# The Intangible Divide:

# Why do so few firms invest in innovation?

By James Bessen (TPRI, Boston University) and Xiupeng Wang (TPRI, Northeastern University)

December, 2024

Abstract: Investments in software, R&D, and advertising have surged, nearing half of U.S. private nonresidential investment. Yet just a few hundred firms dominate this growth. Most firms, including large ones, regularly invest little in capitalized software and R&D, widening this "intangible divide" despite falling intangible prices. Using comprehensive US Census microdata, we document these patterns and explore factors associated with intangible investment. We find that firms invest significantly less in innovation-related intangibles when their rivals invest more. One firm's investment can obsolesce rivals' investments, reducing returns. This negative pecuniary externality worsens the intangible divide, potentially leading to significant misallocation.

Keywords: intangibles, R&D, software, innovation, obsolescence
JEL Codes: E22, O31, O32

Thanks to Erich Denk for research assistance, for comments from Po Hsuan, Gerald Marschke, Mike Meurer, Peter Meyer, Ronja Roettger, Timothy Simcoe, Leo Sveikauskas, Neil Thompson, James Traina, and Xiao Yin, as well as commenters at seminars at TPRI, BU Economics, MIT Futuretech, the NBER Productivity Lunch, CIST, IIOC, CAED, and GEIRC. Thanks to Holden Diethorn for providing data. This work was supported in part through a grant from the Alfred P. Sloan Foundation (Grant #2023-21111). Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2735 (CBDRB-FY23-0284, CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407, CBDRB-FY24-P2735-R11770).

# Introduction

Capital investment in intangible assets, such as research and development (R&D), software, advertising, and artistic originals has grown rapidly, making it a major driver of economic growth in recent decades (Corrado, Hulten, and Sichel 2009). Data from the Bureau of Labor Statistics (BLS) shows that the share of U.S. private nonresidential investment going to these intangible assets has grown dramatically from an estimated 24% in 1987 to 40% today (see Figure 1). Some economists attribute this growth to declining prices of these assets relative to other investment goods (Zhang 2019; Lashkari, Bauer, and Boussard 2024). For the price of software relative to structures has plummeted 5.4% per year while those for R&D and advertising experienced more modest relative declines.[1]

Yet, what aggregate statistics do not show is that this investment is hardly uniform. First, the rise is not about intangible investment generally, but is specifically about growth in R&D and capitalized software developed for internal use. These two assets account for 94 percent of the aggregate growth in Figure 1, and they are distinctly related to *innovation*—the majority of capitalized software is for the development of new applications for internal use that can be properly classified as innovation (see below). Second, as demonstrated in this paper, over ninety percent of the growth in software and R&D capital expenditures is accounted for by just a few hundred firms, while most firms make no detectable formal capital investments in any given year. These two phenomena underscore a substantial "intangible divide" between firms' innovation investments and, surprisingly, this divide has deepened even as the relative prices of intangible assets have fallen—investment grew sharply for top firms but not for others and fewer firms invest at all in any year. Importantly, this skewness is not observed in other intangible assets such as advertising. Given well-established evidence showing that software and R&D improve firm productivity, why do so few firms invest?[2] And why haven't sharp drops in asset

[1] Thanks to Corby Garner of BLS for providing the data. Official price indices for software may sharply understate the rate of decline (Fleming 2023).

[2] A large literature finds substantial private returns to R&D (see review in Hall, Mairesse, and Mohnen 2010). Other research has found a positive link between firm productivity or firm value and information technology generally (Brynjolfsson and Hitt 1996; Brynjolfsson, Hitt, and Yang 2002) and

prices induced more firms to invest in innovation?

Large scale non-investment is not necessarily at odds with the standard workhouse model of invention (Nordhaus 1967; Scherer 1972). In this model, inventors face an "invention possibilities frontier" that is initial convex but becomes concave at higher levels of investment. This convexity could be because of fixed costs of conducting R&D or other reasons such as alternative sources of knowledge (see below). Because firms vary in the returns they earn from R&D, they face different "invention possibilities frontiers." As a result, at a given price of R&D, some firms may find investing in it unprofitable and hence choose not to invest, while others do. However, a drop in the price of R&D should make investment profitable for more firms.[3] And a substantial drop—such as the fall in software prices—should significantly increase the share of firms that invest. We observe no such increase, posing a puzzle.

This paper first documents this puzzle and the skewness of innovation investment generally. We next present a model that explains the observed behavior as the result of strategic interaction between firms. We then run regressions to test the importance of this strategic interaction as well as a range of other factors that might influence innovation investment decisions.

We begin by presenting descriptive statistics about the intangible divide, looking at different types of intangibles in comprehensive samples of firms from the US Census across most industries over a decade or longer. We find: 1) the top 250 firms ranked by intangible investment account for most of it, 2) most firms do not regularly invest in capitalized software and R&D in our sample, but most do invest in advertising, 3) almost all of the substantial growth in intangible investments over the last decade can be attributed to the top firms, and 4) the share of firms (and their share of revenues) that do not invest in R&D and capitalized software has actually grown over the last decade, despite falling relative asset prices.

Our descriptive analysis can rule out some simple explanations for

---

with own software investment specifically (Tambe and Hitt 2012). Advertising has also been related to firm market value (Villalonga 2004; Peters and Taylor 2017).

[3] The drop in the price of performing R&D is only relative to other asset prices or to GDP; the drop in the price of software is absolute.

widespread non-investment. While innovation investments differ substantially by industry, industry differences do not explain the divide—we see substantial non-investment even in industries that have heavy investors. Nor can non-investment be explained by outsourcing of R&D and software development, nor by under-reporting of intangible investments by smaller firms.

We argue that the individual inventor model is incomplete because it ignores strategic interaction. Firms consider the risk of obsolescence when they make investment decisions. In Schumpeterian models of growth (Aghion and Howitt 1992; Grossman and Helpman 1991), technological advance by some firms corresponds to the "creative destruction" of others. This means that heavy investments in innovation by some firms can raise rivals' obsolescence risk. But if an innovation is more likely to be obsolesced, then its expected return will necessarily be less. Facing lower expected returns, firms will invest less in innovation, all else equal. This negative pecuniary externality means that firms may diverge in their responses to falling asset prices.

We develop a model that can explain the divergence. Faced with technological competition, firms may invest in developing their own proprietary technology to achieve competitive advantage. But one firm's competitive advantage is another firm's disadvantage, implying "business stealing." The model shows that faced with falling asset prices, some firms will increase their investment, but that increased rivalry may induce their rivals to invest less or to not invest at all under some conditions.

We test the role of rivalry by running regressions against a set of explanatory variables on the likelihood that firms will invest in innovations and on the amount that they invest if they choose to do so. Consistent with our model, the two strongest predictors of R&D and software investment are firm size, which is strongly positive, and investments in these assets by other firms in the same industry, which is strongly negative. Using an instrumental variable based on changes in R&D tax treatment, we find that the negative impact of rivals' R&D is plausibly causal.

Thus, investment in innovation is highly skewed, there is widespread non-investment, and our evidence supports the notion that negative pecuniary externalities shape the innovation landscape. Indeed, negative pecuniary externalities imply misallocation of investment—some firms, facing a "tax," will invest too little. While it is beyond this paper to examine overall welfare effects, the resulting investment skew may be important. If firms face diminishing returns to their intangible investments, then the intangible divide may potentially raise concerns about a significant misallocation that might slow productivity growth and undermine entrepreneurial dynamism.

## Heterogenous obsolescence

Our key insight is that firm decisions to invest in innovation are affected by rivals' investments that raise the risk that a firm's own investments will soon become obsolete. A firm facing a higher risk of obsolescence will invest less, all else equal, because obsolescence reduces the expected future profit stream. Even within industries, firms will face heterogeneous obsolescence risks and innovation incentives because each firm faces different rivals. Asset price reductions increase innovation incentives for firms with low obsolescence risk, but that increased investment raises obsolescence risk for other firms. Some firms, facing higher obsolescence risk, may then reduce investment, leading to divergent investment behavior.

Of course, obsolescence is widely recognized. Economists typically treat obsolescence by assigning fixed, uniform depreciation rates to R&D and software, often at very high levels. For example, the BEA assigns R&D depreciation rates as high as 40 percent for some industries and software depreciation rates as high as 55 percent. Unlike tangible assets, knowledge assets do not "wear out," becoming less productive with time. Instead, they become less valuable for generating profits. For instance, quasi-rents are reduced when rivals use spillover knowledge or as rivals develop alternative technologies making the original knowledge obsolete. Yet note that both these sources of depreciation depend on the actions of *rivals*. A firm's risk of obsolescence depends significantly on the technological competition it faces in its product markets. Firms facing many rivals who invest in innovation are likely to experience much more rapid obsolescence, all else equal. While some obsolescence occurs exogenously—for example, antibiotics become less valuable as microbes naturally develop resistance—much obsolescence arises endogenously from Schumpeterian rivalry.

However, the prevailing assumption is that knowledge assets experience entirely exogenous and uniform rates of depreciation *independently* of product market rivalry.[4] This is a strong assumption that does not appear to have been explicitly discussed in the empirical literature. While this may have been an innocuous assumption in early R&D research where economists assumed low or insignificant rates of depreciation (under 10 percent),[5] with today's much higher rates of depreciation and obsolescence, it may be one abstraction too many for firm level analysis.

We contend, in contrast, that the risk of obsolescence is endogenous in the sense that it depends on rivals' actions. And this means that firms face different obsolescence risks in different industries and even within narrowly defined industries. A small firm facing a rival that invests heavily in innovation will face a greater risk of obsolescence, all else equal, than the large firm will face. This asymmetry can explain divergent investment behavior and it has important implications for econometric analysis, for policy, and for the distribution of innovative activity.

Consider, for example, the rivalry between Walmart and Dollar General. Walmart has become the world's biggest retailer using proprietary information technology (Basker 2007; Bessen 2022 chapter 1). Walmart spends over $10 billion a year on IT, hires many hundred software developers each year, and uses its proprietary systems (also involving hardware and organization) to stock its stores with an unparalleled selection of products—140,000 stock keeping units (SKUs) in its Supercenter stores—that are delivered efficiently and at low cost. Low prices and large selection are powerful draws for consumers who prefer "one-stop shopping." In contrast, Dollar General invests almost nothing in developing its own software

[4] The BEA and Li and Hall (2020) assign different R&D depreciation rates to some industries, but the assumption is still uniform depreciation rates within industries. Papers using patent data have measured individual obsolescence rates for different firms (Caballero and Jaffe 1993; Ma 2021).

[5] The early researchers on R&D commonly ignored R&D depreciation altogether (Griliches 1980, 424) or they assumed very low values, e.g., 4-7 percent (Mansfield 1968). Assuming these low levels of depreciation to be independent of industry conditions was of no great consequence. However, at the high rates of depreciation used today, the assumption seems less tenable.

although it purchases prepackaged software for some functions.[6] Dollar General has many more stores than Walmart, but these stores are much smaller and offer a limited selection (10-12,000 SKUs). Dollar General competes against Walmart on geography rather than on product—it has many stores, especially in rural towns, that are closer to many consumers than the Walmart Supercenters.

Why doesn't Dollar General invest more in IT systems so that it could offer greater selection and lower prices? Because the returns to adding new products might be low. A new product line offered by Dollar General is more likely to run into competition from new or existing products offered by Walmart, dissipating possible rents. While Dollar General might have good payoffs to other sorts of R&D or software development, Walmart's innovative advantage reduces Dollar General's incentives to invest in new products. New products introduced by Walmart do not face comparable obsolescence risk from Dollar General. The risks are asymmetric. Similar negative pecuniary externalities are also a feature of quality ladder growth models (Grossman and Helpman 1991; Aghion and Howitt 1992) as well as in Sutton's models of vertical differentiation (1991; 2001).

In addition to asymmetric obsolescence risk, firm investment patterns are affected by knowledge spillovers. Substantial spillovers mean that many firms can acquire the knowledge needed to engage in production without formally investing in innovation themselves. In other words, formal investments in R&D and software are not *essential* to production or even profitability. Hence, while almost all firms invest in tangible assets, we see many firms that do not invest in R&D or software.

While non-investment in software has been little studied, economists have long been aware of non-investment in R&D. Beginning in the 1950's and 1960's, researchers using NSF survey data documented that many firms did not invest in R&D even in R&D-intensive industries (Villard 1958; Hamberg 1964; Nelson, Peck, and Kalachek 1967; Bound et al. 1982; see Mezzanotti and Simcoe 2023 for a recent

[6] From 2013-18, Dollar General advertised online for only 18 software developer jobs; during the same period, Walmart advertised for 4,131 (Burning Glass data). A Bank of America study finds that Walmart offers lower prices than Dollar General in groceries. See https://www.supermarketnews.com/retail-financial/walmart-trumps-dollar-general-when-it-comes-price.

reprise). Some economists explored why firms might not report R&D and whether non-performing firms introduced selection bias in estimates of the returns to R&D (Bound et al. 1982; Hall, Mairesse, and Mohnen 2010 p. 18). A related literature looks at how firm size affects R&D spending among those firms that do invest (for a review see Cohen 2010). But most of the research on returns to R&D has simply overlooked the large numbers of non-investing firms (but see Doraszelski and Jaumandreu 2013; Cappelen et al. 2023). Following Griliches (1979), some of these researchers have also considered spillover externalities of R&D (for a review and meta-analysis see Hall, Mairesse, and Mohnen 2010; Ugur et al. 2016). A few of these papers have also included negative externalities from product market rivalry that our study highlights (Bloom, Schankerman, and Van Reenen 2013; Lucking, Bloom, and Van Reenen 2019; Arora, Belenzon, and Sheer 2021). In contrast to these papers, which study select samples of public firms, we use representative samples that can capture the effects of product market rivalry on small and private firms.

