# The Intangible Divide:

# Why do so few firms invest in innovation?

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

August, 2025

Abstract: Investments in software, R&D, and advertising have grown rapidly, now approaching half of U.S. private nonresidential investment. Yet just a few hundred firms account for almost all this growth. Most firms, including many large ones, regularly invest little in capitalized software and R&D, and this "intangible divide" has surprisingly deepened as intangible prices have fallen relative to other assets. Using comprehensive US Census microdata, we document these patterns and explore a variety of 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 contributes to the intangible divide. Under common assumptions about diminishing returns, it could imply 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, Mike Meurer, Peter Meyer, Ronja Roettger, 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, CBDRB-FY25-P2735-R12155).

# 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 contributor to economic growth in recent decades (Corrado et al. 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).

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?[1] And why haven't sharp drops in asset prices induced more firms to invest in innovation?

This paper begins 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 the investment, and 2) most firms do not

[1] A large literature finds substantial private returns to R&D (see review in Hall et al. 2010). Other research has found a positive link between firm productivity or firm value and information technology generally (Brynjolfsson and Hitt 1996; Brynjolfsson et al. 2002) and with own software investment specifically (Tambe and Hitt 2012). Advertising has also been related to firm market value (Villalonga 2004; Peters and Taylor 2017).

regularly invest in capitalized software and R&D in our sample, but most do invest in advertising. Our descriptive analysis can rule out some simple explanations for 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.

Next, we look at the growth of intangibles, focusing on software and R&D, the sources of most of the growth. Almost all of the growth in these assets from 2008/9 to 2018 can be attributed to the top 250 firms in each. The top firms account for 93 percent of the growth in R&D and for 99 percent of the growth in software. Furthermore, 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.

This seems surprising because the relative asset prices of software and R&D have declined sharply. The BEA price index for software investment relative to the GDP deflator declined 30 percent between 2008 and 2018; the relative after-tax price of R&D declined 11 percent over the same period.[2] Moreover, the output elasticity of R&D has been growing substantially (Ando et al. 2025). In our model below, we show that this is equivalent to a drop in the *efficiency* price of R&D, in this case as large as an additional 30 percent decline.

One might expect sharply lower asset prices to increase the share of firms investing in software and R&D. In the workhouse model of invention (Nordhaus 1967; Scherer 1972), lower asset prices should increase the range of firms that invest all else equal. 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. 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. And a substantial drop should

[2] The BEA price index for business R&D grew 1.6 percent relative to the GDP deflator, but the average tax parameter (over all states, see the Appendix discussion of R&D user cost) fell by 13 percent. Fleming (2023) argues that the decline in the price of software investment is even larger.

significantly increase the share of firms that invest, all else equal.[3] The growing skewness of software and R&D poses a puzzle.

We propose that the growing intangible divide arises because of technological rivalry. Software and R&D are long lived knowledge assets in the tradition of the knowledge production function (Griliches 1979). Firms invest in these assets to improve their productivity, but in doing so, they may make the knowledge stocks of rivals obsolete. We hypothesize that rivals' technology investments increases a firm's *expected* obsolescence risk, hence reducing expected returns to innovation and, all else equal, reducing R&D or software investment. Mezzanotti and Simcoe (2023b) find that firm R&D decisions are affected by technology competition. It is quite common for firms to drop R&D in the face of superior innovations from rivals. For example, in 2014 Toyota backed away from years of development of a battery electric vehicle in the face of deeper and more successful development investments by Nissan and Tesla.[4]

Importantly, we hypothesize that firms differ in how they are affected by obsolescence risk. A substantial literature finds that a firm's customers switch to a more productive rival, but this tends to be a slow and extended process that varies with firm characteristics.[5] In general, consumers face switching costs arising for a variety of reasons including reputation (brand loyalty), complementary products, sunk investments in training or learning through use, network effects and more (Klemperer 1995). To the extent that one firm's customers switch more slowly, that firm's obsolescence rate will be lower, all else equal.

---

[3] Of course, any fixed price of investing might also have increased by 30 percent or more, but this seems unlikely without further evidence.

[4] Toyota had a joint venture with Tesla to develop an all-electric RAV4. That ended in 2014 and production on the RAV4 EV was stopped. Toyota Parts Center Blog, "Did The Rav4 EV Get Overlooked?," https://parts.olathetoyota.com/blog/2060/will-drivers-want-a-rav4-ev#:~:text=If%20you%20didn't%20have,after%20the%202014%20model%20year.

[5] Caves (1998) provides a review. Dunne, Roberts, and Samuelson (1988), Cabral and Mata (2003), Foster et al. (2016) provide evidence from micro-data. A variety of theoretical models attribute slow demand growth to informational frictions (Radner 2003; Rob and 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, et al. 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.

Estimating a model of investment demand, we find robust evidence that firm investment in software and R&D is substantially and negatively related to rivals' investments in these assets.[6] We find that this result is robust to the inclusion of many other factors including firm size, firm age, firm scope and geographical reach, outsourcing, management practices, rivals' patents, market share, and markups. It is robust to using a binary dependent variable that equals one if the firm invests and zero otherwise. It is robust to using instrumental variables for R&D and to a variety of other considerations.

And we find evidence of significant heterogeneity in firm sensitivity to obsolescence risk. The negative effect of rivals' R&D and software is substantially *smaller* for large (and older) firms. In our model, this heterogeneity can explain the divergent reactions to falling asset prices. Firms with lower obsolescence risk (e.g., higher customer switching costs) will increase their investment faster, all else equal. Conversely, firms with higher obsolescence risk will grow intangible investment more slowly, perhaps even decreasing their investment (possibly to zero). Thus, the intangible divide grows, explaining the puzzle.

Our evidence highlights the negative pecuniary externality arising from rivals' R&D (and software) via obsolescence risk. Of course, rivals' (and other firms') R&D also has a positive externality from knowledge spillovers. In our empirical analysis, the benefits of positive knowledge spillovers are implicitly included in our control variables.

