Growth is Getting Harder to Find, Not Ideas*

Teresa C. Fort† Nathan Goldschlag‡ Jack Liang§ Peter K. Schott¶
Nikolas Zolas‖

May 1, 2026

Abstract

Relatively flat US productivity growth versus rising R&D expenditures is often interpreted as evidence that ideas are getting harder to find. We build a new 45-year panel tracking the universe of US firms' patenting to investigate the micro underpinnings of this conclusion, separately examining the relationships between research inputs and ideas (patents) versus ideas and growth. We find that average patents per R&D input are increasing, the elasticity of patents to R&D inputs is flat or rising, and there is not systematic evidence of a secular decline in patenting after controlling for research inputs. We then document a positive, significant, and fairly steady relationship between firms' patent and labor productivity growth rates. Average firm growth after controlling for patent growth, however, declines. Together, these results suggest that firms' innovative efforts play a key role in sustaining growth that has not diminished over the last four decades.

Keywords: Innovation, Productivity, R&D, Patents, Firm Growth
JEL Codes: O31, O32, O33, O47, D24

*We thank Nick Bloom, Santiago Caicedo, Cheryl Grim, Chad Jones, Samuel Kortum, Ezra Oberfield, Bitsy Perlman, Michael Peters, Thomas Sampson, Chad Syverson, Fabrizio Zilibotti and numerous conference and seminar participants for helpful comments and suggestions. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the US Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7083300: CBDRB-FY25-CES002-001 and CBDRB-FY25-CED006-012). Parts of the paper were written while Fort, Goldschlag, and Zolas were US Census Bureau employees.

†Tuck School at Dartmouth, NBER, CEPR, CESifo.
‡Economic Innovation Group.
§Temple University.
¶Yale School of Management, NBER, CEPR, CESifo.
‖US Department of State.

1 Introduction

A central question for understanding long-run growth is how firms' profit-maximizing investments in R&D affect idea creation and productivity growth. Romer (1990) shows that firms' R&D investments can sustain long-run growth when new ideas rise linearly with the knowledge stock. If instead a larger stock makes discovery harder, growth will stall. Romer acknowledges this crowding-out possibility and argues that there is "no evidence from recent history to support the belief that opportunities for research are diminishing (p. S84)."

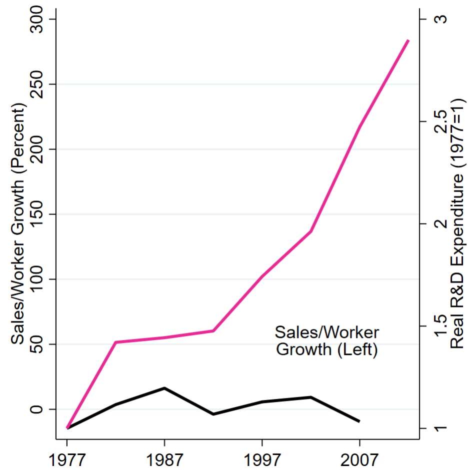
While Romer (1990) predicts higher output growth from greater R&D efforts, the US economy has maintained an average two percent real growth rate over the last 200 years, despite substantial increases in its R&D spending. These patterns have led some researchers to conclude that ideas are getting harder to find (Jones, 1995; Kortum, 1997). Most recently, Bloom et al. (2020) infer declining researcher productivity from the divergence between rising R&D efforts versus flat US productivity growth for various samples of industries and firms. These trends can be rationalized by models in which researchers' marginal products exhibit decreasing returns in the aggregate stock of ideas. In those models, there is no role for innovation policy to affect long-run growth, since improvements in living standards depend solely on population growth.

Although constant productivity growth amid rising research efforts is consistent with ideas getting harder to find, existing papers do not provide direct evidence of this mechanism. Moreover, this inference requires the common but strong assumption that idea growth maps one-to-one with output growth. This assumption rules out a time-varying or nonlinear relationship between ideas and growth, and leaves no role for other factors such as changing market structure or regulation.

In this paper, we analyze whether ideas are getting harder to find by studying these two relationships separately: that between research inputs and ideas, versus that between ideas and growth. To pinpoint the underlying mechanisms at play in macro-models, we construct a novel firm-level panel from 1977 to 2021 with a rich set of R&D inputs using the universe of US private-sector employer businesses in the Census Bureau's Longitudinal Business Database (LBD). Figure 1 depicts R&D inputs and aggregate growth in our data, and delivers the same message highlighted in Bloom et al. (2020): US sales per worker growth since 1977 has remained relatively flat, while R&D inputs have risen threefold. To assess whether the divergence in Figure 1 arises from declining research efficiency, we merge the Census microdata to US patent grants, which we take as a key manifestation of ideas.

Our first contribution is a set of new facts about US patenting that emerges from studying the universe of US firms and quality-adjusted patent measures. Early models with dimin-

Figure 1: Aggregated Micro Data on Growth and Research Inputs



Source: LBD, EC, RADS, and authors' calculations. *Notes*: Figure replicates Bloom et al. (2020) Figure 1 using aggregated US Census Bureau micro data analyzed in this paper. Left axis reports growth in aggregate sales per worker for firms in the LBD by semidecade, where each point is average growth for t to $t+5$. Right axis reports aggregate real R&D expenditure across firms in the LBD, by semidecade, indexed to 1 for the 1977 to 1981 semidecade.

ishing returns to knowledge stocks were motivated by trends of declining patent grants per researcher (Evenson, 1984; Kortum, 1993; Lanjouw and Schankerman, 2004). We also observe this pattern among the subset of publicly traded firms tracked by Compustat, but find that patents per real R&D dollar rise from 1977 to 2017 among the universe of firms. Quality-adjusted measures – external citations (Akcigit and Kerr, 2018) and breakthrough patents (Kelly et al., 2021) – show similar increases, including among Compustat firms. These results do not support a first definition of "ideas are getting harder to find": that average patents per R&D dollar have fallen since the 1970s.

We also document shifts in the composition of US patenting. Manufacturing firms' shares of patent activities fall from a majority in the 1970s to less than a third by the 2010s, with their breakthrough share dropping from 73 to 12 percent. By contrast, shares of citations and breakthroughs by firms in Information, Management, and Professional Services increase steadily, accounting for 45 and 66 percent, respectively, by the 2010s. Although mega-firms (10k+ workers) patent grant and citation shares decrease roughly 10 percentage points, their breakthrough share starts and ends at about 70 percent. These patterns are accompanied by falling patent coverage in Compustat, which drops from 63 to 55 percent over the period. Overall, the data indicate that work focused solely on manufacturing or Compustat firms misses a growing share of US innovative output.

Our second contribution is to estimate time-varying within-firm elasticities of patents to R&D inputs. In contrast to the notion that the marginal product of R&D declines over time, we estimate rising elasticities of patent grants, citations, and breakthroughs to R&D

expenditure. For patent grants, elasticities almost double from 0.26 to 0.48 from 1977 to 2016. Although these estimates are not causal, our use of firm fixed effects ensures that we capture a rise in subsequent patents when a particular firm increases its R&D spending. Moreover, any shocks to the levels of R&D and patents would need to grow over time to match the rising elasticities we document, a pattern which itself would seem to contradict the notion that ideas are getting harder to find.[1]

Census R&D surveys disproportionately sample a subset of large manufacturing firms, which our descriptive evidence shows is less relevant for patenting over time. To address selection into R&D surveys, we construct new measures of firms' research inputs using their payroll in innovation-related establishments, such as R&D labs (as a narrow measure) and Management, Professional, Scientific, and Technical Services establishments (as a broader measure).[2] Whereas patent coverage by firms in the R&D surveys falls from 65 to 59 percent over our period, firms with Management and Professional Services establishments account for about 77 percent of patent grants throughout. We also consider firms' total payroll as a proxy for firms' R&D inputs. Although broad, this measure allows us to estimate relationships among all US patenting firms. Estimated patent elasticities across all three patent activities and all four measures of research inputs are generally flat or rising over time. Estimates among Compustat firms are also flat or rising. These results do not support a second way of defining "ideas are getting harder to find," *i.e.,* that the marginal R&D dollar is less effective at generating an idea over time.

The process of patenting might change for various reasons over our decades-long sample, *e.g.*, in terms of the quality of the ideas firms seek to protect, or how the US Patent and Trademark Office (USPTO) grants approvals. We include semidecade fixed effects in our specifications to ensure that the patenting elasticities are not contaminated by common institutional or macroeconomic shocks. These fixed effects are also important in their own right, since they measure average firm patents after controlling for changes in their R&D. A secular flattening or decline in these fixed effects, even as US patents increase, could be another manifestation of ideas getting harder to find. We document statistically significant secular declines in just two of our 18 specifications, whereas estimates are flat, rising, or mixed in 15 of them. These trends do not provide systematic support for a third way of defining "ideas are getting harder to find," that idea creation has decreased universally across

[1] An important caveat to this conclusion is that patents represent only a subset of ideas. While patent ideas might be getting easier to find, it could be that non-patent ideas are the more important driver of growth, and that they are increasingly elusive. This concern is somewhat mitigated by our finding of a steady relationship between patents and growth in Section 6. We discuss this issue further in Section 7.

[2] R&D labs are classified in NAICS 5417. Management is NAICS 55, Professional, Scientific, and Technical Services is NAICS 54, which includes R&D labs as one industry.

all patenting firms independent of their research efforts over the last four decades.

Our third contribution is to estimate a time-varying relationship between idea and output growth for all US patenting firms, including non-manufacturers. We regress five-year sales per worker growth on patent-stock growth, using size and age bins to proxy for capital. We also consider specifications using sales as the dependent variable and controlling for employment growth. We find positive and statistically significant relationships for both outcomes in most periods that do not weaken over time. Indeed, the relationship between patent growth and sales per worker growth is strongest in the later years. These results do not support a fourth definition of "ideas getting harder to find", a weakening link between ideas and growth. They also address concerns that patent quality has deteriorated over time, since that should manifest as a weakening association between patents and growth.

We use the regression estimates to decompose firms' sales growth into the components predicted by changes in their patents and employment, as well as the size, age, and time fixed effects. While ideas predict fairly consistent growth throughout our sample period, the other factors tend to predict *lower* overall growth in later periods. Notably, the semidecade fixed effects and employment coefficients predict relatively low aggregate growth rates after 2002. Indeed, firms' own ideas predict positive sales growth in the last decade, even as other factors drag it down.

Together, our results indicate that the ratio of patents per R&D dollar is rising, the marginal R&D dollar is at least as effective in the 2010 decade as it was in the 1970s, there is not a systematic secular decline in average firm patents after controlling for their R&D input use, and that, if anything, the link between ideas and productivity growth is rising. These patterns are hard to square with the claim that ideas are getting harder to find. On the other hand, we find that common macro factors predict relatively lower growth rates in the 2000s that are independent of firms' idea growth. One possibility is that (at least some of) these lower growth rates reflect declining spillovers from other firms' ideas into firm-level productivity growth.

This paper contributes to three strands of the literature. First, we add to a large literature that estimates the elasticity of patents to R&D expenditure, with early work summarized by Griliches (1990) and further explored by Blundell et al. (2002). Most related to our analysis on changing elasticities, Hausman et al. (1984) estimate declining patent elasticities for a sample of 128 firms from 1968 to 1974. More recent papers exploit exogenous variation in R&D funding to assess its impact on firm innovation and patenting (Howell, 2017; Meyers and Lanahan, 2022). Our contribution is to estimate flat or rising within-firm elasticities of patent grants, external citations, and breakthroughs to comprehensive measures of R&D

inputs for all firms in the United States over the last four decades.[3]

We also expand on a relatively small set of papers studying the relationship between firm-level patenting and growth. While a long line of research in industrial organization estimates the effect of R&D expenditure on firm productivity (Griliches, 1979; Hall et al., 1998; Aw et al., 2011; Peters et al., 2017; Chen and Xu, 2023; Ando et al., 2025), there is less evidence on whether and how a firm's patents relate to subsequent growth. Balasubramanian and Sivadasan (2011) link manufacturing firms' patenting to increases in their size and productivity. Kline et al. (2019) show that a firm's first patent grant leads to higher profits. Patents grants are also associated with increased stock market value (Hall et al., 2005), though with substantial variation in those returns, with higher returns correlating with future growth (Kogan et al., 2017). Akcigit and Kerr (2018) document a positive relationship between the log of patents and employment growth for US firms over the 1982 to 1997 period. We document a positive and significant relationship between US firms' patent and labor productivity growth that is fairly stable over the last four decades.

Finally, we build on a growing body of empirical work that studies whether flat (or falling) US productivity growth coupled with rising research efforts is due to negative crowding-out effects of a growing knowledge stock.[4] Several influential papers document the rising divergence between productivity growth and research inputs and rationalize it in a framework in which a larger stock of knowledge makes it harder to find new ones (Jones, 1995; Kortum, 1997; Bloom et al., 2020).[5] Jones (2009) provides microevidence that it is harder to innovate as a particular field matures, which he models as decreasing output growth from innovative effort. Most recently, Ekerdt and Wu (2024) show that declining aggregate researcher productivity occurs when heterogeneous workers increasingly select into innovation rather than production, but that such declines are transient. Our contribution is to show that within-firm research productivity, both in terms of translating R&D into ideas and in terms of turning those ideas into output, has not declined over the last 40 years.

This paper also facilitates future empirical work by providing a new crosswalk between the LBD and the USPTO patent data that is available to all researchers with approved projects through the Federal Statistical Research Data Centers (FSRDCs). Our bridge provides the longest period of matched data, spanning 1976 to 2021, matches 92 percent of US-based patent-assignee records, and ensures longitudinal consistency in the matching process. To our knowledge, we are the first to develop and share such longitudinal matching techniques,

[3]The rising elasticities are indirectly suggestive that idea growth is not linear as suggested by Philippon (2022).

[4]Books on this topic include (Cowen, 2011; Brynjolfsson and McAfee, 2014; Gordon, 2017).