Our paper is also related to recent empirical work that ties some intangible investment to rising industrial concentration and markups (Bessen 2020; Calligaris, Criscuolo, and Marcolin 2018; Crouzet and Eberly 2018; Brynjolfsson, Jin, and Wang 2023; Mouel and Schiersch 2023; Lashkari, Bauer, and Boussard 2024) and a variety of theoretical explanations have been advanced for such a link (De Ridder 2023; Aghion et al. 2019; Hsieh and Rossi-Hansberg 2019; Haskel and Westlake 2018). To the extent that firm market shares are influenced by firm intangible investments, the rising skewness of intangible investment contributes to the rise in industry concentration. Moreover, our finding that an elite group of firms dominate intangible investment may be related to the "superstar" firms and "megafirms" found by other researchers (Autor et al. 2020; Song et al. 2019).

# Descriptive statistics

## Data

We use confidential microdata collected by the US Census to explore intangible capital expenditures in R&D and software, and in purchased advertising and promotional services. The investments we measure might best be viewed as proxies for broader investments in new systems that might also include complementary investments in unmeasured activity, such as inhouse marketing expenditures, investments in complementary organizational changes, and informal experimentation.

For each of the investments we measure, we construct a sample over roughly a decade ending in 2017 or 2018, covering companies with employees and revenues in the US nonfarm for-profit sector, including small firms. For survey-based data, we use sampling weights to obtain representative samples. Our analysis is at the firm level and because of data limitations, we construct unbalanced panels. For this reason, most of our analysis is repeated cross-sectional, but we do conduct robustness checks using lagged variables.

Because we focus on studying the effects of innovation-related intangibles at improving the firm's products and services, we exclude industries where the product is the intangible itself.[7] Our samples also consist of firms with positive revenues. This restriction makes the samples consistent over time and revenue-generating firms are arguably more representative of economic activity. In any case, the share of intangible investment by non-revenue firms is small and our general results are similar if we include them (results not disclosed).

Other than these restrictions, for software and R&D, our data cover all private non-farm industries. For advertising, we study just the manufacturing sector due to data limitations. Although manufacturing accounts for just a fraction of all advertising, the Census of Manufactures provides advertising data for a comprehensive sample of establishments.

Our intangible data come from three main sources:

1. The Business R&D and Innovation Survey (BRDIS), 2009 – 2018, is an annual survey of roughly 40,000 for-profit, nonfarm businesses with five or more employees done in collaboration with the National Science

[7] This follows the practice of the statistical agencies. For example, NIPA includes software development as R&D in the software publishing industry (software is the product) but it is counted as software investment in the retail industry. Our excluded industries are NAICS 5112, software publishing, 514191, online information services, 54151, computer design, 541511, custom programming, 5417, R&D, and 54181, advertising agencies. Using BLS and NSF data, we estimate that the excluded industries account for less than 23% of R&D, 18% of software, and 8% of advertising in 2018.

Foundation's National Center for Science and Engineering Statistics.[8] This survey provides data on each firm's worldwide R&D expenditures (see further details at Foster, Grim, and Zolas 2016; Mezzanotti and Simcoe 2023).

2. The Annual Capital Expenditures Survey (ACES), 2008 – 2018, surveys 50,000 companies with employees annually. This survey provides data on capital spending on structures and equipment as well as on three types of software: prepackaged, vendor-customized, and internally developed. ACES only captures capitalized software spending, but much software spending is expensed rather than capitalized and some software is free (Open Source) or is purchased as a service. In the Appendix, we discuss software capitalization and argue that, given the accounting rules, capitalized software is a good measure of software *innovation*. Most capitalized software represents new applications for internal use that are internally developed or customized by vendors. In addition, much of the remaining capitalized software (prepackaged) is used for major new systems and is correlated with internally developed software (correlation coefficient of .324). While major software innovations also involve other expenses such as expensed software, training, adaptation, and maintenance, capitalized software serves as a proxy of this broader investment.
3. The Census of Manufactures, 2007 – 2017, covers all manufacturing establishments every 5 years. We aggregate data to the firm level. This includes data on purchased advertising and promotional services as well as various measures of outsourcing including purchased data processing and other computer services, purchased professional and technical services, and materials, parts, and supplies that might capture embodied new technology. In addition, we link in data on management practices from the Management

[8] The BRDIS survey begins reporting data in 2008, however, the data in that year appear to be incomplete.

and Organizational Practices Survey, a supplement to the Annual Survey of Manufactures (ASM) conducted in 2010 and 2015.[9]

The establishment level Longitudinal Business Database (LBD) provides the industry code and the zip code of each establishment (Jarmin and Miranda 2002), which provides us the number of industries and zip codes a firm operates in. We also use the Revenue-enhanced Longitudinal Business Database (LBD-rev) that provides firm level data on sales as well as data on employees and firm age (Haltiwanger et al. 2016). Unfortunately, revenue data cannot be matched for all firms, so our sample, while still broadly inclusive, is not complete. Haltiwanger et al. (2016) find that the firms in LBD-rev data are broadly similar to the full set of firms on a range of observed characteristics. Below we check the robustness of results using revenue data from other sources or we use the full LBD sample when revenue data are not needed. Finally, we include counts of patents granted using a crosswalk to patent office data developed by Dreisigmeyer et al. (2018).

Our base analysis identifies rivals as firms in the same 6-digit NAICS industry, using time consistent industry definitions (Fort, Klimek, and others 2016). We also use a weighted distance measure to account for firms that produce in multiple industries. These measures are crude, and they do not account for geographic and other variations that affect true product market rivalry. However, to the extent that we mismeasure rival firm investments, our results will be diluted, but we find significant externalities despite this mismeasurement.

## Findings

Table 1 reports descriptive statistics on the top 250 firms ranked by investment for each asset type for each year.[10] These small groups of firms dominate

[9] As we are interested in the firm's management practice in general rather than the change over time, we use the management score from 2015 MOPS only. Following Bloom et al.(2019), we adopt the Management Score which is constructed by taking unweighted average of the 16 MOPS management questions in MOPS 2015. The Management Score variable is prepared by Scott Ohlmacher. For details, please see Bloom et al.(2020)2020 CES-20-41. Please see detailed questionnaire of 2015 MOPS here: https://www2.census.gov/programs-surveys/mops/technical-documentation/questionnaires/ma-10002_15_final_3-2-16.pdf (accessed September 23, 2024).

[10] For brevity, we just report results for total software and own-account software investment. Prepackaged and custom software are correlated with own account software. The correlation coefficients are .324 and .205 respectively.

intangible investment. For instance, column (1) shows top 250 R&D investors contribute over 80% of R&D investment. A similar pattern is observed for self-developed software, as shown in column (3). In total software and in advertising, the top 250 investors in each respective category contribute about 2/3 of the investment.[11] By contrast, the top 250 firms ranked by investment in tangible assets, that is structures plus equipment, account for only 55% of tangible investment. While the rank of 250 is arbitrary, this impression of dominance is consistently reinforced when using analyses (not shown) with different cutoffs.

Below we will see that although industry is an important determinant of intangible investment, the story is not as straightforward as a few industries dominating these top firms. As the second row shows, the top firms in each column are found in over 100 different six-digit NAICS industries.

While the firms in the top groups are large on average—in software and R&D the mean revenue is over $10 billion (row 4)—megafirms (those with $10 billion or more in revenues) comprise a minority of these groups and in total, these top groups account for a minority of revenue overall (row 5).

Finally, the composition of these top groups changes somewhat each year. Do firms nevertheless tend to stay in the top 250? Although we are limited by the unbalanced nature of our panel, row 6 of the table suggests a high degree of persistence for R&D and software: of the firms listed in the top 250 in 2018 and which were also present in the data a decade earlier, over half were in the top 250 then. Furthermore, the top software firms seem to be a largely distinct group from the top R&D firms—only 21% are in both.

Table 2 shows the shares of firms that report zero investment in each intangible (missing values are excluded). Over 90% of firms report zero investment in capitalized software and R&D (a lower share of firms report zero expensed software; see Appendix). The picture is quite different for advertising where only

[11] We exclude observations that do not report the intangible investments. These missing values account for 40 percent in the R&D data and 12 percent in the software data. If these observations represent zero investments, our measures of skewness will be understated. Koh et al. (2022) explore factors related to non-reporting of R&D data in financial statements.

32% of firms don't advertise each year. This might be because advertising investments do not spill over, as noted above.

The second row shows the share of revenue of these zero-investment firms. Over half of all firm revenue is accounted by firms that invest zero in R&D and self-developed software; that share is less for total software and even less for advertising. However, these figures may overstate the extent of non-investment because much investment tends to be "lumpy" (Doms and Dunne 1998; Haltiwanger, Cooper, and Power 1999). Smaller firms might not invest every year because of this lumpiness. Using a panel of firms that we observe for 5 or more consecutive years, we still find a large share of firms that do not invest in R&D and own-account software over 5-year periods (not shown). Moreover, the table shows a sharp contrast with tangible capital even looking at single years—only 11% of revenue comes from firms that invest nothing in tangible capital in any given year. This difference also appears in Figure 2 which shows the share of firms that invest in each intangible by the revenue size of the firm. Even among quite small firms, the majority invest in tangible capital and advertising; however, firms must be well over $1 billion in sales before a majority invest in R&D or own-account software in any given year.

The bottom rows of Table 2 show that non-investment, while varying by industry, is not largely determined by industry. Row 3 repeats row 2 from the previous table, showing the number of distinct industries with firms that are in the top 250. The next two rows show the share of firms with zero investment in just those industries and the share of revenue for those non-investing firms. Clearly, non-investment in R&D and own-account software is still substantial even in those industries where some firms invest heavily.

One concern about the estimates in both Tables 1 and 2 is possible under-reporting. First, to be clear, we are not counting firms that fail to respond to survey questions about investment; the non-investors are firms that do respond but answer "0." Also, we are only concerned about *formal* investments. Firms may tinker and experiment informally, but that is different. Nevertheless, perhaps some firms, especially small firms, might not have the accounting procedures in place to track formal R&D or software investments. Also, R&D tax credits might provide incentives for firms to categorize some non-R&D expenses as R&D, thus exaggerating reported R&D. In the Appendix we compare R&D personnel costs reported in the BRDIS survey to the earnings of scientists and engineers from the Current Population Survey. We conclude that although there may be some reporting bias, the basic picture of the dominance of large firms and substantial non-investment still holds.

Another concern about the findings in these two tables is that they do not use the full survey sample—to obtain revenue data, we match the sample to the LBD-rev, however, not all firms are listed in LBD-rev. To check our findings, we use an alternative source of revenue data (from BRDIS, see Appendix), finding that scaling by firm size is broadly similar. Finally, the Appendix also shows results for software that includes both capitalized software (used in Tables 1 and 2) and expensed software. This broader measure of software is not as dominated by the top 250 firms, but it also includes software that is not necessarily innovation.

Table 3 looks at how these investment patterns have changed over time. The top row shows the total deflated investment for each asset type for two separated years and the difference between them. The second row shows the corresponding investment by just the top 250 firms and the third row shows the investment of other firms. The fourth row shows the share of investment from the top 250 firms; in each case, the share of the top firms increased. The fifth row shows the share of the increase accounted for collectively by the top 250 firms. The top firms account for almost all of the increase. The difference between the top and the rest is particularly striking for software. In the face of a 16% price decline from 2008 to 2018, investment by the top 250 rose 48% while investment by the rest rose only 6%.[12]

The bottom two rows show the increases for the share of non-investing firms and the revenue-weighted share of non-investing firms. These increase for all investment types except for a small decrease in non-investing firms for own accounts

[12] Note that the figures in the top of the table are rounded. Also note that the groups of top 250 firms are not the same firms in 2008 as in 2018, although the majority are (Table 1). Also, compositional changes by firm size do not appear to be substantial. The total number of firms with employees increased by 1% during this period, and there were slightly more medium size firms. The share of firms with 1-19 employees fell 0.4%, the share of firms with 20-499 employees rose 0.3%, and the share of larger firms did not change significantly. Data from US Census, County Business Patterns.

software and a decrease for non-advertisers' share of sales. Thus, generally, the intangible divide for R&D and software seems to be growing wider, top firms sharply increasing their investment, other firms hardly increasing investment, and mostly fewer firms investing at all.

Yet this poses a puzzle because it is occurring against falling relative prices for intangible assets. Also, output shares of intangibles have been rising, suggesting rising output elasticities. For standard market-based assets, one might think falling prices and rising output elasticity would increase investment across the board.

# A Model of Obsolescence

## Investment demand

We present a model both to guide our empirical analysis and to also demonstrate conditions where firms may respond divergently to falling investment prices. The key notion we wish to capture in our model is that a firm's decision to invest in innovation depends in part on the risk of obsolescence that the resulting innovations are expected to experience. If firm *i* expects an annual profit stream of $\pi_i$ from an innovation, we can define an expected obsolescence rate, $\delta_i$, such that the net expected profit at time *t* is $e_i^{-\delta t}\pi_i$.

Obsolescence risk might vary across firms for at least two reasons. First, if a firm's rivals invest more in innovation, then the probability that rivals might develop a superior technology is greater, all else equal. Then $\delta_i$ would increase with rivals' innovation investments. Second, firms may differ in the rate at which a superior rival technology diminishes the profit stream. The introduction of a more productive technology does not immediately eliminate profits of incumbent firms—obsolescence can take time. A substantial literature finds that more productive firms grow faster than less productive firms (and are less likely to exit) but this appears to be an extended process (Caves 1998 provides a review). For instance, Foster et al. (2016) find that industry entrants have greater technical efficiency than older firms, but much smaller size of demand. Demand grows over time as firms build "customer bases."[13] But this suggests, conversely, that firms with large customer bases (or other

[13] See also Dunne, Roberts, and Samuelson (1988) and Cabral and Mata (2003). A variety of theoretical models attribute slow demand growth to informational frictions (Radner 2003; Rob and

idiosyncratic factors that increase demand) might lose demand more slowly in the face of a superior rival technology. That is, we might expect the rate of obsolescence, $\delta_i$, to be smaller for firms with large customer bases, all else equal. Then, if $S_i = \sum_{j \neq i} I_j$ measures investment by other firms in the same industry, and if $R_i$ measures the firm's customer base, $\delta_i = \delta(S_i, R_i)$. Below we will parameterize and test both hypothetical sources of heterogeneous obsolescence risk.

