggested by the early conglomerate literature focusing on excessively diversified firms. Further consistent with these views, we documented earlier that few firms actually choose to operate across markets that are spatially distant, especially in more recent years. Lastly, the the table shows that there is more VC backed entry and product market fluidity around firms with nearby scope, consistent with increases in innovation not only within the firm, but also external to the firm itself in these nearby scope markets. This further supports the confluence of the aforementioned works by Teece (1980) and Henderson and Cockburn (1996), alongside the outsourcing-of-innovation predictions of Phillips and Zhdanov (2013).

We next examine how our main results vary with firm scale, and we thus sort firms into annual terciles based on firm size using ex ante total assets. Table 16 displays the results using the same tercile-based format as Table 15. We thus report regression coefficients separately for large, mid-size and small firm subsamples. As before, our instrumented variable of interest is a measure of scope (FIC-Scope), and for parsimony, we only report this coefficient for each regression. All regressions include firm and year fixed effects and controls for size and age.

The results in Table 16 indicate that results are stronger for small firms. They do more acquisitions, more R&D, have higher sales growth and have higher valuations as scope increases. On the other hand, we see a negative effect (not quite significant) on valuation for the largest tercile firms, which is in sharp contrast to the significantly positive valuation result for smaller firms. This suggests diminishing returns to scope as scale increases, and suggests that scope increases might eventually reach a natural limit. Yet we note that these results are suggestive, and we believe further research examining whether an optimal scale exists (and what this optimal scale is) could be fruitful. We also note that our results are regression-based and therefore give insights regarding the average firm. Any one individual firm might have unique capabilities that allow it to expand scope further (or less) than what is implied by the average.

We find that VC funding and product fluidity are significantly higher around smaller firms with high scope incentives, but this coefficient is still significantly positive for larger firms. This is consistent with the idea in Phillips and Zhdanov (2013) that large firms are perhaps ideal candidates to purchase outsourced innovation specifically from startups. Overall, our results in this section are consistent with the impact of scope on many policies being stronger for firms focused on nearby scope expansion and for smaller firms. The stock market also values scope expansion for these firms significantly more.

We conclude this section with a note on robustness. Readers might be concerned that our results are excessively driven by ultra-large companies such as Amazon, which are known to have experienced substantial growth in scope. Table IA8 thus considers a specification where we drop the 50 largest companies from the sample in each year based on lagged assets and our results are fully robust. We conclude that the impact of scope is much more broad-based as it impacts firms of all sizes as documented above.

# 7 Implications for Concentration

Although our paper's primary objective is to document the rise in scope and examine the role of scope in corporate finance strategies and performance, we also present some initial evidence that increasing scope has implications for industry concentration. Given this evidence is suggestive, we discuss it briefly here and provide details in Section 2 of the Online Appendix.

Intuitively, if the number of firms in the economy were held fixed, and every firm expanded its scope to serve twice the number of product markets, it follows that pure horizontal competition would increase economy-wide as more firms would be serving each market, and consumers would have twice as many options in each market. However, if these expanded firms are (incorrectly) assigned only to their historical industry classifications, industry concentration ratios would be overstated. One remedy is to use the Compustat segment tapes (Hoberg and Phillips (2010b), Grullon, Larkin, and Michaely (2019)), which allow the researcher to assign firms to more than one industry. However, as we note earlier, SFAS 131 decoupled segment reporting from the actual industries firms operate in starting in 1997. Likely as a consequence, Figure 1 shows that segments do not capture increases in scope.

We present and discuss results in the Section 2 of the Online Appendix that show that the trend of increasing industry concentration ratios documented in the literature (see Grullon, Larkin, and Michaely (2019) for example) can be explained by the scope increases we report.

We consider two approaches to adjust concentration ratios for increasing scope. The first method computes HHIs using the scope segments that we construct using text, thus replacing the potentially problematic Compustat Segment database that has been used in the literature. This approach shows no increase in industry concentration since 1997 (see Figure 4). A limitation of this approach is that we don't have sales weights by segment and therefore we use textual intensity weights, where the weights are the similarity of the firm's product description to the vocabulary that we used to construct each corresponding FIC-300 segment. We also find similar results when we equally weight across segments. This method is discussed in detail in Section 2.1 of the Online Appendix. Our second method does not rely on segment weights. Rather, we compare each firm's product description similarity to narrower SIC-3 product words and broader SIC-2 product words. We justify this approach by showing that, over time, firm product descriptions load more on the broader SIC-2 industry vocabulary and less on the narrower SIC-3 vocabulary, indicating that firms operate at a coarser granularity in later years. Figure 5 reinforces our above finding using granularity-weighted HHIs and documents that HHI indices do not increase since 1997. This method is discussed in detail in Section 2.2 of the Online Appendix.

The goal of these suggestive results is to show that increases in scope can lead to decreases in concentration. A limitation of our concentration analysis common to studies in this area is that we only have scope data for U.S. publicly traded firms and we are unable to account for private or foreign firms. We thus plan to share our data and methods for future researchers to extend. However, we also note that stylized facts suggest that accounting for both foreign competitors and private firms should reinforce our finding that concentration is not strongly increasing over time. For example, globalization is increasing (see Hoberg and Moon (2017)) and accounting for foreign competition would likely further reduce the growth rate of concentration over time. In addition, studies including Ewens and Farre-Mensa (2020) suggest that larger firms are staying private longer, and accounting for larger private firms should further reduce the measured growth rate of concentration. Yet we advocate caution in drawing overly strong conclusions given these limitations.

Our results suggest that purely horizontal concentration does not increase when scope is taken into account. Increases in scope, however, might motivate different concerns. Although scope increases reduce horizontal concentration as firms serve more markets, scope may increase anti-competitive conduct through product bundling, by increasing barriers to new entry, or it can induce kill zones in the market for innovation (see Kamepalli, Rajan, and Zingales (2020)). Firms achieving high levels of both scale and scope might realize particularly elevated levels of market power through these channels.

# 8 Conclusions

We use textual analysis of firm 10-Ks to compute novel firm-year measures of firm scope and likely operating segments. Using our new measures, we find that the scope of U.S. firms has increased dramatically during our sample period from 1989 to 2017. Our findings illustrate the rise of a new 21st century high-scope firm, which increases product scope using innovation and acquisitions. These firms are capable of servicing multiple product markets without increasing the number of operating segments. Indeed, analogous tests using the traditional Compustat segment tapes are generally uninformative regarding the scope of U.S. public firms. These findings support our thesis that modern firms build multiple-industry related product portfolios while maintaining a simple single segment organizational form.

We find that firms increase scope by acquiring more, divesting less, and increasing innovation spending in R&D, but they do not increase CAPX. The increased innovation spending is consistent with developing increased flexibility in production. Firms increasing scope also realize higher valuations and higher sales growth. Scope expansion is financed using equity rather than debt, consistent with intangibles and asset redeployment not creating material amounts of new collateral. We document that related-markets scope expansion is also highly valued by the market, which illustrates the novelty of our results given the discount for conglomerates previously documented in the earlier conglomerate literature.

We conclude our analysis with exploratory evidence that the increase in scope that we report might explain why other studies have documented a trend of increasing concentration over time. We compute adjusted HHIs that account for the fact that increased scope can increase competition as more firms operate in more overlapping markets. We find that these adjusted HHIs are essentially unchanging since 1997. These results suggest that an extended narrative that considers the growth in scope can help to contextualize the previous reports of increasing HHIs and the high levels of M&A activity. In particular, much M&A has been targeted at increasing scope, a business strategy that can produce positive net present value as our findings indicate. Yet these results must be interpreted with care, as increased scope can lead to significant antitrust concerns in the form of product bundling and increased bargaining power within supply chains or decreased entry into new markets that larger already established firms have moved into. We believe that future research examining scope and market power in detailed product areas could be particularly impactful given the importance of these issues to regulators and society at large.

# References


Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, 2020, The fall of the labor share and the rise of superstar firms, *The Quarterly Journal of Economics* 135, 645–709.

Bena, Jan, and Kai Li, 2014, Corporate innovations and mergers and acquisitions, *Journal of Finance* 69, 1923–1960.

Berger, Phillip, and Eli Ofek, 1995, Diversification's effect on firm value, *Journal of Financial Economics* 37, 39–65.

Blei, David, A Ng, and M Jordan, 2003, Latent dirichlet allocation, *Journal of Machine Learning Research* 3, 993–1002.

Boehm, Johannes, Swati Dhingra, and John Morrow, 2022, The comparative advantage of firms, *Journal of Political Economy* forthcoming.

Braguinsky, Serguey, Atsushi Ohyama, Tetsuji Okazaki, and Chad Syverson, 2020, Product innovation, product diversification, and firm growth: Evidence from japan's early industrialization, University of Maryland Working Paper.

Bramoulle, Yann, Habiba Djebbari, and Bernard Fortin, 2009, Identification of peer effects through social networks, *Journal of Econometrics* 113, 41–55.

Chandler, Alfred D, and Takashi Hikino, 1994, *Scale and Scope* (Harvard University Press).

Cohen-Cole, Ethan, Andrei Kirilenko, and Eleonora Patacchini, 2014, Trading networks and liquidity provision, *Journal of Financial Economics* 113, 235–251.

Crémer, Jacques, Luis Garicano, and Andrea Prat, 2007, Language and the theory of the firm, *Quarterly Journal of Economics* 122, 373–407.

Custodio, Claudia, 2012, Mergers and acquisitions accounting and the diversification discount, *Journal of Finance* 69, 219–240.

Doidge, Craig, Andrew Karolyi, and Rene Stulz, 2017, The u.s. listing gap, *Journal of Financial Economics* pp. 464–487.

Ewens, Michael, and Joan Farre-Mensa, 2020, The deregulation of the private equity markets and the decline in ipos, *The Review of Financial Studies* 33, 5463–5509.

Fresard, Laurent, 2010, Financial strength and product market behavior: The real effects of corporate cash holdings, *Journal of Finance* 65, 1097–1122.

———, Gerard Hoberg, and Gordon Phillips, 2020, Innovation activities and integration through vertical acquisitions, *Review of Financial Studies* 33, 2937–2976.

Grullon, Gustavo, Yelena Larkin, and Roni Michaely, 2019, Are us industries becoming more concentrated?, *Review of Finance* p. 697–743.

Hansen, Lars Peter, 1988, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Hart, Oliver, and John Moore, 1990, Property rights and the nature of the firm, *Journal of Political Economy* 98, 1119–1158.

Henderson, Rebecca, and Iain Cockburn, 1996, Scale, scope, and spillovers: the determinants of research productivity in drug discovery, *Rand Journal of Economies* 27, 32–59.

Hoberg, Gerard, and Craig Lewis, 2017, Do fraudulent firms produce abnormal disclosure?, *Journal of Corporate Finance* 43, 58–85.

Hoberg, Gerard, and S Katie Moon, 2017, Offshore activities and financial vs operational hedging, *Journal of Financial Economics* 125, 217–244.

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

——— , 2010b, Real and financial industry booms and busts, *Journal of Finance* 65, 45–86.

——— , 2016, Text-based network industry classifications and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

——— , 2018, Industry choice and product language, 64, 3735–3755.

——— , and Nagpurnanand Prabhala, 2014, Product market threats, payout, and financial flexibility, *Journal of Finance* 69, 293–314.

Hund, John, Donald Monk, and Sheri Tice, 2020, A manufactured diversification discount, *Critical Finance Review forthcoming.*

Kamepalli, Sai Krishna, Raghuram Rajan, and Luigi Zingales, 2020, Kill zone, .

Kim, Hyunseob, and Howard Kung, 2017, The asset redeployability channel: How uncertainty affects corporate investment, *Journal of Financial Economics* 30, 245–280.

Kleibergen, Frank, and Richard Paap, 2006, Generalized reduced rank tests using the singular value decomposition, *Journal of Econometrics* 133, 97–126.

Kwon, Spencer Yongwook, Yueran Ma, and Kaspar Zimmermann, 2021, 100 years of rising corporate concentration, *Available at SSRN 3936799.*

Lang, Larry, and Rene Stulz, 1994, Tobin's q, corporate diversification, and firm performance, *Journal of Political Economy* 102, 1248–1280.

Li, Feng, Russell Lundholm, and Michael Minnis, 2013, A measure of competition based on 10-k filings, *Journal of Accounting Research* 51, 399–436.

Maksimovic, Vojislav, and Gordon Phillips, 2002, Do conglomerate firms allocate resources inefficiently across industries? theory and evidence, *Journal of Finance* 57, 721–767.

Matvos, Gregor, Amit Seru, and Rui Silva, 2018, Financial market frictions and diversification, *Journal of Financial Economics* 127, 21–50.

Panzar, J., and R. Willig, 1977, Economies of scale in multi-output production, *Quarterly Journal of Economics* 91, 481–93.

——— , 1981, Economies of scope, *American Economic Review* 71, 268–272.

Peters, Ryan, and Luke Taylor, 2017, Intangible capital and the investment-q relation, *Journal of Financial Economics* 122, 251–272.

Philippon, Thomas, and German Gutierrez, 2017, Declining competition and investment in the u.s., *working paper.*

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.

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

Sargan, John, 1958, The estimation of economic relationships using instrumental variables, *Econometrica* 26, 393–415.

Song, Shiwon, 2020, The informational value of segment data disaggregated by underlying industry: Evidence from the textual features of business descriptions, *working paper*.

Teece, David J., 1980, Economies of scope and the scope of the enterprise, *Journal of Economic Behavior and Organization* 1, 223–247.

Villalonga, Belen, 2004, Does diversification cause the diversification discount, *Financial Management* 33, 5–27.

Williamson, Oliver E, 1975, Markets and hierarchies, *New York* 2630.