[5]Recent papers propose that R&D labor and R&D capital are complementary inputs into knowledge production, with declines in the latter making ideas harder to find (Growiec et al., 2022; Ekerdt, 2024).

which can be generalized to improve merges of other external datasets to the Census Bureau's micro data. The crosswalk is also used to produce the Census Bureau's public-use Business Dynamics Statistics of Patenting Firms tables, which are publicly available tabulations of US job creation and destruction by firms' patent status.

The remainder of this paper is structured as follows. Section 2 provides theoretical background, Section 3 describes the data, Sections 4 provides four new facts, Section 5 estimates patent elasticities, Section 6 estimates the relationship between patents and growth, Section 7 connects our empirical findings back to theory, and Section 8 concludes.

2 From Macro Growth Theory to Firm Regressions

Romer (1990) models how firms' profit-maximizing R&D drives long-run growth. The knowledge stock, or number of ideas, A evolves according to

$$\dot{A} = \delta H_A A, \tag{1}$$

where $\dot{A}$ is the flow of new ideas, H_A is total human capital (or researchers) devoted to innovation, and δ is a productivity parameter. Ideas are nonrival and only partially excludable, so that the flow of future ideas is increasing in the aggregate knowledge stock.

Final output is produced by a representative firm using the following technology

$$Y(H_Y, L, x) = H_Y^{\alpha} L^{\phi} \int_{i \in A} x_i^{1-\alpha-\phi}\, di, \tag{2}$$

where H_Y is human capital devoted to final-good production, L is labor services, and x_i is a durable capital-good input. Each input x_i, requires an idea i (or blueprint) to produce, with monopolistically competitive input producers converting ω units of foregone consumption into one unit of inputs. Although there is an infinite potential number of ideas, only those that have been discovered – denoted by A in equation (1) – can be used to produce. Crucially, and in contrast to prior growth models, there is no decreasing returns to such capital as new varieties of x_i are added. The aggregate knowledge stock thus determines the number of capital-good inputs in the economy, which in turn affects output. In particular, equation (2) can be rewritten as

$$Y(H_Y, L, x) = (H_Y A)^{\alpha} (LA)^{\phi} (K)^{1-\alpha-\phi} \omega^{\alpha+\phi-1}, \tag{3}$$

where K denotes physical capital of durable inputs and is thus the physical embodiment of the stock of ideas.[6] Differentiating equation (3) with respect to time, holding labor fixed, and assuming that ω is time-invariant, the steady-state growth rate of output is determined by the growth rate of ideas, which in turn is determined solely by research productivity and the amount of human capital in knowledge production, *i.e.*,

$$\frac{\dot{Y}}{Y} = \frac{\dot{A}}{A} = \delta H_A. \tag{4}$$

The model thus predicts a higher growth rate when the number of innovation workers rises.

Subsequent macro models have focused on the fact that this prediction is severely counterfactual, as the number (and share) of researchers in the US economy has increased substantially, while aggregate growth has remained fairly constant. Jones (1995) proposes a 'semi-endogenous' growth model with diminishing returns to new ideas

$$\dot{A} = \delta H_A A^{1-\beta} \tag{5}$$

as a way to reconcile Romer (1990) with this macro fact. In this formulation, $\beta > 0$ implies that the growth rate of ideas ($\dot{A}/A$) will decline over time as the existing stock of knowledge expands.[7] More recently, Bloom et al. (2020) define *Researcher Productivity* $:= \frac{\dot{A}_t/A_t}{H_{At}} = \delta A_t^{-\beta}$ and infer the size of β under the null hypothesis that this ratio (and δ) remain constant. They measure $\dot{A}_t/A_t$ as US TFP growth (as well as using various sector measures of productivity growth, such as greater crop yields or higher life expectancy) and H_A as increases in researchers in the US economy (or a particular sector). To the extent that their measure of $\frac{\dot{Y}_t/Y_t}{H_{At}}$ falls over time, the authors infer declining researcher productivity, which they attribute to $\beta > 0$.[8] The formulation in equation (5) is elegant in that they can always 'match' the data perfectly by adjusting β for different time periods or sectors. On the other hand, this approach entails a time-period and sector-specific β, which does not quite align with equation (5).

At least three alternatives to a time-varying β deserve attention. First, researcher productivity might change over time for a variety of reasons, including policy. To the best of our knowledge, the empirical growth literature does not directly tie declining output growth to increases in A in the data.[9] A second possibility is that the relationship between idea

[6]An intermediate step between equations (2) and (3) is $Y(H_Y, L, x) = AH_Y^{\alpha} L^{\phi} \bar{x}^{1-\alpha-\phi}$, where $K = \omega \int_A x_i \, di$, from which it is clear that equation (4) follows.

[7]Romer (1990) notes that "[l]inearity in A is what makes unbounded growth possible," and that researchers would shift towards manufacturing as A grows, thereby lowering growth, if equation (1) were concave in A. Note that in Jones (1995), the $1 - \beta$ above is instead denoted by ϕ. Here we follow the notation in Bloom et al. (2020).

[8]Specifically, they measure $\beta = \frac{g_{H_A}}{g_Y}$, where g denotes growth rates.

[9]Jones (2009) finds that researchers' age at first invention and size of teams both rise as knowledge grows.

growth and output growth is not one-to-one or constant over time. Indeed, Romer (1990) notes that durable inputs may not have additively separable effects on final-good output, "An investigation of complementarity as well as of mixtures of types of substitutability is left for future work" (p. S81). Finally, growth depends not only on ideas, but also on other factors, such as changing regulation (Djankov et al., 2002), market structure (Aghion et al., 2005), climate (Dell et al., 2012), and institutions (Acemoglu et al., 2001), among others. It is possible that idea growth is expanding, and ideas continue to generate productivity improvements, but other factors work against these forces.

To investigate the empirical support for the first two channels – a decline in researchers' ability to generate ideas or a decline in the impact of these ideas on growth – we place firms center stage and study each relationship separately. We first focus on evidence that speaks to researchers' ability to generate new ideas. We begin with a simple measure of research productivity that is readily observable in the data and free of functional-form assumptions:

$$\text{Research Productivity} \coloneqq \dot{A}_t / H_{At} = \frac{\text{new patents}}{\text{research expenditure}},$$

which we will measure using patents per real research expenditures. Observed declines in this measure of research productivity motivated a number of early papers in the semi-endogenous growth literature, including Jones (1995) and Kortum (1997). This definition of research productivity is increasing in the aggregate stock of ideas under Romer (1990), whereas there are diminishing returns to aggregate ideas when $0 < \beta < 1$ in equation (5).

As a second metric on how researchers' ability to generate ideas has evolved, we estimate the elasticity of patents to research inputs. To do so, we adapt equation (5) to:

$$\text{Patents}_{ft} = \delta \text{Research Inputs}_{ft}^{\eta_t} A_t^{1-\beta}, \tag{6}$$

where η_t is the elasticity of firm f's patents (ideas) to its research inputs (H_A) up through year t. We measure research inputs using traditional R&D expenditures, as well as with more comprehensive measures based on firm payroll in innovation-related establishments. Although the exponent on research inputs, η_t, is absent in Romer, both Jones (1995) and Bloom et al. (2020) consider its impact, referring to it as the 'stepping on toes' effect. The distinction here is that we estimate a firm-level elasticity, whereas they consider diminishing returns to aggregate research. An estimate of this patent elasticity less than one implies diminishing returns to R&D efforts for a given firm in a particular time period that would in turn dampen the impact of additional researchers on aggregate growth. In particular, the

He links these patterns to aggregate growth via a model in which research productivity declines as knowledge grows, but he does not link changing inventor age and team size to aggregate output in the data.

lines in Figure 1 would still diverge, though the productivity growth rate would nevertheless rise as research inputs increase as long as $\beta = 0$.

In our formulation, the term $A_t^{1-\beta}$ is the aggregate stock of non-depreciated ideas (patents) that affects all firms equally. Treating A_t as an aggregate rather than firm-specific stock is crucial for capturing the main mechanism in Romer: ideas are (at least partially) non-rival and non-excludable, such that any given firm's efforts to generate new ideas will depend on the total stock of ideas in the economy. As such, it will be captured by year fixed effects in our empirical implementation. The year fixed effects also capture any changes in spillovers across firms (Bloom et al., 2013), effects of declining public R&D expenditures on innovation (Bergeaud et al., forth), and the discovery of new technology "classes," e.g, the PC-internet revolution of the late 1980s and 1990s (Ribeiro, 2025). If Romer's productivity parameter (δ) were time-varying, it would similarly be captured by the year fixed effects. We will estimate and plot these year fixed effects as a third way to assess whether ideas may be getting harder to find, *independent* of firm's research inputs. Although we might have included the firm's own past realizations of ideas (A_{ft}) as a separate term, in practice we will subsume various potential lag structures for the firm's research-input flows using a depreciated stock of its research inputs, which is empirically (and theoretically) correlated with its patent stock, thus making it infeasible to identify separately from its research inputs.

We next consider the second fundamental relationship in Romer: how ideas map to growth. If taken literally, Romer predicts growth solely through firm entry, with each new idea generating a new input producer. The fundamental mechanism at play, however, is that a new idea increases output from a fixed amount of labor and human capital. We capture this spirit of Romer by adapting equation (3) to

$$\text{Growth}(Y_{ft}) = \mu_t \text{Growth}\left(PatentStock_{ft}\right) + \alpha \text{Growth}(L_{ft}) + (1-\alpha)\text{Growth}(K) + \phi A_t, \quad (7)$$

where Y_{ft} is real sales, $PatentStock_{ft}$ is the firm total patents up to year t, L_{ft} is firm employment in t, and K_{ft} is firm capital in t. In this adaptation, we combine labor services (L) and human capital (H_Y) into L. We control for changing labor at the firm level in order to isolate the relationship of patents on growth. Because we lack observable measures of capital for non-public firms outside manufacturing, we control for changes in firm-level capital using firm size and age bins. An estimate of $0 < \mu_t < 1$ indicates positive, but less than proportional returns to ideas for a particular firm and time period. We also posit the potential for the aggregate stock of ideas to influence an individual firm's growth by including A_t. As with idea creation in equation (6), any changes in the relationship between firm growth and the aggregate stock of ideas, or in the effect of knowledge spillovers on firm growth, are captured by year fixed effects.

Our firm-level approach pinpoints where any break occurs — in the mapping from R&D to ideas, or from ideas to growth — rather than collapsing both into a single R&D -to-growth relationship. This two-step approach also mitigates noise from the uncertainty of translating R&D into ideas, with the second step conditioning only on successful R&D as measured by patents.[10]

3 Data

In this section, we explain the construction of our new 45-year firm-level panel dataset of US firms' patents and research-input use.

3.1 Patent Data

We measure a firm's output of ideas using patents. The US Patent and Trademark Office's (USPTO) PatentsView (PV) database provides information on the identity and location of all granted US patents from 1977 through 2021 and their corresponding assignees and inventors, dates of application and granting, and citations. To avoid truncation due to the lag between a patent's application and grant date, we use granted patents through 2016. This five-year cut off captures about 90 percent of granted patents.[11] We restrict our analyses to "domestic" patents, which we define as patents with at least one assignee located in the United States. We exclude "foreign" patents because the LBD contains no information on firms without US establishments, and we would potentially overstate researcher research productivity by including foreign inventions in their idea counts.[12] All subsequent references to patents refer to domestic patents, unless otherwise specified.

Griliches (1998) highlights patents as a measure of innovation, but notes two limitations relevant here. First, not all inventions are patented, for instance because firms protect them as trade secrets instead. Since this paper relies on patents as a measure of innovative output, our analysis may underestimate the extent to which research inputs generate ideas. Second, patents differ greatly in their quality or economic impact (Bessen, 2008). To address potential variation in patent quality and impact, we rely on two commonly used measures:

[10] For example, König et al. (2022) find that Chinese firms that perform R&D and actually obtain patents experience larger TFP gains than R&D performers without patents.

[11] The mean lag between patent application and granting is 2.6 years, with median, 90th, 95th, and 99th percentiles at 2, 5, 6, and 8 years, respectively.

[12] See Figure A1 for a breakout of foreign and domestic granted patents by application year. Our domestic patents include patents by foreign firms when they list a US establishment as an assignee.

forward "external" citations and breakthrough patents.

To compute external citations, we follow the recommendation from Jaffe and de Rassenfosse (2017) and sum the total number of forward citations received by a patent within the first five years of the patent's grant date.[13] We exclude the firm's own "self" citations, since these may be more related to patent thickets aimed at preventing competition than to increases in new ideas (Cohen et al., 2000). We measure breakthrough patents from Kelly et al. (2021), hereafter KPST. They estimate patent "novelty" as the ratio of a patent's forward textual similarity to its backward textual dissimilarity with other patents over five-year intervals. In our analysis, we define breakthrough patents as those whose novelty is in the top five percent across all patents from 1836 to 2021, net of year fixed effects.[14] Since citations and breakthroughs are both forward-looking, these measures are available through 2011.

3.2 US Census Establishment- and Firm-level Data

We construct a firm-level panel of employment, payroll, sales, and R&D expenditure by combining multiple datasets from the US Census Bureau. We start with the Longitudinal Business Database (LBD), which provides employment, payroll, industry, and geography for all private, non-farm employer establishments from 1976 to 2021 (Jarmin and Miranda, 2002; Chow et al., 2021). We supplement the LBD with establishment-level sales using the quinquennial Economic Census data conducted in years that end in 2 or 7.[15] We restrict the analysis to establishments that are in scope for the Business Dynamics Statistics (BDS) data, which ensures that we use research-quality observations and have sample that aligns with published official statistics on the US economy. This limitation excludes establishments in Government and those that the Census Business Register considers out of scope. We also drop Public Administration (NAICS 92) and Agriculture (NAICS 11) establishments, as well as establishments with zero payroll (i.e., "non-payroll active" establishments).

We aggregate the establishment-level data to the firm level using the Census Bureau's cross-sectional firm identifier (*firmid*), which we correct for spurious longitudinal breaks that occur when firms transition between single and multi-unit status. A benefit of the Census

[13]Windowing ensures that all patents have the same time to accumulate citations. We classify citations from patents for which we do not have a matched firm identifier (about 8 percent of patents) to be an external citation.