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 et al. 1967; Bound et al. 1982; see Mezzanotti and Simcoe 2023a for a recent 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 et al. 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

[6] While our data do not permit a similar analysis for advertising and marketing expenditures, a survey of the literature finds that firms tend to increase their advertising as rivals' advertising increases (Bagwell 2007).

considered spillover externalities of R&D (for a review and meta-analysis see Hall et al. 2010; Ugur et al. 2016). A few of these papers have also included negative externalities from product market rivalry (Bloom et al. 2013; Lucking et al. 2019; Arora et al. 2021). These papers study select samples of public firms; in contrast, we use representative samples that can capture the effects of technological 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 et al. 2018; Crouzet and Eberly 2018; Brynjolfsson et al. 2023; Mouel and Schiersch 2023; Lashkari et al. 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).

A negative pecuniary externality might imply misallocation of investment—some firms, facing a "tax," will invest too little. While it is beyond this paper to examine the overall welfare effects of the intangible divide, the resulting investment skew is an important subject for future research.

We begin by describing our data and the main descriptive statistics. We then present a model of knowledge competition to both guide our empirical analysis and to understand a possible cause for the growth in the intangible divide. We then conduct regression analysis and conclude.

# 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 Foundation's National Center for Science and Engineering Statistics.[8] This survey provides data on each firm's

[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.

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

worldwide R&D expenditures (see further details at Foster, Grim, et al. 2016; Mezzanotti and Simcoe 2023a).

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

[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).

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 et al. 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 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

[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.

[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.

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 32% of firms don't advertise each year.

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 et al. 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 in our sample 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 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 software and R&D. 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. In the next section we develop a model that might explain why the distribution of knowledge assets is highly skewed even in the face of declining prices.

# A Model of Knowledge Competition

## Investment demand

R&D and software capex are distinct from assets such as advertising in that they are long lived knowledge assets. Firms invest to develop new knowledge but in doing so they may make rivals' knowledge obsolete. In this discussion we will discuss only R&D for

[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.

brevity, but the analysis applies equally to software. To analyze investment demand, we start with the "knowledge production function" (Griliches 1979; for a review see Hall et al. 2010),

$$R = AX^{\alpha}K^{\beta} \quad (1)$$

where $R$ is deflated firm revenue, $X$ is a composite of input factors (labor, physical capital, and intermediate inputs), $K$ is the firm R&D stock (defined below), and $A$ captures other sources of productivity, knowledge spillovers from other firms, knowledge gained from learning-by-doing, and the quality of researchers.[13] We assume $\alpha < 1$ to capture diminishing returns to scale and/or markups.

The knowledge stock represents the accumulated knowledge accrued from the sequence of past R&D investments. But as rivals develop superior technology, some of this knowledge becomes obsolete, that is, it is no longer commercially valuable. To capture the effect of obsolescence on the knowledge stock and hence on productivity, we follow the literature and construct R&D capital stocks using the perpetual inventory method with an obsolescence rate, $\delta$, and R&D investment $I_t$,

$$K_{t+1} = I_t + (1 - \delta)K_t. \quad (2)$$

In the empirical R&D literature, the depreciation of R&D is, by analogy to physical capital, assumed to be exogenous, fixed, and uniform across firms. But we want to explicitly consider the role of rivals' R&D on obsolescence.[14] When firms compete strategically using their technical knowledge, then rivals' new knowledge can reduce the commercial value of the focal firm's knowledge, that is, obsolescence becomes endogenous. To understand this behavior, we need to develop a model of knowledge competition.

While the empirical R&D literature largely ignores interfirm rivalry, a large industrial organization literature addresses innovation competition in product markets (Reinganum 1989 provides a review). But this literature does not address knowledge obsolescence. Typically, firms use R&D to "race" to be first to patent a single, commonly known innovation, but R&D is "memoryless," only pertaining to a race to dominate a particular

[13] Sometimes a term capturing knowledge spillovers is explicitly modeled using a sum of other firms' R&D weighted by technological proximity as in Bloom et al. (2013) where technological proximity is calculated using patent similarity. Instead, we capture the effect of knowledge spillovers in the revenue term; see discussion below.

[14] "Obsolescence" is a preferable term since knowledge does not "wear out" like physical capital; it just loses some of its commercial value.

product market.[15] Knowledge competition is different, affecting many product markets, including future potential markets that are yet to be conceived. Consider an example. Dupont began investing in polymer research in the 1920s (Hounshell and Smith 1988), developing a knowledge base that served as a platform to develop new products. Beginning in the 1930s, Dupont introduced neoprene, nylon, and a whole series of new materials, they refined the chemical engineering processes to produce these materials, and they developed many specific applications in a long list of specific product markets. Knowledge obsolescence occurs when new technologies supersede old, e.g., nylon superseded rayon on some applications. Obsolescence did not destroy the ideas of rayon technology; it just reduced their *commercial value*. In some market niches, nylon completely replaced rayon; in others, rayon just became less profitable; in others, rayon continued to be profitable. The knowledge stock $K$ represents the stock of commercially valuable knowledge; the obsolescence rate $\delta$ represents the reduction of knowledge that is no longer commercially valuable.[16]

Assuming that $\delta$ is a function of rivals' R&D and that each firm takes its expected obsolescence rate as given, we can derive the firm's R&D demand. First, assume that conventional inputs, $X$, are statically optimized each period so that, given a common cost of inputs, $p_X$, we can write a profit function,

$$\Pi(K_i) = (1-\alpha)\hat{R}(K_i) \tag{3}$$

where is $\hat{R}(K_i)$ revenue after maximizing profits with respect to $X$ (see Appendix).[17]
The Bellman equation for the value of the firm (suppressing the firm subscript) is then

$$V(K_t) = \max_{K_{t+1}} \left[ \Pi(K_t) - p_K(K_{t+1} - (1-\delta)K_t) + \frac{1}{1+r} V(K_{t+1}) \right] \tag{4}$$

where the first term captures the flow of profits based on the existing knowledge stock, the second term, using (2), represents the cost of R&D at price $p_K$, and the third term captures

[15] An important exception is Doraszelski (2003), but even this is a more limited role for knowledge than we consider below. Quality ladder models (Grossman and Helpman 1991; Aghion and Howitt 1992) do consider *product* obsolescence but R&D is memoryless and there is no obsolescence of long lived knowledge assets.

[16] Note that the obsolescence of knowledge is not quite the same as product obsolescence. For instance, knowledge obsolescence can reflect the foreclosure of future opportunities.