We assume that innovations create new products as determined by a version of the "knowledge production function" (Griliches 1979) that depends on the firm's own investment and also on "spillovers" from other firms, also measured by $S_i$.[14] Let the number of new products (treated as a continuous variable) that firm *i* introduces in year 0 be

$$n_i = S_i^{\alpha_i} I_i^{\beta}, \quad 0 < \alpha, \beta < 1 \tag{1}$$

where $I_i$ is the firm's investment in R&D (or software development), $S_i$ is investment by other firms in the same industry, $\alpha_i$ is the spillover elasticity, which is allowed to vary across firms, and $\beta$ is the investment elasticity. In this setup, rivals' investments have two opposing effects: knowledge spillovers provide a greater return on investment while greater obsolescence risk reduces expected returns. Given a cost $w$ of innovation, the total cost of innovation is $w \cdot I$. The firm chooses an optimal level of investment to maximize the present value of the expected profit stream less this cost. Then, given a discount rate of $r$, the net expected present value of current innovations is

$$V_i(I) = \int_{t=0}^{\infty} n_i \pi_i e^{-(\delta_i + r)t} dt - wI_i = \frac{S_i^{\alpha_i} I_i^{\beta} \pi_i}{r + \delta_i} - wI_i. \tag{2}$$

---

Fishman 2005; Bar-Isaac and Tadelis 2008; Arkolakis 2010; Dinlersoz and Yorukoglu 2012; Drozd and Nosal 2012; Perla 2013; Gourio and Rudanko 2014; Foster, Haltiwanger, and Syverson 2016). In these models, consumers generally lack information about the quality of a firm's product or brand, but they learn this information over time by experience with the products or by various types of communication. Firms can build "customer bases" of informed consumers by investing in customer acquisition (e.g., offering discounts to new customers) and by disseminating information via sales, marketing, and advertising.

[14] This model treats spillovers and own-investment as complement; previous evidence suggest that firms need to invest in order to use spillovers (Mansfield, Schwartz, and Wagner 1981; Cohen and Levinthal 1989). Below we find evidence that some firms increase their own investment with greater investment by rivals; this would not happen to substitutes.

Taking the first order maximizing condition with respect to $I$, assuming an interior solution, and solving for $\ln I$, we obtain an investment demand equation,

$$\ln \hat{I}_i = \frac{1}{1-\beta}\left[\ln \frac{\beta \pi_i}{w} + \alpha_i \ln S_i - \ln(\delta_i + r)\right]. \tag{3}$$

This describes an interior solution. We assume that if optimal demand is too small, the firm will choose to invest nothing, that is, a corner solution will obtain. An investment threshold might arise because the firm can gain sufficient knowledge via spillovers to meet its production needs; then investment may produce too little additional benefit to exceed the cost, that is, there is an opportunity cost. Alternatively, an indivisibility in knowledge production (e.g., a firm might not be able to hire a fraction of a scientist) might impose a minimum threshold. Generally, if $\hat{I}_i < T$, the firm will invest nothing. Then, given a distribution of firms over the various parameters in (3), something that raises (lowers) $\hat{I}_i$ will increase (decrease) the share of firms that make positive investments. Thus, equation (3) provides a basis for empirical analysis of whether firms invest and, if so, how much.

## Equilibrium

Before proceeding to the empirical analysis, it is helpful to explore how this model might explain divergent responses to falling investment costs. Each firm sets its investment according to (3), taking rivals' investments, $S_i$, as given. Assuming a Nash equilibrium, we analyze how levels of investment change with the cost of innovation, $w$. To make the calculation analytically tractable, we examine the case of two rival firms, *i* and *j*. Also, without loss of significant generality, we assume that $r = 0$.

We wish to solve for

$$\frac{d \ln \hat{I}_i}{d \ln w} = \frac{\partial \ln \hat{I}_i}{\partial \ln w} + \frac{\partial \ln \hat{I}_i}{\partial \ln S_i}\frac{d \ln S_i}{d \ln w}. \tag{4}$$

With only two firms, note that $S_i = I_j$ and $S_j = I_i$. Using this and assuming a Nash equilibrium, we find (see Appendix)

$$\frac{d \ln \hat{I}_i}{d \ln w} = \frac{\gamma_i - \alpha_i - (1-\beta)}{(1-\beta)^2 - (\gamma_i - \alpha_i)(\gamma_j - \alpha_j)}, \quad \gamma_i \equiv \frac{\partial \ln \delta_i}{\partial \ln S_i}, \tag{5}$$

where $\gamma_i$ is the elasticity of the obsolescence rate with respect to rival investment. This implies that $\frac{d \ln \hat{I}_i}{d \ln w} > \frac{d \ln \hat{I}_j}{d \ln w}$ if $\gamma_i - \alpha_i > \gamma_j - \alpha_j$.[15] That is, because firm *i* has a high $\gamma$ relative to $\alpha$, it is more sensitive to rivals' investment and hence has a more elastic response to changes in the investment price. Given sufficient heterogeneity in these parameters, we can identify a necessary condition for divergent investment behavior:

$$\frac{d \ln \hat{I}_i}{d \ln w} > 0 > \frac{d \ln \hat{I}_j}{d \ln w} \quad \text{if} \quad \gamma_i - \alpha_i > 1 - \beta > \frac{(\gamma_i - \alpha_i)(\gamma_j - \alpha_j)}{1 - \beta}. \tag{6}$$

In this case, a decrease in the cost of innovation, $w$, will induce less investment by firm *i* and greater investment by firm *j*. The intuition behind this result is that firms differ in their response to increased rivals' investment. To summarize, differences in firm response to rivals' investment mean that falling prices of innovation investment widen differences in investment, and, given sufficient dispersion in $\gamma - \alpha$, some firms will reduce investment when prices fall, or, if their target investment level falls below the threshold, they might no longer invest at all. Thus, this model can explain the behavior pattern we observe. The bottom line of this analysis is that not only do differences in firm investment persist in the face of falling investment prices, but falling prices might even *widen* the investment differences between firms.

Why might $\alpha$ and $\gamma$ differ across firms? First, as discussed above, firms with large customer bases might respond to rivals' innovations more slowly; that is, these firms might have a lower elasticity $\gamma_i$. Similarly, firms with large customer bases might have greater "absorptive capacity" (Cohen and Levinthal 1989). That is, large firms may have relatively more R&D resources devoted to monitoring and adapting rivals' innovations.[16] Thus $\alpha$ might be larger for firms with large customer bases. Below we test for these effects, and we find significant evidence for them.

---

[15] Assuming the denominator is positive.

[16] And large firms tend to have larger absorptive capacity because they can spread these costs over a larger product base (Cohen and Klepper 1996a).

## Regression Analysis

### Empirical implementation

Is there evidence that firm investment in intangibles is, in fact, significantly influenced by rival firm investments? We explore this hypothesis using investment demand regressions based on (3), both a linear probability model for whether firms choose to invest or not and a regression on log investment for those firms that do invest.

We can adapt the investment demand equation, (3), parameterizing several terms as first order functions of $S_i$ and $R_i$. We use revenue to proxy the firm's customer base and in the Appendix we explore alternative revenue measures, including domestic and worldwide revenue. We interpret $\ln(\delta_i + r) = f(\ln S_{it-1}, \ln R_{it}, t)$ as a single function that can be locally approximated with a with a translog form and time dummies. Following the production function literature, an arbitrary function can be locally approximated using a translog form (Christensen, Jorgenson, and Lau 1973), $f(\ln S_{it-1}, \ln R_{it}, t) \approx \phi_S \ln S_{it-1} + \phi_{SR} \ln S_{it-1} \cdot \ln R_{it} + \phi_R \ln R_{it} + \nu_t$.[17] To this we add $\alpha_i \ln S_i$ and a $\ln R_i$ term to capture possible size variation in $\pi_i$. Combining these terms, and adding industry fixed effects,

$$\ln \hat{I}_{it} = \theta_S \ln S_{it-1} + \theta_R \ln R_{it} + \theta_{SR} \ln S_{it-1} \cdot \ln R_{it} + \theta_X X_{it} + \mu_k + \nu_t + \epsilon_{it} \quad (7a)$$

where $X_{it}$ is vector of other firm characteristics including firm age, product diversity, geographic spread, outsourcing, and management practices, $\mu_k$ is a fixed effect for the 6-digit NAICS industry, $\nu_t$ is a year fixed effect, and $\epsilon_{it}$ is an error term. In some specifications we also include patents granted to other firms in the same industry. We use the lag of $\ln S$ both to avoid contemporaneous shocks that also affect the dependent variable and because it is natural to expect a lag (we also instrument this variable, see below). We test our hypotheses using simple tests on coefficients: a negative pecuniary externality corresponds to $\theta_S < 0$; such effect is weaker for larger firms if $\theta_{SR} > 0$.

---

[17] We drop the squared terms to improve estimates.

Specification (7a) serves to estimate investment demand conditional on firms choosing to invest. As above, the decision to invest can be thought of as condition that the target demand exceeds some threshold, $\hat{I}_{it} > T$. This condition can be estimated with a linear probability model,

$$Y_{it} = \theta_S \ln S_{it-1} + \theta_R \ln R_{it} + \theta_{SR} \ln S_{it-1} \cdot \ln R_{it} + \theta_X X_{it} + \mu_k + \nu_t + \epsilon_{it} \quad (7b)$$

where the dependent variable, $Y_{it}$, is a binary variable that is one if firm *i* invests at time *t* and is zero if the firm does not invest then.

We construct two different measures of $S_{it-1}$ based on the intangible investments of rival firms. In the R&D literature, the effect of rivals' investments is captured as a weighted sum of rivals' R&D capital stocks, where R&D stocks are constructed from investment flows over many years using the perpetual inventory method (Bloom, Schankerman, and Van Reenen 2013; Lucking, Bloom, and Van Reenen 2019; Arora, Belenzon, and Sheer 2021). Given our data, we cannot construct capital stocks without severely limiting our sample and, in any case, assuming constant depreciation rates is problematic. Since we want to analyze a broad representative sample, we instead construct a measure based on just the past year's intangible investments for firms in industry *J*:[18]

$$S_{i,t-1}^1 \equiv \sum_{\substack{j \neq i \\ j \in J}} I_{j,t-1}$$

where the industry is defined as the firm's primary 6-digit NAICS industry. Note that there is a possible bias estimating (7a) using this rivalry measure because the dependent variable is implicitly related to $S_{i,t-1}^1$. In the Appendix, we analyze this bias and estimate it in our data, showing that does not meaningfully affect our conclusions.

This simple metric is our preferred measure of intangible rivalry. However, while the overwhelming majority of firms sell in only one industry, there are many multi-industry firms. While industry activity is typically identified at the establishment, investment in R&D and software in our data is recorded at the firm

[18] Regressions using capital stocks typically use last year's end-of-year stock to avoid confounding the installed capital with current year investment. Also, gestation lags mean that only rivals' earlier investments will affect the subject firm's prospects (Pakes and Schankerman 1979). In any case, current R&D is highly correlated with R&D capital stocks (Hall, Griliches, and Hausman 1984).

level. If firms direct their innovative activity predominately to their main industry, $S_i^1$ is a reasonable measure where it is calculated using the firm's primary industry and the regression then captures the effect of rivalry in that industry. Otherwise, the firm's intangible investment might be divided proportionally across the various industries where the firm is active. To capture this case, we define an alternative measure of spillover/rivalry:

$$S_{i,t-1}^2 = \sum_k e_{ik,t-1} \cdot \sum_{j \neq i} e_{jk,t-1} I_{j,t-1}$$

where $e_{ik}$ is the share of firm *i*'s employment in industry *k*.[19]

It is possible that both of these measures might be endogenous. For instance, an increase in "technological opportunity" in an industry might both increase the dependent variable and also increase the spillover/rivalry measure, biasing coefficient $\theta_S$ upwards. For R&D, we are able to use an instrumental variable approach developed by Bloom et al. (2013) and described in the Appendix. The instrumental variable is the log of the tax cost of R&D capital. The tax cost of R&D is the after-tax cost of each dollar of R&D investment, taking into account state and Federal tax rates, R&D tax credits, and depreciation allowances. The tax cost varies with changes in the state and Federal tax treatments of R&D that are plausibly exogenous from specific national industry conditions. It's worth noting that, although this instrument has been used in the literature, it does raise a possible violation of the exclusion restriction: if rival firms perform R&D in the same state as the focal firm, then that state's R&D tax cost will be correlated with both the dependent variable and the

[19] In calculating this measure, we exclude establishments that are management facilities (NAICS 55) or intangible producing establishments (NAICS 54). We also limit the calculation to the firm's four largest industries. Another approach used by Bloom et al. (2013) measures rivalry effects by weighting other firm investments by their "closeness" to firm *i*, measured as the cosine similarity of their shares in different industries (see also Lucking, Bloom, and Van Reenen 2019; Arora, Belenzon, and Sheer 2021). For example, consider three firms: two have 75% of their activity in automobile manufacture and 25% in auto loans; the third has 75% of its activity in automobile manufacture and 25% in tractor manufacture. The first two are "closer" to each other, but for our purposes, the effect of automobile innovations resulting from intangible investment would be the same across all three. Hence, we prefer $S_i^2$ which will reflect that difference. Also, the cosine similarity measure is prohibitively time-consuming to calculate for our large sample using 6-digit industries; Bloom et al. use 4-digit industry segments in a small sample of R&D-performing firms.

instrument. In the Appendix we show that any such effect is not significant, and we find very similar results using an instrument that excludes rivals in the same state.