[14]Their updated dataset includes over 11.4 million patents granted from 1836 to 2022. They define a "breakthrough" patent as one whose novelty is in the top 5 or 10 percent across all patents over that interval, net of year fixed effects. Intuitively, these patents represent "breaks" with the past that are also widely emulated going forward.

[15]The Economic Censuses include the Census of Manufactures, Wholesale Trade, Retail Trade, Services, Finance, Insurance, and Real Estate, Construction, Transportation, Communication and Utilities, and Mining.

establishment-level data is that it records every establishment's primary industry code in each year. We construct annual measures of each firm's employment and payroll across sectors using the vintage-consistent North American Industrial Classification (NAICS) codes developed by Fort and Klimek (2018). These codes ensure that we assign establishments consistently to sectors and industries over the period, even as the underlying classification systems change. We assign firms to their primary industry in each year based on the majority of their payroll. The mix of firm employment across establishments in different industries is also crucial for the construction of alternate measures of innovation inputs, described below. We follow the BDS and define firm age as the difference between the current year and the first year of positive employment of the firm's oldest establishment in the year that the $firmid$ first appears.

The majority of past studies on research efficiency measure innovation inputs using R&D expenditures. We therefore compile R&D expenditures over the 1977 to 2021 period for the subset of firms in the National Science Foundation's R&D surveys, collectively referred to as RADS. These surveys provide R&D expenditures for a rotating sample of approximately 45 thousand firms each year, though the sample size and frame varies over time. We aggregate R&D expenditures to the firm level and merge them to the LBD panel (see Appendix D for details).

3.3 Compustat Firm Panel

Since much past work relies on publicly traded firms, we construct a similar panel using Compustat, which contains financial information for publicly traded firms derived from companies' Security and Exchange Commission filings. The data include firms' global sales, employment, General and Administrative (SG&A) expenses, and R&D expenditure.

3.4 New Measures of Research Inputs

A key ingredient necessary to estimate research efficiency is a measure of knowledge or innovation inputs. While R&D expenditure is a natural starting point, the rotating nature of the R&D surveys, changes in their sample-selection criteria, and their tilt towards large manufacturing firms pose serious limitations on the sample of firms we can study. Past work also documents "missing R&D" expenditures among public firms that could bias estimates of the impact of R&D on patenting (Koh and Reeb, 2015).

To mitigate potential biases from sample selection or variation in R&D reporting over time, we propose three new measures of 'research inputs' in the LBD panel based on a firm's payroll in innovation-related establishments. Our most narrow measure calculates a firm's wage bill in R&D labs (NAICS 5417). While such labs are appealing due to their clear focus on innovation, they may miss considerable research efforts to the extent that plants whose major activity is not R&D nevertheless contain workers engaged in these activities. We therefore construct a second measure based on the firm's wage bill in Professional, Scientific, and Technical Services as well as Management establishments (NAICS 54-55). These sectors cover establishments engaged in high-skill, more knowledge-intensive tasks such as engineering and computer system design.[16] Finally, we provide an upper bound on firms' R&D employment expenditures using their total payroll across all establishments. Although clearly a crude measure, total payroll covers all potential patenting firms in our sample.

For Compustat firms, we use both capitalized selling, general and administrative expenses (SG&A) – a broad measure of innovation inputs that is often used in the construction of firms' intangible assets – as well as their reported R&D expenditures.

We deflate the R&D variables, as well as all other nominal variables, e.g., firm sales and payroll in the LBD, using the CPI. While some researchers deflate these series using the average skilled-worker wage (e.g., Bloom et al., 2020), we view the CPI as conservative for our purposes insofar as it captures increases in worker skill and ability that would be deflated away using skilled-worker wages.

3.5 Matching Patents to Firm-level Panel Data

We merge the USPTO patent-assignees to firm names and geography in the Census Bureau's Business Register (CBPBR). The CBPBR is built from the universe of administrative Internal Revenue Services payroll tax filings and provides the frame for the LBD. The resulting crosswalk (BDSPF-Long, described in Appendix A) makes four contributions for future work: it spans the longest period of existing patent-to-Census firmid matches (1977 to 2021), it prioritizes longitudinal consistency (92 percent match rate, 94 percent precision), it is available with its code to FSRDC researchers, it is used to create public-use tabulations of patenting firms' business dynamics statistics, and it will be updated annually.[17] To the best of our knowledge, it is the first bridge to incorporate longitudinal consistency when matching external data to a Census firm definition.

Although the BDSPF-Long has higher match rates than previous efforts, it is important

[16] NAICS 54 encompasses nine 4-digit NAICS industries: legal (5411), accounting (5412), engineering (5413), design (5414), computer system design (5415), consulting (5416), R&D (5417), advertising (5418), and other (5419). We include "Management" establishments (NAICS 55), since they often employ R&D workers, and because establishments that perform two or more Professional, Scientific, or Technical Services for other establishments of their firm are classified in this sector.

[17] The BDSPF tabulations are available at `https://www.census.gov/programs-surveys/bds.html`. Appendix A provides a comparison of BDSPF-Long to existing matches by Balasubramanian and Sivadasan (2011), (1975 to 1997), Kerr and Fu (2008) (1976 to 2000) and Graham et al. (2018) (2000 to present).

to note that the match rates of all bridges between patents and firms decline over time, in our case from 95 percent in the early 1970s to 90 percent in the late 2010s (Appendix Figure A1). This decline is driven by increases in the shares of patent-assignee records that do not match to the CBPBR at all, as well as those that match based upon fewer and noisier characteristics, trends that suggest patents are increasingly granted to small (i.e., single-establishment) firms.[18] Since we use patents as our measure of ideas, the secular decline in match rates is most likely to exert a downward bias in our estimates of research productivity.

We use the BDSPF-Long (patent to CBPBR) crosswalk to assign patents to *firmid*s in our LBD panel. We create a separate match of US patents to the publicly traded firms in Compustat by combining two recent patent-to-firm mappings developed by Kogan et al. (2017) and Dyevre and Seager (2024). We exploit these two mappings to increase the overall number of patents we can match to Compustat, and to cross-validate patents encompassed by each approach. See Appendix B for further details.

In both panels, we match patents to firms by grant year, but use the patent *application* year in our analyses so that they are most proximate in time to innovation investments.[19] As noted at the beginning of this section, because of the typical gap between patents' application and grant years, our analysis sample period runs from 1977 to 2016 for patent grants, and 1977 to 2011 for external citations and breakthrough patents.

We report match rates for patent grants and breakthroughs by semidecade in Appendix Table B1.[20] As the firms in our LBD panel are a subset of the entities in the CPBPR (see Section 3.2), the match rate for the LBD is lower than quoted above — 89 percent of all domestic patents in the first semidecade versus 85 percent by the final semidecade. We note, however, that the beginning-to-end 4 percentage point decline in the match rate in the LBD is slightly less than the 5 percentage point decline in the overall crosswalk rates in the CBPBR. By contrast, the match rates in Compustat start lower and decline relatively more over the same period, falling from 63 to 55 percent. The LBD panel thus covers 26 to 34 percent more patents than Compustat, with steadier match rates throughout.

Match rates for breakthrough patents are also higher in the LBD panel relative to Compustat. The LBD panel covers 91 percent of breakthroughs in the first period and 87 percent

[18]Appendix Figure A3 shows that match rates would have remained at about 94 percent if the raw match-pass rates had held constant. Finding a match between a patent assignee and a single-establishment firm is probabilistically more challenging because such firms have only one observation in each year in the CBPBR. The Census collects more detailed, numerous, and frequent information for multi-unit firms.

[19]Our LBD panel also misses any matches in the crosswalk for which the matched LBD firm is not active in the patent's application year, though we allow for pre-birth patenting in the growth regressions in Section 6 when constructing firms' patent stocks.

[20]We do not report rates for external citations, since these are defined only for matched firms, or for patent value since this measure only exists for publicly traded firms.

by the end, whereas Compustat coverage is about 73 to 72 percent in those periods. Compustat coverage is also more volatile over the four decades: in the 1987 to 1991 period, its coverage falls to just 60 percent of breakthrough patents, while the LBD panel is steady at 87 percent.

The share of patent activity in Compustat may decline for at least two reasons. First, as a higher share of firms chooses to stay private, Compustat's coverage of economic activity will naturally decline (Doidge et al., 2017; Schlingemann and Stulz, 2022). Second, as we will show in the next section, the composition of US patenting firms shifts towards smaller, non-manufacturing firms, both of which are less likely to be in Compustat. Regardless of the underlying reasons for the decline, the lower and more volatile match rates in Compustat suggest that those data may not provide as accurate a representation of US research efficiency as the LBD, since falling coverage may manifest as falling efficiency.

4 New Facts on US Patenting Activity

In this section, we exploit our long time series on the universe of private-sector firms to provide 4 new facts about US patenting activity that are relevant to understanding the evolution of US research inputs, outputs, and productivity.

4.1 Trends in US Patent Activity

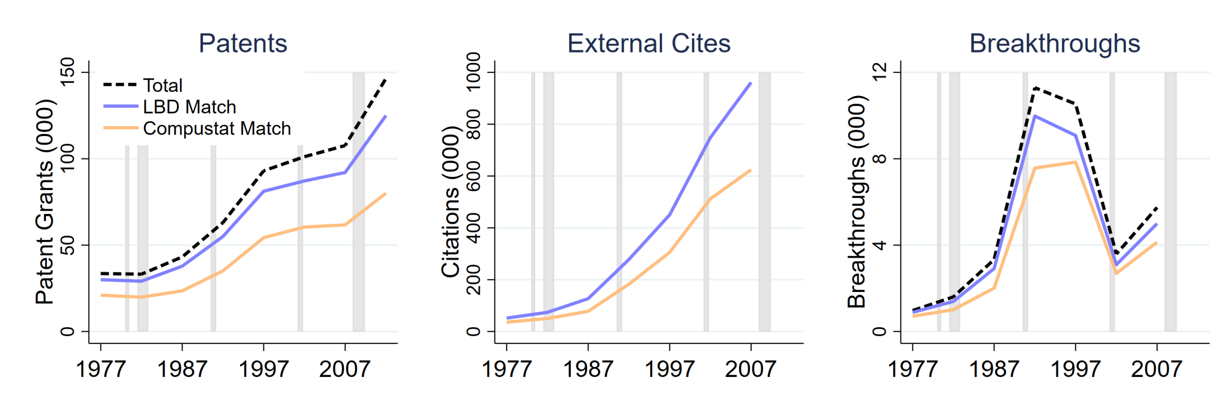
Figure 2 depicts patent grants, citations, and breakthroughs in the LBD and Compustat panels, where each point is a five-year forward-looking average.[21] Patent grants and citations rise over the period, more so among the universe of firms than among Compustat firms. Breakthroughs surge in the middle of the sample period, plummet, and then begin to recover. This breakthrough pattern is strikingly similar to an inverted-U shape of new technologies identified by Bloom et al. (2025) in top-cited US patents over the same period.

4.2 Patents per R&D Input

A natural first metric for analyzing whether ideas are getting harder to find is research productivity, measured as average patents per R&D input. Indeed, some of the key growth papers in the 1990s developed models to match a salient fact on declining patent productivity: the number of patents per R&D researcher (or R&D dollar) declined both in the US and across countries starting in the 1970s (Evenson, 1984; Kortum, 1993; Jones, 1995; Kortum,

[21] Here, and throughout the rest of the paper, we report results by semidecade to minimize Census disclosure burden.

Figure 2: Patent Activity in the LBD and Compustat

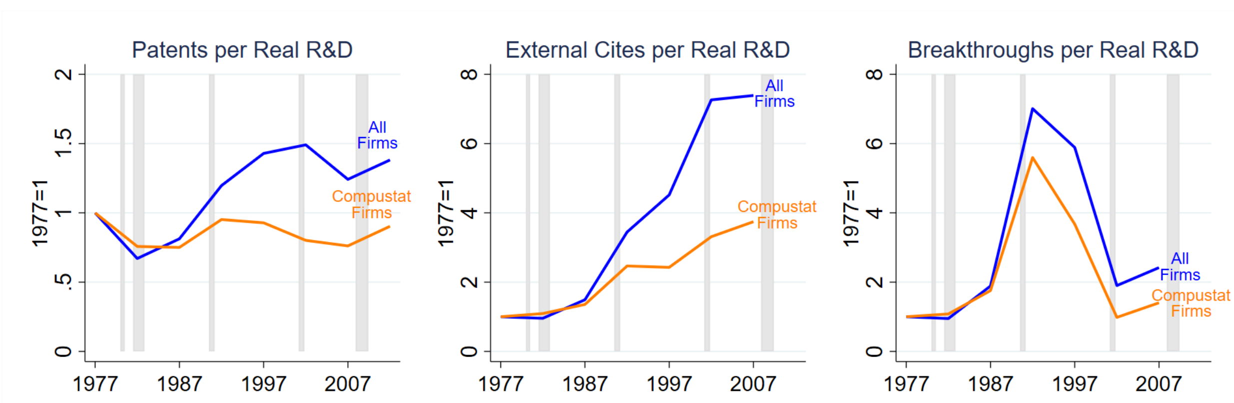


Source: PV, LBD, KPST, BDSPF-Long, DS, Compustat and authors' calculations. *Notes:* Figure compares LBD- and Compustat-matched patent grants, external citations, and breakthrough patents to their totals in the PV data. Each point represents a 5-year forward-looking average, *e.g.*, the data point for 2012 is the average across 2012 to 2016. Data for breakthroughs and external citations end in 2007 - 2011 because these are 5-year forward-looking measures. External citations are defined only for matched patents, so no totals are reported. All series are plotted by patent application year.

1997). We replicate past work to depict how such research productivity evolves in the LBD and Compustat firm panels and over a longer time period.