[17] $\hat{R} = \left( A\, K^{\beta} \frac{\alpha}{p_X} \right)^{\frac{1}{1-\alpha}}$.

the continuation value of the firm with the increased knowledge stock. The discount rate is $r$. Taking the first order condition and assuming steady state growth path, $I = gK$, we derive an investment demand equation with fixed effects of the form (see Appendix)

$$\ln \hat{I} = \ln R - \ln(r + \delta) - \ln \frac{p_K}{\beta} + FE + \epsilon. \quad (5)$$

This is the firm's reaction function to a given $\delta$. Note that the third term is efficiency price of R&D, $p \equiv \frac{p_K}{\beta}$. Our related paper finds that the output elasticity of R&D, $\beta$, has risen substantially, so we want to explore the impact of declining efficiency price.

## Nash Equilibrium

We can do this by exploring the Nash equilibrium. We consider two configurations: symmetric firms and two asymmetric firms. To do this we need to specify a functional relationship between rivals' R&D, which we label $S$, and the obsolescence rate. A simple flexible form is to let $\ln(r + \delta) = \gamma \ln S$, where we specify the measurement of $S$ for empirical analysis below. For the moment, consider $\gamma$, the elasticity of rivals' R&D, to be a constant; below we will allow firms to have different elasticities.

Then, expanding the expression for $\hat{R}(K_i)$ and dropping the fixed effects to simplify the analysis, we have a reaction function[18]

$$\ln \hat{I}_i = \frac{\theta}{1-\alpha} \ln A - \theta \ln p - \theta\gamma \ln S\,, \quad \theta \equiv \frac{1-\alpha}{1-\alpha-\beta}.$$

### Symmetric equilibrium

In the symmetric case with *n* firms, let $S_i = (n-1)I_i$ and the equilibrium firm R&D level, $\tilde{I}$, decreases with the efficiency price of R&D

$$\frac{d \ln \tilde{I}}{d \ln p} = \frac{\partial \ln \tilde{I}}{\partial \ln p} + \frac{\partial \ln \tilde{I}}{\partial \ln S} \frac{d \ln S}{d \ln p} = -\theta - \theta\gamma \frac{d \ln \tilde{I}}{d \ln p}$$

which solves to

$$\frac{d \ln \tilde{I}}{d \ln p} = -\frac{\theta}{1+\gamma\theta} < 0.$$

[18] We assume $\alpha + \beta < 1$.

As the efficiency price of R&D falls, in the symmetric Nash equilibrium, firm R&D investment rises, rivals' R&D rises, and the obsolescence rate rises.

### Asymmetric equilibrium, 2 firms

Now let there be two firm, *i* and *-i*, and let $S_i = I_{-i},\ S_{-i} = I_i$. Also, let the two firms differ in their elasticities to rivals' R&D, $\gamma$, such that $\gamma_i < \gamma_{-i}$. These different values of $\gamma$ could arise, as above, from different switching costs.

With these different elasticities, we have two reaction functions,

$$\ln I_i = \frac{\theta}{1-\alpha}\ln A - \theta \ln p - \theta\gamma_i \ln I_{-i}$$

$$\ln I_{-i} = \frac{\theta}{1-\alpha}\ln A - \theta \ln p - \theta\gamma_{-i} \ln I_i.$$

Rivals' R&D investments are strategic substitutes; an increase by one firm decreases the response of its rival. Solving, we get the Nash equilibrium R&D

$$\ln \tilde{I}_j = \frac{1-\theta\gamma_j}{1-\theta^2\gamma_i\gamma_{-i}}\left(\frac{\theta}{1-\alpha}\ln A - \theta \ln p\right), \quad j = i, -i.$$

Firms with high customer switching costs (low $\gamma_j$) will tend to have greater R&D investment in equilibrium and

$$\frac{d \ln \tilde{I}_j}{d \ln p} = -\frac{1-\theta\gamma_j}{1-\theta^2\gamma_i\gamma_{-i}}\theta.$$

Assuming $\theta^2\gamma_i\gamma_{-i} < 1$ it follows that

$$\frac{d \ln \tilde{I}_i}{d \ln p} < \frac{d \ln \tilde{I}_{-i}}{d \ln p}, \quad \gamma_i < \gamma_{-i}$$

so that a *decrease* in the efficiency price of R&D raises R&D proportionately more at firm *i* than at firm *-i*. That is, a decrease in the efficiency price of R&D widens the relative gap between the R&D investments of firms; high R&D firms (those with high customer switching costs) will tend to increase their R&D proportionally more than low R&D firms.

Note also that if $\theta\gamma_{-i} > 1$, then the low R&D firm's $\tilde{I}_{-i}$ could actually decrease with a falling efficiency price. The intuition is that the equilibrium occurs between two opposing forces: the lower price make investment more attractive but rising R&D rivalry makes it more costly. With heterogeneous $\gamma_j$, some firms are more sensitive to the effects of rivalry. If we further assume that firms will only invest in R&D above a certain threshold (the initial

convexity), then a decline in the efficiency price of R&D might cause fewer firms to invest in R&D at all; some low R&D firms might drop out of R&D.

To summarize, our model can explain the puzzles of intangible growth assuming that firms differ in their switching costs. A falling efficiency price can induce those firms that invest heavily in R&D to invest proportionally more. And it can also induce some firms to stop performing R&D. While other forms of investment can exert negative pecuniary externalities on rivals, R&D and software are distinguished by this substantial heterogeneity in the response to obsolescence risk. Our regression analysis below shows that these effects are indeed substantial.

## Empirical implementation

We use equation (5) to derive an investment demand function that we estimate. Also, assuming that demand must exceed a threshold before firms will invest, we can use that equation to estimate a linear probability model where the dependent variable is one if the firm invests in a given year and zero otherwise.

We begin with a baseline definition of rivals' intangible investment as the sum of the investments of other firms in the same industry, for firm $i$ in industry $J$ let $S_{it} \equiv \sum_{j \neq i, j \in J} I_{jt}$. By industry we mean a group of firms whose knowledge stocks might obsolesce some knowledge of other firms in the industry. Rivals, in this sense, might not be product market competitors, but they potentially obsolesce some knowledge of the focal firm. In practice, we define industries as the firms whose primary industries are in the same 6-digit NAICS class. This is a broad classification; we surely miss some firms that might be a source of obsolescence, and we include some firms that are not. To the extent that we mis-measure industries, our coefficients on rivals' R&D will be attenuated.