# Appendix A. Variable definitions

Table A1: Variable definitions Table A1

| Variable | Definition | Source |
|---|---|---|
| FIC-Scope | The number of TNIC (see Hoberg and Phillips 2016) industries (using the FIC-300 classification based on 1997 industry clusters) that each firm's 10-K product description is similar to. The classification from firm to segments is based on a 2% granularity, and firm-segments similarities are deemed to be pairs for the 2% highest textual similarities between each firm and the text describing the 300 FIC industries. | |
| NAICS-Scope | This is computed in a similar way to the FIC-scope variable. The NAICS scope is based on the text describing NAICS industries (using the highly detailed 963 page 2017 NAICS manual) instead of TNIC FIC industries. The classification from firm to segments is based on a 2% granularity, and firm-segment similarities are deemed to be pairs for the 2% highest textual similarities between each firm and the text describing the 311 4-digit NAICS industries. | |
| Product Breadth | Using metaHeuristica queries, we count the number of paragraphs that mention "product lines" or "product categories". These phrases indicate high levels of product breadth. Product Breadth is this number of paragraphs scaled by the total number of paragraphs in the 10-K. | |
| Prod/Svc Breadth | This variable is computed analogous to Product Breadth, but we also count paragraphs that mention either "service lines" or "service categories". | |
| Product Breadth Detail | Same as Product Breadth, as described above, except that we only count paragraphs that additionally mention a specific clarifying term in the following list: {breadth, broad, broader, wide, multiple, numerous, diverse, categories, divisions} | |
| Prod/Svc Breadth Detail | Same as Product/Svc Breadth, as described above, except that we only count paragraphs that additionally mention a specific clarifying term in the following list: {breadth, broad, broader, wide, multiple, numerous, diverse, categories, divisions} | |
| Sectoral Redeployment Potential | This is an instrument indicating a shift in the incentives for firms to increase scope. The data draws from information in the Bureau of Economic Analysis, Compustat, and the research paper Kim and Kung (2017). This variable is the average cosine similarity between the asset utilization vector of the focal firm's NAICS industry and that of the NAICS industry of the focal firm's TNIC-2 peers that are not also TNIC-3 peers. This latter step ensures that the measure is based on product market peers that are close but a bit more distant in product space than are near peers, which further increases the extent of exogenous content in the measure. When this value is high, it indicates that expansion in scope for the focal firm is likely to have low cost regarding expansion into neighboring product markets in the given year. | |
| Sectoral Opportunity Set Potential | This variable is similar to the above except that it is based on product offerings rather than the inputs to production (asset vectors). This variable is computed as the HHI, or concentration ratio, of companies that are in the focal firm's TNIC-2 industry but not in the most proximate TNIC-3 industry. The HHI calculation is based on the NAICS codes of the firms in these near but slightly more distant product markets. When this value is high, it indicates high growth opportunities to scope expansion for the focal firm in the given year. | |
| Logassets | Natural logarithm of total assets of the firm Compustat | |
| Log Age | Natural logarithm of one plus the current year of observation minus the first year the firm appears in the Compustat database Compustat | |
| Valuation Ratio | This ratio is computed as the market value of the firm (book assets minus book equity plus market equity), all divided by book assets. Market equity is Compustat shares outstanding times the share price at the end of the fiscal year PRCC. Book equity is shareholders equity (Compustat SEQ), plus TXDITC minus preferred stock (PSTKRV, and if missing, then PSTKL, and if missing then UPSTK). Shareholders equity iS SEQ, but if missing, is Compustat CEQ plus UPSTK, and if missing, is assets less long term assets. | |
| 10K Size | Natural logarithm of one plus the total number of paragraphs in the focal firm's 10-K report. | |
| TNIC HHI | The concentration ratio based on TNIC industries as computed in Hoberg and Phillips (2016). | |
| NAICS HHI | The concentration ratio based on NAICS industries as computed in Grullon, Michaely, and Larkin (2019). | |

| Variable | Definition |
|---|---|
| Acquirer Dummy | A dummy equal to one if the given firm had an acquisition become effective in the current year according to the SDC Platinum database. |
| Target Dummy | A dummy equal to one if the given firm had a sale of assets or a merger become effective in the current year according to the SDC Platinum database. |
| R&D/Assets | Compustat XRD divided by total assets AT, winsorized at the 1/99% level. This variable is set to zero if XRD is missing. |
| CAPX/Assets | Compustat CAPX divided by total assets AT, winsorized at the 1/99% level. This variable is set to zero if it is missing. |
| Sales Growth | Natural logarithm of total sales in the current year $t$ divided by total sales in the previous year $t-1$. |
| Asset Growth | Natural logarithm of total assets in the current year $t$ divided by total assets in the previous year $t-1$. |
| OI/Assets | Compustat OIBDP divided by total assets AT, winsorized at the 1/99% level. |
| Equity Issuance/Assets | Computed as Compustat (SSTK - PRSTKC) divided by total assets AT, winsorized at the 1/99% level. |
| Debt Issuance/Assets | Computed as Compustat DLTIS divided by total assets AT, winsorized at the 1/99% level. |
| Equity Repurchases/Assets | Computed as Compustat PRSTKC divided by total assets AT, winsorized at the 1/99% level. |
| Dividends/Assets | Computed as Compustat DVC divided by total assets AT, winsorized at the 1/99% level. |

Additional stop words dropped from the NAICS manual vocabularies: THE, AND, OF, COMPANIES, IN, SERVICES, CLASSIFIED, OR, INCLUDES, EXCLUDES, PRODUCTS, PRIMARILY, INCLUDING, NOT, PROVIDING, DIVERSIFIED, OTHER, THAT, TO, ENGAGED, GAS, MANAGEMENT, OPERATORS, RELATED, OWNERS, A, PRODUCERS, CONSUMER, ELSEWHERE, PROVIDERS, ALSO, FOR, COMPONENTS, DEVELOPMENT, PRODUCTION, AS, BUT, CENTERS, WITH, ARE, PRODUCING, LARGE, NON, OPERATING, OPERATIONS, USING, FROM, IT, MULTI, EITHER, EMPLOYMENT, THREE, UNDER, WHOSE, ACTIVITY, CORPORATE, DO, END, HELD, HIGH, MORE, WHICH, THEIR, WIDE, ACROSS, ASSETS, AT, OPERATE, INDUSTRY, MANUFACTURING, ESTABLISHMENTS, THIS, COMPRISES, EXCEPT, CROSS, REFERENCES, MERCHANT, SUCH, GROUP, EXAMPLES, ALL, PRODUCT, ILLUSTRATIVE, THESE, ACTIVITIES, MAY, NEW, PURCHASED, TYPE, MADE, SUPPORT, SECTOR, ONE, SUBSECTOR, WITHOUT, BASIS, INCLUDED, WORK, KNOWN, PROCESSING, PROVIDE, DIRECT, ORGANIZATIONS, PREPARATION, SELLING, GROWING, INTO, OTHERS, FOLLOWING, BUSINESS, COMBINATION, MISCELLANEOUS, SALE, INDUSTRIES, USE, MAKING, ORDER, PROGRAMS, THEY, BENEFICIATING, SIMILAR, STOCK, CONTRACT, BASED

## Table 1: Summary Statistics

Summary statistics are reported for our sample of 100,525 observations based on annual firm observations from 1988 to 2017. Our main variables of interest, TNIC-scope and NAICS-scope, are based on scoring each firm's Item 1 business description based on how similar it is to the product text of specific fixed industries. For FIC-scope, fixed industries are based on the TNIC FIC-300 industries (see Hoberg and Phillips 2016), and for NAICS-scope, it is based on 4-digit NAICS industries. All variables are described in detail in the variable list in Appendix A and in Section 3 of the paper.

| Variable | Mean | Std. Dev. | Minimum | Median | Maximum | # Obs |
|---|---|---|---|---|---|---|
| | | ***Panel A: Scope and Segment Variables*** | | | | |
| TNIC-Scope | 6.923 | 5.482 | 0.000 | 6.000 | 30.000 | 100,525 |
| NAICS-Scope | 6.269 | 7.519 | 0.000 | 4.000 | 47.000 | 100,525 |
| # Compsutat Segments | 1.452 | 0.862 | 1.000 | 1.000 | 11.000 | 100,525 |
| | | ***Panel B: Accounting Variables*** | | | | |
| R&D/Assets | 0.056 | 0.127 | 0.000 | 0.000 | 2.944 | 100,525 |
| CAPX/Assets | 0.058 | 0.070 | -0.000 | 0.037 | 0.725 | 100,525 |
| Acquisition Dummy | 0.287 | 0.452 | 0.000 | 0.000 | 1.000 | 100,525 |
| Target Dummy | 0.126 | 0.332 | 0.000 | 0.000 | 1.000 | 100,525 |
| Valuation (M/B) | 1.758 | 2.156 | 0.064 | 1.204 | 203.313 | 99,938 |
| Sales Growth | 0.110 | 0.437 | -6.177 | 0.076 | 9.383 | 100,112 |
| Asset Growth | 0.077 | 0.356 | -4.294 | 0.050 | 5.529 | 100,494 |
| Equity Issuance | 0.051 | 0.152 | -0.002 | 0.004 | 2.895 | 100,525 |
| Debt Issuance | 0.103 | 0.207 | 0.000 | 0.002 | 1.999 | 100,525 |
| Dividends/Assets | 0.008 | 0.028 | 0.000 | 0.000 | 1.248 | 100,424 |
| Equity Repurchase | 0.016 | 0.042 | 0.000 | 0.000 | 0.518 | 93,049 |
| Log Assets | 5.446 | 2.122 | 0.694 | 5.327 | 13.590 | 100,525 |
| Log Age | 2.622 | 0.766 | 0.693 | 2.565 | 4.220 | 100,525 |

Table 2: Pearson Correlation Coefficients

Pearson Correlation Coefficients are reported for our sample of 100,525 observations based on annual firm observations from 1988 to 2017. Our main variables of interest, TNIC-scope and NAICS-scope, are based on scoring each firm's Item 1 business description based on how similar it is to the product text of specific fixed industries. For FIC-scope, fixed industries are based on the TNIC FIC-300 industries (see Hoberg and Phillips 2016), and for NAICS-scope, it is based on 4-digit NAICS industries. All variables are described in detail in the variable list in Appendix A and in Section 3 of the paper.

| Row | Variable | Fic-Scope | NAICS-Scope | # CS Segments | Log Assets | Log Age | R&D/ Assets | CAPX/ Assets | Acquisition Dummy | Target Dummy | Sales Growth |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | NAICS-Scope | 0.579 | | | | | | | | | |
| | # Compsutat Segments | 0.186 | 0.198 | | | | | | | | |
| | Log Assets | 0.269 | 0.273 | 0.312 | | | | | | | |
| | Log Age | 0.038 | -0.047 | 0.309 | 0.400 | | | | | | |
| | R&D/Assets | 0.010 | -0.039 | -0.155 | -0.250 | -0.147 | | | | | |
| | CAPX/assets | -0.016 | 0.045 | -0.043 | 0.026 | -0.097 | -0.092 | | | | |
| | Acquisition Dummy | 0.067 | 0.088 | 0.133 | 0.301 | 0.075 | -0.089 | -0.015 | | | |
| | Target Dummy | 0.054 | 0.057 | 0.167 | 0.248 | 0.168 | -0.056 | -0.006 | 0.156 | | |
| | Sales Growth | 0.030 | 0.055 | -0.032 | -0.013 | -0.166 | 0.032 | 0.099 | 0.104 | -0.065 | |
| | TNIC HHI | -0.133 | -0.299 | 0.091 | -0.182 | 0.171 | -0.149 | -0.124 | -0.042 | -0.011 | -0.071 |

## Table 3: Disney Scope vs Time

The table displays the LDA-based scope allocations of Disney in 1990 (Panel A), 2003 (Panel B), and 2017 (Panel C).

| Year | Topic | Amount | Word List |
|---|---|---|---|
| | | | Panel A: Disney Scope Allocations in 1990 |
| 1990 | 296 | 275.0 | series, production, warner, live, feature, studio, distribution, company, release, produced |
| 1990 | 180 | 155.2 | event, team, garden, super, champion, league, collectible, sporting, arena, concert |
| 1990 | 100 | 128.4 | programming, broadcast, network, cable, satellite, program, time, channel, broadcasting, household |
| 1990 | 222 | 127.3 | motion, picture, screen, movie, production, cinema, company, theater, creative, sound |
| 1990 | 290 | 43.2 | publishing, adult, company, publisher, toy, character, gift, imperial, english, preview |
| 1990 | 122 | 41.9 | music, audio, disc, theater, studio, content, videocassette, licensors, digital, consumer |
| | | | Panel B: Disney Scope Allocations in 2003 |
| 2003 | 100 | 745.3 | programming, broadcast, network, cable, satellite, program, time, channel, broadcasting, household |
| 2003 | 296 | 641.1 | series, production, warner, live, feature, studio, distribution, company, release, produced |
| 2003 | 180 | 448.8 | event, team, garden, super, champion, league, collectible, sporting, arena, concert |
| 2003 | 245 | 316.9 | radio, market, broadcasting, broadcast, ownership, company, rule, interest, communication, local |
| 2003 | 222 | 294.5 | motion, picture, screen, movie, production, cinema, company, theater, creative, sound |
| 2003 | 198 | 92.6 | cable, system, service, regulation, local, rate, programming, television, authority, ownership |
| 2003 | 127 | 88.3 | resort, vacation, grand, valley, lift, ownership, white, company, located, owner |
| 2003 | 167 | 86.3 | company, food, guest, menu, quality, concept, location, service, beverage, item |
| 2003 | 285 | 75.2 | florida, development, resident, county, residential, company, alaska, residence, miami, area |
| 2003 | 217 | 72.6 | group, room, reservation, eagle, brand, leisure, travel, occupancy, company, service |
| | | | Panel C: Disney Scope Allocations in 2017 |
| 2017 | 100 | 736.4 | programming, broadcast, network, cable, satellite, program, time, channel, broadcasting, household |
| 2017 | 296 | 620.8 | series, production, warner, live, feature, studio, distribution, company, release, produced |
| 2017 | 180 | 416.7 | event, team, garden, super, champion, league, collectible, sporting, arena, concert |
| 2017 | 245 | 237.4 | radio, market, broadcasting, broadcast, ownership, company, rule, interest, communication, local |
| 2017 | 222 | 193.4 | motion, picture, screen, movie, production, cinema, company, theater, creative, sound |
| 2017 | 127 | 138.6 | resort, vacation, grand, valley, lift, ownership, white, company, located, owner |
| 2017 | 148 | 117.5 | service, launch, satellite, programming, primestar, channel, system, echostar, company, directv |
| 2017 | 190 | 115.3 | cost, ferc, order, service, pipeline, settlement, restructuring, interstate, transportation, regulatory |
| 2017 | 217 | 101.8 | group, room, reservation, eagle, brand, leisure, travel, occupancy, company, service |
| 2017 | 167 | 84.2 | company, food, guest, menu, quality, concept, location, service, beverage, item |
| 2017 | 290 | 62.2 | publishing, adult, company, publisher, toy, character, gift, imperial, english, preview |

Table 4: Scope Statistics vs Compustat Segment Counts

The table reports scale and scope statistics separately for firms based on how many operating segments the firm reports in the Compustat database. Our main variables of interest, TNIC-scope and NAICS-scope, are based on scoring each firm's Item 1 business description based on how similar it is to the product text of specific fixed industries. For FIC-scope, fixed industries are based on the TNIC FIC-300 industries (see Hoberg and Phillips 2016), and for NAICS-scope, it is based on 4-digit NAICS industries. Assets are from Compustat (variable AT).

| # Segments | FIC-SCope | NAICS-Scope | Assets | # Obs. |
|---|---|---|---|---|
| 1 segment | 6.41 | 5.56 | 1365 | 71,575 |
| 2 segments | 7.53 | 7.02 | 3255 | 17,939 |
| 3 segments | 8.63 | 8.57 | 6192 | 7,447 |
| 4 segments | 9.90 | 10.21 | 10084 | 2,353 |
| 5+ segments | 12.21 | 15.17 | 30776 | 1,211 |

Table 5: Scope Statistics vs Firm Size

The table reports scale and scope statistics separately for firms sorted into size quintiles. Sorts are annual and are based on Compustat assets (variable AT). Our main variables of interest, FIC-scope and NAICS-scope, are based on scoring each firm's Item 1 business description based on how similar it is to the product text of specific fixed industries. For FIC-scope, fixed industries are based on the TNIC FIC-300 industries (see Hoberg and Phillips 2016), and for NAICS-scope, it is based on 4-digit NAICS industries. Assets are from Compustat (variable AT). We report statistics for the full sample and separately for single segment firms only (see column headers).