Figure 3 presents research productivity – patent activity per real R&D expenditure flows over time in the LBD and Compustat. Prior work divides R&D expenditures by R&D worker wages to infer the number of researchers in these calculations. We deflate expenditures by the CPI to allow for heterogeneity in researcher ability that would likely be captured in greater real expenditures. We again minimize Census disclosure burden by reporting results by semidecade.

Figure 3: Research Productivity in the LBD and Compustat



Source: PV, LBD, KPST, RADS, BDSPF-Long, DS, Compustat and authors' calculations. *Notes:* Figure reports patents, breakthroughs, and external citations per real R&D expenditure for firms in the LBD and Compustat. Ratios are averages across 5-year semidecades from 1977 to 2012, and these averages are indexed to 1 in the first, 1977 to 1981 semidecade. Patents are assigned to firm by application year. In the LBD panel, the set of firms included in the ratio is restricted to those in the R&D surveys in each year.

The first panel of Figure 3 shows patents per real R&D dollar. Among the universe of LBD firms, and in contrast to previous research, this measure rises to roughly 1.5 times its initial level by the final period. Among the subset of publicly traded firms in Compustat, research productivity tracks the trend for all firms during the first decade but diverges sharply thereafter, ending the sample below its starting point.

Citations per R&D dollar grow steadily until they flatten in the final decade. Breakthroughs per R&D dollar exhibit an inverted-U driven by the aforementioned surge, but finish the period above their starting level. There is no evidence of a secular decline in either of these activities among both groups of firms. These trends are summarized in our first fact:

Fact 1. *Research productivity – measured as patent grants, external citations, or breakthroughs per real R&D dollar – all rise over the last four decades for the universe of firms. By contrast, Compustat firms' patent grants per R&D dollar are lower at the end of the sample period than the beginning. Their citations and breakthroughs per R&D dollar, however, both rise, albeit less sharply than for the universe of firms.*

Fact 1 demonstrates the importance of examining the universe of firms over a long time horizon using multiple measures of ideas to draw conclusions about research productivity.

4.3 Changing Composition of Patenting Firms

To understand the evolution of US patenting and the sharp divergence in research productivity documented above, we decompose LBD firms' patenting activity along three dimensions: the firm's major sector of activity (based on payroll), the firm's size according to employment, and the firm's age. We show how these activities have changed over time with separate decompositions for each decade from the 1970s to the 2010s, where the first and last decades consist of 1977 to 1979 and 2010 through 2016, respectively.

The top row of Figure 4 reveals a steady decline in manufacturing firms' shares of patent grants, citations, and breakthroughs, which decrease from above 70 percent in the 1970s to 30, 23, and 12 percent, respectively, by the 2010s.[22] Manufacturers are supplanted by firms in Wholesale, Professional Services, and Management, which account for about 45 percent of each patent activity by 2010s. Firms in Information, which includes software, telecommunications, and data processing, are the most prominent breakthrough patenters in later years, accounting for 37 percent of breakthroughs in the last decade. We summarize these stark patterns in a second fact:

[22] See Appendix Table G2 for exact numbers.

Figure 4: Patent Activity by Firm Sector, Size, and Age (LBD Panel)

Source: PV, LBD, KPST, BDSPF-Long, and authors' calculations. *Notes:* Figure provides a breakdown of patent grants, external citations, and breakthroughs by LBD firms' major 2-digit NAICS sector, size, and age. Firms in the LBD cannot be observed prior to 1977, so age is left-censored in early years for most firms.

Fact 2. *Manufacturing firms accounts for the majority of US patenting until the 2000s, after which firms in Wholesale, Professional Services, Management, and Information become most important.*

Fact 2 demonstrates that a complete evaluation of US researcher efficiency must expand beyond the manufacturing sector. The rise in patenting among Wholesale and Professional Services firms resonates with work on the increasing prevalence of factoryless goods producers – firms that design goods and coordinate the production process, but outsource physical transformation tasks to other firms that are often in other countries (Bernard and Fort, 2015; Kamal, 2023; Fort, 2023). Firms with Management as their primary sector may reflect US multinationals' strong role in US innovation, given their disproportionate employment shares in that sector (Kamal et al., 2022). The evidence here indicates that these firms, some of which may have transitioned from the manufacturing sector (Ding et al., 2022), are increasingly important for US innovation, but likely to be missed by studies with the traditional focus on manufacturing.

The middle row in Figure 4 (see also Appendix Table G3) depicts patent activity shares by firm size categories. Firms with more than 10 thousand employees, *i.e.*, "mega firms", see their share of patent grants and citations fall from 58 percent in the late 1970s to about 48 percent by the 2010s. Small firms (1-19 workers) see the largest share gains – roughly 3 percentage points – despite aggregate employment shifting toward mega-firms over this period (see Appendix Table G3).

Mega-firms' share of breakthroughs evolves differently, as it is U-shaped over time.[23] Firms with 100 to 500 employees, by contrast, almost double their share of breakthroughs in the 1980s, while firms with 2,500 to 10 thousand workers see their portion of breakthroughs jump in the 1990s. These changes likely reflect a combination of strong growth by a cohort of innovative firms and a shift in the types of firms that innovate. For example, the 1990s witnessed an explosion of electrical engineering patents during the personal computer and internet boom (see Appendix Figure F1), with a small subset of these firms growing into the tech giants that persist today.

The bottom row of Figure 4 illustrates how patents are distributed across firms of different ages. Since we construct firm-age bins using the minimum year of the firm's oldest establishment, age is left-censored in different years for different categories making time-series comparisons difficult. We therefore focus on comparing shares within an age group across patent activities.

Young firms (0-10 years) have disproportionate citation shares relative to their patents, while 11 to 15 year-old firms have outsized breakthrough shares, particularly in the 2010s. Firms born around 1995 to 2000 – just after the breakthrough boom – are especially strong innovators 10 to 15 years later.

Fact 3. *Mega-firms (10k+ employees) dominate US patenting until the 1990s, but their shares of patents and citations fall to less than 50 percent by the 2010s. By contrast their share of breakthroughs dips to 59 percent in the 1990s compared to about 70 percent at the*

[23] Braguinsky et al. (2025) define novel patents as those that combine two technology classes for the first time and similarly find that mega-firms' importance in such patents exhibits a U-shaped pattern over this period.

beginning and end of the sample period. 11 to 15 year-old firms exhibit a rise in breakthrough shares from the 2000s to the 2010s.

These changing patterns of firm size and age in US patenting highlight the need to control for both factors when evaluating research output and its impact on growth. They also suggest that innovation may come in bursts from new discoveries, followed by growth and consolidation by a subset of those firms with breakthroughs. The recovery of breakthrough shares by a small set of mega-firms may also signal a changing market structure for new technology fields in which new innovators grow and then stifle further entry.

4.4 Trends in Research Input Coverage and Growth

In the next section, we will estimate the elasticity of patents to research inputs. The R&D expenditures used as research inputs in most studies are derived from NSF surveys that tend to target large manufacturing firms, which Facts 2 and 3 indicate are less relevant over time. The first column of Table 1 shows that firms reporting positive R&D expenditure in these surveys account for cover 65 percent of US patents in the 1970s and 1980s, but only 59 percent thereafter. This decline occurs even as the share of firms participating in the survey triples from 0.0004 to 0.0012.

Table 1: Shares of Patents and Firms by Firms with Research Input Flows

	R&D Expenditure		R&D Lab Pay		Prof & Management Pay	
	Patents	Firms	Patents	Firms	Patents	Firms
1970	0.65	0.0004	0.54	0.0012	0.77	0.07
1980	0.64	0.0004	0.55	0.0014	0.77	0.09
1990	0.59	0.0004	0.48	0.0017	0.75	0.11
2000	0.59	0.0007	0.49	0.0021	0.77	0.13
2010	0.59	0.0012	0.48	0.0024	0.77	0.14

Source: PV, LBD, BDSPF-Long and authors' calculations. *Notes:* Table reports shares of total patents and firms covered by firms with positive R&D expenditure, payroll in R&D labs, or payroll in Professional Services or Management establishments.

The trend is similar for firms with payroll in R&D labs (middle panel): their patent coverage falls from 54 to 48 percent, while their share of all firms doubles from 0.0012 to 0.0024. Firms with positive payroll in Professional Services and Management – which include R&D labs as well as other innovation-related activities such as Computer Systems Design and Engineering – represent about 77 percent of patents throughout the sample period. Their share of firms also doubles, from .07 to .14.

Fact 4. *The share of patents covered by firms in the Census R&D survey falls from 65 percent in the late 1970s to 59 percent in the 2010 decade. Firms with payroll in Professional, Scientific, and Technical Services and Management establishments cover 75% of US patents from 1977 through 2016.*

Fact 4 motivates our use of payroll in innovation-related establishments as additional measures of research inputs. We compare these measures in Figure 5, which depict average aggregate innovation-input flows across firms. As above, each point contains a five-year forward-looking average, and each series is normalized to 1 in the first period. The first panel shows that R&D expenditure among all firms rises substantially, by a factor of 2.9 over the entire period. R&D lab payroll growth aligns closely with R&D expenditure, while payroll in Professional Services and Management establishments grows the most, rising by a factor greater than 3. By contrast, total real payroll does not quite double over the 40-year period. The second panel shows that Compustat firms' R&D expenditure grows more quickly (by a factor of more than 4), while SG&A expenditure tracks fairly closely with the LBD payroll in Professional and Management Services.

The last two panels of Figure 5 convert these flows into firms' depreciated stocks, which we will use to subsume potential lag structures of R&D when estimating patent elasticities.[24] These stock measures grow substantially more than the flows alone. The most notable difference is that the flattening in R&D expenditure among LBD firms evident in the flows in 1987 to 2002 is muted and appears later for stocks.

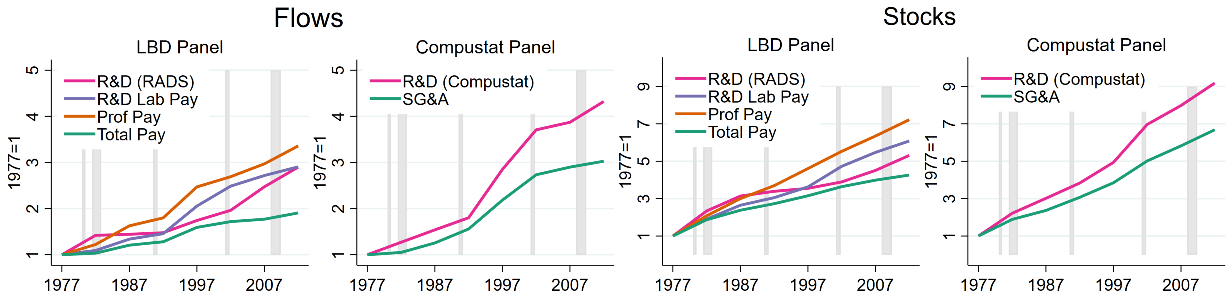
In sum, patent grants, citations, and breakthroughs per real R&D expenditure rise among the universe of firms. These increases coincide with a shifting composition of patenting toward smaller, non-manufacturing firms with lower Compustat and R&D -survey coverage. These patterns highlight the need to expand beyond manufacturing to study US innovation, and motivate our construction of alternative research-input measures.

5 Patent Elasticity Estimates

In this section we estimate within-firm elasticities of patents to R&D inputs, controlling for firm and time fixed effects. The time-varying elasticity estimates and semidecade fixed effects provide our second and third measures of research productivity over time.

[24] We use a 15 percent depreciation rate. See Appendix Section E for details.

Figure 5: Research-Input Flows and Stocks



Source: PV, LBD, BDSPF-Long, DS, RADS, Compustat and authors' calculations. *Notes:* First two panels report average annual real aggregate research-input flows across firms in the LBD and Compustat by semidecade. Third and fourth panels report depreciated stocks of these measures. In the LBD panel, research inputs are firms' RADS R&D expenditure, R&D lab (NAICS 5417) payroll, Professional Services and Management (NAICS 54-55) payroll, and total payroll. Series are normalized to 1 in the first semidecade. Reporting is done by semidecade to minimize disclosure concerns, *e.g.*, a data point for 2012 is the average value across firms from 2012 to 2016.

5.1 Empirical Specification

We estimate patent elasticities by adapting equation (6) to

$$Y_{ft} = exp\left(\sum_{j=1977(5)k} \eta_j d[j]_{ft} \times ln(ResearchInputs_{ft}) + \sum_{j=1977(5)k} \gamma_j d[j]_{ft} + \gamma_f\right) + \epsilon_{ft}, \quad (8)$$

where Y_{ft} is firm f's patent output (patents, citations, or breakthroughs) in year t. The first term on the right-hand-side interacts the firm's R&D inputs (H_A in Romer's notation) with semidecade dummies that are equal to 1 for years t in semidecade j. For example, $j = 1977$ encompasses years 1977 to 1981. η_j is the elasticity of interest, capturing the marginal change in patent output with respect to a change in firm f's real research-inputs during semidecade j. γ_j are semidecade fixed effects that capture average idea creation in a particular five-year period that is common to all firms, after controlling for variation in their research inputs.[25] Firm fixed effects γ_f ensure that our estimates are identified from variation within a firm over time.

We use each of the four measures of research inputs described in Section 3.4 in four separate specifications. Although research inputs are a flow in the growth models, there may be a lag between inputs and outputs in our firm-level setting. In practice, we find that depreciated stocks of firms' past research-input expenditures depicted in Figure 5 capture a range of lags well, while reducing disclosure burden and enabling us to estimate time-varying elasticities.[26]

[25] We use annual data to estimate semidecade elasticities and fixed effects to minimize Census disclosure burden. Specifications with yearly elasticities and year fixed effects yield similar conclusions.

[26] In line with our findings, Hausman et al. (1984) show that the current flow of R&D expenditures generally

We estimate equation (8) via Poisson-Pseudo Maximum Likelihood (PPML). The set of firms included in each regression is restricted to those with at least one patent, citation, or breakthrough over the period and for which we observe an input $flow$ in that year, as we set stocks to missing if flows are not present. For the payroll-based stocks, this refinement is equivalent to dropping firms in any year in which they do not have those establishments. For the LBD R&D stock, firms are dropped when they do not appear in the RADS with positive R&D expenditures. Standard errors are clustered by firm.