Rather than use a capital stock for rivals' investment, we use lagged investment expenditure. Because of the limited longitudinal dimension of our data, calculating capital stocks would severely reduce our sample. While the lagged expenditure also reduces our sample, the reduction is less. Below we use inverse propensity weights to check for the robustness of our results to the associated sample selection. The R&D literature finds that stocks and flow yield very similar results (Hall et al. 2010). We could use current firm

expenditures, but this risks spurious correlation due to contemporaneous shocks that affect all firms in the industry and because there is evidence of a gestation lag (Hall et al. 2010).

We use the firm's primary industry to identify rivals. While most firms only produce in a single 6-digit NAICS industry, many firms produce in multiple industries. To check the robustness of our rivalry measure, we define an alternative measure:

$$S_{it}^2 = \sum_J e_{iJt} \cdot \sum_{j \neq i} e_{jJt} I_{jt}$$

where $e_{iJt}$ is the share of firm *i*'s employment in industry *J* at time *t*.[19] The Appendix reports regressions using this measure.

Another concern is that the baseline measure may potentially be biased because it is constructed as a leave-one-out measure: for any year, $S$ is the industry total minus the focal firm investment, and focal firm investment is the dependent variable, possibly allowing for spurious correlation. In the Appendix we estimate this bias and find that it is too small to materially affect our results. Also, we instrument $S$ (see below) and we use the multi-industry measure and find our results are robust to these substitutions.

Finally, we want our rivalry measure to take account of firm heterogeneity in switching costs as above. We take firm revenue as a proxy for customer switching cost (we also use firm age in an instrumental variable estimation) so that $\ln(r+\delta) = \gamma_S \cdot \ln S_{it-1} + \gamma_x \cdot \ln R_{it} \cdot \ln S_{it-1}$. Inserting this in equation (5) and using inverse hyperbolic sine in lieu of logarithms (in the data some values of $S$ are quite small), we get

$$\ln I_{it} = \gamma_t + \gamma_J + \gamma_S \cdot \operatorname{asinh} S_{it-1} + \gamma_x \cdot \operatorname{asinh} S_{it-1} \cdot \ln R_{it} + \gamma_R \cdot \ln R_{it} + \epsilon_{it} \qquad (6)$$

with year and industry fixed effects. This is the specification we estimate in Table 4 with additional controls. We expect that $\gamma_S < 0,\ \gamma_x, \gamma_R > 0$ as we find. Defining a binary

[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). It is not clear that the distribution over industries in the cosine similarity is informative about obsolescence risk. In any case, cosine similarity is prohibitively time-consuming to calculate for roughly 1,000 6-digit industries and tens of thousands of firms each year. Bloom et al. use a small number of industries in a small sample of Compustat firms.

dependent variable $y_{it} = \mathrm{I}[I_{it} > 0]$, we use the same right-hand side (plus additional controls) in Tables 5 and 6.

It is possible that our rivalry measures might be endogenous. For instance, an increase in "technological opportunity" in an industry might both increase the dependent variable and also increase the rivalry measure, biasing coefficient $\gamma_S$ 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 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. It is also possible that revenue is endogenous; we instrument revenue with firm age to check this concern.

Finally, the reader might note that while we estimate the negative pecuniary externality of rivals' intangible investment we do not explicitly measure the positive externalities that come from knowledge spillovers. In the literature, some researchers use a version of the knowledge production function that explicitly includes external R&D (not necessarily from rivals) that provides a positive spillover benefit, $R = AX^{\alpha}K^{\beta}Z^{\mu}$ using a weighted sum of R&D, $Z_i \equiv \sum w_j I_j$, where the weights measure technological proximity as derived from patent data (Jaffe 1986; Bloom et al. 2013). Since only 18 percent of R&D performing firms patent (Mezzanotti and Simcoe 2023a), this would severely limit our sample. Instead, we implicitly include knowledge spillovers in the productivity measure $A$ so that spillovers are implicitly included in $R$ that we use in equation (6). While $Z$ and $S$ are likely correlated, that correlation does not bias our estimates.

# 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. Standardized variables let us conduct a "horse race" to see which variables are most influential in the regression. 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]

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

[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.

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. We also instrument for firm revenue and find that our results are robust to concerns about endogenous firm sales.

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. This suggests that switching costs may differ significantly between large and small 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 $\hat{\gamma}_S \overline{\ln S_{it-1}} + \hat{\gamma}_x \overline{\ln S_{it-1} \cdot \ln R_{it}}$, where the bar signifies sample mean. The last row of Table 4 shows the externality at sample means which are substantial. The magnitude of the coefficients $\hat{\gamma}_S$ and $\hat{\gamma}_x$ imply firms of different sizes will respond differently to falling asset prices; large firms will increase their intangible investments significantly, small firms less so or possibly even decrease their investments. Thus, the estimates can account for the puzzle of the growing intangible divide.

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.

Our measure of rivals' intangible investment appears to be negatively associated with focal firm investment and instrumented variables suggest that this relationship is plausibly causal. But perhaps it is not the rivals' intangible investment that is the driving factor here; perhaps there is some other aspect of rivals that lowers focal firm investment, but which is also correlated with rivals' intangible investment. For instance, a rival with a large market share might tend to invest more in R&D and that market share might also reduce the prospects of the focal firm, leading it to reduce R&D. Of course, for R&D the instrumental variable estimation suggests that the relationship between rivals' R&D and focal firm R&D is independent of these confounding issues. Nevertheless, in Appendix Table A10 we run the R&D regression for the manufacturing sector with a variety of additional controls that relate to the relative market position of the focal firm: the focal firm's market share (which is one minus rivals' aggregate market share), the share of the firm's business that is outside of manufacturing, the firm's markup, and the firm's mean R&D tax cost.[22] The significant coefficients on these control variables suggest that market competition affects R&D investment. But these controls do not diminish the coefficients on rivals' R&D, suggesting that rivals' R&D independently affects focal firm R&D and is more than just a proxy for rivals' market share, etc. We also tested specifications (not shown) with controls for a variety of IO competition concerns.[23] All of these specifications still showed a statistically significant increase in the (negative) coefficient on rivals' R&D.