| | | All Firms | | | | | Single Segment Firms Only | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Row | Firm Size Quintile | # Segments | FIC-Scope | NAICS-Scope | Assets | # Obs. | # Segments | FIC-Scope | NAICS-Scope | Assets | # Obs. |
| | Small Firms | 1.22 | 5.65 | 4.02 | 23 | 20,094 | 1 | 5.50 | 3.88 | 19 | 14,304 |
| | Quintile 1 | 1.26 | 6.16 | 5.04 | 102 | 20,110 | 1 | 6.07 | 4.72 | 74 | 14,321 |
| | Quintile 2 | 1.35 | 6.69 | 5.93 | 305 | 20,113 | 1 | 6.28 | 5.39 | 203 | 14,320 |
| | Quintile 3 | 1.50 | 7.50 | 7.18 | 936 | 20,111 | 1 | 6.76 | 6.18 | 604 | 14,321 |
| | Big Firms | 1.94 | 8.62 | 9.17 | 11728 | 20,097 | 1 | 7.41 | 7.64 | 5926 | 14,309 |

Table 6: Scope Statistics vs Industry-Pair-Relatedness

The table reports the distribution of the industries spanned by single firms (scope) across all industry pairs, sorted by how similar are the industries in the given pair regarding horizontal relatedness using TNIC similarities (Panel A) or vertical relatedness using vertical TNIC (VTNIC) relatedness (Panel B). We explain the methodology for panel A based on horizontal relatedness but note that the methodology for Panel B is exactly parallel but uses pairwise vertical relatedness scores from Fresard, Hoberg, and Phillips (2020) instead of horizontal relatedness scores from Hoberg and Phillips (2016). For each pair of FIC-300 industries in each year, we first tabulate the number of firms that operate in both industries in the pair based on the FIC-scope variable's construction. A firm is thus designated as operating in both industries if the given firm's business description is highly similar to the text of both industries in the pair. The result is a panel database of industry-pair-years indicating the number of firms operating in each pair. We then sort industry pairs into deciles based on the average horizontal TNIC similarity score of all firms in the first industry relative to those in the second. Industries that score highly are spatially close in the horizontal sense in the TNIC space (Panel B is similar but is based on vertical relatedness). Finally, we sum the firm-operating-pairs in each decile and report the fraction of operating pairs in each decile. We report this fraction for all firms, only for single segment firms and only for multi-segment firms. Finally, we report the average TNIC distance of the industry pairs in each decile and the number of industry pairs in each group in the final columns.

| Industry-Pair Similarity Decile | Fraction Scope Pairs (All Firms) | Fraction Scope Pairs (Single-Seg) | Fraction Scope Pairs (Multi-seg) | Average TNIC-pair Similarity | # Obs. |
|---|---|---|---|---|---|
| | | Panel A: Horizontal Relatedness | | | |
| Least Similar | 0.037 | 0.039 | 0.031 | 0.001 | 1,157,496 |
| Decile 2 | 0.044 | 0.047 | 0.036 | 0.002 | 1,157,739 |
| Decile 3 | 0.046 | 0.048 | 0.040 | 0.003 | 1,157,830 |
| Decile 4 | 0.053 | 0.055 | 0.047 | 0.005 | 1,157,492 |
| Decile 5 | 0.055 | 0.055 | 0.056 | 0.006 | 1,157,013 |
| Decile 6 | 0.069 | 0.069 | 0.068 | 0.008 | 1,158,488 |
| Decile 7 | 0.084 | 0.084 | 0.086 | 0.011 | 1,157,162 |
| Decile 8 | 0.096 | 0.091 | 0.107 | 0.014 | 1,157,686 |
| Decile 9 | 0.135 | 0.131 | 0.145 | 0.021 | 1,157,433 |
| Most Similar | 0.381 | 0.381 | 0.385 | 0.047 | 1,157,813 |
| | | Panel B: Vertical Relatedness | | | |
| Least Similar | 0.082 | 0.086 | 0.067 | 0.002 | 1,157,796 |
| Decile 2 | 0.071 | 0.075 | 0.056 | 0.003 | 1,157,737 |
| Decile 3 | 0.055 | 0.055 | 0.058 | 0.004 | 1,157,524 |
| Decile 4 | 0.061 | 0.062 | 0.062 | 0.004 | 1,157,942 |
| Decile 5 | 0.072 | 0.071 | 0.079 | 0.005 | 1,157,118 |
| Decile 6 | 0.079 | 0.079 | 0.085 | 0.006 | 1,157,482 |
| Decile 7 | 0.084 | 0.083 | 0.090 | 0.006 | 1,157,327 |
| Decile 8 | 0.107 | 0.107 | 0.110 | 0.008 | 1,157,779 |
| Decile 9 | 0.127 | 0.124 | 0.131 | 0.009 | 1,157,775 |
| Most Similar | 0.263 | 0.257 | 0.262 | 0.015 | 1,157,565 |

## Table 7: Risk Profile Regressions

The table reports the results of OLS regressions where measures of risk are regressed on measures of scope decomposed into components that are spatially near versus far from the focal firm. The first dependent variable is stock market risk used in Panels A and C, which is computed as the standard deviation of daily stock returns in year $t$. The second is cashflow volatility used in Panels B and D, which is computed as the standard deviation of a firm's quarterly operating income scaled by assets, computed over the 8 quarters of year $t$ and $t+1$. We express both as percentages for ease of interpretation. All RHS variables are computed using data from year $t-1$. We report results for three tercile-based subsamples within each panel and as noted in the first column. The subsamples are constructed in each year by sorting firms based on the average product market distance of the FIC industry segments they are assigned to when constructing our FIC-scope variable (firms assigned to just one FIC segment have a distance of zero). Firms in the close-scope tercile subsample are operating across industries that are highly related, whereas firms in the far-scope tercile are operating across industries that are more distant in the product space. Our main variable of interest, FIC-scope, is based on scoring each firm's Item 1 business description based on how similar it is to the product text of specific fixed industries. The fixed industries are based on the TNIC FIC-300 industries (see Hoberg and Phillips 2016). All regressions include year fixed effects, and controls for firm size, and firm age. The regressions in Panels C and D additionally include firm fixed effects. $t$-statistics for both panels are clustered by firm and shown in parentheses.

| Row | Dependent Variable | FIC-Scope | Log Assets | Log Age | ARSQ | # Obs |
|---|---|---|---|---|---|---|
| | | ***Panel A: Dependent Variable is Stock Volatility (OLS with Year Fixed Effects)*** | | | | |
| (1) | Close-Scope Tercile | 0.018 | –0.546 | –0.519 | 0.36 | 34,781 |
| | | (2.970) | (–42.940) | (–16.930) | | |
| (2) | Medium-Scope Tercile | 0.014 | –0.507 | –0.351 | 0.39 | 27,413 |
| | | (4.220) | (–42.970) | (–13.340) | | |
| (3) | Far-Scope Tercile | 0.003 | –0.550 | –0.381 | 0.37 | 29,627 |
| | | (0.910) | (–42.280) | (–13.100) | | |
| | | ***Panel B: Dependent Variable is Cashflow Volatility (OLS with Year Fixed Effects)*** | | | | |
| (4) | Close-Scope Tercile | 0.039 | –0.465 | –0.277 | 0.17 | 33,781 |
| | | (4.820) | (–30.090) | (–8.440) | | |
| (5) | Medium-Scope Tercile | 0.025 | –0.505 | –0.233 | 0.19 | 26,798 |
| | | (5.650) | (–29.560) | (–6.910) | | |
| (6) | Far-Scope Tercile | 0.019 | –0.484 | –0.204 | 0.18 | 28,846 |
| | | (4.550) | (–28.820) | (–6.720) | | |
| | | ***Panel C: Dependent Variable is Stock Volatility (OLS with Firm and Year Fixed Effects)*** | | | | |
| (7) | Close-Scope Tercile | 0.024 | –0.264 | –0.306 | 0.66 | 34,781 |
| | | (3.040) | (–8.010) | (–3.500) | | |
| (8) | Medium-Scope Tercile | 0.008 | –0.251 | –0.467 | 0.67 | 27,413 |
| | | (1.660) | (–7.560) | (–5.410) | | |
| (9) | Far-Scope Tercile | 0.007 | –0.230 | –0.285 | 0.68 | 29,627 |
| | | (1.580) | (–6.450) | (–3.430) | | |
| | | ***Panel D: Dependent Variable is Cashflow Volatility (OLS with Firm and Year Fixed Effects)*** | | | | |
| (10) | Close-Scope Tercile | –0.005 | –0.322 | –0.097 | 0.60 | 33,781 |
| | | (–0.610) | (–8.320) | (–1.080) | | |
| (11) | Medium-Scope Tercile | 0.012 | –0.364 | –0.169 | 0.63 | 26,798 |
| | | (2.190) | (–7.810) | (–1.850) | | |
| (12) | Far-Scope Tercile | 0.007 | –0.400 | –0.059 | 0.65 | 28,846 |
| | | (1.490) | (–8.540) | (–0.680) | | |

## Table 8: High Product Breadth Validation Regressions

The table reports validation regressions in which the dependent variable is a direct text-based measure of companies indicating that their products are broad. We consider four query-based measures obtained using the metaHeuristica software platform, based on product and service breadth. "Product Breadth" is the number of 10-K paragraphs containing the phrases {product lines, product categories}. "Prod/Svc Breadth" is analogously defined based on the search phrases {product lines, product categories, service lines, service categories}. "Product Breadth Detail" runs the same query as "Product Breadth" but is more stringent and additionally requires that the paragraph include one word from the following list: {breadth, broad, broader, wide, multiple, numerous, diverse, categories, divisions}. "Prod/Svc Breadth Detail" is a parallel more stringent version of the baseline "Prod/Svc Breadth" query. All four variables are scaled by the number of paragraphs in the 10-K overall. All regressions include firm and year fixed effects, and controls for firm size, age, 10-K size, M/B, and the TNIC HHI. Results are robust to dropping any of the controls. All regressions include firm and year fixed effects, coefficients are multiplied by 100 for ease of viewing, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | FIC-Scope | NAICS-Scope | # Segments | Log Assets | Log Age | Log 10K Size | M/B | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|---|---|---|
| (1) | Product Breadth | | | | 1.650 | 0.400 | –8.422 | –0.209 | –2.119 | 72,280 |
| | | | | | (3.640) | (0.260) | (–7.300) | (–2.520) | (–1.760) | |
| (2) | Prod/Svc Breadth | | | | 1.774 | 0.642 | –8.869 | –0.209 | –2.063 | 72,280 |
| | | | | | (3.840) | (0.410) | (–7.470) | (–2.490) | (–1.680) | |
| (3) | Prod Breadth Detail | | | | 0.763 | –1.124 | –3.327 | 0.016 | –1.558 | 72,280 |
| | | | | | (3.820) | (–1.700) | (–7.270) | (0.500) | (–3.260) | |
| (4) | Prod/Svc Breadth Detail | | | | 0.791 | –0.917 | –3.459 | 0.017 | –1.497 | 72,280 |
| | | | | | (3.920) | (–1.370) | (–7.290) | (0.540) | (–3.070) | |
| (5) | Product Breadth | | | 1.381 | 1.513 | 0.140 | –8.501 | –0.205 | –2.124 | 72,280 |
| | | | | (2.800) | (3.320) | (0.090) | (–7.370) | (–2.470) | (–1.760) | |
| (6) | Prod/Svc Breadth | | | 1.299 | 1.646 | 0.397 | –8.943 | –0.205 | –2.068 | 72,280 |
| | | | | (2.570) | (3.550) | (0.260) | (–7.530) | (–2.440) | (–1.680) | |
| (7) | Prod Breadth Detail | | | 0.430 | 0.720 | –1.205 | –3.352 | 0.017 | –1.559 | 72,280 |
| | | | | (1.920) | (3.590) | (–1.820) | (–7.310) | (0.540) | (–3.260) | |
| (8) | Prod/Svc Breadth Detail | | | 0.415 | 0.750 | –0.996 | –3.483 | 0.019 | –1.498 | 72,280 |
| | | | | (1.820) | (3.700) | (–1.480) | (–7.330) | (0.580) | (–3.070) | |
| (9) | Product Breadth | 0.640 | | 0.959 | 1.075 | 0.569 | –9.263 | –0.224 | –0.410 | 72,280 |
| | | (8.430) | | (1.960) | (2.390) | (0.380) | (–7.980) | (–2.740) | (–0.340) | |
| (10) | Prod/Svc Breadth | 0.663 | | 0.861 | 1.192 | 0.841 | –9.732 | –0.225 | –0.293 | 72,280 |
| | | (8.520) | | (1.720) | (2.600) | (0.550) | (–8.150) | (–2.720) | (–0.240) | |
| (11) | Prod Breadth Detail | 0.216 | | 0.288 | 0.572 | –1.060 | –3.609 | 0.011 | –0.981 | 72,280 |
| | | (6.360) | | (1.300) | (2.910) | (–1.610) | (–7.810) | (0.340) | (–2.010) | |
| (12) | Prod/Svc Breadth Detail | 0.228 | | 0.265 | 0.594 | –0.843 | –3.754 | 0.012 | –0.888 | 72,280 |
| | | (6.610) | | (1.170) | (2.990) | (–1.260) | (–7.840) | (0.380) | (–1.780) | |
| (13) | Product Breadth | | 0.309 | 1.063 | 1.253 | 0.391 | –9.075 | –0.210 | –0.700 | 72,280 |
| | | | (6.020) | (2.150) | (2.770) | (0.260) | (–7.780) | (–2.530) | (–0.580) | |
| (14) | Prod/Svc Breadth | | 0.327 | 0.962 | 1.371 | 0.662 | –9.550 | –0.210 | –0.562 | 72,280 |
| | | | (6.240) | (1.900) | (2.970) | (0.430) | (–7.940) | (–2.500) | (–0.450) | |
| (15) | Prod Breadth Detail | | 0.132 | 0.294 | 0.609 | –1.098 | –3.596 | 0.015 | –0.952 | 72,280 |
| | | | (5.290) | (1.310) | (3.090) | (–1.660) | (–7.760) | (0.480) | (–1.960) | |
| (16) | Prod/Svc Breadth Detail | | 0.137 | 0.275 | 0.635 | –0.884 | –3.737 | 0.017 | –0.867 | 72,280 |
| | | | (5.360) | (1.200) | (3.190) | (–1.320) | (–7.770) | (0.520) | (–1.750) | |