5.2 Patent Elasticity Results

The top row of Figure 6 plots our estimates of η_j for the LBD panel. The elasticity of patents to R&D expenditure (left panel) increases from 0.26 to 0.48 between the first and last periods, indicating that a 10 percent increase in R&D expenditure is associated with 2.5 percent more patents in the beginning period versus 4.6 percent more at the end. These magnitudes are in line with those in Hausman et al. (1984)'s examination of 128 public firms from 1968 to 1974, but our increasing trend contrasts with their finding of lower elasticities over time.[27] Citation and breakthrough elasticities to R&D, reported in the second and third panels, also rise over the period, though are not always statistically significantly different.

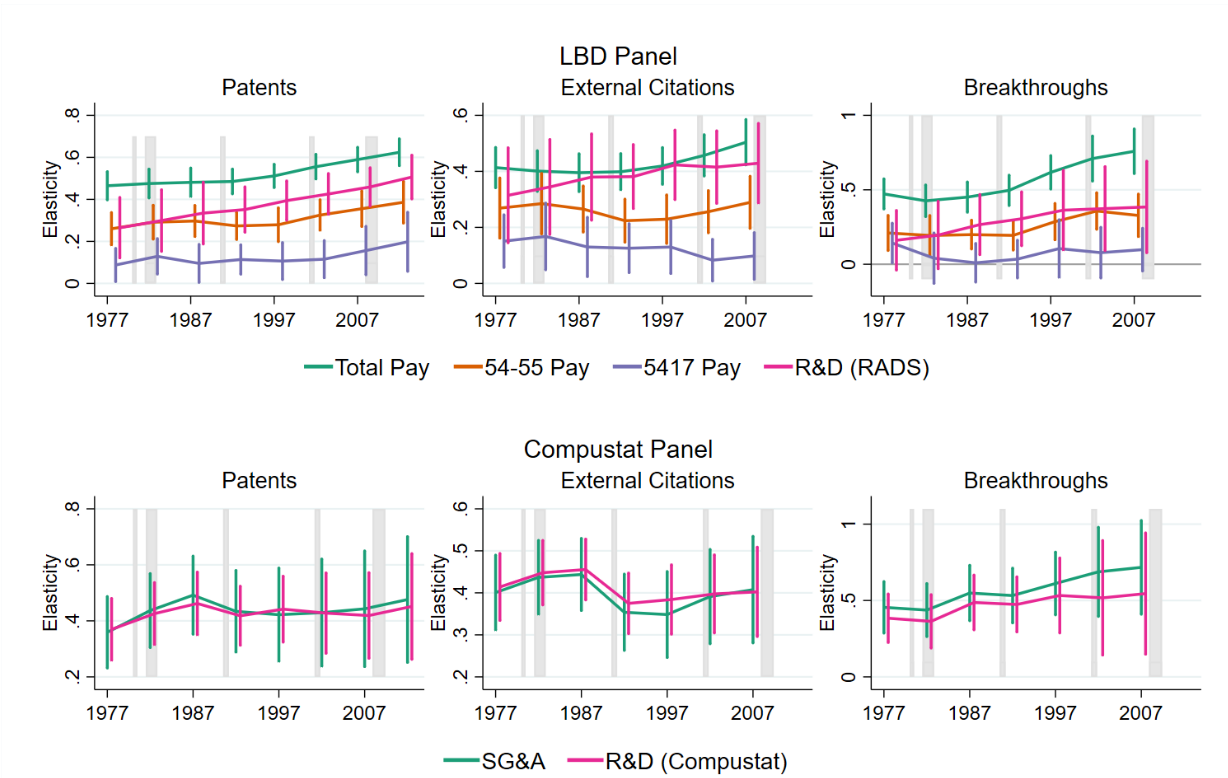
The remaining estimates in the top row of Figure 6 expand beyond the R&D survey sample and report elasticities with respect to Professional Services and Management pay, R&D lab pay, and total payroll. Across all but one specification, elasticities are rising or flat. The exception is the elasticity of citations to R&D lab pay, which declines slightly though not in a statistically significant sense.

Across the four research inputs, patent elasticities are generally highest for the total payroll stock and lowest for R&D lab pay. Some of this variation may arise from the very different levels in patenting across firm types. Firms in the total pay sample average 0.22 patents per year in the late 1970s, while R&D labs average 4.3 in that period (Appendix Table G1). Another possibility is that R&D labs may increasingly provide services to other firms, whereas in the past they focused on in-house R&D. Such outsourcing will reduce our ability to link their R&D inputs to idea outputs. Consistent with this explanation, the R&D labs sample grows five-fold over the period, while their average number of patents per firm falls almost in half (Appendix Table G1) and their patent coverage declines (Table 1). This changing sample composition for R&D labs is also evident in the firm fixed effects, which we

subsume lags once firm fixed effects are included.

[27] Hausman et al. (1984) use a Poisson specification with firm fixed effects and a time trend. They find a same-year elasticity of 0.31 when lags for the previous 5 years are included, and 0.35 when lags are excluded (see their Table 2). However, when also including an interaction of the time trend with current R&D expenditures, their estimates imply a *decline* in the elasticity from 0.48 to 0.34 over their eight-year period.

Figure 6: Estimated Patent Elasticities by semidecade (η_j)



Source: PV, LBD, KPST, RADS, BDSPF-Long, DS, Compustat and authors' calculations. Notes: Figure reports estimates of η_j's from equation (8) by patenting activity and research input for the LBD and Compustat panels. Whiskers denote 95 percent confidence intervals. Standard errors are clustered at the firm-level. Underlying coefficients and standard errors are reported in Appendix Tables H1 and H2. Firms counts reported in G1.

discuss below in Section 5.3.

The top row of Figure 6 indicates that the marginal R&D dollar is at least as effective at producing ideas by the 2010s as they were in the late 1970s for the universe of firms. The second row of Figure 6 depicts the same message among publicly traded firms, where the elasticities of patents and breakthroughs to both R&D and SG&A expenditure are flat or rising. For external citations, they rise initially, fall during the early 1990s and then mostly recover by the end of the period.

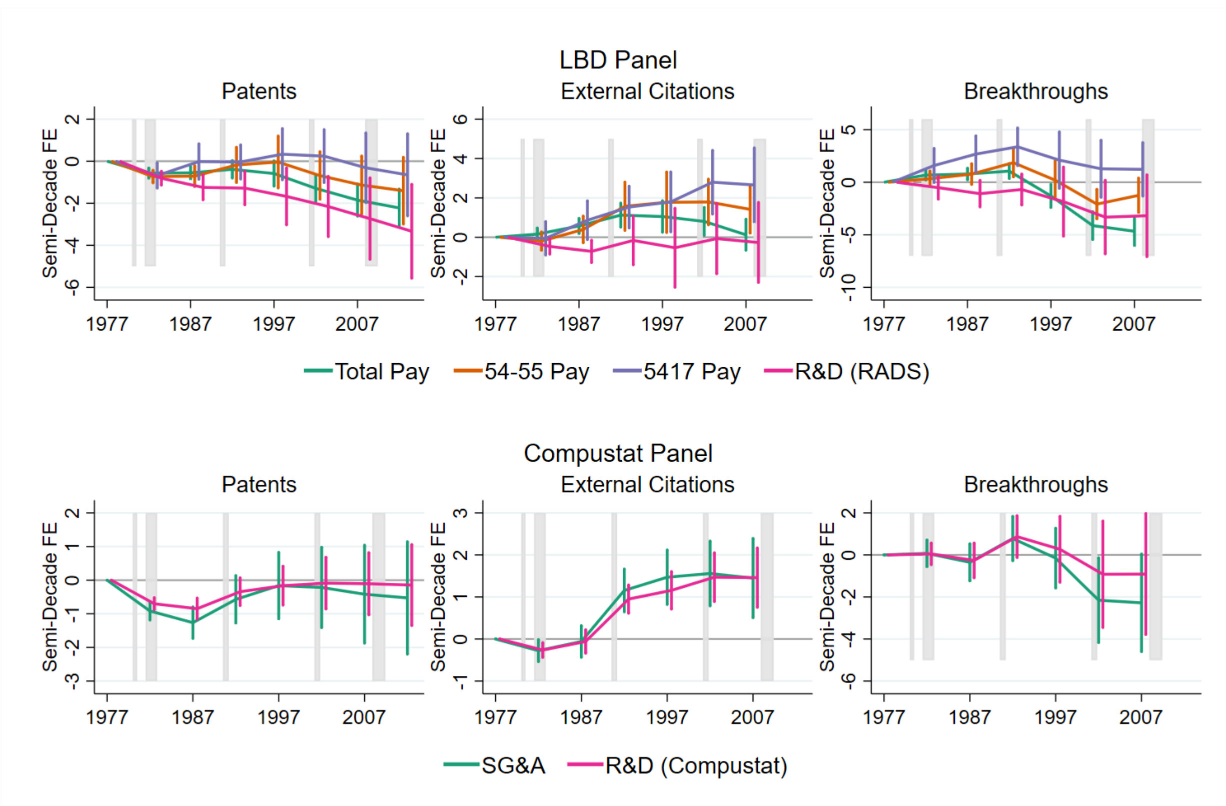
These within-firm estimates directly map R&D changes to patent output, controlling for the possibility that more productive firms spend more on R&D and patent more effectively. These estimates may be biased if time-varying, firm-level productivity shocks increase both R&D and patents simultaneously. For such a bias to drive rising elasticities, however, these shocks would need to be increasing over time, a pattern that in and of itself would seem to contradict the notion that ideas are getting harder to find.

5.3 Semidecade Fixed-Effect Estimates

A third way in which ideas might be getting harder to find is a decline in firms' idea generation independent of their research efforts. In principle, such a decline is captured by the semidecade fixed effects in equation (8), which measure average firm patents per period after controlling for variation their research inputs. In practice, these fixed-effects also include any other tends in innovation common to all firms, such as changes in patenting laws or macroeconomic fluctuations.

Figure 7 reports the estimated semidecade fixed effects ($\hat{\gamma}_j$) using the same format as the previous figure. Overall, across all 18 specifications summarized in the figure, we document statistically significant secular declines in just two instances: average grants after controlling for changes in R&D expenditure or changes in total payroll. The trends are flat, rising, or mixed in 15 of the 18. We thus do not find systematic evidence that the average number of patents per firm declines over time when examining high-quality patent measures and a range of research inputs.

Figure 7: Estimated Semidecade Fixed Effects from Patent Elasticity Regressions (γ_j)



Source: PV, LBD, KPST, RADS, BDSPF-Long, Compustat and authors' calculations. *Notes:* Figure reports estimates of γ_j from equation (8) by patenting activity and research-input stock for the LBD and Compustat panels. Whiskers denote 95 percent confidence intervals. Standard errors are clustered by firm. Underlying coefficients and standard errors are reported in Appendix Tables H1 and H2.

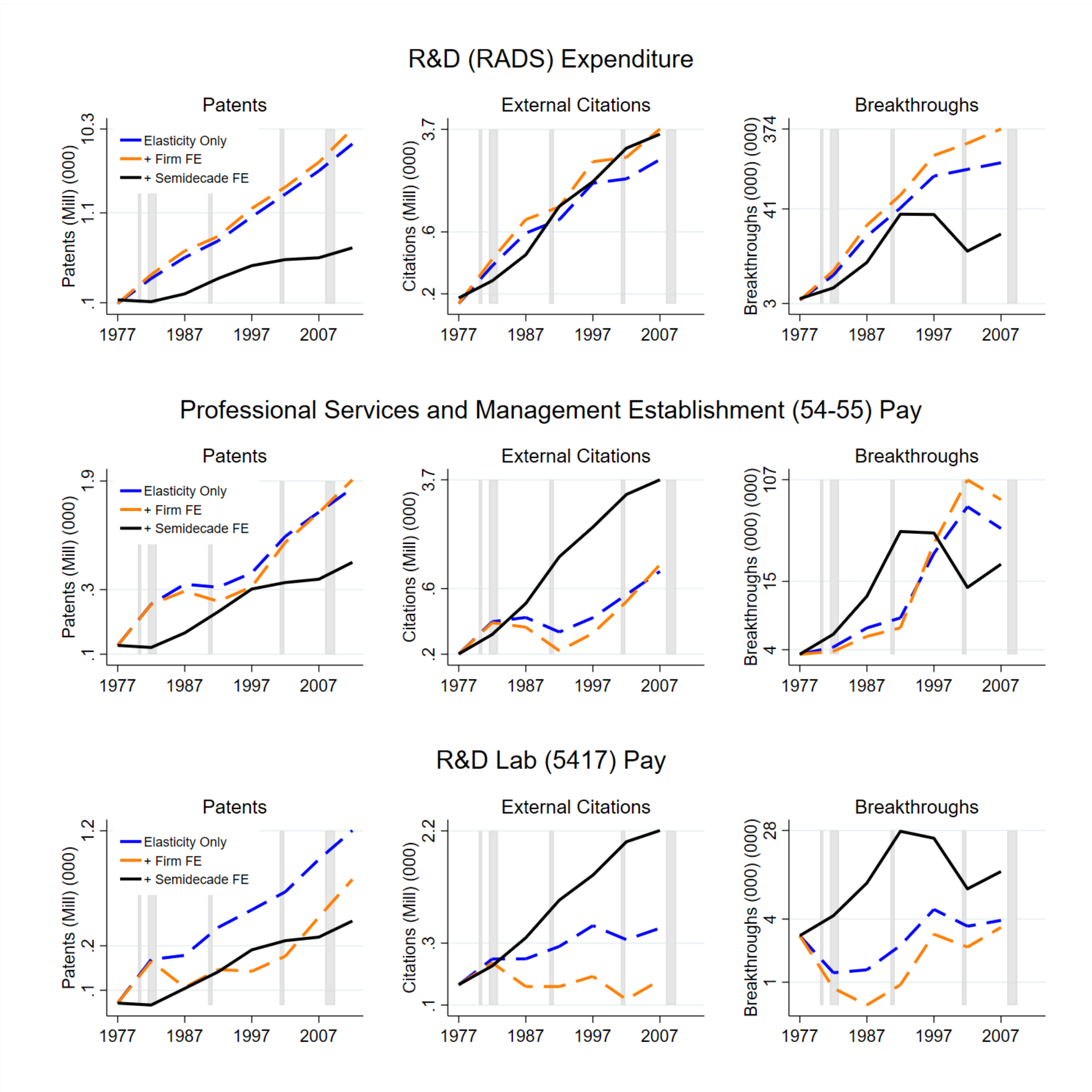
Figure 8 illustrates the economic significance of our estimates by plotting total predicted patents, citations, and breakthroughs for each research input using our coefficient estimates from equation (8). We construct each prediction at the firm level and then sum over firms to obtain totals. The first line in each panel (blue dash) plots predicted patents using only firms' observed research inputs and the estimated elasticities. The second (orange dash) adds in the role of the firm fixed effects. While time-invariant for each firm, their contribution to overall predicted patenting varies as the sample composition changes. The final line (solid black) adds the contribution of the semidecade fixed effects, thereby matching actual US patents by construction. Each line is indexed to the actual number of patents, citations, or breakthroughs in 1977, so that the lines provide information about how the contribution of each factor changes over time, but cannot be compared in levels.