## Likelihood of investing

Table 5 shows the results for the basic linear probability model where the dependent variable is 1 if the firm invests and 0 otherwise. 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

[22] Several of these variables depend on data that we only have for the manufacturing sector. Firm market share is the output-weighted market share of the firm's manufacturing establishments. Non-manufacturing share is the share of the firm's total employment outside of manufacturing. Markup is the output-weighted markup of the firm's establishments calculated using the Raval method for intermediate inputs. R&D tax cost is R&D-weighted mean of the firm's tax cost (state and Federal) from Barth et al. (Barth, Davis, Diethorn, et al. 2023).

[23] These included a flag for market leader; the market share of the industry's top four firms; flags for incumbent and entrant firms (< 5 year of age); flags for laggard and leader firms (below/above median TFP); and flags for neck & neck competition vs. "unlevelled" competition (below/above median industry markup; also, below/above industry TFP dispersion).

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, 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 software 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.

# Conclusion

Investment in intangibles has been growing rapidly but this growth is highly uneven. First, most of the growth comes from R&D and capitalized software, not other intangible assets. Second, for each of these two assets, growth is dominated by a small number of firms. The top 250 firms in each make most of the investment and the top 250 firms are responsible for almost all the *increase* in investment. At the same time, the share of firms that do not invest anything in software or R&D each year has fallen. And these developments have occurred as the effective prices of software and R&D have fallen substantially. The intangible divide appears substantial and persistent, and, if anything, it has been growing.

Based on standard models, one might expect lower asset prices to broaden investment in these assets, so this does seem to be a puzzle. We present a model that can explain the distinctive observed behavior. Software and R&D are long lived knowledge assets; firms invest in these assets to build their knowledge to increase productivity. But the increase of knowledge by one firm can obsolesce some knowledge of other firms. Greater risk of expected obsolescence reduces firm incentives to invest, all else equal. And while lower asset prices increase the incentives for firms to invest in software and R&D, they also increase the incentives for *rivals* to invest, raising obsolescence risk and partly offsetting the benefits of lower prices. In a symmetric equilibrium, lower asset prices raise intangible investment overall. But firms may differ in the degree to which obsolescence risk affects them. For instance, customers of established firms may face high costs to switch to a rival firm, reducing the impact of obsolescence. Critically, our model shows that when firms differ in the degree to which obsolescence affects them, lower asset prices can widen the intangible divide: more of the investment growth comes from firms that already invest heavily while other firms invest less or even stop investing in the intangibles at all. So, our model can account for the growth of the intangible divide in software and R&D investment.

To explore the empirical relevance of technological rivalry, we explore a wide variety of factors that possibly influence firm investment in R&D and software, including firm size, firm age, firm scope and geographical reach, outsourcing, management practices, rivals'

patents, and rivals' investments in software and R&D. We find robust evidence of a substantial negative relationship between rivals' intangible investment and focal firm investment. For R&D, instrumental variable estimation suggests this relationship is plausibly causal. Moreover, we find important heterogeneity: large firms are much less affected by rivals' intangible investment. Thus, the persistence and growth of the intangible divide appear to be intimately related to technological rivalry and obsolescence.

Is a persistent intangible divide 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 et al. 1987; Scherer 1991; Cohen and Klepper 1996a; 1996b; Ferreira et al. 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 et al. 2018; Crouzet and Eberly 2018; Brynjolfsson et al. 2023; Mouel and Schiersch 2023; Lashkari et al. 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 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.

Ando, Yoshiki, James Bessen, and Xiupeng Wang. 2025. *The Rising Returns to R&D: Ideas Are Not Getting Harder to Find*.

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.

Bagwell, Kyle. 2007. "The Economic Analysis of Advertising." *Handbook of Industrial Organization* 3: 1701–844.
Bar-Isaac, Heski, and Steven Tadelis. 2008. *Seller Reputation*. Now Publishers Inc.
Barth, Erling, James C Davis, Holden A. Diethorn, Gerald R. Marschke, and Andrew Wang. 2023. "Worker Mobility, R&D Human Capital, and Firm Productivity." *Working Paper*.
Barth, Erling, James C Davis, Richard B Freeman, and Kristina McElheran. 2023. "Twisting the Demand Curve: Digitalization and the Older Workforce." *Journal of Econometrics* 233 (2): 443–67.
Bessen, James. 2020. "Information Technology and Industry Concentration." *Journal of Law and Economics* 63 (3): 531–55.
Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, et al. 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.
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.
Cohen, Wesley M. 2010. "Fifty Years of Empirical Studies of Innovative Activity and Performance." In *Handbook of the Economics of Innovation*, vol. 1. Elsevier.
Cohen, Wesley M, and Steven Klepper. 1996a. "A Reprise of Size and R & D." *The Economic Journal* 106 (437): 925–51.
Cohen, Wesley M, and Steven Klepper. 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.
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." *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. 2003. "An R&D Race with Knowledge Accumulation." *Rand Journal of Economics*, 20–42.
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.
Grimm, Bruce T., Brent R. Moulton, and David B. Wasshausen. 2005. "Information-Processing Equipment and Software in the National Accounts." In *Measuring Capital in the New Economy*. University of Chicago Press. https://www.nber.org/books-and-chapters/measuring-capital-new-economy/information-processing-equipment-and-software-national-accounts.
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, Jacques Mairesse, and Pierre Mohnen. 2010. "Measuring the Returns to R&D." In *Handbook of the Economics of Innovation*, vol. 2. 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*. 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.
Hounshell, David A, and John Kenley Smith. 1988. *Science and Corporate Strategy: Du Pont R and D, 1902-1980*. Cambridge University Press.
Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2019. *The Industrial Revolution in Services*. National Bureau of Economic Research.
Jaffe, Adam B. 1986. "Technological Opportunity and Spillovers of R&D: Evidence from Firms, Patents, Profits and Market Value." *American Economic Review* 76: 984–1001.
Jarmin, Ron S, and Javier Miranda. 2002. "The Longitudinal Business Database." *Available at SSRN 2128793*.
Klemperer, Paul. 1995. "Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade." *The Review of Economic Studies* 62 (4): 515–39.
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): 1769–815.
Lucking, Brian, Nicholas Bloom, and John Van Reenen. 2019. "Have R&D Spillovers Declined in the 21st Century?" *Fiscal Studies* 40 (4): 561–90.
Mezzanotti, Filippo, and Timothy Simcoe. 2023a. "Innovation and Appropriability: Revisiting the Role of Intellectual Property." *NBER Working Paper*, no. 31428.
Mezzanotti, Filippo, and Timothy Simcoe. 2023b. *Research and/or Development? Financial Frictions and Innovation Investment*. National Bureau of Economic Research.
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.
Moylan, Carol. 2001. *Estimation of Software in the U.S. National Income and Product Accounts*. Bureau of Economic Analysis, U.S. Department of Commerce.
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 1.
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.
Reed, Kathleen. 2015. "Capitalization of Software Development Costs." Office of Chief Counsel, Internal Revenue Service, December 4. https://www.irs.gov/pub/irs-wd/201549024.pdf.
Reinganum, Jennifer F. 1989. "The Timing of Innovation: Research, Development, and Diffusion." *Handbook of Industrial Organization* 1: 849–908.