# Table 9: First-Stage Regressions

The table reports the results of first-stage regressions where measures of scope (FIC-scope and NAICS-scope) are regressed on our two instruments in addition to all controls and fixed effects. Our first instrument is "Sectoral Redeployment Potential" which is a product market spatially localized version of the asset redeployability measure in Kim and Kung (2017). In particular, we use the BEA capital flows table and represent the assets of each 4-digit NAICS industry as a vector. For each focal's local product market, we compute the average redeployability between the focal firm's nearest peer NAICS industries and the focal firm's distant peer NAICS industries. Intuitively, when the assets of near peers are easily redeployed to the market of more distant peers, the focal firm faces a low cost to expanding its market outward in space (lower cost of increasing scope). The second instrument "Sectoral Opportunity Set Potential" is simply the concentraiton ratio of 4-digit NAICS industries that the moderately distant peers in the focal firm's product market reside in. When this concentration ratio is high, it indicates that the focal firm's peers all tend to operate in the same market, which in turn implies there are few opportunities for scope expansion by the focal firm (as there are fewer related product markets that are spatially close). For success in the first stage, we predict that the former measure will be positively related to our scope variables, and the second will be negatively related. All regressions include firm and year fixed effects, and controls for firm size, age, 10-K size, M/B, and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Sectoral Redeployment Potential | Sectoral Opportunity Set Potential | Log Assets | Log Age | # Obs |
|---|---|---|---|---|---|---|
| (1) | FIC-Scope | 1.251 | 6.941 | 0.882 | –0.830 | 99,506 |
| | | (4.330) | (11.910) | (19.290) | (–6.030) | |
| (2) | NAICS-Scope | 1.074 | 12.772 | 1.279 | –0.976 | 99,506 |
| | | (2.380) | (14.680) | (17.730) | (–5.060) | |
| (3) | # Segments | 0.019 | 0.293 | 0.107 | 0.161 | 99,506 |
| | | (0.300) | (2.900) | (12.340) | (6.260) | |

## Table 10: Investment Regressions

The table reports the second stage results of 2-stage instrumental variable regressions where the dependent variable is a firm investment policy such as acquisitions, divestitures (target of an acquisition), R&D/assets or CAPX/assets. Our instrumented variable of interest is a measure of scope (FIC-Scope or NAICS-Scope) as indicated in the panel headers. The first-stage regressions are displayed in Table 9 and include two instruments for scope (explained in detail in Table 9). The first is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. We also include controls for size, age, and in Panel C, we additionally include controls for market to book and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Scope Variable | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|
| | | | ***Panel A: FIC-Scope is Scope Variable*** | | | | |
| (1) | Acquirer Dummy | 0.026 | –0.006 | –0.029 | | | 98,196 |
| | | (4.660) | (–0.920) | (–2.710) | | | |
| (2) | Target Dummy | –0.009 | 0.038 | 0.039 | | | 98,196 |
| | | (–2.270) | (8.570) | (5.390) | | | |
| (3) | R&D/Assets | 0.003 | –0.016 | 0.009 | | | 98,196 |
| | | (3.900) | (–12.690) | (4.370) | | | |
| (4) | CAPX/Assets | 0.000 | –0.002 | –0.014 | | | 98,196 |
| | | (0.450) | (–2.630) | (–9.820) | | | |
| (5) | Vertical Integration | 0.003 | –0.002 | 0.003 | | | 98,151 |
| | | (13.680) | (–8.350) | (6.950) | | | |
| | | | ***Panel B: NAICS-Scope is Scope Variable*** | | | | |
| (6) | Acquirer Dummy | 0.016 | –0.003 | –0.035 | | | 98,196 |
| | | (4.860) | (–0.580) | (–3.540) | | | |
| (7) | Target Dummy | –0.005 | 0.036 | 0.042 | | | 98,196 |
| | | (–2.060) | (9.060) | (6.240) | | | |
| (8) | R&D/Assets | 0.001 | –0.015 | 0.008 | | | 98,196 |
| | | (3.570) | (–12.860) | (3.980) | | | |
| (9) | CAPX/Assets | 0.000 | –0.002 | –0.014 | | | 98,196 |
| | | (0.570) | (–2.870) | (–10.190) | | | |
| (10) | Vertical Integration | 0.002 | –0.001 | 0.002 | | | 98,151 |
| | | (14.740) | (–8.000) | (6.470) | | | |
| | | | ***Panel C: FIC-Scope is Scope Variable (extra controls added)*** | | | | |
| (11) | Acquirer Dummy | 0.036 | –0.007 | –0.017 | 0.009 | 0.107 | 97,624 |
| | | (4.360) | (–0.840) | (–1.500) | (9.180) | (4.560) | |
| (12) | Target Dummy | –0.012 | 0.039 | 0.036 | –0.002 | –0.028 | 97,624 |
| | | (–2.130) | (7.120) | (4.630) | (–3.400) | (–1.660) | |
| (13) | R&D/Assets | 0.004 | –0.017 | 0.009 | –0.001 | 0.002 | 97,624 |
| | | (3.190) | (–11.910) | (4.190) | (–2.100) | (0.610) | |
| (14) | CAPX/Assets | 0.000 | –0.001 | –0.013 | 0.002 | –0.003 | 97,624 |
| | | (–0.480) | (–0.790) | (–8.720) | (9.090) | (–1.240) | |
| (15) | Vertical Integration | 0.004 | –0.002 | 0.003 | 0.000 | 0.008 | 97,579 |
| | | (10.610) | (–7.530) | (5.960) | (–4.320) | (7.720) | |

# Table 11: Investment Regressions (by Fama-French-5 Major Sectors)

The table reports the second stage results of 2-stage instrumental variable regressions where the dependent variable is a firm investment policy such as acquisitions, divestitures (target of an acquisition), R&D/assets or CAPX/assets. These regressions use the same specification as in Table 10 except we now run them using 5 industry subsamples based on the Fama-French-5 industry groupings (see Panel headers). Our instrumented variable of interest is a measure of scope (FIC-Scope). The first-stage regressions are displayed in Table 9 and include two instruments for scope (explained in detail in Table 9). The first is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Scope Variable | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|
| | | | ***Panel A: FIC-Scope (Tech-Industry Subsample)*** | | | | |
| (1) | Acquirer Dummy | 0.000 | 0.045 | –0.065 | 0.00 | 26,252 | |
| | | (–0.020) | (3.110) | (–2.220) | | | |
| (2) | Target Dummy | 0.003 | 0.021 | 0.073 | 0.00 | 26,252 | |
| | | (0.180) | (2.120) | (3.660) | | | |
| (3) | R&D/Assets | 0.009 | –0.027 | 0.027 | 0.00 | 26,252 | |
| | | (2.250) | (–8.360) | (4.120) | | | |
| (4) | CAPX/Assets | 0.000 | 0.000 | –0.008 | 0.00 | 26,252 | |
| | | (–0.090) | (–0.330) | (–2.510) | | | |
| | | | ***Panel B: FIC-Scope (Manufacturing-Industry Subsample)*** | | | | |
| (5) | Acquirer Dummy | 0.034 | –0.059 | –0.042 | 0.00 | 20,519 | |
| | | (3.310) | (–3.490) | (–1.720) | | | |
| (6) | Target Dummy | –0.008 | 0.066 | 0.038 | 0.00 | 20,519 | |
| | | (–1.090) | (5.430) | (2.230) | | | |
| (7) | R&D/Assets | 0.000 | –0.002 | 0.005 | 0.00 | 20,519 | |
| | | (–0.510) | (–2.030) | (2.420) | | | |
| (8) | CAPX/Assets | 0.000 | –0.002 | –0.012 | 0.00 | 20,519 | |
| | | (0.050) | (–0.970) | (–3.420) | | | |
| | | | ***Panel C: FIC-Scope (Consumer-Industry Subsample)*** | | | | |
| (9) | Acquirer Dummy | 0.032 | –0.020 | –0.028 | 0.00 | 21,432 | |
| | | (2.320) | (–1.260) | (–1.290) | | | |
| (10) | Target Dummy | –0.015 | 0.043 | 0.022 | 0.00 | 21,432 | |
| | | (–1.370) | (3.700) | (1.560) | | | |
| (11) | R&D/Assets | 0.003 | –0.004 | 0.001 | 0.00 | 21,432 | |
| | | (2.420) | (–2.920) | (0.370) | | | |
| (12) | CAPX/Assets | 0.002 | –0.007 | –0.013 | 0.00 | 21,432 | |
| | | (1.510) | (–4.100) | (–4.500) | | | |
| | | | ***Panel D: FIC-Scope (Health-Industry Subsample)*** | | | | |
| (13) | Acquirer Dummy | 0.015 | 0.012 | –0.030 | 0.00 | 12,528 | |
| | | (1.320) | (1.060) | (–0.990) | | | |
| (14) | Target Dummy | –0.032 | 0.043 | –0.026 | 0.00 | 12,528 | |
| | | (–3.020) | (3.750) | (–1.100) | | | |
| (15) | R&D/Assets | 0.013 | –0.045 | 0.016 | 0.00 | 12,528 | |
| | | (2.110) | (–6.710) | (1.070) | | | |
| (16) | CAPX/Assets | –0.001 | 0.001 | –0.010 | 0.00 | 12,528 | |
| | | (–0.460) | (0.580) | (–2.570) | | | |
| | | | ***Panel E: FIC-Scope (Misc-Industry Subsample)*** | | | | |
| (17) | Acquirer Dummy | 0.041 | –0.052 | –0.046 | 0.00 | 15,583 | |
| | | (2.750) | (–2.960) | (–1.540) | | | |
| (18) | Target Dummy | –0.001 | 0.027 | 0.068 | 0.00 | 15,583 | |
| | | (–0.160) | (2.430) | (4.240) | | | |
| (19) | R&D/Assets | 0.001 | –0.003 | 0.003 | 0.00 | 15,583 | |
| | | (1.920) | (–2.130) | (1.560) | | | |
| (20) | CAPX/Assets | 0.002 | –0.005 | –0.023 | 0.00 | 15,583 | |
| | | (1.310) | (–2.200) | (–5.250) | | | |

# Table 12: Outcomes Regressions

The table reports the second stage results of 2-stage instrumental variable regressions where the dependent variable is a firm outcome variable such as the market to book ratio (market value of firm divided by total assets), sales growth, asset growth or profitability. Our instrumented variable of interest is a measure of scope (FIC-Scope or NAICS-Scope) as indicated in the panel headers. The first-stage regressions are displayed in Table 9 and include two instruments for scope (explained in detail in Table 9). The first is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. We also include controls for size, age, and in Panel C, we additionally include controls for market to book and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Scope Variable | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|
| | | | ***Panel A: FIC-Scope is Scope Variable*** | | | | |
| (1) | Valuation (M/B) | 0.080 | –0.532 | –0.344 | | | 97,625 |
| | | (4.310) | (–17.410) | (–7.360) | | | |
| (2) | Sales Growth | 0.031 | –0.098 | –0.182 | | | 97,825 |
| | | (5.790) | (–14.990) | (–17.860) | | | |
| (3) | Asset Growth | 0.047 | –0.206 | –0.038 | | | 98,193 |
| | | (8.830) | (–30.220) | (–3.550) | | | |
| (4) | OI/Assets | –0.001 | 0.012 | 0.000 | | | 97,995 |
| | | (–0.540) | (3.510) | (0.010) | | | |
| | | | ***Panel B: NAICS-Scope is Scope Variable*** | | | | |
| (5) | Valuation (M/B) | 0.040 | –0.512 | –0.373 | | | 97,625 |
| | | (3.730) | (–17.880) | (–8.230) | | | |
| (6) | Sales Growth | 0.018 | –0.094 | –0.191 | | | 97,825 |
| | | (5.960) | (–15.990) | (–20.450) | | | |
| (7) | Asset Growth | 0.028 | –0.200 | –0.050 | | | 98,193 |
| | | (9.500) | (–33.210) | (–5.180) | | | |
| (8) | OI/Assets | –0.001 | 0.011 | 0.000 | | | 97,995 |
| | | (–0.490) | (3.650) | (0.100) | | | |
| | | | ***Panel C: FIC-Scope is Scope Variable (extra controls added)*** | | | | |
| (9) | Valuation (M/B) | 0.095 | –0.448 | –0.156 | 0.222 | 0.246 | 97,391 |
| | | (3.770) | (–14.290) | (–3.850) | (10.340) | (2.690) | |
| (10) | Sales Growth | 0.042 | –0.091 | –0.150 | 0.032 | 0.129 | 97,267 |
| | | (5.460) | (–11.480) | (–13.030) | (12.100) | (5.740) | |
| (11) | Asset Growth | 0.056 | –0.192 | 0.002 | 0.044 | 0.137 | 97,624 |
| | | (7.140) | (–23.500) | (0.130) | (16.340) | (6.100) | |
| (12) | OI/Assets | –0.002 | 0.014 | 0.004 | 0.005 | 0.000 | 97,430 |
| | | (–0.470) | (3.800) | (0.690) | (4.780) | (0.030) | |

# Table 13: Venture Capital Funding Similarity and Fluidity Regressions

The table reports the second stage results of 2-stage instrumental variable regressions where the dependent variable is a measure of early-stage startup financing or innovative activity in the given firm's product market. Both dependent variables are developed in Hoberg, Phillips, and Prabhala (2014). VC funding similarity is the cosine similarity of the given focal firm's 10-K product description to the average vocabulary used by all startups in the given year (where the startup vocabulary is obtained from Venture Expert business descriptions of all startups receiving their first round of financing in the given year). Product market fluidity is the average market-wide change (fluidity) in the use of the given firm's 10-K product description vocabulary by all other firms. A high value indicates a large amount of product innovation by competing firms in the focal firm's product markets. Our instrumented variable of interest is a measure of scope (FIC-Scope or NAICS-Scope) as indicated in the panel headers. The first-stage regressions are displayed in Table 9 and include two instruments for scope (explained in detail in Table 9). The first is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. We also include controls for size, age, and in Panel C, we additionally include controls for market to book and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Scope Variable | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|
| | ***Panel A: FIC-Scope is Scope Variable*** | | | | | | |
| (1) | VC Funding Similarity | 1.926 | –1.170 | 0.824 | | | 98,186 |
| | | (15.040) | (–8.190) | (3.020) | | | |
| (2) | Product Market Fluidity | 1.251 | –0.773 | 0.262 | | | 97,090 |
| | | (14.640) | (–8.180) | (1.470) | | | |
| | ***Panel B: NAICS-Scope is Scope Variable*** | | | | | | |
| (3) | VC Funding Similarity | 1.093 | –0.868 | 0.265 | | | 98,186 |
| | | (17.460) | (–7.920) | (1.240) | | | |
| (4) | Product Market Fluidity | 0.750 | –0.634 | –0.077 | | | 97,090 |
| | | (20.240) | (–9.220) | (–0.580) | | | |
| | ***Panel C: FIC-Scope is Scope Variable (extra controls added)*** | | | | | | |
| (5) | VC Funding Similarity | 2.575 | –1.588 | 1.115 | –0.034 | 5.515 | 97,614 |
| | | (11.030) | (–7.130) | (3.010) | (–1.410) | (8.190) | |
| (6) | Product Market Fluidity | 1.483 | –0.900 | 0.329 | –0.015 | 2.716 | 96,523 |
| | | (10.650) | (–6.770) | (1.520) | (–1.140) | (6.920) | |