The first row in Figure 8 provides some support for the notion that granted *patents* are getting harder to find: actual patent growth inclusive of the time effects is substantially flatter than that predicted by firms' R&D expenditure (left panel). On the other hand, such a dampening only occurs in the second half of the sample period for breakthroughs (third panel), and not at all for citations (second panel).

The next two rows in Figure 8 show that firms' Professional Services and Management pay (row 2) and R&D Lab Pay (row 3) tend to predict fairly steady increases in firms' patent grants, citations, and breakthroughs. Predicted patenting inclusive of the time effects is somewhat lower for grants (though only in certain years), substantially higher for citations, and changing over time for breakthroughs. For firms with R&D labs, the firm fixed effects increasingly pull down predicted patents, consistent with the changing composition of that sample towards lower patenting firms. Overall, we do not find systematic evidence that the time fixed effects dampen predicted patenting growth across all our measures of patents and research inputs.

Intriguingly, the surge and subsequent decline in breakthrough patents midway through our sample period suggests that the emergence of new technologies or fields – in this case the PC/internet boom of the late 1980s and 1990s – can deliver a temporary burst of ideas. The subsequent leveling off of idea creation within such areas as they mature accords well with the micro evidence presented in Jones (2009), the sector-level analyses in Bloom et al. (2020), and the theoretical models of Evenson and Kislev (1976) and Ribeiro (2025), where applied research within a particular technology has diminishing returns over time but basic research opens up new frontiers. In more recent work, Fieler and Lee (2026) identify idea bursts as the emergence of new patent classes, while Teng (2025) shows that in the pharmaceutical industry, the declining rate of discovery of new functional molecular groups is offset by new drugs that combine such groups. In line with that work, we note that breakthroughs

Figure 8: Predicted Patenting Activity Using Patent Regression Estimates (LBD Panel)



Source: PV, LBD, KPST, RADS, BDSPF-Long, and authors' calculations. *Notes:* Figure reports aggregate predicted patenting activity by semidecade over the 1977 to 2007 (external citations, breakthroughs) or 2012 (patents) sample period. In each case, the aggregate prediction is the sum of predicted firm-level outcomes. Predictions are decomposed into the contribution of research inputs, that plus the contribution of the firm fixed effects, and that plus the contribution of the semidecade fixed effects (see equation 8). The final line corresponds to the actual growth of each activity in our regression samples. Y-axis uses log scale and plots the minimum, maximum and one-tenth the distance between them.

start growing again by the end of the period, perhaps as firms increasingly combine the IT inventions from the 1990s with other products as shown in Braguinsky et al. (2025).

Taking Stock In investigating the mapping between research inputs and ideas, we document flat or rising research productivity across three measures. First, average patents per R&D input has increased for the universe of firms in the United States (Fact 1). Second, elasticities of patents to R&D inputs are less than one, but flat or rising over time, indicating that the marginal R&D dollar is at least as effective in the 2010s as it was in the late 1970s. Third, average patents per firm after controlling for research inputs display statistically significant secular declines in just two of 18 specifications, whereas they are flat, rising, or mixed in 15 of them. Together, these results suggest that is has not gotten harder for firms to translate research inputs into ideas.[28]

6 Estimating the Link Between Ideas and Growth

If it is not getting harder for firms to find ideas, why aren't US growth rates rising as research expenditures increase? In this section, we examine how ideas map to growth, starting with simple descriptive statistics before moving to formal estimates.

6.1 Aggregate Growth in the LBD Panel

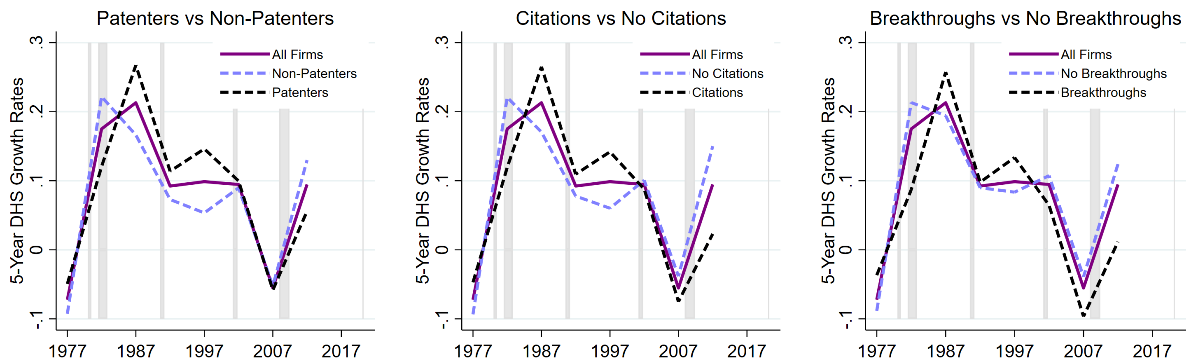
We first document patenting firms' outsized contribution to aggregate US sales growth. Despite representing less than 1 percent of all firms, these innovators account for 28 to 72 percent of total US growth over the period. As demonstrated in Figure 9, firms with at least one patent grant (left panel), citation (middle) or breakthrough (right) exhibit higher 5-year sales growth than non-patenters for most of the first half of our sample period, from 1987 through 2002. These positive growth premia disappear in 2002, and turn negative by the 2010s. This reversal raises the possibility that the relationship between ideas and growth might be weakening towards the end of our sample. On the other hand, the sample of patenting firms may be differentially affected by other macro factors in later years. We investigate these possibilities formally in the next section.

We note that the spike in aggregate growth rates during the 1987 to 1992 semidecade coincides with the addition of the Economic Census of Finance, Insurance, and Real Estate in 1992. In undisclosed results, we drop establishments in year $t + 5$ if they are in the LBD but missing from any of the Economic Censuses in year t, and vice-versa. This correction This correction addresses any spurious changes in growth rates that would arise from establishments being present in the LBD but not the Economic Censuses.[29]

[28] We note that this result does not necessarily imply that an additional research dollar generates an increase in the growth rate of ideas, a point we discuss further in Section 7.

[29] The Census Bureau imposes limits on the number of estimates that can be disclosed for particular

Figure 9: Aggregate Sales Growth By Patenting Activity



Source: LBD, EC, PV, KPST, and authors' calculations. *Notes*: Figure reports the aggregate sales growth from t to $t + 5$ across all firms, and by their patenting activity, *i.e.*, those with at least one patent grant, external citation, or breakthrough over the entire period, versus not. The black dashed lines in this figure are equivalent to those in Figures 10 and 13. The value for 1977 represents growth from 1977 to 1982, etc. .

Figure 10 decomposes the contribution of continuers, entrants and exiters to patenting and non-patenting firms' sales growth. We follow Davis et al. (1996) (DHS) in defining growth as

$$\text{Growth}\left(Sales_f^{t:t+5}\right) = \frac{Sales_f^{t+5} - Sales_f^t}{\frac{1}{2}\left(Sales_f^{t+5} + Sales_f^t\right)}. \tag{9}$$

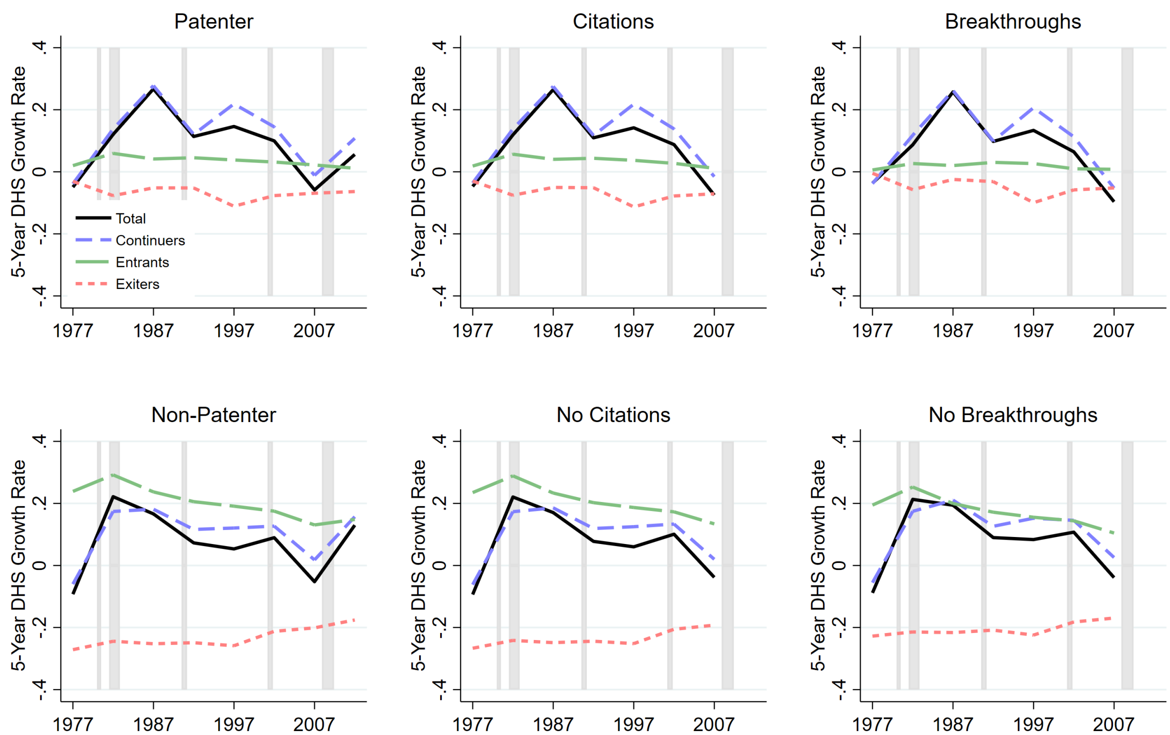
This growth rate is comparable to a log change for continuers, and equal to 2 and -2 for entrants and exiters. As revealed in the figure, continuers account for the bulk of patenting firms' growth, particularly during the 1997 to 2002 semidecade. Among non-patenters, by contrast, entrants make the largest contribution to total growth, though this contribution exhibits a clear secular decline over the period. Our sales-per-worker regression in the next section focuses on continuers, while the sales specifications also include entrants and exiters.

6.2 Growth Regression Specification

We analyze the relationship between productivity growth and patent growth by adapting equation (7) to estimate:

samples, so we are waiting to obtain further comments and suggestions to retain as much disclosure flexibility as possible to respond to those comments. We remain confident in the paper's conclusions.

Figure 10: Aggregate Sales Growth Contributions By Patenting Activity and Margin



Source: LBD, EC, PV, KPST, and authors' calculations. *Notes*: Figure reports the aggregate sales growth from t to $t+5$ across all firms, and by their patenting activity, *i.e.*, those with at least one patent grant, external citation, or breakthrough over the entire period, versus not. The black dashed lines in this figure are equivalent to those in Figures 10 and 13. The value for 1977 represents growth from 1977 to 1982, etc. DHS sales growth is defined in equation 9.

$$\begin{aligned}\text{Growth}\left(X_f^{t:t+5}\right) = &\sum_{j=1977(5)k} \mu_j d[j]_{ft} \times \text{Growth}\left(IdeaStock_f^{t:t+5}\right) + \sum_{j=1977(5)k} \phi_j d[j]_{ft} \\ &+ \sum_{b=1:5} \rho_b^S \{SizeBin_{fb}\} + \sum_{b=1:5} \rho_b^A \{AgeBin_{fb}\} + \nu_{ft},\end{aligned} \tag{10}$$

where the left-hand side is the DHS growth rate of firm f's sales per worker between t and $t+5$, as defined in equation (9). μ_j is the coefficient of interest, capturing the relationship between new ideas and productivity growth. $\text{Growth}\left(IdeaStock_f^{t:t+5}\right)$ is the DHS growth rate in a firm's *stock* of patents, external citations, or breakthroughs in each semidecade.[30] The final two terms are indicators for firm size and age bins, which we use as proxies for capital.[31] These controls also address potential variation in average firm growth rates arising

[30]We assume that firms' ideas do not depreciate when constructing these stocks.

[31]While Census data have an explicit measure of capital for manufacturing firms, Fact 2 demonstrates

from changes in the composition of firms over time (Fact 3). As in our earlier regressions, ϕ_j represent semidecade fixed effects, which capture secular trends in sales per worker growth driven by the other factors that are not included in the specification.

We estimate equation (10) via Ordinary Least Squares on the subset of continuing firms in each of the patent samples (*i.e.*, the dashed blue line in Figure 10). We also estimate a variant of equation (10) with firm sales as the dependent variable in which we also control for firm employment growth, $\text{Growth}\big(Employment_f^{t:t+5}\big)$. These specifications include all firms with patenting changes over the period, *i.e.,* the same samples used in Section 5.1. We follow Davis et al. (1996) and weight the regressions by firms' average sales in t and $t+5$, *i.e.,* the denominator in equation (9).