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.
Scherer, Frederic M. 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, 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.
Wilson, Daniel J. 2009. "Beggar Thy Neighbor? The in-State, out-of-State, and Aggregate Effects of R&D Tax Credits." *The Review of Economics and Statistics* 91 (2): 431–36.

# Figures

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

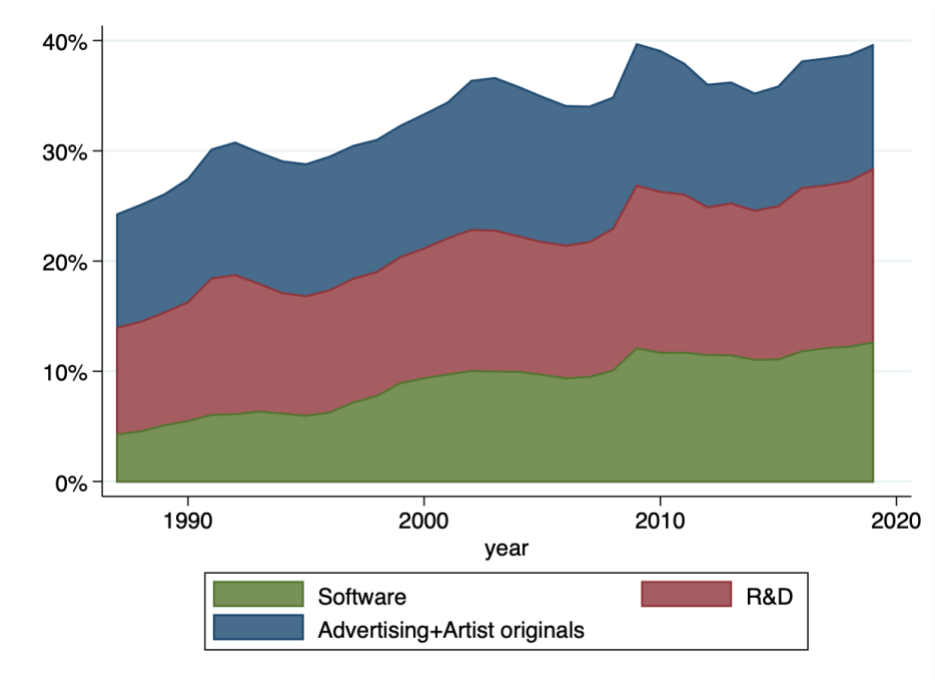
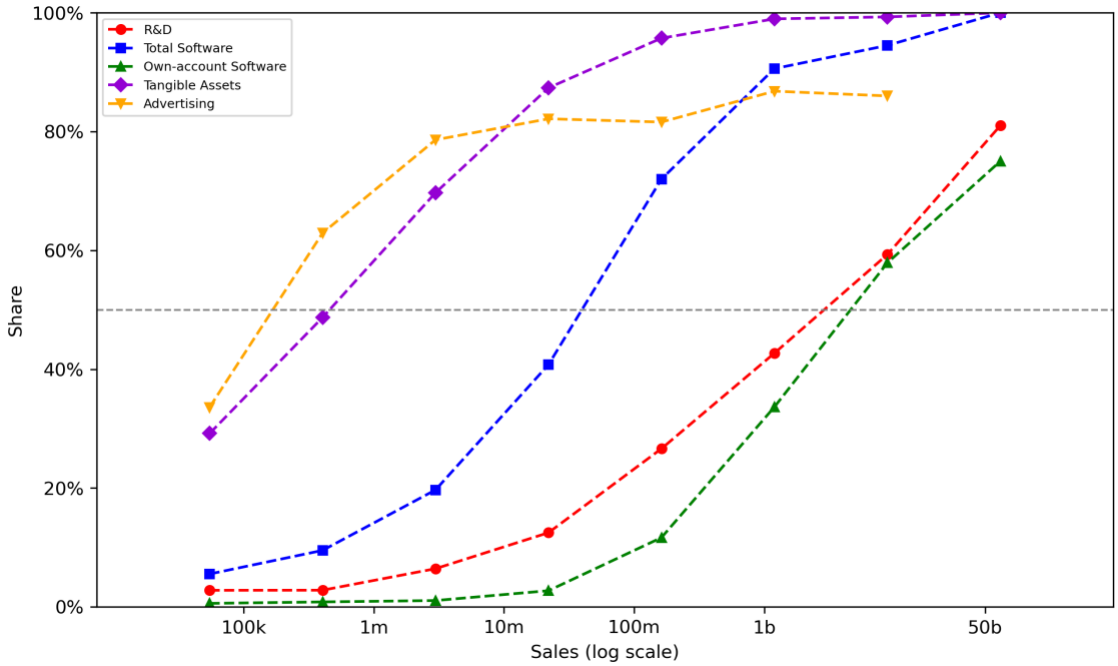



**Figure 2. Share of firms that reports positive investment in asset type by firm real sales ($2012)**

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 | △ | 2008 | 2018 | △ | 2008 | 2018 | △ | 2007 | 2017 | △ |
| **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 |

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

## Model

We assume that the firm statically optimizes profits with respect to inputs $X$ each period, $\max_X [AX^\alpha K^\beta - p_X X]$ where $p_X$ is the price of inputs. The first order maximizing condition gives us

$$\hat{X} = \frac{\alpha}{p_X}\hat{R}.$$

Inserting this in the expression for revenue and rearranging

$$\hat{R} = \left(A\, K^\beta \frac{\alpha}{p_X}\right)^{\frac{1}{1-\alpha}}.$$

Then we can write

$$\Pi(K_i) = \hat{R} - p_x \frac{\alpha}{p_X}\hat{R} = (1-\alpha)\hat{R}(K_i)$$

and

$$\Pi'(K_i) = (1-\alpha)\rho, \quad \rho \equiv \frac{\partial R_i}{\partial K_i} = \beta\frac{\hat{R}}{K}.$$