Finally, in our data a few percent of the estimates of $S_i$ are either very small or zero. These observations turn out to be influential outliers and, we suspect, they reflect incomplete capture in the surveys. To avoid outlier problems, we replace $\ln S_{it-1}$ with the inverse hyperbolic sine, $\operatorname{asinh} S_{it-1} = \ln\left(S_{it-1} + \sqrt{1 + S_{it-1}^2}\right)$.

## Regression results

### Log investment

Table 4 shows regressions on the log of real intangible investment in R&D and total software, including both OLS and IV estimates for R&D, for those firms that choose to invest.[20] We standardize the independent variables by dividing each by its standard deviation in the sample. Then the coefficient represents the effect of a standard deviation increase in the independent variable. Standard errors are clustered by firm. The coefficients are similar across all columns. These regressions include 6-digit industry fixed effects and year fixed effects.

Naturally, this regression sample excludes firms that do not invest, that is, where $\hat{I}_i < T$. This implies some truncation bias in our regression estimates. However, assuming that errors are normally distributed, the standard result finds that this bias *attenuates* coefficient estimates. Hence, the true coefficients should have larger absolute magnitudes and any coefficient tests that reject the null hypothesis should still hold true.

Numerous coefficients are statistically and economically significant, including firm age (strongly negative) and the number of industries the firm sells in (strongly positive). The geographic dispersion is strongly positive for software but not for R&D. Rivals' patents do not have a significant association.[21]

[20] We do not include own account software in this table because the sample is too small (see row 1, Table 2).

[21] In the table we use the lagged number of patents granted. As a robustness check, we also used the inverse hyperbolic sine of lagged patent granted. Results were similar.

However, the most important independent variables appear to be lagged rivals' investments and firm sales. Even though we include industry fixed effects, rivals' investments are associated with a large decrease in firm investment. The coefficients imply that a standard deviation increase in rivals' investment is associated with investment falling by about one half (-.76 and -.58 log points) for small firms on average. These large coefficients imply a substantial negative pecuniary externality. However, these estimates might be biased because of endogeneity as some exogenous factor can affect both the dependent variable and rivals' investments--biasing the coefficient upward. In this case, our estimates would be understated. For R&D, we are able to construct an instrumental variable based on plausibly exogenous changes in state and federal R&D tax credits (see Appendix). The estimated IV coefficient (column 2) is negative and somewhat larger in magnitude. This suggests that the negative impact of rivals' investment on investment is significant and plausibly causal.

Also, the interaction term between rivals' investment and firm revenue is positive, meaning that the negative effect of rivals' investment is more pronounced on small firms and less so on large firms. In the Appendix, we check that these effects are robust to different measures of firm size. We can measure the size of the net externality at the sample mean as $\hat{\theta}_S \overline{\ln S_{it-1}} + \hat{\theta}_{SR} \overline{\ln S_{it-1} \cdot \ln R_{it}}$, where the bar signifies sample mean. This measure of the net externality can be positive or negative, depending on firm size. The last two rows of Table 4 show the sample means of this measure for all firms and for firms with more than 500 employees. For all firms, the mean, taken using sample weights, is negative; for larger firms only, it is significantly positive. Large firms experience net positive spillovers; small firms experience net negative product market rivalry externalities.

This finding can be compared to research by Bloom, Schankerman, and Van Reenen that looks at both the positive and negative externalities of R&D investment (2013; see also Lucking, Bloom, and Van Reenen 2019). Their analysis has the advantage of separating the positive spillover effects from negative pecuniary externalities; lacking patent and R&D data for most firms, we can only report the combined effect. However, Bloom et al. achieve this separation by focusing on mostly large, publicly listed, R&D-performing and patenting firms. While our analysis for large firms confirms such positive net externalities, by taking the advantage of a more representative sample using the Census data, we show a negative externality on average, which seems to be predominantly driven by small firms.[22] A recent meta-analysis of the literature suggests that spillover effects may be rather small (Ugur, Churchill, and Luong 2020).

As previously discussed, firms may invest proportionally across the various industries where the firm operate. Additionally, a multi-industry firm may face rivals from each of these industries. In the Appendix (Table A3), we repeat the regression using the alternative measure of industry rival investments $S^2$ as defined in the previous section. The estimates are similar. The coefficient on rivals' investment is a bit smaller in magnitude for R&D and a bit larger for software. But the negative externality appears robust to changes in industry definition.

Another concern is sample selection bias. Because we use lagged rivals' investment and because many firms are not in our survey samples every year, the regression samples differ from the original samples, especially for small firms. To correct for this selection, in the Appendix we use inverse propensity weights so that our regression results reflect the original sample more closely. We find that the negative externalities are greater with these weights.

**Likelihood of investing**

Table 5 shows the results for the basic linear probability model (7b). Because the sample size is larger, we are able to also include a column for own account software. We also add a column for purchased advertising using data from the quinquennial Census of Manufactures. Because the advertising sample does not have consecutive years for measuring lagged rival investments, we use an alternative specification using log sales and log sales squared. These regressions have fixed effects for 6-digit industry and year. Column 2 provides IV estimates for R&D, which are similar to the OLS estimates in column 1.

As in the previous table, rival investment and firm sales are important coefficients. The coefficients on rivals' investments are negative and significant,

[22] Also, they include positive spillovers from firms that are technologically close, but which are not necessarily rivals in product markets; we only look at positive spillovers *within* industries.

although not as large as in the previous table. However, compared to the baseline mean probability of investing, shown at the bottom of the table, the coefficients are quite significant. The estimated IV coefficient (column 2) is negative and somewhat larger in magnitude. This suggests that the impact of rivals' investment on the likelihood of investment is significant and plausibly causal. The other variables show similar coefficients as in the previous table. Younger firms are slightly more likely to invest; firms that invest in software tend to sell in more industries and more geographic areas, consistent with some macro models. The number of patents held by rival firms has, at best, a small positive effect on investment.

Some firms might choose to outsource innovation activity: they can outsource R&D and software development, or they can purchase intermediates that have embodied R&D or software. It is possible that many firms, especially small ones, do not invest because they outsource instead. We can check whether outsourcing substitutes for own investment by including measures of outsourcing in our linear probability model. The Census of Manufactures records various measures of outsourcing including purchased data processing and other computer services and purchased professional and technical services. Therefore, we restrict our sample to the manufacturing sector and present such estimates in Table 6. Because we use the quinquennial Census of Manufactures to gain a complete look at manufacturing firms, we cannot construct our lagged measures of rivals' investment, so we use an alternative specification that includes a quadratic in log sales. Outsourced professional and technical services have positive coefficients, suggesting that they complement own investment instead of substituting. Outsourced software and data processing services have positive coefficients that are not statistically significant except for the total software regression which has a small and marginally significant negative coefficient. For intermediates, the coefficients are not significant except for the total software regression, which has a significant positive coefficient. Overall, outsourcing seems to serve more as a complement to own investment than as a substitute.

This regression also includes a measure of advanced management practices. These have a substantial positive association with R&D and advertising investment.

### Summary

Why do so few firms invest in innovation, especially as asset prices have fallen? The evidence largely rules out several possible explanations for widespread non-investment in R&D and software. It is not just a matter that firms fail to report investment on Census surveys. Nor does widespread non-investment arise because firms outsource R&D or software development rather than make their own investments; instead, while many firms do outsource this activity, outsourcing appears to complement own investment rather than to act as a substitute. And while industry is an important factor influencing these investments, it is not dispositive. Even in industries that include heavy investors, many other firms do not invest.

We find a variety of factors associated with firm decisions to invest in intangibles and with the level of investment chosen by investing firms, including firm age, product diversity, and management practices. But the two strongest factors associated with firm investment decisions are firm size, which is positively correlated with investment, and rival firm intangible investments, which are negatively correlated. Moreover, for R&D, instrumental variable estimation suggests that the investment-suppressing effect of rivals' investments is plausibly causal.

While many factors affect firm decisions whether to invest in intangibles or not, these last two factors are important because they can help explain the high degree of concentration in investment that we observe in the data, namely, that a small number of mostly large firms account for most of the investment in capitalized software and R&D and almost all growth in these investments, even as asset prices have dropped. Technological rivalry appears central to understanding the intangible divide. Firms compete fiercely by introducing innovations that obsolesce rivals' competitive advantages. But risk of obsolescence affects firms unevenly, tending to reinforce differences in investment.

## Conclusion

The intangible divide is substantial and persistent. A few hundred firms account for the lion's share of investment in R&D and capitalized software, and this divide has persisted or even grown in the face of substantial declines in the price of developing software. If trends continue, it seems likely that the distribution of investment in innovation will remain highly skewed, mostly conducted by a small elite of firms.

Is this a problem for social welfare? For tangible investments, a skewed distribution is not too troubling because market prices may serve to allocate these assets to their best uses. However, with innovation-related assets, we find significant externalities and large non-market activity—firms mainly build their own knowledge assets. That means that the allocation to firms is not necessarily the most advantageous. Asymmetric negative pecuniary externalities imply that some firms may invest more than is socially optimal because they gain rents from business stealing; other firms will invest too little because they face excessive risks of obsolescence. While the former firms may prefer investing in areas that help maintain their market power, these latter firms may have valuable technological opportunities that go undeveloped.

There are several reasons to be concerned about highly skewed investment in innovation. First, a skewed distribution of investment in innovation might be an ineffective way to innovate. For example, Cohen, reviewing the literature on R&D (2010, 140), argues that the rate of technical advance depends not just on the level of aggregate R&D investment but also on the distribution of innovative activity across firms of different sizes and capabilities. Indeed, if individual firms face diminishing returns to investments in R&D and software, as is widely assumed, then the social planner will optimally prefer a distribution that includes many small investments. And, indeed, Agrawal et al. (2014) compare innovation across regions and find that the most innovative regions have both large R&D labs and a sizeable population of small ones. Moreover, firms of diverse sizes make heterogeneous sorts of innovations that might be complementary: large firm innovations may be more oriented to processes or exploiting known technologies while small firms might be more oriented to new products and experimenting with new technologies (Pavitt, Robson, and Townsend 1987; Scherer 1991; Cohen and Klepper 1996a; 1996b; Ferreira, Manso, and Silva 2014). No firm has a monopoly on good ideas, so a high concentration of firms that invest in innovation may imply an inefficiently narrow "gene pool." And while small firms may face a larger negative externality from technological rivalry, they nevertheless make important disruptive innovations.

Second, if investments in innovation provide firms with competitive advantages, then a skewed distribution might raise industry concentration and markups. Indeed, some research relates growing intangible investment to rising industrial concentration and markups (Bessen 2020; Calligaris, Criscuolo, and Marcolin 2018; Crouzet and Eberly 2018; Brynjolfsson, Jin, and Wang 2023; Mouel and Schiersch 2023; Lashkari, Bauer, and Boussard 2024). Greater market power may imply social welfare losses and excessive political power.

Third, to the extent that the intangible divide is driven by the asymmetric impact of rival firm investments, greater skewness in innovation investment may limit the growth of new and small firms, even new firms with highly productive technologies. The growth of these firms is an important component of aggregate productivity growth (Haltiwanger et al. 2016) and also may be an important channel of social mobility, especially for minorities or women.

Thus, skewed investment and widespread non-investment of intangibles might create real problems, but the impact of skewed investment in intangibles is not well studied, making definitive conclusions difficult. More research can reveal the extent to which the intangible divide affects innovation, productivity growth, and industry dynamism.

# References


Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J Klenow, and Huiyu Li. 2019. "A Theory of Falling Growth and Rising Rents." National Bureau of Economic Research.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica: Journal of the Econometric Society*, 323–51.

Agrawal, Ajay, Iain Cockburn, Alberto Galasso, and Alexander Oettl. 2014. "Why Are Some Regions More Innovative than Others? The Role of Small Firms in the Presence of Large Labs." *Journal of Urban Economics* 81:149–65.

Arkolakis, Costas. 2010. "Market Penetration Costs and the New Consumers Margin in International Trade." *Journal of Political Economy* 118 (6): 1151–99.

Arora, Ashish, Sharon Belenzon, and Lia Sheer. 2021. "Knowledge Spillovers and Corporate Investment in Scientific Research." *American Economic Review* 111 (3): 871–98.

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 (2): 645–709.

Bar-Isaac, Heski, and Steven Tadelis. 2008. *Seller Reputation*. Now Publishers Inc.

Basker, Emek. 2007. "The Causes and Consequences of Wal-Mart's Growth." *Journal of Economic Perspectives* 21 (3): 177–98.