6.3 Growth Regression Results

Figure 11 presents the estimated coefficients from equation (10).[32] The top panel depicts estimates for the coefficients on firms' ideas ($\hat{\mu}_j$). We estimate a positive relationship between idea growth and sales per worker growth across all three measures that is statistically significant in all but one case (breakthroughs in the 1977 semidecade). Even more striking, we find that these estimates tend to rise over time, *i.e.*, that the relationship between ideas and growth is *strengthening*. For patent grants, a 10 percentage point increase in the firm's patent stock growth rate is associated with a 1.2 percentage point increase in its sales per worker growth rate in 1977, versus a 3.7 percentage point increase in the 2012 to 2017 semidecade.

The bottom panel of Figure 11 reports the semidecade fixed effects ($\hat{\phi}_j$) from the same specifications. The 1977 to 1982 period is the omitted category, so each estimate shows average firm growth over the period relative to growth in the initial period, after controlling for changes in a firm's ideas and its size and age. Across the three specifications and samples, these residual growth rates rise until the early 1990s, after which they exhibit a steady decline. By the end of our sample period, firms' average labor productivity growth rates are close to their 1977 levels (and not statistically different from them).[33] These falling semidecade fixed effects reveal that the decline in growth rates documented in Figure 9 relate to common aggregate shocks rather than a weakening of the link between patent and output growth.

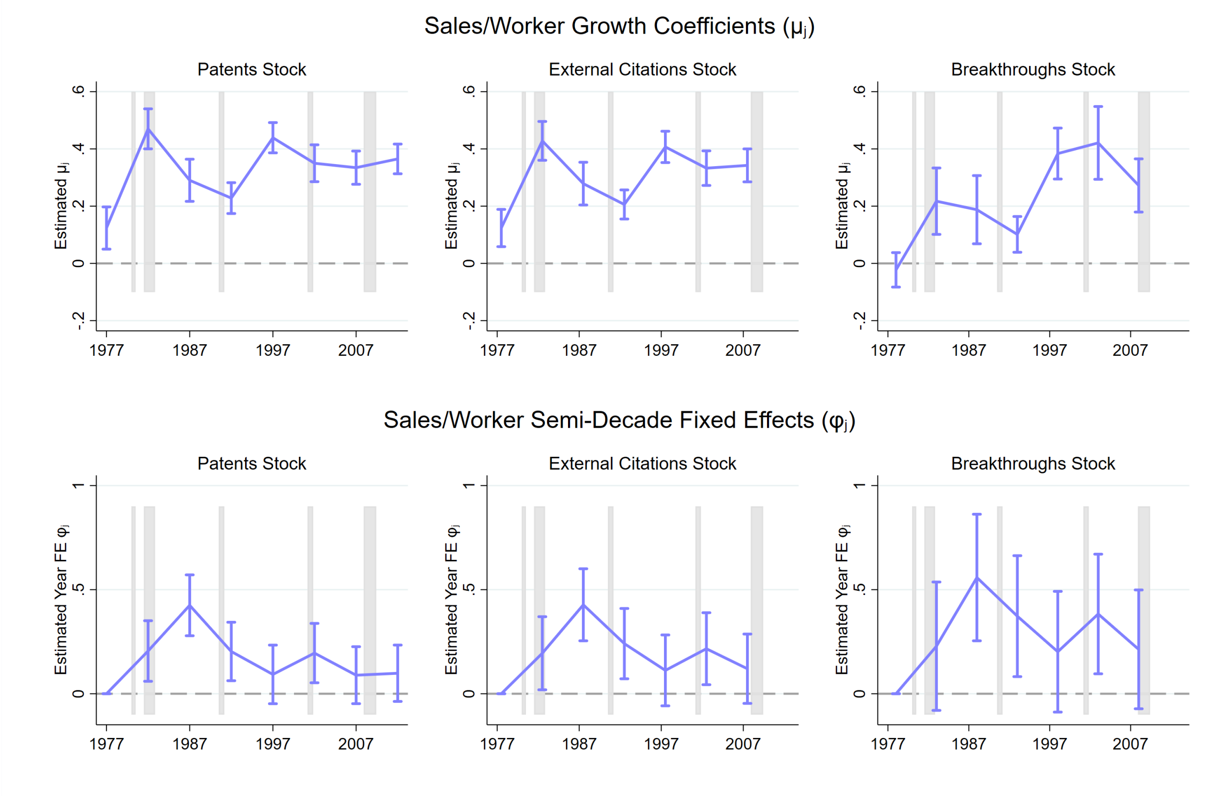
We also estimate equation (10) using the growth rate of firm sales as the dependent variable. This specification allows us to include entrants and exiters along with continuing

that a complete analysis US patenting over our sample period must extend beyond that sector. We use the same firm size and age bins depicted in Figure 4.

[32] Appendix Table I1 reports the estimates.

[33] As noted above, there is a spike in growth rates and thus also in these semidecade fixed effects for 1987 to 1992, some of which is due to the fact that the Census of Finance is first available in 1992.

Figure 11: Estimated Sales/Worker Growth Coefficients (LBD Panel)



Source: LBD, EC, PV, KPST, and authors' calculations. *Notes:* Figure reports relationship between LBD firms' sales per worker and idea stock growth between years t and t+5. Standard errors are clustered at the firm-level with 95 percent confidence intervals depicted. Appendix Table I1 reports corresponding estimates. .

firms, such that we cover essentially all growth by patenting firms over the period (Figure 10). To capture the spirit of Romer, in which new ideas increase output from a fixed set of inputs, we control for firm employment growth, $Growth\left(Employment_f^{t:t+5}\right)$. This addition may attenuate our estimates of the relationship between patent and sales growth if, for example, new ideas induce firms to expand product portfolios by hiring more workers. On the other hand, it mitigates concerns that our estimates of μ_j simply reflect firm growth due to unobserved demand or productivity shocks that increase firm size.

Figure 12 presents these results. Firm sales growth is positively associated with patent grant and citation growth in all semidecades, and is statistically significant in most of them. The estimates for breakthrough patents follow an analogous pattern, though are noisier and only statistically significant in the 1997 to 2002 and 2002 to 2007 semidecades. The semidecade fixed-effect trajectories are similar to those in the sales per worker specifications:

they rise after the first 1977 to 1982 period, are fairly flat through the 2002 to 2007 period, and then decline.[34] A notable difference between the sales per worker versus sales regressions is that the coefficients on patent growth in the sales regressions are substantially smaller than the time fixed-effect estimates, as well as smaller than the patent growth coefficients in the sales per worker specifications. This smaller magnitude may, at least in part, reflect our inclusion of firm employment growth as an additional control variable that co-moves with ideas in driving sales.[35]

Figure 12: Estimated Sales Growth Coefficients (LBD Panel)

Source: LBD, EC, PV, KPST, and authors' calculations. *Notes:* Figure reports relationship between LBD firms' sales and idea stock growth between years t and $t+5$. Standard errors are clustered at the firm-level with 95% CIs depicted. Appendix Table I1 reports corresponding estimates.

.

[34]Again, here we ignore the spike for the 1987 to 1992 period.

[35]In undisclosed results, we estimate equation (10) with firm sales as the dependent variable and *without* controlling for employment growth. This approach is motivated by the insight in Hsieh and Klenow (2009) that differences in revenue productivity across firms may simply reflect misallocation due to distortions. We focus on the specifications here with employment growth controls as a more conservative approach that provides a lower bound on the relationship between ideas and growth, but are happy to request disclosure of this alternative upon request.

Figure 13 uses the coefficient estimates from the sales growth regression to illustrate the contribution of each variable to aggregate growth by patenting firms. The first, red line in each panel reports total predicted sales growth based only on firms' actual patent-stock growth, *i.e.,* the weighted average of each firm's predicted sales growth using $\hat{\mu}_j$ and their patent-stock growth. For patents (left panel) this contribution is positive in all periods, with a marked increase for the 1997 to 2002 semidecade. This predicted increase aligns with the time period in which patenting firms grow relatively more than non-patenters (Figure 9), suggesting that new ideas played a key role in their disproportionate performance.

The blue lines in Figure 13 depict predicted sales growth from firm's employment growth. This contribution is negative and sizable starting in the 2002 to 2007 semidecade across all idea stocks, indicating that firms reduced employment growth considerably in those years, and sales growth followed suit. This period corresponds to the disappearance of patenters' growth premia (Figure 9), suggesting that the impact of ideas on growth may have been swamped by macroeconomic factors in those years.

The green lines in each panel represent the contribution of firms' size and age. These factors play a fairly small role explaining aggregate sales growth that is lower at the end of the sample than the beginning. This trend is consistent with a shift of economic activity towards large and old firms that have comparatively lower growth rates (Haltiwanger et al., 2013).

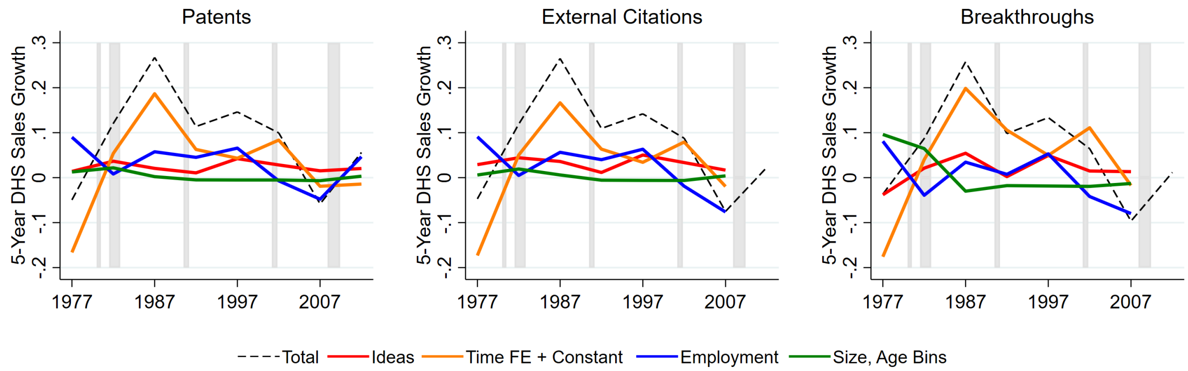
Finally, the semidecade fixed effects (orange lines) predict substantial increases in sales growth after the initial 1977 to 1982 period, but declining growth in later years.[36] These declines capture all aggregate shocks that are common across firms, including any reductions in spillovers across patenting firms.

Table 2 reports the most important numbers underlying Figure 13, including total growth for the patenting-firm sample in each semidecade, along with the predicted growth from firms' actual patent-stock and employment growth. The remaining component ("Other") comprises predicted net growth from the time, size, and age fixed effects. Predicted growth from ideas in each five-year period ranges from 1 to 4 percentage points. It is highest in the 1982 to 1987 and 1997 to 2002 periods. Predicted growth from ideas is equal to just over a quarter of total growth by patenting firms over four decades.

The next two panels in Table 2 report the predicted growth from citation and breakthrough patent-stock growth, along with their shares of the total growth in those samples. Predicted growth from citations is generally the same or higher than predicted growth from patent grants. It ranges from 1 to 5 percentage points and is equal to more one third of total

[36] As noted above, a portion of the spike in these semidecade fixed effects for 1987 to 1992 is due to the fact that the Census of Finance is first available in 1992.

Figure 13: Contributions of Firm Sales Growth Regression Terms to Aggregate Sales Growth



Source: LBD, EC, PV, KPST, and authors' calculations. *Notes:* Figure depicts total predicted sales growth from patent-stock growth (red line), employment growth (blue line), size and age bins (green line), and time fixed effects and constant (orange line). The sum of these predictions matches total growth among patenting firms by construction (black dashed), which is also displayed in Figure 9.

.

growth among the sample of firms with external patent citations. Predicted growth from breakthroughs is more volatile, in line with the uncertainty in generating transformative ideas. Indeed, it is negative in the first period, but then 5 percentage points in the third and fifth periods. Across all measures, patent growth predicts the highest sales growth during the 1997 to 2002 period, which corresponds to a period in which both the elasticity estimates and actual patent growth are relatively high. Predicted growth from patenting is lower after this burst, and especially for breakthroughs.

Table 2: Predicted Growth from Ideas vs Other Factors

	Patent Sample					Citation Sample		Breakthrough Sample	
	Total	Predicted Growth From:			Patent	Predicted Growth From:		Predicted Growth From:	
	Growth	Patents	Emp	Other	Share	Citations	Share	Breakthroughs	Share
1977-1982	-0.05	0.01	0.09	-0.15	-0.20	0.03	-0.60	-0.04	1.00
1982-1987	0.12	0.04	0.01	0.08	0.33	0.04	0.33	0.02	0.22
1987-1992	0.27	0.02	0.06	0.19	0.07	0.04	0.15	0.05	0.19
1992-1997	0.11	0.01	0.05	0.06	0.09	0.01	0.09	0.00	0.00
1997-2002	0.15	0.04	0.07	0.04	0.27	0.05	0.36	0.05	0.38
2002-2007	0.10	0.03	-0.01	0.08	0.30	0.03	0.33	0.01	0.17
2007-2012	-0.06	0.02	-0.05	-0.03	-0.33	0.02	-0.25	0.01	-0.10
2012-2017	0.06	0.02	0.05	-0.01	0.33				
1977-2017	0.70	0.19	0.26	0.25	0.27	0.22	0.35	0.12	0.23

Source: PV, EC, LBD, KPST, BDSPF-Long and authors' calculations. *Notes:* Table reports total growth by period, as well as predicted growth using the estimated coefficients depicted in Figure 12 along with firms' actual patent growth, employment growth, and the estimated time and size-age bin fixed effects. Patent Share is the fraction of total growth over the period that is predicted from firms' actual growth in patent grants. Citation and breakthrough samples report the contributions of citation and breakthrough growth only, and their shares of total growth in those samples.