The firm's dynamic optimization defines the Bellman equation

$$V(K_t) = \max_{K_{t+1}}\left[\Pi(K_t) - p_K(K_{t+1} - (1-\delta)K_t) + \frac{1}{1+r}V(K_{t+1})\right].$$

Taking the first order condition, using the envelope theorem, and assuming steady state growth path, $I = gK$, we derive an investment demand equation with fixed effects of the form

$$\ln \hat{I} = \ln R - \ln(r+\delta) - \ln\frac{p_K}{\beta} + FE + \epsilon.$$

## Robustness checks

### Under/over reporting of R&D

Figure 2 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.[24] 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

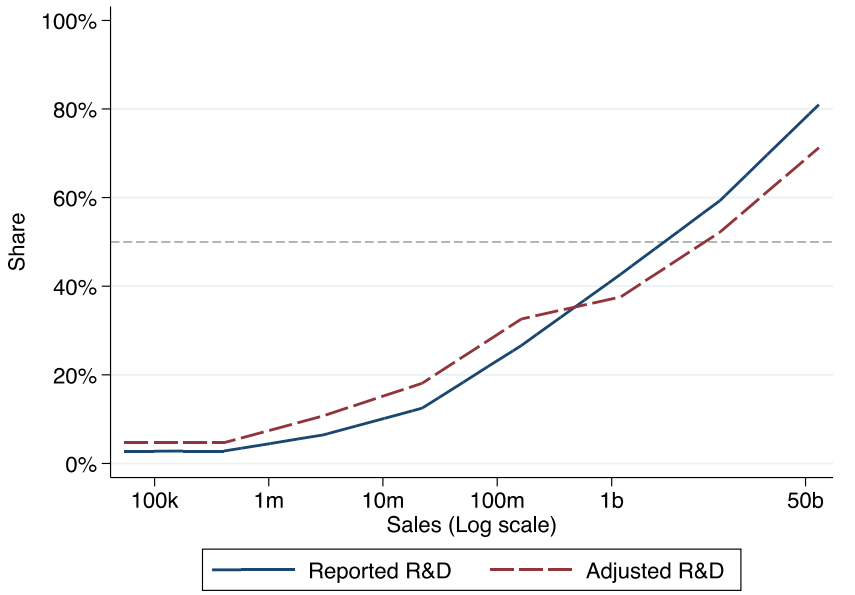
| Firm size (employees) | Share of Domestic R&D Performed by Company | Share of earnings of scientists and engineers |
|---|---|---|
| 10-99 | 7.3% | 12.2% |
| 100-499 | 8.0% | 11.6% |
| 500-999 | 4.0% | 4.9% |
| 1000+ | 80.8% | 71.3% |
| Total | 100% | 100% |

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 2 to correct for this possible bias by multiplying the share of firms reporting R&D investment by the ratio of science/engineering

[24] 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.

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.

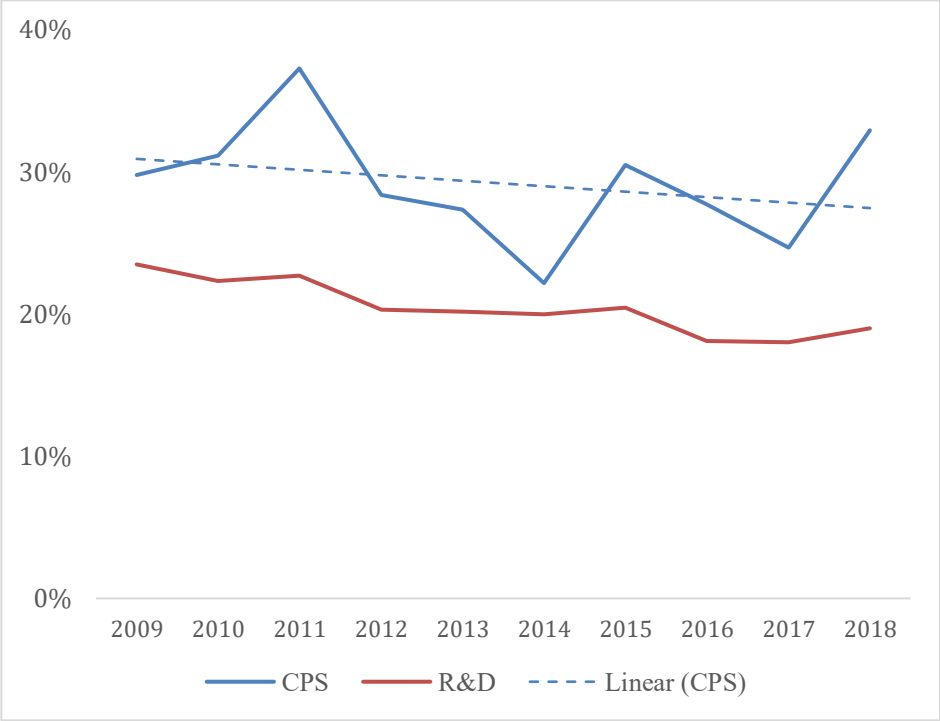
Figure A1. Share of firms that reports positive R&D investment by firm real sales




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 et al. 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.

### 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.

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 et al. 2019). The instrument we use is the log of the tax term in Jorgensonian R&D user cost as developed by Wilson (2009),

$$\rho_{it} = \frac{1 - A^c_{st} - A^c_{ft} - A^d_{st} - A^d_{ft}}{1 - \tau^e_{st} - \tau^e_{ft}}, \quad s = state, f = Federal$$

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 with columns 3 and 4. We then added the log R&D tax cost on the right-hand side (coefficient not shown). 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, in column 4, 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).

Another concern is that firm sales itself is endogenous.