## Table 14: Financing Regressions

The table reports the second stage results of 2-stage instrumental variable regressions where the dependent variable is a firm financing policy such as equity issuance, debt issuance, dividends, or equity repurchases. Our instrumented variable of interest is a measure of scope (FIC-Scope or NAICS-Scope) as indicated in the panel headers. The first-stage regressions are displayed in Table 9 and include two instruments for scope (explained in detail in Table 9). The first is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. We also include controls for size, age, and in Panel C, we additionally include controls for market to book and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Scope Variable | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|
| | | | ***Panel A: FIC-Scope is Scope Variable*** | | | | |
| (1) | Equity Issuance | 0.009 | –0.050 | –0.009 | | | 98,196 |
| | | (7.220) | (–25.060) | (–2.950) | | | |
| (2) | Debt Issuance | 0.002 | –0.011 | 0.010 | | | 98,196 |
| | | (0.760) | (–3.960) | (1.990) | | | |
| (3) | Dividends/Assets | –0.001 | 0.001 | 0.002 | | | 98,097 |
| | | (–2.870) | (1.900) | (2.440) | | | |
| (4) | Repurchases/Assets | –0.001 | 0.005 | 0.005 | | | 90,680 |
| | | (–1.460) | (7.430) | (5.070) | | | |
| | | | ***Panel B: NAICS-Scope is Scope Variable*** | | | | |
| (5) | Equity Issuance | 0.005 | –0.049 | –0.012 | | | 98,196 |
| | | (7.530) | (–26.340) | (–4.010) | | | |
| (6) | Debt Issuance | 0.001 | –0.011 | 0.010 | | | 98,196 |
| | | (0.970) | (–4.390) | (2.050) | | | |
| (7) | Dividends/Assets | –0.001 | 0.001 | 0.002 | | | 98,097 |
| | | (–2.710) | (1.520) | (3.300) | | | |
| (8) | Repurchases/Assets | 0.000 | 0.004 | 0.006 | | | 90,680 |
| | | (–1.290) | (7.860) | (5.590) | | | |
| | | | ***Panel C: FIC-Scope is Scope Variable (extra controls added)*** | | | | |
| (9) | Equity Issuance | 0.010 | –0.047 | 0.000 | 0.010 | 0.021 | 97,624 |
| | | (5.580) | (–21.180) | (–0.130) | (10.190) | (3.890) | |
| (10) | Debt Issuance | 0.001 | –0.010 | 0.010 | 0.001 | –0.005 | 97,624 |
| | | (0.160) | (–2.930) | (1.840) | (1.990) | (–0.510) | |
| (11) | Dividends/Assets | –0.002 | 0.001 | 0.002 | 0.000 | –0.004 | 97,528 |
| | | (–2.790) | (2.380) | (2.330) | (4.760) | (–1.990) | |
| (12) | Repurchases/Assets | –0.001 | 0.005 | 0.006 | 0.001 | –0.001 | 90,146 |
| | | (–1.350) | (6.960) | (5.240) | (5.780) | (–0.660) | |

## Table 15: Near versus Far Scope Tercile Subsamples

The table reports the second stage results of 2-stage instrumental variable regressions for various dependent variables as noted in the first column. These regressions use the same IV specification as in earlier tables such as Table 10, except we now run these regressions using 3 scope-distance tercile subsamples as indicated in the column headers for columns two to four. The subsamples are constructed in each year by sorting firms based on the average product market distance of the FIC industry segments they are assigned to when constructing our FIC-scope variable (firms assigned to just one FIC segment have a distance of zero). Firms in the close-scope tercile subsample are operating across industries that are highly related, whereas firms in the far-scope tercile are operating across industries that are more distant in the product space. Our instrumented variable of interest is a measure of scope (FIC-Scope), and we only report this coefficient for each regression for parsimony. Hence each coefficient below is this coefficient for a separate regression based on the subsample indicated in the column header and the dependent variable indicated in the first column. All regressions include firm and year fixed effects, controls for size and age (not reported) and *t*-statistics for the FIC-scope coefficient are clustered by firm and are shown in parentheses.

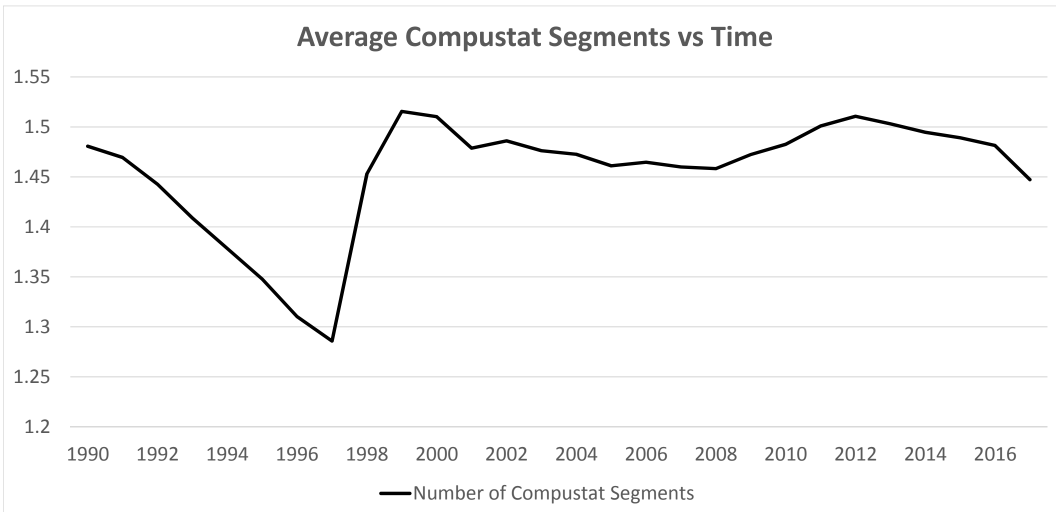
| Row | Dependent Variable | Near Scope Tercile Subsample | Middle Scope Tercile Subsample | Far Scope Tercile Subsample |
|---|---|---|---|---|
| (1) | Acquirer Dummy | 0.050 | 0.041 | 0.039 |
| | | (2.060) | (2.570) | (3.810) |
| (2) | R&D/Assets | 0.011 | 0.002 | 0.001 |
| | | (3.030) | (1.100) | (0.640) |
| (3) | CAPX/Assets | 0.002 | 0.002 | -0.001 |
| | | (0.970) | (1.060) | (-0.930) |
| (4) | Valuation (M/B) | 0.249 | 0.114 | -0.008 |
| | | (2.980) | (2.380) | (-0.290) |
| (5) | Sales Growth | 0.069 | 0.062 | 0.036 |
| | | (2.780) | (4.140) | (3.680) |
| (6) | Asset Growth | 0.131 | 0.070 | 0.051 |
| | | (5.340) | (4.690) | (5.220) |
| (7) | Equity Issuance | 0.028 | 0.013 | 0.006 |
| | | (4.730) | (4.200) | (2.840) |
| (8) | VC Funding Score | 3.163 | 1.743 | 1.179 |
| | | (6.950) | (7.910) | (8.280) |
| (9) | Prod Mkt Fluidity | 5.515 | 2.479 | 1.830 |
| | | (8.280) | (8.190) | (8.200) |

## Table 16: Firm Size Tercile Subsamples

The table reports the second stage results of 2-stage instrumental variable regressions for various dependent variables as noted in the first column. These regressions use the same IV specification as in earlier tables such as Table 10, except we now run these regressions using 3 firm size tercile subsamples as indicated in the column headers for columns two to four. The subsamples are constructed in each year by sorting firms based on past-year assets and hence we have large, mid-size and small firm subsamples. Our instrumented variable of interest is a measure of scope (FIC-Scope), and we only report this coefficient for each regression for parsimony. Hence each coefficient below is this coefficient for a separate regression based on the subsample indicated in the column header and the dependent variable indicated in the first column. All regressions include firm and year fixed effects, controls for size and age (not reported) and $t$-statistics for the FIC-scope coefficient are clustered by firm and are shown in parentheses.

| Row | Dependent Variable | Large-Firm Tercile Subsample | Mid-Size Tercile Subsample | Small-Firm Tercile Subsample |
|---|---|---|---|---|
| (1) | Acquirer Dummy | 0.015 | 0.044 | 0.042 |
| | | (1.950) | (3.460) | (3.420) |
| (2) | R&D/Assets | 0.000 | 0.003 | 0.005 |
| | | (-0.150) | (2.840) | (1.920) |
| (3) | CAPX/Assets | 0.000 | 0.002 | 0.000 |
| | | (1.010) | (1.690) | (-0.260) |
| (4) | Valuation (M/B) | -0.022 | 0.108 | 0.202 |
| | | (-1.560) | (3.640) | (3.280) |
| (5) | Sales Growth | 0.007 | 0.034 | 0.088 |
| | | (2.490) | (3.280) | (4.340) |
| (6) | Asset Growth | 0.012 | 0.063 | 0.104 |
| | | (3.760) | (5.110) | (5.500) |
| (7) | Equity Issuance | 0.000 | 0.007 | 0.022 |
| | | (1.200) | (3.900) | (4.530) |
| (8) | VC Funding Score | 0.988 | 1.403 | 1.730 |
| | | (9.800) | (7.050) | (7.920) |
| (9) | Prod Mkt Fluidity | 1.441 | 2.017 | 2.777 |
| | | (9.600) | (7.380) | (7.980) |

Figure 1: Measures of scope versus time. The upper figure plots the average number of Compustat segments per firm over time. The lower figure plots the average values of FIC-scope and NAICS-scope over our sample period. TNIC-scope and NAICS-scope are based on scoring each firm's Item 1 business description based on how similar it is to the product text of specific fixed industries. For FIC-scope, fixed industries are based on the TNIC FIC-300 industries (see Hoberg and Phillips 2016), and for NAICS-scope, it is based on 4-digit NAICS industries.



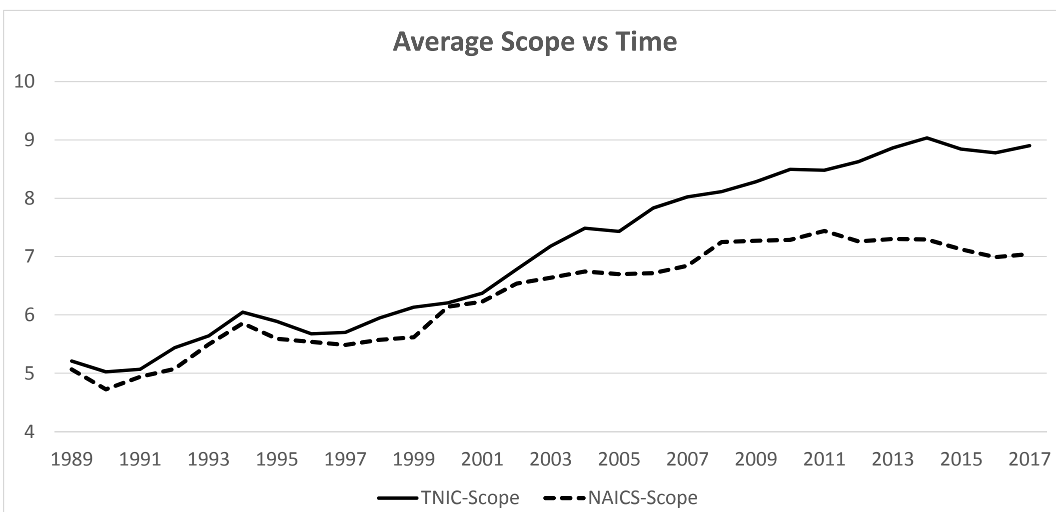
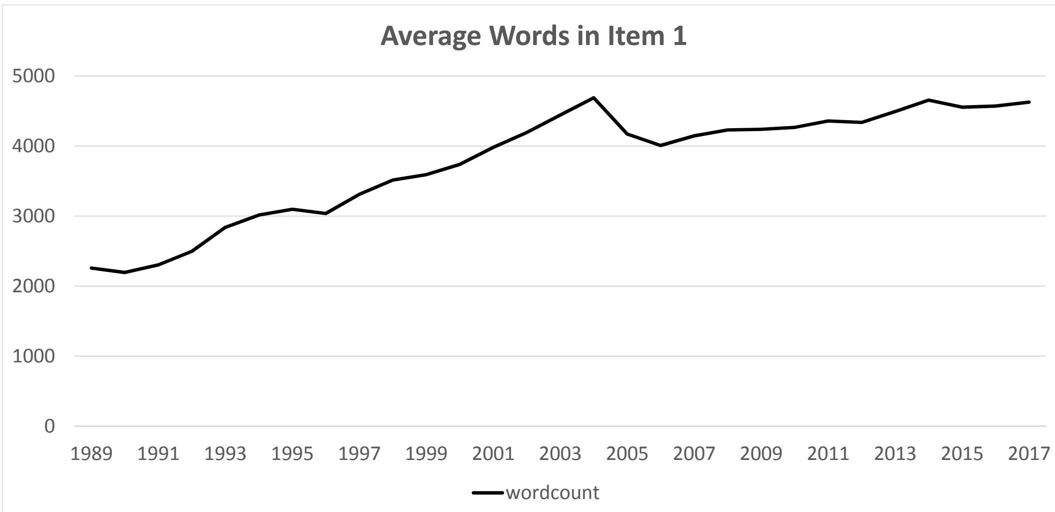



Figure 2: The upper figure reports the average number of words in the firm's 10-K Item 1 business description divided by the number of industries (FIC-based or NAICS-based scope) the firm likely operates in. The goal is to measure the average degree of product variety within industries over time. The lower figure displays the average size of the 10-K Item 1 over time.



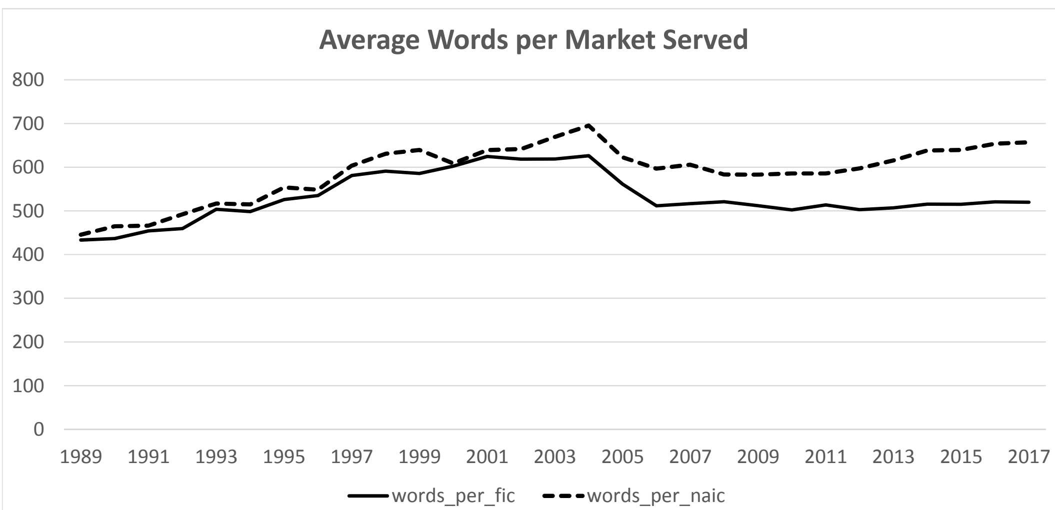



Figure 3: Firm size versus time. The figure displays firm size (measured as Compustat assets, both nominal and inflation adjusted) over time.