Bessen, James. 2020. "Information Technology and Industry Concentration." *Journal of Law and Economics* 63 (3): 531–55.
———. 2022. *The New Goliaths: How Corporations Use Software to Dominate Industries, Kill Innovation, and Undermine Regulation*. New Haven: Yale University Press.
Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen. 2019. "What Drives Differences in Management Practices?" *American Economic Review* 109 (5): 1648–83.
Bloom, Nicholas, Steven J Davis, Lucia Foster, Brian Lucking, Scott Ohlmacher, and Itay Saporta-Eksten. 2020. "Business-Level Expectations and Uncertainty." *US Census*, no. Working Paper Number CES-20-41. https://www.census.gov/library/working-papers/2020/adrm/CES-WP-20-41.html.
Bloom, Nicholas, Mark Schankerman, and John Van Reenen. 2013. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica* 81 (4): 1347–93.
Bound, John, Clint Cummins, Zvi Griliches, Bronwyn H Hall, and Adam B Jaffe. 1982. "Who Does R&D and Who Patents?" National Bureau of Economic Research.
Brynjolfsson, Erik, and Lorin Hitt. 1996. "Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending." *Management Science* 42 (4): 541–58.
Brynjolfsson, Erik, Lorin M Hitt, and Shinkyu Yang. 2002. "Intangible Assets: Computers and Organizational Capital." *Brookings Papers on Economic Activity* 2002 (1): 137–81.
Brynjolfsson, Erik, Wang Jin, and Xiupeng Wang. 2023. "Information Technology, Firm Size, and Industrial Concentration." National Bureau of Economic Research.
Caballero, Ricardo J, and Adam B Jaffe. 1993. "How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth." *NBER Macroeconomics Annual* 8:15–74.
Cabral, Luis, and Jose Mata. 2003. "On the Evolution of the Firm Size Distribution: Facts and Theory." *American Economic Review* 93 (4): 1075–90.
Calligaris, Sara, Chiara Criscuolo, and Luca Marcolin. 2018. "Mark-Ups in the Digital Era." OECD.
Cappelen, Ådne, Pierre Mohnen, Arvid Raknerud, and Marina Rybalka. 2023. "A New Approach to Estimating Private Returns to R&D."
Caves, Richard E. 1998. "Industrial Organization and New Findings on the Turnover and Mobility of Firms." *Journal of Economic Literature* 36 (4): 1947–82.
Christensen, Laurits R, Dale W Jorgenson, and Lawrence J Lau. 1973. "Transcendental Logarithmic Production Frontiers." *The Review of Economics and Statistics*, 28–45.
Cohen, Wesley M. 2010. "Fifty Years of Empirical Studies of Innovative Activity and Performance." In *Handbook of the Economics of Innovation*, 1:129–213. Elsevier.
Cohen, Wesley M, and Steven Klepper. 1996a. "A Reprise of Size and R & D." *The Economic Journal* 106 (437): 925–51.

———. 1996b. “Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D.” *The Review of Economics and Statistics*, 232–43.

Cohen, Wesley M, and Daniel A Levinthal. 1989. “Innovation and Learning: The Two Faces of R & D.” *The Economic Journal* 99 (397): 569–96.

Corrado, Carol, Charles Hulten, and Daniel Sichel. 2009. “Intangible Capital and US Economic Growth.” *Review of Income and Wealth* 55 (3): 661–85.

Crouzet, Nicolas, and Janice Eberly. 2018. “Intangibles, Investment, and Efficiency.” In *AEA Papers and Proceedings*, 108:426–31.

De Ridder, Maarten. 2023. “Market Power and Innovation in the Intangible Economy.” *American Economic Review*.

Dinlersoz, Emin M, and Mehmet Yorukoglu. 2012. “Information and Industry Dynamics.” *American Economic Review* 102 (2): 884–913.

Doms, Mark E., and Timothy Dunne. 1998. “Capital Adjustment Patterns in Manufacturing Plants.” *Review of Economic Dynamics* 1 (2): 409–29. https://doi.org/10.1006/redy.1998.0009.

Doraszelski, Ulrich, and Jordi Jaumandreu. 2013. “R&D and Productivity: Estimating Endogenous Productivity.” *Review of Economic Studies* 80 (4): 1338–83.

Dreisigmeyer, David, Nathan Goldschlag, Marina Krylova, Wei Ouyang, Elisabeth Perlman, and others. 2018. “Building a Better Bridge: Improving Patent Assignee-Firm Links.” Center for Economic Studies, US Census Bureau.

Drozd, Lukasz A, and Jaromir B Nosal. 2012. “Understanding International Prices: Customers as Capital.” *American Economic Review* 102 (1): 364–95.

Dunne, Timothy, Mark J Roberts, and Larry Samuelson. 1988. “Patterns of Firm Entry and Exit in US Manufacturing Industries.” *The RAND Journal of Economics*, 495–515.

Ferreira, Daniel, Gustavo Manso, and André C Silva. 2014. “Incentives to Innovate and the Decision to Go Public or Private.” *The Review of Financial Studies* 27 (1): 256–300.

Fleming, Martin. 2023. “Enterprise Information and Communications Technology: Software Pricing and Developer Productivity Measurement.”

Fort, Teresa C, Shawn D Klimek, and others. 2016. “The Effects of Industry Classification Changes on US Employment Composition.” *Tuck School at Dartmouth*.

Foster, Lucia, Cheryl Grim, and Nikolas Zolas. 2016. “A Portrait of Firms That Invest in R&D.” Working Papers 16–41. Center for Economic Studies, U.S. Census Bureau. https://ideas.repec.org/p/cen/wpaper/16-41.html.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. “The Slow Growth of New Plants: Learning about Demand?” *Economica* 83 (329): 91–129.

Gourio, Francois, and Leena Rudanko. 2014. “Customer Capital.” *Review of Economic Studies* 81 (3): 1102–36.

Griliches, Zvi. 1979. “Issues in Assessing the Contribution of Research and Development to Productivity Growth.” *The Bell Journal of Economics*, 92–116.

———. 1980. “Returns to Research and Development Expenditures in the Private Sector.” In *New Developments in Productivity Measurement*, 419–62. University of Chicago press.

Grossman, Gene M, and Elhanan Helpman. 1991. "Quality Ladders in the Theory of Growth." *The Review of Economic Studies* 58 (1): 43–61.
Hall, Bronwyn H, Zvi Griliches, and Jerry A Hausman. 1984. "Patents and R&D: Is There a Lag?" National Bureau of Economic Research.
Hall, Bronwyn H, Jacques Mairesse, and Pierre Mohnen. 2010. "Measuring the Returns to R&D." In *Handbook of the Economics of Innovation*, 2:1033–82. Elsevier.
Haltiwanger, John, Russell Cooper, and Laura Power. 1999. "Machine Replacement and the Business Cycle: Lumps and Bumps." *American Economic Review* 89 (4): 921–46.
Haltiwanger, John, Ron S. Jarmin, Robert Kulick, and Javier Miranda. 2016. "High Growth Young Firms: Contribution to Job, Output, and Productivity Growth." In *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*, 11–62. University of Chicago Press. https://www.nber.org/books-and-chapters/measuring-entrepreneurial-businesses-current-knowledge-and-challenges/high-growth-young-firms-contribution-job-output-and-productivity-growth.
Hamberg, Daniel. 1964. "Size of Firm, Oligopoly, and Research: The Evidence." *Canadian Journal of Economics and Political Science/Revue Canadienne de Economiques et Science Politique* 30 (1): 62–75.
Haskel, Jonathan, and Stian Westlake. 2018. *Capitalism without Capital: The Rise of the Intangible Economy*. Princeton University Press.
Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2019. "The Industrial Revolution in Services." National Bureau of Economic Research.
Jarmin, Ron S, and Javier Miranda. 2002. "The Longitudinal Business Database." *Available at SSRN 2128793*.
Koh, Ping-Sheng, David M Reeb, Elvira Sojli, Wing Wah Tham, and Wendun Wang. 2022. "Deleting Unreported Innovation." *Journal of Financial and Quantitative Analysis* 57 (6): 2324–54.
Lashkari, Danial, Arthur Bauer, and Jocelyn Boussard. 2024. "Information Technology and Returns to Scale." *American Economic Review* 114 (6): 1769–1815.
Li, Wendy CY, and Bronwyn H Hall. 2020. "Depreciation of Business R&D Capital." *Review of Income and Wealth* 66 (1): 161–80.
Lucking, Brian, Nicholas Bloom, and John Van Reenen. 2019. "Have R&D Spillovers Declined in the 21st Century?" *Fiscal Studies* 40 (4): 561–90.
Ma, Song. 2021. "Technological Obsolescence." National Bureau of Economic Research.
Mansfield, Edwin. 1968. *Industrial Research and Technological Innovation: An Econometric Analysis*. New York: Norton for the Cowles Foundation for Research in Economics, Yale University.
Mansfield, Edwin, Mark Schwartz, and Samuel Wagner. 1981. "Imitation Costs and Patents: An Empirical Study." *The Economic Journal* 91 (364): 907–18.
Mezzanotti, Filippo, and Timothy Simcoe. 2023. "Innovation and Appropriability: Revisiting the Role of Intellectual Property." *NBER Working Paper*, no. 31428.

Mouel, Marie Le, and Alexander Schiersch. 2023. "Intangible Capital and Productivity Divergence." *Review of Income and Wealth* n/a (n/a). https://doi.org/10.1111/roiw.12653.

Nelson, Richard R, Merton J Peck, and Edward D Kalachek. 1967. "Technology, Economic Growth, and Public Policy; a Rand Corporation and Brookings Institution Study."

Nordhaus, William D. 1967. "The Optimal Life of a Patent." *Cowles Foundation Discussion Papers*, November.

Pakes, Ariel, and Mark Schankerman. 1979. "The Rate of Obsolescence of Knowledge, Research Gestation Lags, and the Private Rate of Return to Research Resources." National Bureau of Economic Research.

Pavitt, Keith, Michael Robson, and Joe Townsend. 1987. "The Size Distribution of Innovating Firms in the UK: 1945-1983." *The Journal of Industrial Economics*, 297–316.

Perla, Jesse. 2013. "Product Awareness and the Industry Life Cycle." *Unpublished Manuscript, New York University*.

Peters, Ryan H, and Lucian A Taylor. 2017. "Intangible Capital and the Investment-q Relation." *Journal of Financial Economics* 123 (2): 251–72.

Radner, Roy. 2003. "Viscous Demand." *Journal of Economic Theory* 112 (2): 189–231.

Rob, Rafael, and Arthur Fishman. 2005. "Is Bigger Better? Customer Base Expansion through Word-of-Mouth Reputation." *Journal of Political Economy* 113 (5): 1146–62.

Scherer, Frederic M. 1972. "Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation." *The American Economic Review* 62 (3): 422–27.

———. 1991. "Changing Perspectives on the Firm Size Problem, In ZJ Acs and DB Audretsch (Eds.), Innovation and Technological Change: An International Comparison, Pp. 24-38." University of Michigan Press.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming up Inequality." *The Quarterly Journal of Economics* 134 (1): 1–50.

Tambe, Prasanna, and Lorin M Hitt. 2012. "The Productivity of Information Technology Investments: New Evidence from IT Labor Data." *Information Systems Research* 23 (3-part-1): 599–617.

Ugur, Mehmet, Sefa Awaworyi Churchill, and Hoang M Luong. 2020. "What Do We Know about R&D Spillovers and Productivity? Meta-Analysis Evidence on Heterogeneity and Statistical Power." *Research Policy* 49 (1): 103866.

Ugur, Mehmet, Eshref Trushin, Edna Solomon, and Francesco Guidi. 2016. "R&D and Productivity in OECD Firms and Industries: A Hierarchical Meta-Regression Analysis." *Research Policy* 45 (10): 2069–86.

Villalonga, Belén. 2004. "Intangible Resources, Tobin'sq, and Sustainability of Performance Differences." *Journal of Economic Behavior & Organization* 54 (2): 205–30.

Villard, Henry H. 1958. "Competition, Oligopoly, and Research." *Journal of Political Economy* 66 (6): 483–97.

Zhang, Lichen. 2019. "Intangibles, Concentration, and the Labor Share." Mimeo.

# Figures

Figure 1 Share of US Private Nonresidential Investment, 61 Industries
Source: BLS

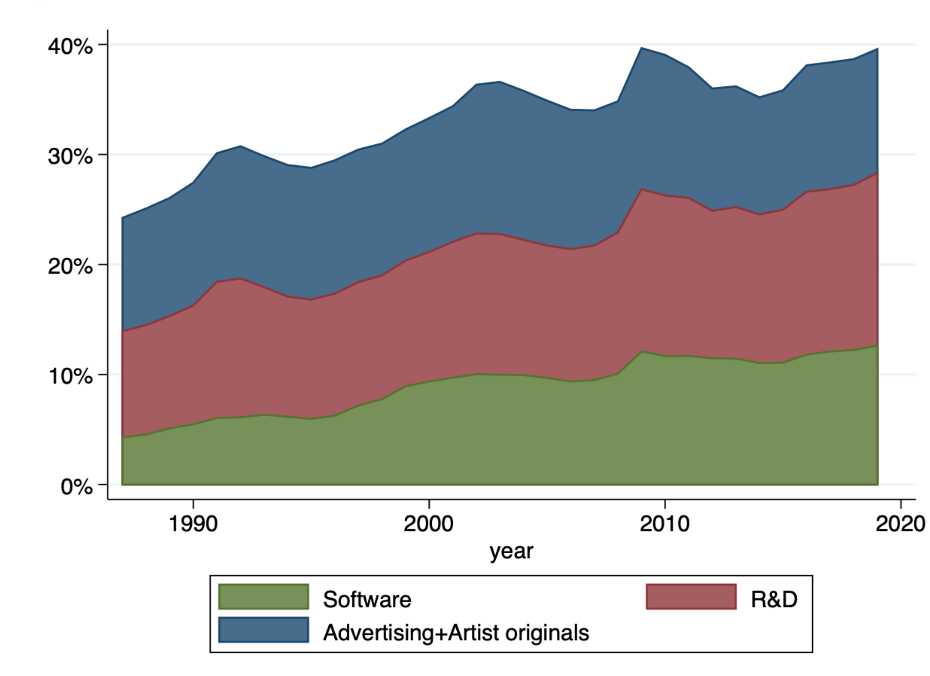
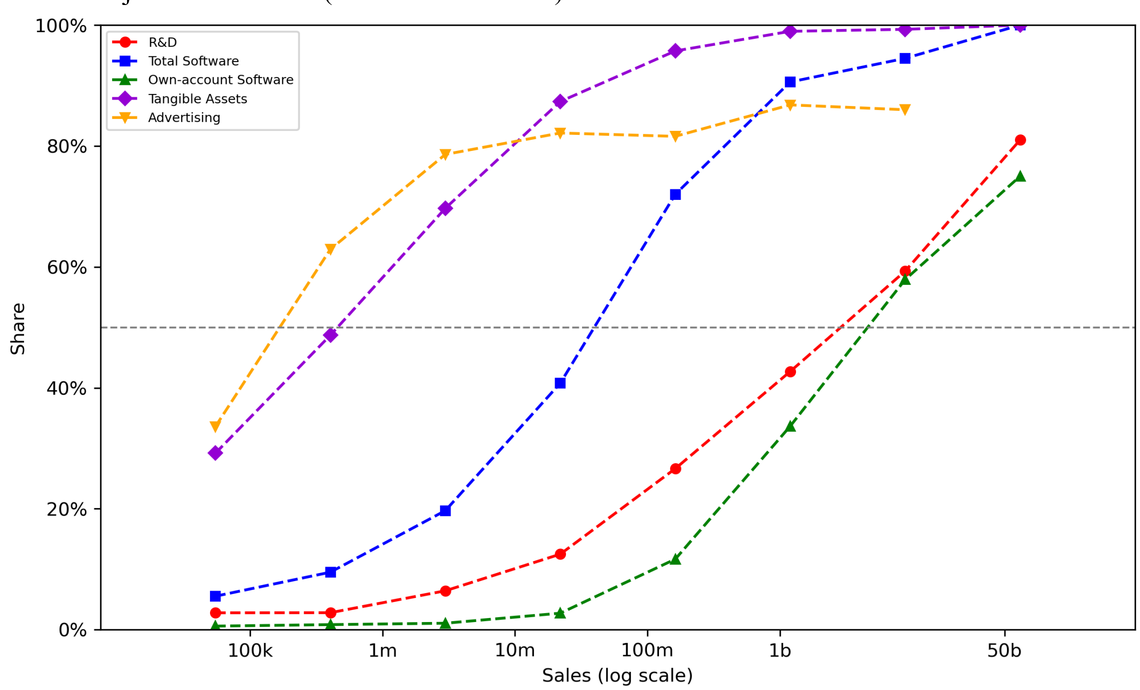