Table A7. Instrumenting Firm Revenue, 2009-2018

| | (1) Log R&D | (2) Log Total Software |
|---|---|---|
| Asinh rivals' investment$_{t-1}$ | -0.765*** | -0.645*** |
| | (0.133) | (0.189) |
| Log sales (Pred.) | 0.519*** | 1.237*** |
| | (0.0449) | (0.176) |
| Asinh rivals' investment$_{t-1}$ x Log sales (Pred.) | 0.0606*** | 0.132*** |
| | (0.0106) | (0.0407) |
| No. of industries | 0.122*** | 0.238*** |
| | (0.0148) | (0.0210) |
| No. of zip codes | -0.00717 | 0.116*** |
| | (0.0254) | (0.0146) |
| Rivals' patents | 0.014 | 0.0185 |
| | (0.0269) | (0.0212) |
| Adjusted $R^2$ | 0.629 | 0.776 |
| N (rounded) | 35000 | 64000 |

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. For log sales we use a value predicted by firm age with firm and year fixed effects. Column 1 also uses predict rivals' investment as above. FSRDC Project Number 2735 (CBDRB-FY24-P2735-R11074, CBDRB-FY24-P2735-R11407, CBDRB-FY25-P2735-R12155).

### 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 A8. 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 A9 repeats two regressions from Table 4 both with 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 A9. 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 |

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).

### Possible confounding variables

To test whether the coefficient on rivals' R&D is driven by confounding characteristics of market competition, we add several variables to capture these effects.

Table A10

| | |
|---|---|
| Asinh(rival R&D$_{t-1}$) | -2.430*** |
| | (0.375) |
| Log(revenue) x asinh(rival R&D$_{t-1}$) | 0.147*** |
| | (0.026) |
| Log(revenue) | 0.166*** |
| | (0.052) |
| Firm market share | 7.478*** |
| | (0.895) |
| Non-manufacturing share | 1.647*** |
| | (0.197) |
| Markup | 0.225*** |
| | (0.047) |
| R&D tax cost | -1.548** |
| | (0.677) |
| Adjusted $R^2$ | 0.612 |
| Observations | 20000 |

Note: Standard errors are shown in parentheses and are clustered by 6-digit industry (*** p<0.01, ** p<0.05, * p<0.10). Regressions include fixed effects for 6-digit industry and year and use adjusted sample weights. Rivals' investment is lagged investment of all other firms in the same 6-digit industry. Firm market share is the output-weighted market share of the firm's establishments. Non-manufacturing share is the share of the firm's total employment outside of manufacturing. Markup is the output-weighted markup of the firm's establishments calculated using the Raval method for intermediate inputs. The number of observations is rounded to protect confidentiality. FSRDC Project Number 2735 (CBDRB-FY25-P2735-R12155).

## Additional Summary Statistics

Table A11. 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).

## Regression Bias from Construction of Rivalry Measure

In the investment demand equation (6), the dependent variable, $\ln I_{it}$, is implicitly related to the rivalry measure 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

$$\ln I_{it} = \beta S_{it} + \epsilon_{it}, \quad S_{it} = \ln(T_{t-1} - I_{it-1}), \quad T_t = \sum_j I_{jt} \tag{A1}$$

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,

$$\hat{\beta} = \frac{cov(\log I_{it}, S_{it})}{var(S_{it})}. \tag{A2}$$

To see the bias, it is helpful to decompose the error term in (A1) into two parts,

$$\epsilon_{it} = \phi_i + \theta_{it} \tag{A3}$$

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

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

Plugging (A1) and (A3) into (A2), the OLS estimate is

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

$$B \equiv \frac{cov(\phi_i, S_{it})}{var(S_{it})} + \frac{cov(\theta_{it}, S_{it})}{var(S_{it})}. \tag{A4}$$

*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:

$$cov(\phi_i, S_{it}) = cov\left(\phi_i, \log T_{t-1} + \log\left(1 - \frac{I_{it-1}}{T_{t-1}}\right)\right) \approx -cov\left(\phi_i, \frac{I_{it-1}}{T_{it-1}}\right) < 0 \tag{A5}$$

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[25]

$$\frac{cov(\phi_i, S_{it})}{var(S_{it})} \approx -\frac{cov\left(\overline{\log I_\iota}, \frac{I_{it-1}}{T_{it-1}}\right)}{var(S_{it})} + \beta \frac{var\left(\overline{\frac{I_\iota}{T_\iota}}\right)}{var(S_{it})}. \tag{A6}$$

We estimate the first term on the right and find it is small relative to the coefficient estimates for both R&D and software (see Table). We also estimate $var\left(\overline{\frac{I_\iota}{T_\iota}}\right)/var(S_{it})$ and

---

[25] From (A1) and (A3), $\log I_{it} = \beta S_{it} + \phi_i + \theta_{it}$ so that $\overline{\log I_\iota} = \beta \bar{S}_\iota + \phi_i$ since $E[\theta_{it}] = 0$. Further, following the analysis in (A5), $\bar{S}_\iota \approx \overline{\log T_\iota} - \overline{\left(\frac{I_\iota}{T_\iota}\right)}$ Note also that $cov\left(\overline{\left(\frac{I_\iota}{T_\iota}\right)}, \frac{I_{it-1}}{T_{it-1}}\right) = var\left(\overline{\frac{I_\iota}{T_\iota}}\right)$ so that $cov\left(\overline{\log I}, \frac{I_{it-1}}{T_{it-1}}\right) \approx cov\left(\phi_i, \frac{I_{it-1}}{T_{it-1}}\right) - \beta var\left(\overline{\frac{I_\iota}{T_\iota}}\right)$.

find that it, too, is small but positive, reducing the net negative bias further. Relative to the estimates of $\hat{\beta}$ from Table 4, the negative bias in (A4) is less than 10% except for the first period. Moreover, it is offset by the positive bias term, $cov(\theta_{it}, S_{it})$. Hence, we do not think that the net bias substantially affects our results, so we ignore it in the main paper.

Table A12. Bias estimates

| | $-\frac{cov\left(\overline{\log I_\iota}, \frac{I_{it-1}}{T_{it-1}}\right)}{var(S_{it})}$ | $\frac{var\left(\overline{\frac{I_\iota}{T_\iota}}\right)}{var(S_{it})}$ |
|---|---|---|
| R&D, 2009-2018 | -.043 | .060 |
| Software, 2008-2018 | -.047 | .027 |

Note: These estimates are calculated on the samples for Table 4. FSRDC Project Number 2735 (CBDRB-FY25-P2735-R12155).