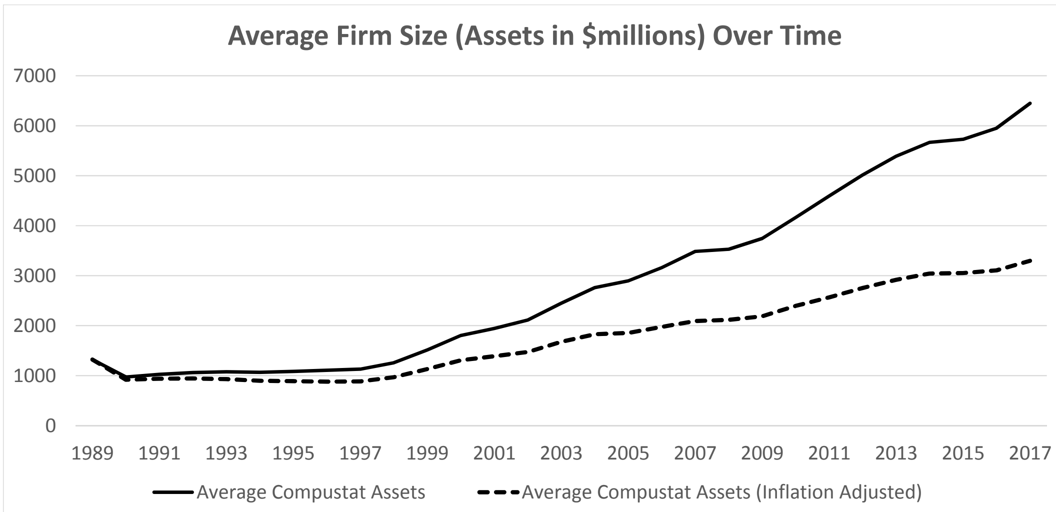
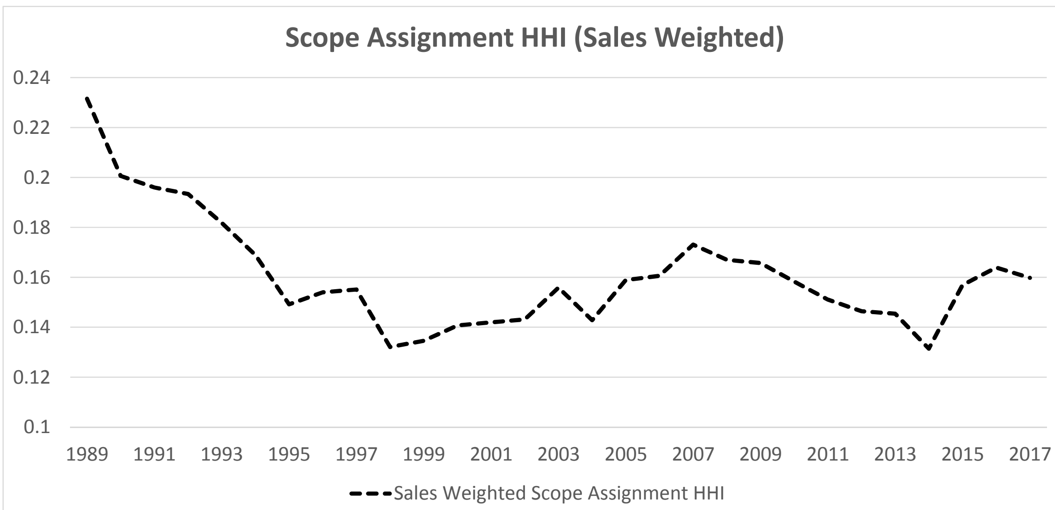


Figure 4: The FIC-Scope Implied HHI computes the HHI after allowing firms to have a presence in multiple industries, as is identified during the derivation of FIC-scope itself. Market shares are based on sales. Each firm's sales are allocated across the multiple sectors each firm is assigned to using similarity weights (similarity weights are defined as $Q_{i,j,t,FIC}$ in equation (1)). HHIs are then computed at the FIC industry level using these allocated sales where firms operate in multiple sectors. We then aggregate these HHIs back to the firm level by computing weighted averages over the sectors each firm operates in (again using weights $Q_{i,j,t,FIC}$). Note that our results are similar if we use equal weights instead. We then aggregate HHIs to the economy-wide annual level by computing a sales weighted average of the firm HHIs or an equal weighted average of the firm HHIs. The upper figure reports the sales-weighted average HHI over time and the lower figure reports the equal weighted average HHI over time.



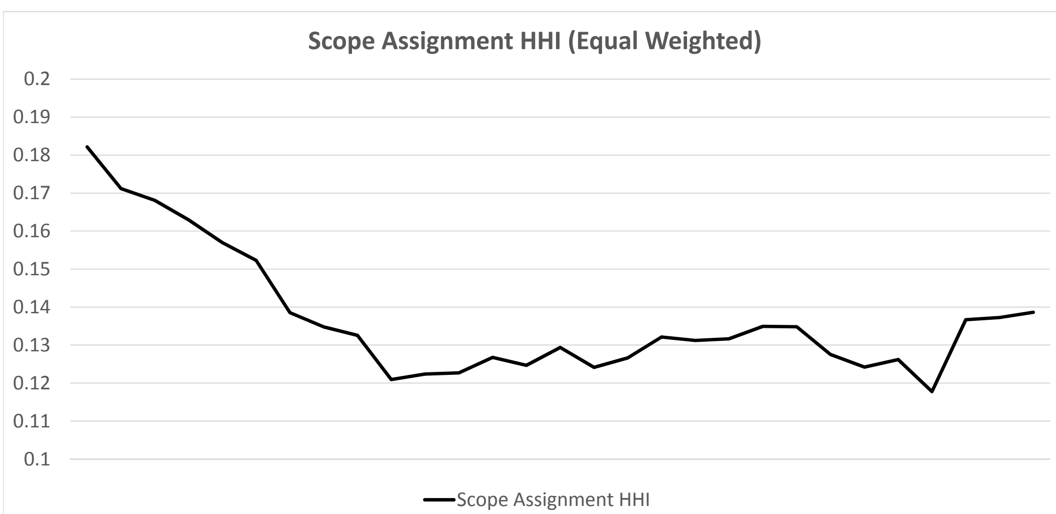
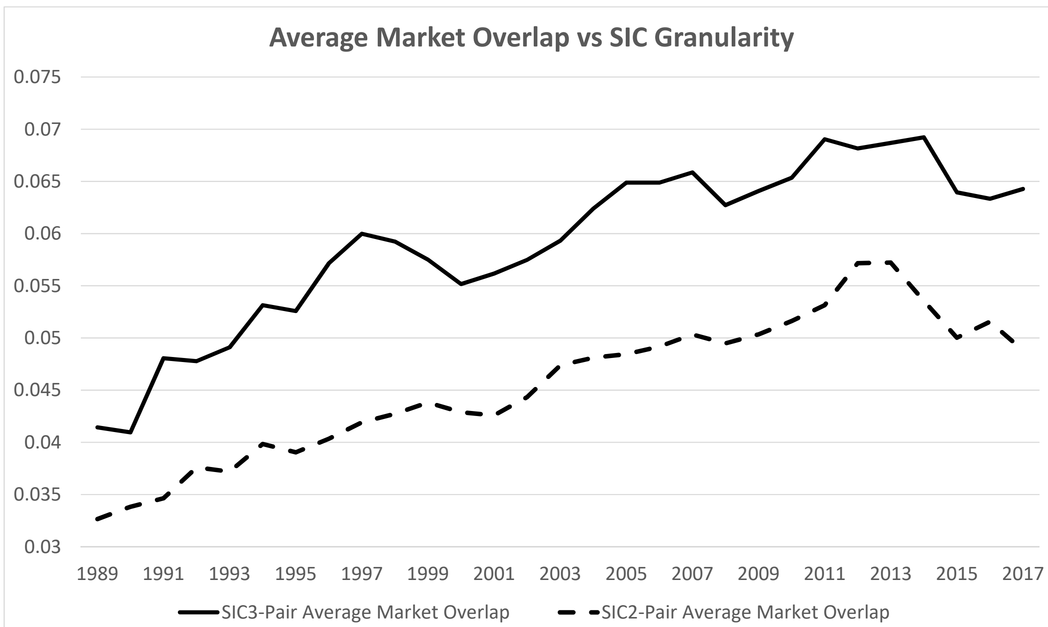
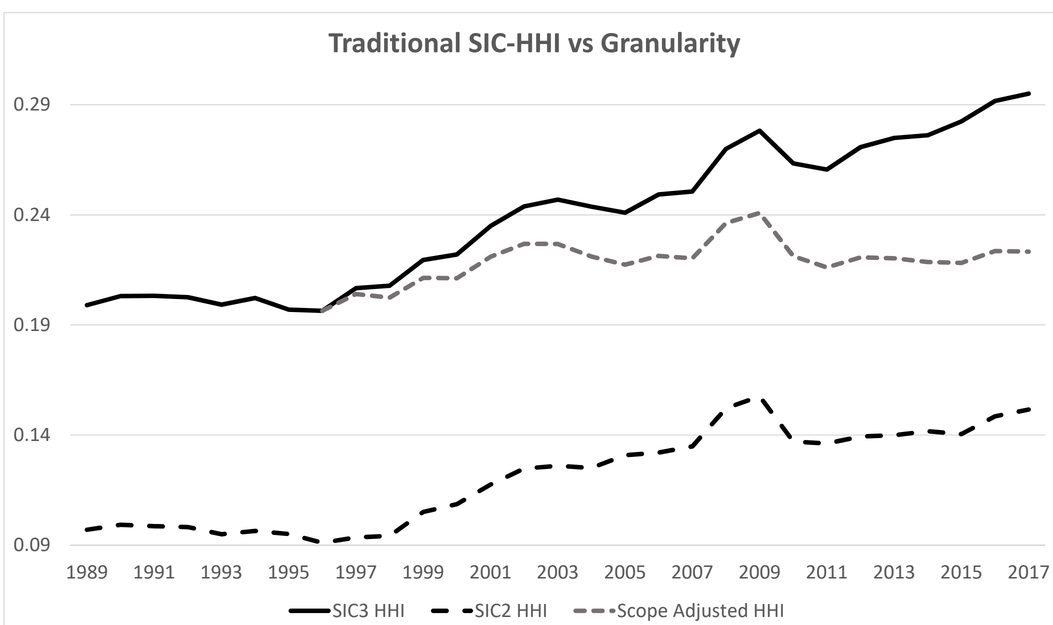



Figure 5: The upper figure reports the average market overlap of pairs of firms that are in the same SIC3 or SIC2 industries. Market overlap at the industry pair level is the average of firm-pair market overlap for all permutations of firms where one firm is in each of the industries being compared at the industry-pair level. Firm-pair overlap is the intersection of industries the two firms in the pair likely operate in divided by the union of industries they operate in (based on the industries assigned to each firm as indicated by the construction of the FIC-scope variable). This market overlap score ranges from zero to unity and is one if the firms operate in exactly the same industry and zero if they have no overlaps. The lower figure reports the two and three digit SIC HHI over time. The scope-adjusted HHI is the average the SIC2 and SIC3 HHI, where the weights start at zero in 1996 and grow linearly until they reach 50% by the end of our sample in 2017.



# Online Appendix:
# Scope, Scale and Competition:
# The 21st Century Firm

Gerard Hoberg and Gordon M. Phillips

(not for publication)

# 1. Measuring Scope Using Latent Dirichlet Allocation

Our primary measures of firm-scope are based on measuring each firm's textual overlap with the vocabularies that define various industry classifications including the FIC-300 classification as outlined in Hoberg and Phillips (2016) and the 4-digit NAICS classification. These baseline specifications have the benefit of simplicity and ease of interpretation: a firm with greater scope uses vocabulary in its Item 1 that overlaps more with the specialized vocabulary of more than one well-defined industry.

Our first set of Online Appendix tests examine robustness to defining industries using a topic model based on Latent Dirichlet Allocation (LDA). LDA is a computational linguistics model developed by Blei, Ng, and Jordan (2003). As an earlier example in finance, Hoberg and Lewis (2017) apply LDA on 10-K MD&A content (using 100 topics) to identify textual factors that associate with fraud.

To develop an LDA-based measure of scope, we run LDA on 10-K Item 1's at a granularity of 300 topics. We choose 300 topics to match the granularity of the FIC-300 industry classification used in our baseline FIC scope measure. Also to remain consistent with our baseline specifications, we fit the 300-topic LDA model using only the 10-K business descriptions from 1997 (although our results are robust to fitting the model separately in each year). Once we fit the LDA model using 1997 data, we then apply the model on all years of our sample (1989 to 2017). The result of this procedure is that we have a separate set of 300 topic loadings for each firm in each year, which we denote as $TL_{i,k,t}$, where $i$ indicates a firm, $k$ is a topic from one to 300, and $t$ is a year. We drop any topics from our calculations if they are too prevalent (more than 20% of firms in our sample have positive loadings on the given topic), indicating they are likely boilerplate and do not contain industry-specific content (our results are robust if we omit this step).

Topic loadings are derived through a maximum likelihood optimization and can be viewed as estimated probabilities that each topic is prevalent in the given firm-year's 10-K. Therefore, topic loadings $TL_{i,k,t}$ are non-negative and are bounded in $[0, 1]$. To measure scope using topic loadings, we next estimate the total mount of 10-K Item 1 text allocated to each topic ($Q_{i,k,t,LDA}$) by taking the product of the topic loadings and the size of the firm's 10-K Item 1 (measured as number of words $S_{i,t}$): $Q_{i,k,t,LDA} = TL_{i,k,t} \cdot S_{i,t}$.

Finally, as we did in equations (2) and (3) using our baseline methods, we use a 2% threshold $\bar{Q_{LDA}}$ that is fitted using 1997 data to determine if a given firm $i$ is likely operating in the industry identified by topic $k$ in year $t$ as follows:

$$LDA - Scope_{i,t} = \sum_{j=1,\ldots 300} Indicator\{Q_{i,j,t,LDA} > \bar{Q_{LDA}}\} \tag{6}$$

The resulting measure of scope ($LDA - Scope_{i,t}$) has a mean of 5.24, which is similar to that of 6.93 of our main scope measure ($FIC - Scope_{i,t}$). To complete our tests of robustness, we examine if our results in Table 8 to Table 14 are robust to using our LDA-based scope variable. The results are displayed in Tables IA5 to IA7. Side-by-side analysis of these tests indicate that our results are fully robust using our LDA-based measure of scope. LDA scope is also increasing over time as is the case for our baseline measures FIC-scope and NAICS-scope.