**Figure 2. Share of firms reporting positive investment in asset type by firm real sales ($2012)**
FSRDC Project Number 2735 (CBDRB-FY23-0284)



# Tables

### Table 1. Top 250 firms dominate investment, 2017/8

| | (1) R&D | (2) Total software | (3) Own account software | (4) Advertising | (5) Tangible assets |
|---|---|---|---|---|---|
| Top firms, share of investment | 83% | 67% | 85% | 68% | 55% |
| **Across many industries** | | | | | |
| Top firms, # of 6-digit industries | 101 | 109 | 108 | 138 | 97 |
| **Top firms are big, but not all the biggest** | | | | | |
| Top firms, share of revenue | 17% | 23% | 19% | 18% | 26% |
| Top firms, mean sales; (Billion $2012) | 11.23 | 19.27 | 16.39 | 5.84 | 22.13 |
| Share of mega-firms in top 250 | 23% | 38% | 31% | 12% | 49% |
| **Top firms are persistent and in distinct groups** | | | | | |
| Continuity Percent | 72% | 53% | 53% | 51% | 64% |
| Top 250 overlap with top R&D | | 21% | | | |

Note: Data from 2018 except for advertising, which is from 2017. All estimates are weighted using Census population weights for each survey (excluding advertising data). Advertising estimates come from the Census of Manufactures; R&D data come from BRDIS; other data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. FSRDC Project Number 2735 (CBDRB-FY23-0284)

**Table 2. Most firms do not invest in R&D and Software**

| | (1) R&D | (2) Total software | (3) Own account software | (4) Advertising | (5) Tangible assets |
|---|---|---|---|---|---|
| Share of firms with no investment | 97% | 93% | 99% | 32% | 63% |
| Share of sales of non-investors | 62% | 35% | 62% | 18% | 11% |
| **Even in industries that include top 250 firms** | | | | | |
| Top firms, # of 6-digit industries | 101 | 109 | 108 | 138 | 97 |
| Share of firms with no investment | 88% | 94% | 99% | 25% | 67% |
| Share of sales of non-investors | 25% | 21% | 49% | 9% | 3% |

Note: Reporting firms that enter 0 investment (observations with missing data are excluded) in any year. Data from 2018 except for advertising, which is from 2017. For each asset, the top 250 firms are selected annually. All estimates are weighted using Census population weights for each survey (excluding advertising data). Advertising estimates come from the Census of Manufactures; R&D data come from BRDIS; other data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. FSRDC Project Number 2735 (CBDRB-FY23-0284)

Table 3. Top firms grow more dominant while non-investing share increases

| | R&D ($ billions) | | | Total SW Capex ($ billions) | | | Own-account SW Capex ($ billions) | | | Purchased Advertising ($ billions) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Year | 2009 | 2018 | Δ | 2008 | 2018 | Δ | 2008 | 2018 | Δ | 2007 | 2017 | Δ |
| **Intangible investment** | | | | | | | | | | | | |
| All firms | 165 | 227 | 62 | 52 | 66 | 14 | 17 | 31 | 14 | 10 | 15 | 6 |
| Top 250 firms | 132 | 189 | 58 | 30 | 44 | 14 | 14 | 26 | 12 | 6 | 10 | 4 |
| Other firms | 33 | 38 | 4 | 22 | 22 | 0 | 3 | 5 | 2 | 3 | 5 | 1 |
| Top 250 share | 80% | 83% | | 58% | 67% | | 82% | 85% | | 57% | 68% | |
| Top 250 share of increase | | | 93% | | | 99% | | | 89% | | | 76% |
| **Non-investing firms** | | | | | | | | | | | | |
| Share of firms with no investment | 96.8% | 97.3% | 0.5% | 90.0% | 93.4% | 3.4% | 99.6% | 99.2% | -0.4% | 19.0% | 32.1% | 13.1% |
| Share of sales of non-investing firms | 60.2% | 61.7% | 1.5% | 21.7% | 34.6% | 12.9% | 60.7% | 62.3% | 1.6% | 22.2% | 17.6% | -4.6% |

Note: Investment amounts are in billions of 2012 dollars. For each asset, the top 250 firms are selected annually. All estimates are weighted using Census population weights for each survey (excluding advertising data). Advertising estimates come from the Census of Manufactures; R&D data come from BRDIS; other data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. For this reason, totals are less than corresponding totals in the National Accounts. In addition, these are deflated figures and exclude expensed software.
FSRDC Project Number 2735 (CBDRB-FY23-0284)

Table 4. Log investment demand

| | (1) | (2) | (3) |
|---|---|---|---|
| | Log R&D | Log R&D | Log Total software |
| | OLS | IV | OLS |
| Asinh rivals' investment$_{t-1}$ | -0.7592*** | -0.8846*** | -0.5834*** |
| | (0.1444) | (0.1474) | (0.2052) |
| Log sales | 1.229*** | 1.217*** | 1.204*** |
| | (0.1037) | (0.0968) | (0.1758) |
| Asinh rivals' investment$_{t-1}$ x Log sales | 0.1443*** | 0.1623*** | 0.1343*** |
| | (0.0267) | (0.0274) | (0.0395) |
| Firm age | -0.1011*** | -0.1014*** | -0.0805** |
| | (0.0269) | (0.0269) | (0.0400) |
| No. of industries | 0.1392*** | 0.1381*** | 0.2557*** |
| | (0.0154) | (0.0153) | (0.0215) |
| No. of zip codes | -0.0035 | -0.0029 | 0.1239*** |
| | (0.0236) | (0.0235) | (0.0152) |
| Rivals' patents | 0.0288 | 0.0296 | 0.0229 |
| | (0.0255) | (0.0255) | (0.0220) |
| Adjusted $R^2$ | 0.6115 | 0.6118 | 0.7441 |
| N (rounded) | 38000 | 38000 | 71000 |
| Mean net externality | -0.3633 | -0.4717 | -0.6746 |
| Mean net externality (employment>500) | 0.3022 | 0.2073 | 0.2213 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS; software data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include fixed effects for 6-digit industry and year and use sample weights. Rivals' investment is lagged investment of all other firms in the same 6-digit industry. The net externality is calculated as $\hat{\theta}_S \overline{\ln S_{it-1}} + \hat{\theta}_{SR} \overline{\ln S_{it-1} \cdot \ln R_{it}}$ where the bars signify sample means. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407).

**Table 5. Likelihood of investing**

Dependent variable: 1 if invest in year, 0 otherwise

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | R&D | R&D | Total software | Own account software | Purchased advertising |
| | OLS | IV | OLS | OLS | OLS |
| Asinh rivals' investment$_{t-1}$ | -0.0356*** | -0.0653*** | -0.0313*** | -0.0365** | |
| | (0.0127) | (0.0134) | (0.0053) | (0.0166) | |
| Log sales | 0.0959*** | 0.0828*** | 0.0019 | 0.0607*** | 0.7010*** |
| | (0.0107) | (0.0091) | (0.0023) | (0.0161) | (0.0200) |
| (Log sales)$^2$ | | | | | -0.5389*** |
| | | | | | (0.0212) |
| Asinh rivals' investment$_{t-1}$ | 0.0081*** | 0.0137*** | 0.0076*** | 0.0105*** | |
| x Log sales | (0.0030) | (0.0029) | (0.0011) | (0.0039) | |
| Firm age | -0.0107** | -0.0106** | -0.0040** | -0.0062 | 0.0243*** |
| | (0.0052) | (0.0052) | (0.0016) | (0.0041) | (0.0046) |
| No. of industries | 0.0106*** | 0.0096*** | 0.0331*** | 0.0644*** | 0.0377*** |
| | (0.0028) | (0.0028) | (0.0027) | (0.0055) | (0.0019) |
| No. of zip codes | -0.0069* | -0.0054* | 0.0317*** | 0.0140*** | .0039 |
| | (0.0036) | (0.0033) | (0.0036) | (0.0025) | -- |
| Rivals' patents | 0.0102 | 0.0106 | 0.0089*** | 0.0183** | |
| | (0.0102) | (0.0102) | (0.0026) | (0.0072) | |
| Mean dep. Variable 2017/8 | 0.027 | 0.027 | 0.066 | 0.008 | |
| Adjusted $R^2$ | 0.3804 | 0.3809 | 0.1075 | 0.2431 | .0722 |
| N (rounded) | 59000 | 59000 | 118000 | 118000 | 219000 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS; software data from ACES; advertising data from the Census of Manufactures. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include fixed effects for 6-digit industry and year. Rivals' investment is lagged investment of all other firms in the same 6-digit industry. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074).

**Table 6. Likelihood of investing, extended variables for manufacturing**
Dependent variable: 1 if invest in year, 0 otherwise

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | R&D | Total software | Own account software | Purchased advertising |
| Log sales | 0.2872* | 1.046** | -0.8183*** | 0.3586*** |
| | (0.1675) | (0.4895) | (0.1177) | (0.0488) |
| (Log sales)$^2$ | -0.0398 | -0.5272 | 1.047*** | -0.2378*** |
| | (0.1569) | (0.4635) | (0.1289) | (0.0401) |
| Firm age | 0.013 | -0.0761 | -0.021* | 0.0501*** |
| | (0.0226) | (0.0464) | (0.0112) | (0.0072) |
| No. of industries | 0.0186 | 0.0026 | 0.0617** | 0.0240*** |
| | (0.0162) | (0.0298) | (0.0259) | (0.0033) |
| No. of Zip codes | 0.0099 | -0.0169 | 0.0555** | 0.0032 |
| | (0.0133) | (0.0189) | (0.0264) | (0.0028) |
| Prof., technical services/ Sales | 0.0246** | 0.0223** | 0.0046 | 0.2505*** |
| | (0.0115) | (0.0101) | (0.0084) | (0.0662) |
| Computer, SW services/ Sales | 0.0244 | -0.1736* | 0.074 | 0.0332 |
| | (0.0248) | (0.1001) | (0.0691) | (0.0253) |
| Intermediates / Sales | -0.0024 | 0.0709*** | -0.003 | 0.0007 |
| | (0.0087) | (0.0245) | (0.0085) | (0.0021) |
| Adv. management practices | 0.0847*** | 0.0273 | 0.0075 | 0.0266*** |
| | (0.0179) | (0.0323) | (0.0068) | (0.0063) |
| Adjusted $R^2$ | 0.3698 | 0.3268 | 0.2831 | 0.0769 |
| N (rounded) | 7900 | 4800 | 4800 | 29500 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS; software data from ACES; all other data from the Census of Manufactures. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include fixed effects for 6-digit industry and year. Advanced management practices come from the MOPS supplemental survey to the 2012 Census of Manufactures. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11407).

# Appendix for "The Intangible Divide: Why do so few firms invest in innovation?"

By James Bessen (TPRI, Boston University) and Xiupeng Wang (TPRI, Northeastern University)

December, 2024

## Model

We wish to solve for the equilibrium value of

$$\frac{d\ln \hat{I}_i}{d\ln w} = \frac{\partial \ln \hat{I}_i}{\partial \ln w} + \frac{\partial \ln \hat{I}_i}{\partial \ln S_i}\frac{d\ln S_i}{d\ln w}. \tag{A1}$$

From (3) and assuming $r = 0$, we have

$$\frac{\partial \ln \hat{I}_i}{\partial \ln w} = -\frac{1}{1-\beta}$$

$$\frac{\partial \ln \hat{I}_i}{\partial \ln S_i} = \frac{\alpha_i - \gamma_i}{1-\beta}, \quad \gamma_i \equiv \frac{\partial \ln \delta_i}{\partial \ln S_i}.$$

Also, for the two-firm case,

$$\frac{d\ln S_i}{d\ln w} = \frac{d\ln I_j}{d\ln w}.$$

Expanding (A1) and assuming a Nash equilibrium so we can substitute in the expression for $\frac{d\ln I_j}{d\ln w}$,

$$\begin{aligned}\frac{d\ln \hat{I}_i}{d\ln w} &= -\frac{1}{1-\beta} + \frac{\alpha_i - \gamma_i}{1-\beta}\cdot\frac{d\ln I_j}{d\ln w} \\ &= -\frac{1}{1-\beta} + \frac{\alpha_i - \gamma_i}{1-\beta}\left(-\frac{1}{1-\beta} + \frac{\alpha_j - \gamma_j}{1-\beta}\cdot\frac{d\ln I_i}{d\ln w}\right).\end{aligned}$$

Rearranging and simplifying,

$$\frac{d\ln \hat{I}_i}{d\ln w} = \frac{\gamma_i - \alpha_i - (1-\beta)}{(1-\beta)^2 - (\gamma_i - \alpha_i)(\gamma_j - \alpha_j)}.$$

## Robustness checks

### Under/over reporting of R&D

Figure 3 shows striking differences in investment patterns for small firms: for firms with about $500k in annual revenue, about half the firms invest in advertising and tangible assets, while less than 10% invest in software or R&D. Perhaps small firms have a harder time tracking R&D and software investments. On the other hand, perhaps large firms recategorize expenses to exaggerate R&D spending in order to earn bigger tax credits. We can check the general validity of our estimates by comparing R&D spending on personnel to labor compensation costs of scientists and engineers estimated with data from the Current Population Survey.

First, first we compare R&D personnel costs for 2013, with compensation costs for scientists and engineers, calculated as the weighted sum of weekly earnings times average weeks worked times 1.2 to cover fringe benefits.[1] These two estimates do not measure exactly the same thing: the R&D data include personnel who are not scientists and engineers; the CPS data include scientists and engineers who work outside of R&D, for example, in technical sales or quality assurance. Nevertheless, the two estimates are reassuringly similar: $214b from the R&D data, $218b for the CPS data.

Second, we compare the relative shares of R&D and science/engineering earnings across firm size classes. Even though these two estimates measure somewhat different things, we would still expect that the relative share of R&D and earnings in each size class should be similar. Table A1 shows these shares averaged over the years 2009-2018.

Table A1. Shares of expenditure by firm employment size

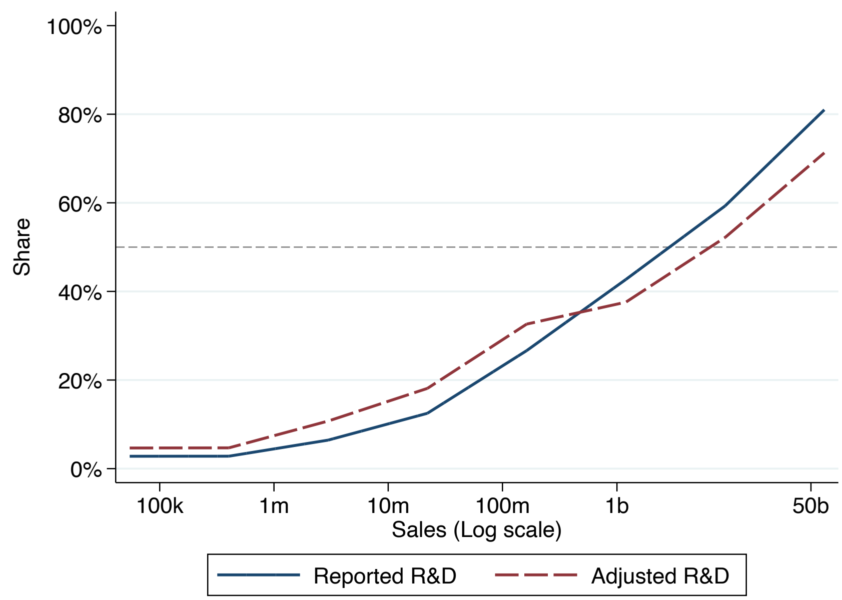
| Firm size (employees) | Share of Domestic R&D Performed by Company | Share of earnings of scientists and engineers |
|---|---|---|
| 10-99 | 7.3% | 12.2% |
| 100-499 | 8.0% | 11.6% |
| 500-999 | 4.0% | 4.9% |
| 1000+ | 80.8% | 71.3% |

[1] We exclude civil engineers, actuaries, and software developers. The R&D personnel data come from NSF public use files, National Center for Science and Engineering Statistics. 2020. Business Research and Development: 2018. NSF 21-312. Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/pubs/nsf21312/ and earlier years.

| Total | 100% | 100% |
|---|---|---|

These numbers suggest that there might, indeed, be some under-reporting of R&D by small firms and/or over-reporting by the largest firms. How significant are these possible biases? We can adjust the curve for R&D in Figure 3 to correct for this possible bias by multiplying the share of firms reporting R&D investment by the ratio of science/engineering earning to R&D expenditures for each size class. That is, assuming that the true share of R&D performed by firms with 10-99 employees is 12.2%, we multiply the share reporting investment by 12.2/7.3. The reported and adjusted shares of firms investing are shown in Figure A1.

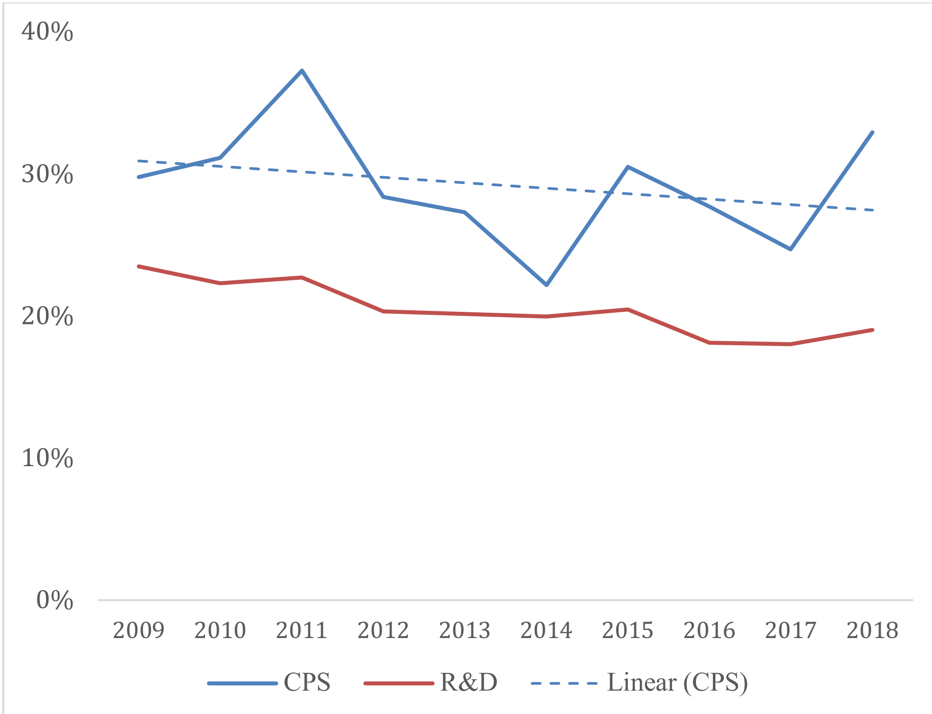
Figure A1. Share of firms that reports positive R&D investment by firm real sales




The adjustment for reporting bias does not qualitatively change the import of the figure in a significant way.

A further check comes from time trends. We compare the share of expenditures made by small and medium firms (those with fewer than 1000 employees) for R&D and scientists/engineers, respectively, and plot them in Figure A2.

Figure A2. Share of expenditures for firms with <1000 employees




While the CPS estimates of science/engineer earnings are higher and more variable than the R&D series, the trends are closely matched, suggesting that the time changes we report in our data are not substantively affected by changing reporting bias.

### Capitalized vs. Expensed Software

The US accounting rules for internal use software (GAAP 350-40) divide software development spending (own and custom) into three phases corresponding to the life cycle of the software:

1. The exploratory or research phase. This includes expenses on determining the needed functionality and exploring the feasibility of the technology.
2. Application development. Once feasibility is demonstrated, this includes the expenses of developing the application, including spending on pre-packaged software and on outside contractors.
3. Maintenance. Once the application is rolled out, this includes further expenses for debugging, upgrades, training, etc.

According to the accounting rules, only the second phase is capitalized; the others are expensed. The second phase also represents spending on developing an application that is new to the firm, so it can be considered a firm-level innovation that is intended to improve firm performance. And this is the kind of software investment we seek to measure. However, software capital expenditure likely understates innovation-related software investment for two reasons. First, some amount of research (phase 1) also constitutes investment in innovation. Second, firms often do not strictly follow the accounting rules, tending to expense software that should be capitalized (Moylan 2001; Grimm, Moulton, and Wasshausen 2005; Reed 2015; Barth, Davis, Freeman, et al. 2023, Appendix A). Small expenditures also tend to be expensed. Also, a minority of capitalized software consists of long-lived prepackaged software that might not represent innovation. However, these investments are correlated with own-account capital spending (coefficient .324), so it seems likely that much of this investment is also related to new, innovative systems. Nevertheless, capitalized software investment provides a measure that closely corresponds to the concept we wish to measure, even if it is understated.

To check the robustness of some of our findings, we can use data on expensed software reported in the Information and Communication Technology Supplement to ACES which is only available from 2003 to 2013 (excluding 2012). The table below presents some statistics comparable to figures in Tables 1-3. Note that these figures are necessarily from an earlier time period, so are not directly comparable.

Table A2. Share of investment in total software (capitalized + expensed)

| | Total software (capitalized + expensed) 2003 | 2013 |
|---|---|---|
| Top firms, share of investment | 56% | 55% |
| Top firms, share of revenue | 31% | 23% |
| Share of firms with no intangible expenditures | 79% | 82% |
| Share of total sales from firms with no intangible expenditures | 20% | 22% |
| Total real intangible (Bill.) | 69.6 | 117.4 |
| Top 250 share of increase | | 53% |
| Continuity Percent | | 50% |

Adding expensed software modestly decreases the share of investment by the top 250 firms and the share of non-investing firms. Only 53% of the increase in capitalized + expensed software is accounted for by the top 250 firms. But the share of firms with no software expenditures increased. The top 250 firms are not as dominant when expensed software is included, suggesting that smaller investors tend to expense software relatively more. This is in line with our finding that smaller firms tend to invest less in software.

### Regression Bias from Construction of Rivalry Measure

In the investment demand equation (7a) the dependent variable, $\ln I_{it}$, is implicitly related to the rivalry measure, $S^1_{i,t-1}$, and that can cause a bias in the coefficient estimate for $S_{it}$. We can see this by "partialing out" these variables, that is, using the residuals of these two variables after regressing them on other independent variables, including fixed effects and the constant term. Then the basic regression can be written in simplified form as

$$\ln I_{it} = \beta S_{it} + \epsilon_{it}, \quad S_{it} = \ln(T_{t-1} - I_{it-1}), \quad T_t = \sum_j I_{jt} \tag{A2}$$

where $I_{it}$ is the real R&D (or software) investment made by firm *i* at time *t* in a given industry, $T_{t-1}$ is the sum of these investments over all firms in that industry at time *t*-1, and $S_{it}$ is our rivalry measure.

There is a potential estimation problem because the construction of $S_{it}$ involves $I_{it-1}$ which likely covaries with the dependent variable. We can see this by looking at the OLS estimator,

$$\hat{\beta} = \frac{cov(\log I_{it}, S_{it})}{var(S_{it})}. \tag{A3}$$

To see the bias, it is helpful to decompose the error term in (1) into two parts,

$$\epsilon_{it} = \phi_i + \theta_{it} \tag{A4}$$

where although $E[\epsilon_{it}] = 0$, possibly

$$cov(\theta_{it}, \theta_{jt}), cov(\theta_{it}, \theta_{it-1}) \neq 0.$$

Plugging (A2) and (A4) into (A3), the OLS estimate is

$$\hat{\beta} = \frac{cov(\beta S_{it}, S_{it}) + cov(\phi_i, S_{it}) + cov(\theta_{it}, S_{it})}{var(S_{it})} = \beta + B,$$

$$B \equiv \frac{cov(\phi_i, S_{it})}{var(S_{it})} + \frac{cov(\theta_{it}, S_{it})}{var(S_{it})}. \tag{A5}$$

*B* is thus the bias. We expect that the first term in the bias is negative, and the second term is positive. Decomposing the first term, we can see that it is negative:

$$cov(\phi_i, S_{it}) = cov\left(\phi_i, \log T_{t-1} + \log\left(1 - \frac{I_{it-1}}{T_{t-1}}\right)\right) \approx -cov\left(\phi_i, \frac{I_{it-1}}{T_{it-1}}\right) < 0 \qquad (A6)$$

since $cov(\phi_i, \log T_{t-1}) = 0$ and $\log(1+x) \approx x$ for small $x$. The second term in the numerator, $cov(\theta_{it}, S_{it})$, is likely greater than zero because firms within an industry experience common shocks in R&D or software demand (shocks to business conditions, technological opportunities) that are serially correlated. Hence the net bias could be either positive or negative. Note that this bias only affects the investment demand equation and does not come into play in the regressions on whether firms choose to invest or not.

Fortunately, we can estimate the magnitude of this bias by proxying $\phi_i$ with the sample mean of $\log I_{it}$ for each firm, $\overline{\log I_\iota} = \frac{1}{T}\sum_t \log I_{it}$. We can show that[2]

$$\frac{cov(\phi_i, S_{it})}{var(S_{it})} \approx \frac{cov\left(\overline{\log I_\iota}, \frac{I_{it-1}}{T_{it-1}}\right)}{var(S_{it})} + \beta \frac{var\overline{\left(\frac{I_\iota}{T_\iota}\right)}}{var(S_{it})}. \qquad (A7)$$

We estimate the first term on the right and find it is small relative to the coefficient estimates for both R&D and software. We also estimate $var\overline{\left(\frac{I_\iota}{T_\iota}\right)}/var(S_{it})$ and find that it, too, is small. We conclude that the negative bias in (A5) is small and, moreover, it is offset by the positive bias term, $cov(\theta_{it}, S_{it})$. This bias, however, should be eliminated in our IV regressions because the instrument is independent of $\theta_{it}$. The IV estimates are also similar to the ordinary OLS estimates. Hence, we do not think that the net bias substantially affects our results, so we ignore it in the main text. Note that this bias is not an issue using the alternative measure of rivals' investment (next).