# 2. Broadening Scope of Businesses

The primary issue with existing classifications is they have fixed granularities, and as firms broaden scope of their operations, concentration in specific product markets cannot be computed using narrow industry assignments of firm sales data. Firms may produce in multiple 3-digit SIC codes and competition is mismeasured if researchers assign their sales to just a small subset of these SIC codes. The extent of this bias could be time-varying. For example, firms may produce in a just one 3-digit SIC industry code early in our sample, while later producing in multiple SIC codes - thus making one or even two 2-digit SIC codes more representative of its scope of production.

We first illustrate this point using a test that provides external validity of this idea. We intentionally avoid using the spatial TNIC representation of industries for this test and instead consider annual OLS regressions where the intensity of managerial competition complaints is the dependent variable (see Li, Lundholm, and Minnis (2013)). We regress this variable on both one minus the Compustat SIC-3 HHI where each firm is assigned to the Compustat 3-digit SIC code it reports and one minus the Compustat SIC-2 HHI - using the 2-digit SIC code the firm reports. We flip the sign on the HHIs for convenience as (1-HHI) is a positive measure of competition. We also use the same sample selection criteria as Grullon, Larkin, and Michaely (2019) for consistency. The results are reported in Table IA9.

The table illustrates an economically large trend toward increasing importance of coarser SIC-2 codes in understanding firm production and competition (reinforcing our conclusion that firms are operating in more markets over time). In the first year of this sample, 1997, only competition measured using the SIC-3 HHI predicts competition complaints. This early result indicates that 3-digit SIC codes well-represented the appropriate granularity of market boundaries at which competition among firms took place. However, throughout our sample, the relative importance of the HHI measured using 2-digit SIC codes increases and the relative importance of the SIC-3 digit HHI decreases. By the end of our sample, the coefficients for both HHIs are roughly equal in size, suggesting that competition is taking place across multiple 3-digit SIC codes and complaints are arising from multiple 3-digit level product markets. In the next section, we will show that market overlaps using our TNIC scope-based framework will indicate the same conclusion, and adjusting HHIs for broadening scope suggests that horizontal concentration is not rising materially in our sample.

## 2.1 Scope Adjustment via Market Overlap Analysis

We now develop an intuitive adjustment of HHIs for scope based on examining how market overlap varies with granularity. We define "Firm-Pair Market Overlap" for a pair of firms $i$ and $j$ using the FIC-300 classification and the firm-specific market definitions implied by equation (2). For example, suppose firm $i$ operates in industry A and B, and $j$ operates in A, C and D. Market Overlap for this pair is the true overlap in their markets, which is $\frac{1}{4}$, as they intersect on just one industry but the union of industries they serve is four. We next define "Industry-Pair Market Overlap" for a pair of industries in any classification as the average Firm-Pair Market Overlap averaged over all permutations of pairs of firms $i$ in the first industry and $j$ in the second industry. If the two industries have high Industry Pair Market Overlap, it follows that the boundary between the two industries is not material and that competition plays out at a more coarse level of industry granularity.

[**Insert Figure 5 Here**]

To illustrate the impact of scope on industry boundaries and the role of granularity, the upper graph in Figure 5 plots the average Industry-Pair Market Overlap for all pairs of 4-digit SIC industries that have the same 3-digit SIC code, and also for all 4-digit SIC industries that have the same 2-digit SIC code but not the same 3-digit SIC code. Both statistics have been increasing rather dramatically throughout our sample. This illustrates that the aforementioned rise in scope is indeed rendering narrow industry boundaries less relevant over time, especially for more fine granularities such as three-digit SIC codes. The more important observation, however, is that the average market overlap for the SIC-2 pairs late in our sample is actually higher than the level of market overlap for SIC-3 pairs early in our sample. This indicates that industry boundaries are as strong today at the two-digit SIC level as they were at the three-digit SIC level 25 years ago.

Existing studies note 1997 as a pivotal year in which the rise in concentration began to accelerate. At the start of this year, SIC-2 market overlap was 4.04% and SIC-3 market overlap was 5.72%. By the end of our sample, SIC-2 market overlap rose to 4.88%, and this

0.84% rise is enough to close 50% of the ex-ante gap of 1.68% between SIC-2 and SIC-3. It follows that between 1997 and present day, the granularity at which competition takes place moved roughly one half of one level of granularity. To show the impact of such a shift on concentration levels, we plot three trend-lines for concentration in the lower graphic of Figure 5. These include the benchmark SIC-2 HHI and the SIC-3 HHI as computed in the existing literature,[11] as well as a mixture of the two that starts at 100% SIC-3 HHI in 1997 and linearly moves to 50% SIC-3 HHI and 50% SIC-2 HHI at the end of our sample. Only the mixed HHI roughly holds market overlap fixed during the crucial post-1996 sample, and hence only this specification is a reasonable scope-adjusted HHI trend line.

The figure illustrates that horizontal concentration is not rising materially when we consider a scope-adjusted HHI. In contrast, we replicate the finding in the existing literature that concentration does appear to be rising dramatically if we do not adjust HHI measures for scope. The scope-adjusted HHI essentially allows granularity to shift with average scope whereas past studies hold granularity fixed over time. Adjusting granularity is necessary because increasing scope broadens industry boundaries, and competition thus occurs over increasingly coarse levels of granularity. The linear adjustment we employ in this section is highly simplified, and we reiterate that our goal here is to show intuition for how scope can impact competitive granularity, which in turn, can impact how competition is changing over time. In the next section, we adopt a more direct scope-adjusted measure of concentration based on our implicit modeling of the multiple industries firms operate in.

## 2.1 Scope Adjustment via Implied Multi-Industry Assignments

The construction of the FIC-scope variable assigns each firm to multiple industries when its Item 1 is similar to more than one industry. We now use this enhanced data structure to compute new HHIs at the FIC-300 industry level that use this multi-industry-assignment-classification directly. To do so, we first allocate each firm's total sales to the multiple industries it is assigned to using the basic similarity weights (see $Q_{i,j,t,FIC}$ in equation (1)) that were used to construct the classification itself. HHIs are then computed at the FIC industry level using these allocated sales where firms operate in multiple sectors. We then aggregate these HHIs back to the firm level by computing weighted averages over the sectors

[11] We compute baseline SIC-2 and SIC-3 HHIs following the sample and weighting scheme used by Grullon, Larkin, and Michaely (2019). We limit the sample to firms with CRSP exchange codes of 1 to 3, CRSP share codes 10 and 11, sales and assets greater than one million, and we exclude financials and utilities. We also compute HHIs based on assigning firms to more than one industry if indicated in the Compustat segment tapes. The annualized average HHIs are also weighted by sales.

each firm operates in (again using weights $Q_{i,j,t,FIC}$). Note that our results are similar if we use equal weights instead.

We then aggregate HHIs to the economy-wide annual level by computing a sales weighted average of the firm HHIs or an equal weighted average of the firm HHIs.[12] We then plot both estimates of the HHI faced by average firm in each year in the Figure 4. The figure illustrates, as was the case with the scope adjustment used in the previous section, that horizontal concentration levels are not rising materially after 1997.

The results in this section suggest that an extended narrative might be relevant to understand the rise in industry concentration reported in the literature. This extended narrative is that scope has been rising rapidly as companies merged and listings declined, and as a consequence, traditional HHIs measured without adjustment are increasing. Yet the rise in these HHIs might not indicate reductions in horizontal competition as they are based on overly rigid classifications that do not account for scope and that assign firms to single industry categories.

[12]The sales-weighted approach is used in Grullon, Larkin, and Michaely (2019).

# Table IA1: Pfizer Scope vs Time

The table displays the LDA-based scope allocations of Pfizer in 1990 (Panel A), 2003 (Panel B), and 2017 (Panel C).

| Year | Topic | Amount | Word List |
|---|---|---|---|
| | | | Panel A: Pfizer Scope Allocations in 1990 |
| 1990 | 123 | 123.4 | drug, supplement, formulation, delivery, company, skin, johnson, oral, nutritional, dietary |
| 1990 | 57 | 93.5 | company, medtronic, healing, abnormality, abnormal, systemic, stimulation, pulmonary, clot, fusion |
| 1990 | 34 | 93.0 | surgical, surgery, medical, catheter, tissue, company, device, disposable, surgeon, needle |
| 1990 | 89 | 91.3 | pharmaceutical, product, generic, prescription, cosmetic, baxter, ammunition, company, bowling, manufacturer |
| 1990 | 211 | 82.5 | health, provider, company, managed, plan, service, medical, benefit, employer, cost |
| 1990 | 251 | 76.4 | cancer, disease, company, treatment, development, drug, clinical, compound, study, trial |
| 1990 | 126 | 71.4 | company, food, corn, seed, feed, fish, fertilizer, flower, grain, production |
| 1990 | 81 | 66.6 | puerto, rico, coating, specialty, polymer, paint, synthetic, used, fiber, dupont |
| 1990 | 241 | 59.7 | patient, care, medical, treatment, facility, physician, company, therapy, service, dialysis |
| 1990 | 260 | 51.8 | blood, heart, system, disease, cardiac, body, approximately, procedure, lung, treatment |
| 1990 | 210 | 44.9 | mineral, dakota, operation, reclamation, mine, south, rock, property, surface, permit |
| | | | Panel B: Pfizer Scope Allocations in 2003 |
| 2003 | 251 | 395.5 | cancer, disease, company, treatment, development, drug, clinical, compound, study, trial |
| 2003 | 89 | 321.5 | pharmaceutical, product, generic, prescription, cosmetic, baxter, ammunition, company, bowling, manufacturer |
| 2003 | 123 | 312.0 | drug, supplement, formulation, delivery, company, skin, johnson, oral, nutritional, dietary |
| 2003 | 244 | 219.9 | reimbursement, medicare, care, health, program, medicaid, company, facility, physician, living |
| 2003 | 241 | 184.6 | patient, care, medical, treatment, facility, physician, company, therapy, service, dialysis |
| 2003 | 211 | 136.5 | health, provider, company, managed, plan, service, medical, benefit, employer, cost |
| 2003 | 57 | 112.5 | company, medtronic, healing, abnormality, abnormal, systemic, stimulation, pulmonary, clot, fusion |
| 2003 | 257 | 72.5 | would, proposal, proposed, state, congress, enacted, bill, predict, change, legislative |
| 2003 | 260 | 67.7 | blood, heart, system, disease, cardiac, body, approximately, procedure, lung, treatment |
| | | | Panel C: Pfizer Scope Allocations in 2017 |
| 2017 | 251 | 403.7 | cancer, disease, company, treatment, development, drug, clinical, compound, study, trial |
| 2017 | 89 | 330.8 | pharmaceutical, product, generic, prescription, cosmetic, baxter, ammunition, company, bowling, manufacturer |
| 2017 | 244 | 222.8 | reimbursement, medicare, care, health, program, medicaid, company, facility, physician, living |
| 2017 | 211 | 172.8 | health, provider, company, managed, plan, service, medical, benefit, employer, cost |
| 2017 | 123 | 161.7 | drug, supplement, formulation, delivery, company, skin, johnson, oral, nutritional, dietary |
| 2017 | 257 | 146.9 | would, proposal, proposed, state, congress, enacted, bill, predict, change, legislative |
| 2017 | 241 | 133.1 | patient, care, medical, treatment, facility, physician, company, therapy, service, dialysis |
| 2017 | 104 | 105.9 | burn, memorial, payer, capitation, provider, median, wale, clinical, hospital, campbell |
| 2017 | 57 | 103.6 | company, medtronic, healing, abnormality, abnormal, systemic, stimulation, pulmonary, clot, fusion |

# Table IA2: Investment Regressions (One-Stage Regressions Using Scope Incentive Variables)

The table reports one-stage results based on the two-stage regression results for investment variables in Table 10. The key difference in this table is that we include the two scope incentive variables directly as key RHS variables instead of using these two variables as instruments. The first scope incentive variable is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. As always, all RHS variables are measurable as of year $t-1$ and all dependent variables are as of year $t$. We also include controls for size, age, market to book, and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Asset Redeployability Scope Incentive | Opportunity Set Scope Incentive | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|---|
| (1) | Acquirer Dummy | –0.003 | 0.223 | 0.017 | –0.051 | | | 99,505 |
| | | (–0.120) | (5.110) | (5.510) | (–5.830) | | | |
| (2) | Target Dummy | –0.032 | –0.039 | 0.029 | 0.046 | | | 99,505 |
| | | (–1.990) | (–1.220) | (12.490) | (7.280) | | | |
| (3) | R&D/Assets | 0.015 | 0.008 | –0.013 | 0.007 | | | 99,505 |
| | | (2.870) | (1.380) | (–13.870) | (3.580) | | | |
| (4) | CAPX/Assets | –0.002 | 0.004 | –0.002 | –0.015 | | | 99,505 |
| | | (–0.560) | (0.790) | (–3.380) | (–10.670) | | | |
| (5) | Acquirer Dummy | –0.010 | 0.241 | 0.023 | –0.043 | 0.011 | 0.018 | 98,936 |
| | | (–0.470) | (5.510) | (7.020) | (–4.910) | (11.280) | (2.190) | |
| (6) | Target Dummy | –0.032 | –0.030 | 0.029 | 0.044 | –0.002 | 0.004 | 98,936 |
| | | (–1.980) | (–0.920) | (11.860) | (6.890) | (–4.720) | (0.700) | |
| (7) | R&D/Assets | 0.015 | 0.001 | –0.014 | 0.007 | –0.001 | –0.008 | 98,936 |
| | | (2.830) | (0.220) | (–14.140) | (3.450) | (–1.680) | (–5.460) | |
| (8) | CAPX/Assets | –0.002 | 0.001 | –0.001 | –0.013 | 0.002 | –0.002 | 98,936 |
| | | (–0.910) | (0.230) | (–2.000) | (–9.400) | (9.180) | (–1.750) | |

# Table IA3: Outcomes Regressions (One-Stage Regressions Using Scope Incentive Variables)

The table reports one-stage results based on the two-stage regression results for investment variables in Table 12. The key difference in this table is that we include the two scope incentive variables directly as key RHS variables instead of using these two variables as instruments. The first scope incentive variable is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. As always, all RHS variables are measurable as of year $t-1$ and all dependent variables are as of year $t$. We also include controls for size, age, market to book, and the TNIC HHI. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Asset Redeployability Scope Incentive | Opportunity Set Scope Incentive | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|---|
| (1) | Valuation (M/B) | 0.413 | 0.226 | –0.460 | –0.404 | | | 98,938 |
| | | (3.990) | (1.420) | (–20.160) | (–9.270) | | | |
| (2) | Sales Growth | 0.025 | 0.235 | –0.071 | –0.208 | | | 99,124 |
| | | (1.140) | (5.540) | (–19.020) | (–25.550) | | | |
| (3) | Asset Growth | 0.042 | 0.350 | –0.165 | –0.077 | | | 99,502 |
| | | (2.160) | (9.610) | (–44.760) | (–10.200) | | | |
| (4) | OI/Assets | –0.006 | –0.003 | 0.011 | 0.001 | | | 99,301 |
| | | (–0.430) | (–0.180) | (4.370) | (0.210) | | | |
| (5) | Valuation (M/B) | 0.315 | 0.149 | –0.370 | –0.217 | 0.226 | –0.002 | 98,705 |
| | | (3.650) | (1.130) | (–17.770) | (–5.800) | (10.520) | (–0.050) | |
| (6) | Sales Growth | 0.016 | 0.247 | –0.057 | –0.181 | 0.034 | 0.024 | 98,567 |
| | | (0.740) | (5.920) | (–15.430) | (–21.400) | (12.550) | (2.730) | |
| (7) | Asset Growth | 0.021 | 0.325 | –0.146 | –0.038 | 0.046 | –0.003 | 98,936 |
| | | (1.120) | (9.210) | (–42.300) | (–5.010) | (16.570) | (–0.380) | |
| (8) | OI/Assets | –0.007 | 0.001 | 0.013 | 0.005 | 0.005 | 0.004 | 98,740 |
| | | (–0.540) | (0.070) | (5.390) | (0.950) | (4.760) | (1.110) | |

## Table IA4: Financing Regressions (One-Stage Regressions Using Scope Incentive Variables)

The table reports one-stage results based on the two-stage regression results for investment variables in Table 14. The key difference in this table is that we include the two scope incentive variables directly as key RHS variables instead of using these two variables as instruments. The first scope incentive variable is a measure of the extent to which the broader product market surrounding a focal firm is characterized by a high degree of outward-directed asset redeployability indicating a low cost to scope expansion by existing firms. The second is a measure of the size of the focal firm's outward-expansion opportunity set. We also include controls for size, age, market to book, and the TNIC HHI. As always, all RHS variables are measurable as of year $t-1$ and all dependent variables are as of year $t$. All regressions include firm and year fixed effects, and $t$-statistics are clustered by firm and shown in parentheses.

| Row | Dependent Variable | Asset Redeployability Scope Incentive | Opportunity Set Scope Incentive | Log Assets | Log Age | Mkt/Book | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|---|
| (1) | Equity Issuance | 0.018 | 0.057 | –0.042 | –0.017 | | | 99,505 |
| | | (2.600) | (6.030) | (–28.500) | (–6.140) | | | |
| (2) | Debt Issuance | –0.011 | 0.028 | –0.010 | 0.008 | | | 99,505 |
| | | (–1.010) | (1.330) | (–5.470) | (1.810) | | | |
| (3) | Dividends/Assets | –0.004 | –0.007 | 0.000 | 0.003 | | | 99,405 |
| | | (–1.840) | (–1.770) | (–1.160) | (4.890) | | | |
| (4) | Repurchases/Assets | –0.004 | –0.003 | 0.004 | 0.006 | | | 92,077 |
| | | (–1.690) | (–0.610) | (10.130) | (6.150) | | | |
| (5) | Equity Issuance | 0.012 | 0.049 | –0.039 | –0.008 | 0.010 | –0.005 | 98,936 |
| | | (1.790) | (5.420) | (–27.240) | (–2.750) | (10.460) | (–2.200) | |
| (6) | Debt Issuance | –0.014 | 0.022 | –0.010 | 0.009 | 0.001 | –0.006 | 98,936 |
| | | (–1.190) | (1.090) | (–5.360) | (1.950) | (2.220) | (–1.440) | |
| (7) | Dividends/Assets | –0.004 | –0.006 | 0.000 | 0.003 | 0.000 | 0.001 | 98,839 |
| | | (–1.930) | (–1.470) | (–0.470) | (5.240) | (4.180) | (1.780) | |
| (8) | Repurchases/Assets | –0.004 | –0.001 | 0.004 | 0.006 | 0.001 | 0.001 | 91,546 |
| | | (–1.770) | (–0.260) | (11.100) | (6.590) | (5.730) | (1.530) | |

# Table IA5: High Product Breadth Validation Regressions (LDA Scope Variable)

The table reports validation regressions using the same model as our baseline results in Table 8. However, in this table, we focus on the key RHS variable LDA scope, which is our measure of scope that is based on an LDA model of the content of firm 10-K Item 1s.

| Row | Dependent Variable | LDA-Scope | # Segments | Log Assets | Log Age | Log 10K Size | M/B | TNIC HHI | # Obs |
|---|---|---|---|---|---|---|---|---|---|
| (1) | Product Breadth | 0.479 | 1.254 | 1.299 | 0.601 | –9.020 | –0.216 | –0.772 | 72,276 |
| | | (5.200) | (2.530) | (2.860) | (0.390) | (–7.710) | (–2.600) | (–0.630) | |
| (2) | Prod/Svc Breadth | 0.528 | 1.158 | 1.410 | 0.905 | –9.515 | –0.217 | –0.578 | 72,276 |
| | | (5.550) | (2.300) | (3.040) | (0.580) | (–7.900) | (–2.580) | (–0.460) | |
| (3) | Prod Breadth Detail | 0.170 | 0.385 | 0.644 | –1.042 | –3.536 | 0.013 | –1.079 | 72,276 |
| | | (4.590) | (1.720) | (3.250) | (–1.570) | (–7.600) | (0.420) | (–2.180) | |
| (4) | Prod/Svc Breadth Detail | 0.186 | 0.366 | 0.667 | –0.817 | –3.684 | 0.015 | –0.974 | 72,276 |
| | | (4.860) | (1.600) | (3.330) | (–1.210) | (–7.630) | (0.460) | (–1.930) | |

## Table IA6: First-Stage Regressions (LDA Scope Model)

The table reports the results of first-stage regressions where our LDA-based measure of scope (LDA-scope) is regressed on our two instruments in addition to all controls and fixed effects. This table is run using the same models as our baseline results in Table 9. However, in this table, we focus on the key RHS scope measure LDA scope, which is our measure of scope that is based on an LDA model of the content of firm 10-K Item 1s.

| Row | Dependent Variable | Sectoral Redeployment Potential | Sectoral Opportunity Set Potential | Log Assets | Log Age | # Obs |
|---|---|---|---|---|---|---|
| (1) | LDA-Scope | 0.550 | 6.339 | 0.653 | –0.771 | 99,501 |
| | | (2.510) | (16.060) | (20.090) | (–8.610) | |
| (2) | # Segments | 0.019 | 0.297 | 0.107 | 0.161 | 99,505 |
| | | (0.300) | (2.910) | (12.330) | (6.270) | |

# Table IA7: Corporate Finance Regressions (LDA scope model)

The table reports the results of second-stage 2-stage instrumental variable regressions where the dependent variable is a firm investment policy, an outcome variable, or a firm financing policy (as noted in the first column). This table is run using the same models as our baseline results in Tables 10, 12, and 14. However, in this table, we focus on the key RHS scope measure LDA scope, which is our measure of scope that is based on an LDA model of the content of firm 10-K Item 1s.

| Row | Dependent Variable | LDA-Scope | Log Assets | Log Age | # Obs |
|---|---|---|---|---|---|
| (1) | Acquirer Dummy | 0.033 | –0.004 | –0.026 | 98,191 |
| | | (4.990) | (–0.720) | (–2.400) | |
| (2) | Target Dummy | –0.010 | 0.036 | 0.039 | 98,191 |
| | | (–2.030) | (8.940) | (5.440) | |
| (3) | R&D/Assets | 0.003 | –0.015 | 0.009 | 98,191 |
| | | (3.570) | (–12.830) | (4.250) | |
| (4) | CAPX/Assets | 0.000 | –0.002 | –0.014 | 98,191 |
| | | (0.550) | (–2.810) | (–9.690) | |
| (5) | Valuation | 0.082 | –0.514 | –0.349 | 97,620 |
| | | (3.760) | (–17.680) | (–7.430) | |
| (6) | Sales Growth | 0.038 | –0.095 | –0.179 | 97,820 |
| | | (6.140) | (–16.040) | (–18.120) | |
| (7) | Asset Growth | 0.057 | –0.202 | –0.034 | 98,188 |
| | | (9.890) | (–33.550) | (–3.370) | |
| (8) | OI/Assets | –0.001 | 0.011 | 0.000 | 97,990 |
| | | (–0.450) | (3.580) | (0.030) | |
| (9) | Equity Issuance | 0.010 | –0.049 | –0.009 | 98,191 |
| | | (7.700) | (–26.190) | (–2.890) | |
| (10) | Debt Issuance | 0.003 | –0.011 | 0.011 | 98,191 |
| | | (0.910) | (–4.280) | (2.060) | |
| (11) | Dividends/Assets | –0.001 | 0.001 | 0.002 | 98,092 |
| | | (–2.600) | (1.670) | (2.560) | |
| (12) | Repurchases/Assets | –0.001 | 0.004 | 0.005 | 90,675 |
| | | (–1.290) | (7.700) | (5.000) | |

## Table IA8: Corporate Finance Regressions (Robustness to Dropping 50 Largest Firms in Each Year)

The table reports the results of second-stage 2-stage instrumental variable regressions where the dependent variable is a firm investment policy, an outcome variable, or a firm financing policy (as noted in the first column). This table is run using the same models as our baseline results in Tables 10, 12, and 14. However, in this table, we drop the 50 largest firms in each year to illustrate that our results are not driven by mega-firms. Our results also remain robust if we drop the 100 largest, 500 largest, or even 1000 largest firms in each year.

| Row | Dependent Variable | FIC-Scope | Log Assets | Log Age | # Obs |
|---|---|---|---|---|---|
| (1) | Acquirer Dummy | 0.028 | –0.008 | –0.031 | 96,755 |
| | | (4.700) | (–1.240) | (–2.880) | |
| (2) | Target Dummy | –0.011 | 0.036 | 0.038 | 96,755 |
| | | (–2.680) | (8.220) | (5.420) | |
| (3) | R&D/Assets | 0.003 | –0.016 | 0.009 | 96,755 |
| | | (3.770) | (–12.480) | (4.190) | |
| (4) | CAPX/Assets | 0.000 | –0.002 | –0.015 | 96,755 |
| | | (0.450) | (–2.460) | (–10.100) | |
| (5) | Valuation | 0.082 | –0.524 | –0.363 | 96,188 |
| | | (4.260) | (–17.100) | (–7.700) | |
| (6) | Sales Growth | 0.032 | –0.099 | –0.186 | 96,385 |
| | | (5.690) | (–14.680) | (–18.020) | |
| (7) | Asset Growth | 0.049 | –0.210 | –0.040 | 96,752 |
| | | (8.790) | (–29.570) | (–3.620) | |
| (8) | OI/Assets | –0.001 | 0.013 | 0.000 | 96,561 |
| | | (–0.530) | (3.700) | (0.010) | |
| (9) | Equity Issuance | 0.010 | –0.052 | –0.009 | 96,755 |
| | | (7.250) | (–25.130) | (–2.790) | |
| (10) | Debt Issuance | 0.002 | –0.011 | 0.009 | 96,755 |
| | | (0.750) | (–3.750) | (1.690) | |
| (11) | Dividends/Assets | –0.001 | 0.001 | 0.002 | 96,659 |
| | | (–2.510) | (1.670) | (2.780) | |
| (12) | Repurchases/Assets | –0.001 | 0.005 | 0.006 | 89,289 |
| | | (–1.460) | (7.530) | (5.270) | |

## Table IA9: Competition Complaints vs HHIs and Granularity

The table reports annual cross sectional OLS descriptive regressions where the dependent variable is the intensity of the firm's competition complaints in its 10-K, which is computed as the number of 10-K paragraphs that mention competition divided by the total number of paragraphs in the 10-K. The two RHS variables are measures of concentration (with the sign reversed so they can be interpreted as positive measures of competition) at different granularities. In particular include the Compustat SIC-3 HHI and the Compustat SIC-2 HHI. Each is computed as the sales-based concentration among firms in the given SIC code defined based on three digit and two digit SIC codes, respectively. Finally, we report the fraction of 2-digit granularity as the (1 - SIC-2 HHI) coefficient divided by the sum of the coefficients for both HHIs (truncated at zero in the first three years). This indicates the fraction of total HHI weights that are attached to the more coarse granularity. A high fraction indicates that, in the given year, the economy is such that competition takes place mostly at the 2-digit granularity rather than at the 3-digit granularity. A low value for this fraction indicates the converse.

| Year | One minus SIC-2 HHI | One minus SIC-3 HHI | Adj $R^2$ | Fraction 2-digit Granularity | # Obs. |
|---|---|---|---|---|---|
| 1997 | -0.001 | 0.011 | 0.014 | 0.000 | 5,521 |
| | (-0.46) | (8.15) | | | |
| 1998 | -0.003 | 0.011 | 0.012 | 0.000 | 5,297 |
| | (-1.34) | (7.63) | | | |
| 1999 | -0.003 | 0.012 | 0.017 | 0.000 | 5,076 |
| | (-1.24) | (8.66) | | | |
| 2000 | 0.006 | 0.010 | 0.022 | 0.369 | 4,827 |
| | (2.46) | (7.90) | | | |
| 2001 | 0.007 | 0.010 | 0.019 | 0.401 | 4,359 |
| | (2.37) | (6.74) | | | |
| 2002 | 0.009 | 0.005 | 0.010 | 0.642 | 3,954 |
| | (3.17) | (3.30) | | | |
| 2003 | 0.009 | 0.006 | 0.007 | 0.600 | 3,631 |
| | (2.34) | (2.90) | | | |
| 2004 | 0.011 | 0.008 | 0.028 | 0.577 | 3,540 |
| | (4.44) | (6.02) | | | |
| 2005 | 0.011 | 0.007 | 0.023 | 0.604 | 3,465 |
| | (4.25) | (5.11) | | | |
| 2006 | 0.007 | 0.006 | 0.019 | 0.565 | 3,378 |
| | (3.34) | (4.91) | | | |
| 2007 | 0.008 | 0.004 | 0.016 | 0.661 | 3,305 |
| | (3.85) | (3.80) | | | |
| 2008 | 0.009 | 0.004 | 0.023 | 0.674 | 3,120 |
| | (4.54) | (4.29) | | | |
| 2009 | 0.008 | 0.005 | 0.024 | 0.636 | 3,006 |
| | (4.21) | (4.78) | | | |
| 2010 | 0.010 | 0.004 | 0.024 | 0.744 | 2,899 |
| | (5.20) | (3.72) | | | |
| 2011 | 0.013 | 0.003 | 0.029 | 0.825 | 2,755 |
| | (6.47) | (2.91) | | | |
| 2012 | 0.005 | 0.003 | 0.009 | 0.655 | 2,665 |
| | (2.51) | (2.92) | | | |
| 2013 | 0.004 | 0.004 | 0.014 | 0.516 | 2,654 |
| | (2.10) | (4.23) | | | |
| 2014 | 0.004 | 0.003 | 0.011 | 0.538 | 2,695 |
| | (1.98) | (3.43) | | | |
| 2015 | 0.003 | 0.004 | 0.016 | 0.445 | 2,643 |
| | (1.82) | (4.57) | | | |
| 2016 | 0.003 | 0.004 | 0.015 | 0.484 | 2,546 |
| | (1.93) | (4.21) | | | |
| 2017 | 0.005 | 0.003 | 0.017 | 0.619 | 2,453 |
| | (2.96) | (3.67) | | | |