### Alternative measure of rivals' investment

As discussed in the text, we also measure rivals' intangible investment using a weighted average (by employment) across different industries. Here we show the investment demand and linear probability models equivalent to Tables 4 and 5 using the alternative measure of rivals' investment.

---

[2] From (A2) and (A4), $\log I_{it} = \beta S_{it} + \phi_i + \theta_{it}$ so that $\overline{\log I_\iota} = \beta \overline{S_\iota} + \phi_i$ since $E[\theta_{it}] = 0$. Further, following the analysis in (A6), $\overline{S_\iota} \approx \overline{\log T_\iota} - \overline{\left(\frac{I_\iota}{T_\iota}\right)}$ Note also that $cov\left(\overline{\left(\frac{I_\iota}{T_\iota}\right)}, \frac{I_{it-1}}{T_{it-1}}\right) = var\overline{\left(\frac{I_\iota}{T_\iota}\right)}$ so that $cov\left(\overline{\log I}, \frac{I_{it-1}}{T_{it-1}}\right) \approx cov\left(\phi_i, \frac{I_{it-1}}{T_{it-1}}\right) - \beta var\overline{\left(\frac{I_\iota}{T_\iota}\right)}$.

Table A3. Log investment, alternative rival measure

| | (1) | (2) |
|---|---|---|
| | Log R&D | Log Total software |
| | OLS | OLS |
| Asinh rivals' investment$_{t-1}$ | -0.5701*** | -0.6318*** |
| | (0.1427) | (0.1434) |
| Log sales | 0.9593*** | 0.6901*** |
| | (0.1661) | (0.2251) |
| Asinh rivals' investment$_{t-1}$ x Log sales | 0.1472*** | 0.1800*** |
| | (0.0295) | (0.0374) |
| Firm age | -0.0905*** | -0.081** |
| | (0.0279) | (0.0400) |
| No. of industries | 0.1309*** | 0.2495*** |
| | (0.0153) | (0.0213) |
| No. of zip codes | -0.0004 | 0.1247*** |
| | (0.0239) | (0.0152) |
| Rivals' patents | 0.0248 | 0.0131 |
| | (0.0251) | (0.0218) |
| Adjusted $R^2$ | 0.6144 | 0.7457 |
| N (rounded) | 34000 | 71000 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS; software data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include fixed effects for 6-digit industry and year and use sample weights. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407).

Table A4. Likelihood of investing, alternative rival investment

Dependent variable: 1 if invest in year, 0 otherwise

| | (1) | (2) | (3) |
|---|---|---|---|
| | R&D | Total software | Own account software |
| | OLS | OLS | OLS |
| Asinh rivals' investment | -0.0013 | -0.0305*** | -0.0448*** |
| | (0.0152) | (0.0092) | (0.0158) |
| Log sales | 0.1091*** | -0.0092*** | 0.0182 |
| | (0.0198) | (0.0036) | (0.0217) |
| Asinh rivals' investment x Log sales | 0.003 | 0.0075*** | 0.0148*** |
| | (0.0037) | (0.0011) | (0.0038) |
| Firm age | -0.0097* | -0.0040** | -0.0063 |
| | (0.0053) | (0.0016) | (0.0041) |
| No. of industries | 0.0107*** | 0.0325*** | 0.0630*** |
| | (0.0029) | (0.0028) | (0.0055) |
| No. of zip codes | -0.0088** | 0.0317*** | 0.0135*** |
| | (0.0038) | (0.0036) | (0.0025) |
| Rivals' patents | 0.011 | 0.0083*** | 0.0175** |
| | (0.0104) | (0.0027) | (0.0073) |
| Adjusted $R^2$ | 0.3659 | 0.1093 | 0.2439 |
| N (rounded) | 54000 | 118000 | 118000 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS; software data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include fixed effects for 6-digit industry and year. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074).

### Instrumental variable

One concern with our investment regressions is that technology shocks might be correlated with both a firm's investment and its rivals' investments, leading to a spurious correlation. Since this would be a positive correlation, it would tend to weaken our results, that is, the true coefficient on rivals' investment would be more negative that our estimate. Nevertheless, to control for this or related endogenous effects, we lag rivals' investment and we instrument it following Bloom et al. (2013; see also Lucking, Bloom, and Van Reenen 2019). The instrument we use is the log of the tax term in Jorgensonian R&D user cost as developed by Wilson (2009),

$$\rho_{it} = \frac{1 - A_{st}^{c} - A_{ft}^{c} - A_{st}^{d} - A_{ft}^{d}}{1 - \tau_{st}^{e} - \tau_{ft}^{e}}, \quad s = state, f = Federal$$

where $A^c$ is the present value of R&D investment tax credits, $A^d$ is the present value of R&D depreciation allowances, and $\tau^e$ is the effective tax rate. The Jorgensonian user cost of R&D is then $\rho_{it} P_t^{R\&D}(r_t + \delta)$, where $P_t^{R\&D}$ is the price of R&D services, $r_t$ is the interest rate, and $\delta$ is the depreciation rate. We use data on state and federal tax credits developed by Barth et al. (2023). We find that log R&D tax cost is a reasonably good predictor of R&D so the first stage regression is strong:

Table A5. First Stage IV regression

| | Log R&D |
|---|---|
| Log R&D tax cost | -1.607*** |
| | (0.0851) |
| Adjusted $R^2$ | 0.8775 |
| N (rounded) | 76000 |
| F Test | 356.7 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). Regression includes firm fixed effects. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2735. (CBDRB-FY24-P2735-R11407).

Following the literature, we use this instrument to predict rivals' investment and we then use these predicted values to calculate $S$ in the IV estimates in Tables 4 and 5 of the paper.

There is a possible bias with this approach however: to the extent that rivals invest in the same state as the subject firm, the exclusion restriction might be violated. That is, that state's R&D tax cost would influence both the dependent variable and our predicted measure of rival investment. To check the robustness of our IV regression, we first checked whether our measure of rival firms' investment is correlated with the focal firm's log R&D tax cost. If it were not correlated, then rival investment in the same state could not bias our estimates. Column 1 in Table A5 shows this regression with industry and year fixed effects. The association is statistically and economically insignificant. Next, we explored alternatives using a common sample. Column 2 reproduces the same regression as in Table 4 Column 2, with a sample that is consistent across all the variations. We then added the log R&D tax cost on the right-hand side. The R&D tax cost variable is significant, but it does little to change the other coefficients, again implying that rival investment in the same state does not bias our estimates. Then, we went a step further and excluded rivals from the same state in the construction of the total rivals' investment. With this change, the coefficient on rivals' investment was significantly larger. Our basic IV results seem robust to these considerations, so we use the same method as in the literature in the main text of the paper.

Table A6. Alternative instrumental variable estimation

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable | Asinh rivals' investment | Log R&D | Log R&D | Log R&D |
| Log R&D tax cost | 0.023 | | | |
| | (0.086) | | | |
| Asinh rivals' investment$_{t-1}$ | | -0.984*** | -0.983*** | -1.275*** |
| | | (0.152) | (0.152) | (0.191) |
| Log sales | | 1.110*** | 1.109*** | 1.149*** |
| | | (0.122) | (0.121) | (0.113) |
| Asinh rivals' investment$_{t-1}$ x Log sales | | 0.142*** | 0.142*** | 0.185*** |
| | | (0.027) | (0.027) | (0.035) |
| Adjusted $R^2$ | 0.871 | 0.618 | 0.619 | 0.619 |
| N (rounded) | 34000 | 32000 | 32000 | 32000 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include controls for firm age, no. of industries, no. of zip codes, rival firm patents, and fixed effects for 6-digit industry and year and use sample weights. Column 2 uses the same IV as in Table 4, col. 2, column 3 adds a control for log R&D tax cost (not shown), and column 4 calculates rivals' investment excluding firms in the same state as the focal firm. Rivals' investment is lagged investment of all other firms in the same 6-digit industry. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407, CBDRB-FY24-P2735-R11770).

### Alternative measures of firm size

The model suggests that obsolescence risk and spillover efficiency might vary with firm size for a variety of reasons. To measure firm size, our main analysis uses firm revenue from LBD-rev. There are two potential problems with this. First, LBD-rev, which obtains revenue data from tax records, does not have matched revenue data for all firms because of specific features of the tax system. Consequently, some firms are excluded from the analysis when we use these revenue data, possibly giving rise to sample selection bias. Second, it is not clear whether domestic or worldwide sales provide the more relevant measure. In theory, we are seeking a measure of the customer base that might be affected by a rival's innovation. Rivals might or might not compete with the subject firm's overseas subsidiaries. Moreover, LBD-rev obtains revenue from US tax forms, but these might not include revenue from unconsolidated foreign subsidiaries. Fortunately, BRDIS provides its own measures of domestic and worldwide sales, so we can test alternatives. These measures are self-reported, so the revenue concept used is not entirely clear. These variables are also missing for some observations, however, the missing values appear to be independent of the unmatched values in LBD-rev. Table A6 shows the same regression as in Table 4, column 1 using firm size measured with LBD-rev, as in the original, firm size measured with BRDIS domestic sales for all firms in BRDIS with non-missing values, and firm size measured with BRDIS worldwide sales. The measure of firm size does not appear to make much difference to our results, although domestic sales appear to have slightly greater explanatory than self-reported worldwide sales.

Table A7. Log R&D regression with different firm size measures

Dependent variable: Log R&D

| | (1) | (2) | (3) |
|---|---|---|---|
| Sales measures: | Domestic sales from BRDIS, LBDrev not missing | Domestic sales from BRDIS | Worldwide sales from BRDIS |
| Asinh rivals' investment$_{t-1}$ | -0.896*** | -0.841*** | -0.984*** |
| | (0.147) | (0.145) | (0.137) |
| Log sales | 1.130*** | 1.154*** | 0.520*** |
| | (0.109) | (0.107) | (0.110) |
| Asinh rivals' investment$_{t-1}$ x Log sales | 0.177*** | 0.166*** | 0.213*** |
| | (0.028) | (0.028) | (0.028) |
| Adjusted $R^2$ | 0.608 | 0.606 | 0.595 |
| N (rounded) | 37000 | 38000 | 30000 |
| Mean net externality | -0.415 | -0.382 | -0.316 |
| Mean net externality (employment>500) | 0.373 | 0.354 | 0.425 |

Note: Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include controls for firm age, no. of industries, no. of zip codes, rival firm patents, and fixed effects for 6-digit industry and year and use sample weights. Rivals' investment is lagged investment of all other firms in the same 6-digit industry. The net externality is calculated as $\hat{\theta}_S \overline{\ln S_{it-1}} + \hat{\theta}_{SR} \overline{\ln S_{it-1} \cdot \ln R_{it}}$ where the bars signify sample means. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407, CBDRB-FY24-P2735-R11770).

### Sample selection of lagged measures

Our sample is not a full panel, but we use lagged rivals' investment as a key right-hand variable. This means that the regression sample differs from the original sample, possibly creating a bias. To test for the significance of this, we correct for sample selection by using inverse propensity weights—we divide the original sample weights by the probability that a randomly selected firm from the same industry will be in the regression sample. We estimate this probability simply as the industry mean probability of selection. Table A7 repeats two regressions from Table 4 both with the original sample weights and the inverse propensity weights. Not surprisingly, because the original sample has relatively more small firms than the regression sample, the negative externality effects are larger.

Table A8. Log investment with inverse propensity weights

| | (1) | (2) |
|---|---|---|
| Dependent variable | Log R&D | Log Total software |
| | OLS | OLS |
| Asinh rivals' investment$_{t-1}$ | -0.921*** | -0.752*** |
| | (0.142) | (0.280) |
| Log sales | 1.064*** | 0.983*** |
| | (0.105) | (0.242) |
| Asinh rivals' investment$_{t-1}$ x Log sales | 0.175*** | 0.170*** |
| | (0.026) | (0.053) |
| Adjusted $R^2$ | 0.598 | 0.701 |
| Mean net externality | -0.460 | -1.050 |
| Mean net externality (employment>500) | 0.324 | 0.219 |

Note: Inverse propensity weights. Standard errors are shown in parentheses and are clustered by firm (*** p<0.01, ** p<0.05, * p<0.10). R&D data come from BRDIS; software data from ACES. The samples exclude firms without revenue and firms in intangible-producing industries. Regressions include controls for firm age, no. of industries, no. of zip codes, rival firm patents, and fixed effects for 6-digit industry and year. Rivals' investment is lagged investment of all other firms in the same 6-digit industry. The net externality is calculated as $\hat{\theta}_S \overline{\ln S_{it-1}} + \hat{\theta}_{SR} \overline{\ln S_{it-1} \cdot \ln R_{it}}$ where the bars signify sample means. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407, CBDRB-FY24-P2735-R11770).

## Additional Summary Statistics

Table A9. Unstandardized Means and Standard Deviations for RHS Variables ($1000, base year=2009)

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | R&D (BRDIS) | | Software (ACES) | |
| | Mean | SD | Mean | SD |
| Rivals' investment | 1,570,000 | 4,835,000 | 295,700 | 816,200 |
| Revenue | 855,300 | 6,004,000 | 336,100 | 3,483,000 |
| Firm age | 27.21 | 11.29 | 22 | 13.39 |
| No. of industries | 2.853 | 4.965 | 1.913 | 3.072 |
| No. of Zip codes | 20.88 | 172.3 | 16.77 | 143.6 |
| No. of rivals' patents | 821.5 | 2554 | 226.2 | 913.3 |

Note: Corresponds to Tables 4 and 5, but the regressions use standardized variables (divided by the standard deviation). Statistics use sample weights. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2735. (CBDRB-FY24-P2735-R11407).