Overall, our growth regressions show that firms' patents are associated with higher growth rates, that there is no systematic weakening of this association over time, and that declines in firms' average growth rates due to other factors work against the contribution of ideas after the early 2000s. In Appendix Section I.2, we show that these patterns are also evident in unweighted regressions.

7 Discussion

Our firm-level analyses provide empirical support for a key channel in macro-growth theory: that firms' profit-seeking R&D investments generate ideas, and that growth in these ideas in turn raises productivity.[37] In contrast to the hypothesis that these channels are weakening, we document robust relationships for both, even as the aggregate stock of knowledge grows.

A natural question, then, is why the rising R&D inputs displayed in Figure 5, which we show predict increasing patents, do not translate into steadily rising predicted sales growth rates from ideas. The first step towards understanding this apparent contradiction is to note that the elasticities we estimate predict increases in the *flow* of patents from higher R&D expenditures, but do not guarantee increases in the growth *rate* of patents. Using the coefficient estimates in Figure 12, predicted sales growth rates from ideas (*i.e.*, the red line

[37] Jones (2019) categorizes these relationships as the second of Romer's three seminal contributions to growth theory.

in Figure 13) should rise steadily if firms' increased R&D investments led to higher firm-level patent-stock growth rates.

Given the focus on idea growth *rates* in the literature, and to be consistent with our productivity growth regressions, it might seem that the dependent variable in our patent regressions should be patent-stock growth rates rather than flows. Note, however, that the prevalence of extensive-margin changes in firm-level patenting constrains us to using DHS growth rates when studying the mapping from ideas to output growth. Since all new patenting firms start with an initial growth rate of 2, their patenting growth rates must decline with subsequent patenting by construction. Our firm-level estimates thus cannot possibly generate an indefinitely increasing growth rate of ideas from more R&D. Such increases might still occur in aggregate, however, if firms' ideas spillover to other firms' idea creation. Indeed, in Appendix Figure F2 we show that aggregate patent and breakthrough stock growth rates both increase over the period, with substantial cyclicality over the last 100 years. These patterns are inconsistent with a steady decline in patent growth rates as the stock of patents grows, and thus broadly align with our firm-level conclusions that ideas do not seem to be getting harder to find.

The distinction between firm versus aggregate growth rates highlights another key insight from Romer: that knowledge is (at least partially) non-rival and non-excludable. In our view, this point is where our two-step approach truly shines. While equation (1) from Romer (1990) features knowledge spillovers in the production function of ideas, our firm-level specifications highlight the possibility of spillovers not only from aggregate ideas into a firm's new ideas, but also from aggregate ideas into a firm's output. The results in Section 6 suggest that declines in the latter may be particularly important post 2007. Predicted growth from "other factors", *i.e.,* the time fixed effects, is negative in the 2007 to 2017 decade, even after controlling for firms' declines in employment. This negative predicted growth corresponds with the rebound in importance of mega-firms for breakthrough patents, which could signal their ability to exclude competitors and potential entrants from their ideas, and/or to preclude them from growing or even entering at all (*e.g.*, as in Cunningham et al., 2021). Changing spillovers through these channels might also relate to the secular decline in entrants' contribution to non-patenters' growth displayed in Figure 10. In terms of existing theory, such declines could arise for many reasons that are unrelated to crowding out from a growing stock of ideas.

In terms of mapping these insights back to theory, the data suggest that it may be fruitful to explore variants of an aggregate production function along the lines suggested by Jones (2019), in which increasing returns in the aggregate production function arise from the non-rivalry of ideas into output. Such returns seem particularly relevant when thinking about how emergent technologies such as AI may have proportional effects on growth, for example, Anthropic's Claude algorithms are used across a wide range of firms and industries to increase worker productivity. Of course this ability to generate productivity improvements across a large number of firms and sectors, coupled with heterogeneity in technology adoption, also makes tracing out such spillovers particularly challenging. Importantly, these increasing returns may also be impeded by changing market structure or technologies.

A second reason that increased R&D expenditure need not map to steadily rising predicted growth from ideas is the stochastic nature of innovation. That is, while R&D's positive impact on growth *rates* may manifest over a long enough time period, the unpredictability of generating ideas may lead to growth rates that rise and fall over shorter time periods, in line with the aggregate patterns in Figure F2 discussed above. This possibility seems especially salient for breakthroughs, whose time fixed effects exhibit a stark inverted-U pattern over the period. Predicted output growth from breakthrough-stock growth is also the most volatile, suggesting not only that truly innovative ideas come in bursts, but also that they may be the hardest to translate into output. Such translation is also subject to random shocks, for instance as firms experiment with how to adopt new technologies or introduce new products (Brynjolfsson et al., 2021). One feature of our two-step approach is that we condition only on successful patents when assessing the relationship between ideas and growth, thus mitigating some of the uncertainty around translating R&D inputs into ideas when we assess how ideas relate to growth.

Finally, a third reason that we may not predict steadily rising growth rates from firms' increased R&D expenditure is that we only measure sales by firms' US establishments. US firms increasingly specialize in the innovation and marketing stages of the production process, while their manufacturing is performed outside the United States (Fort, 2023). In line with these trends, Fact 2 shows that US patents have shifted towards Information, Professional Services, and Managements sectors, which generally provide inputs to downstream stages of production that may be classified in other sectors or located in other countries. Indeed, Kamal et al. (2022) finds that a disproportionate amount of Management employment belongs to US multinational enterprises, and Fort (2023) shows that "factoryless goods producers" employ relatively high shares of Professional Services and Management employment. Both of these firm types (US MNEs and factoryless goods producers) are much more import-intensive. To the extent that these firms also shift their profits outside the US, as documented by Guvenen et al. (2022), we will underestimate their returns to R&D.[38]

[38]To address the potential concern that we are missing firms' foreign R&D inputs, we construct our R&D measures using firms' total R&D, which includes foreign R&D. Foreign R&D is quite small, and in robustness exercises we find no material difference from excluding them.

8 Conclusion

We build a novel 45-year firm-level panel to study the evolution of US patenting. We document rising patents per R&D input, flat or rising elasticities of patents to R&D inputs, and no robust secular decline in average patents per firm after controlling for its R&D input use. Perhaps most exciting for research on growth, we document a positive and significant relationship between firms' patent and output growth that does not weaken over time. These results indicate that the flat US productivity growth coupled with rising research expenditure are not due to declining research productivity within firms.

While the marginal R&D dollar today is at least as effective at generating patents as it was 40 years ago, firms' average growth rates after controlling for their patent growth fall in the later periods. A natural next question is what has led to a decline in average growth rates across firms that is independent of their patenting? One possibility is that spillovers across firms have declined. If, for example, it is easier for firms to exclude rivals from their ideas, then average growth may decrease, even as ideas expand. Indeed, recent work highlights precisely how important declining knowledge diffusion seems to be for the US economy after 1980 (Akcigit and Ates, 2023).

We view a micro-level analysis of changes in such spillovers as the natural next step to unpacking the relationship between ideas and growth. These spillovers may occur across firms within industries or particular geographic regions. They may also manifest within firms or supply chains, but across industries and regions, or even countries. For example, Braguinsky et al. (2025) document new product introductions in seemingly low-tech sectors such as apparel when firms such as Nike add GPS tracking to their clothing. A fruitful avenue for future work is thus to trace out the cross-sectoral and cross-country patterns in which innovation in one area feeds into increased productivity in others and how such spillovers have evolved over time.

References

Acemoglu, D., S. Johnson, and J. A. Robinson (2001): "The Colonial Origins of Comparative Development: An Empirical Investigation," American Economic Review, 91, 1369–1401.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005): "Competition and Innovation: An Inverted-U Relationship," Quarterly Journal of Economics, 120, 701–728.

Akcigit, U. and S. T. Ates (2023): "What Happened to U.S. Business Dynamism?" Journal of Political Economy, 131, 2319–2369.

Akcigit, U. and W. R. Kerr (2018): "Growth Through Heterogeneous Innovations," Journal of Political Economy, 126, 1374–1443.

Ando, Y., J. Bessen, and X. Wang (2025): "The Rising Returns to RD: Ideas are not getting harder to find," Tech. Rep. 5242171, SSRN.

Aw, B. Y., M. J. Roberts, and D. Y. Xu (2011): "R&D Investment, Exporting, and Productivity Dynamics," American Economic Review, 101, 1312–1344.

Balasubramanian, N. and J. Sivadasan (2011): "What Happens When Firms Patent? New Evidence from US Economic Census Data," Review of Economics and Statistics, 93, 126–146.

Bergeaud, A., A. Guillouzouic, E. Henry, and C. Malgouyres (forth): "From Public Labs to Private Firms: Magnitude and Channels of R&D Spillovers," Quarterly Journal of Economics.

Bernard, A. B. and T. C. Fort (2015): "Factoryless Goods Producing Firms," American Economic Review, 105, 518–23.

Bessen, J. (2008): "The Value of US Patents by Owner and Patent Characteristics," Research Policy, 37, 932–945.

Bloom, N., C. I. Jones, J. Van Reenen, and M. Webb (2020): "Are Ideas Getting Harder to Find?" American Economic Review, 110, 1104–44.

Bloom, N., A. Kalyani, T. Hassan, J. Lerner, M. Mello, and A. Tahoun (2025): "The Diffusion of New Technologies," Quarterly Journal of Economics, 140, 1299–1365.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying Technology Spillovers and Product Market Rivalry," Econometrica, 81, 1347–1393.

Blundell, R., R. Griffith, and F. Windmeijer (2002): "Individual Effects and Dynamics in Count Data Models," The Quarterly Journal of Economics, 117, 481–530.

Braguinsky, S., J. Choi, Y. Ding, K. Jo, and S. Kim (2025): "Mega Firms and New Technological Trajectories in the US," Tech. Rep. 31460, NBER.

Brynjolfsson, E. and A. McAfee (2014): The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, New York: W. W. Norton & Company.

Brynjolfsson, E., D. Rock, and C. Syverson (2021): "The Productivity J-Curve: How Intangibles Complement General Purpose Technologies," American Economic Journal: Macroeconomics, 13, 333–372.

Chen, Y. and D. Xu (2023): "A Structural Empirical Model of R&D, Firm Heterogeneity, and Industry Evolution," The Journal of Industrial Economics, LXXI, 323–353.

Chow, M. C., T. C. Fort, C. Goetz, N. Goldschlag, J. Lawrence, E. R. Perlman, M. Stinson, and T. K. White (2021): "Redesigning the Longitudinal Business Database," Tech. Rep. 28839, National Bureau of Economic Research Working Paper Nr 28839.

Cohen, W. M., R. Nelson, and J. P. Walsh (2000): "Protecting their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (Or Not)," Working Paper 7552, NBER.

Cowen, T. (2011): The great stagnation: How America ate all the low-hanging fruit of modern history, got sick, and will (eventually) feel better: A Penguin eSpecial from Dutton, Penguin.

Cunningham, C., F. Ederer, and S. Ma (2021): "Killer Acquisitions," Journal of Political Economy, 129, 649–702.

Davis, S. J., J. Haltiwanger, and S. Schuh (1996): "Small business and job creation: Dissecting the myth and reassessing the facts," Small business economics, 8, 297–315.

Dell, M., B. F. Jones, and B. A. Olken (2012): "Temperature Shocks and Economic Growth: Evidence from the Last Half Century," American Economic Journal: Macroeconomics, 4, 66–95.

Ding, X., T. C. Fort, S. J. Redding, and P. K. Schott (2022): "Structural Change Within versus Across Firms: Evidence from the United States," Working Paper 30127, National Bureau of Economic Research.

Djankov, S., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2002): "The Regulation of Entry," The Quarterly Journal of Economics, 117, 1–37.

Doidge, C., G. A. Karolyi, and R. M. Stulz (2017): "The U.S. Listing Gap," Journal of Financial Economics, 123, 464–487.

Dreisigmeyer, D., N. Goldschlag, M. Krylova, W. Ouyang, and E. Perlman (2018): "Building a Better Bridge: Improving Patent Assignee-Firm Links," Tech. rep., Center for Economic Studies CES-TN-2018-01.

Dyevre, A. and O. Seager (2024): "Matching Patents to Publicly Listed Firms in the US: 1950-2020," Tech. rep., LSE.

Ekerdt, L. K. (2024): "The Role of R&D Factors in Economic Growth," CES Working paper 24-69, Center for Economic Studies.

Ekerdt, L. K. and K.-J. Wu (2024): "Self-Selection and the Diminishing Returns to Research," Working paper, Center for Economic Studies.

Evenson, R. (1984): "International Invention: Implications for Technology Market Ananlysis," in R&D, Patents, and Productivity, ed. by Z. Griliches, University of Chicago Press, 89–126.

Evenson, R. E. and Y. Kislev (1976): "A Stochastic Model of Applied Research," Journal of Political Economy, 84, 265–282.

Fieler, A. C. and S. K. Lee (2026): "Growing Through Vintages," Working paper, Yale University.

Fort, T. C. (2023): "The Changing Firm and Country Boundaries of US Manufacturers in Global Value Chains," Journal of Economic Perspectives, 37, 31–58.

Fort, T. C. and S. D. Klimek (2018): "The Effects of Industry Classification Changes on US Employment Composition," CES Working Paper.

Gordon, R. (2017): The rise and fall of American growth: The US standard of living since the civil war, Princeton university press.

Graham, S. J., C. Grim, T. Islam, A. C. Marco, and J. Miranda (2018): "Business Dynamics of Innovating Firms: Linking US Patents with Administrative Data on Workers and Firms," Journal of Economics & Management Strategy, 27, 372–402.

Griliches, Z. (1979): "Issues in Assessing the Contribution of Research and Development to Productivity Growth," The Bell Journal of Economics, 92–116.

——— (1990): "Patent Statistics as Economic Indicators: A Survey," Journal of Economic Literature, 28, 1661–1707.

——— (1998): "Patent Statistics as Economic Indicators: a Survey," in R&D and Productivity: the Econometric Evidence, University of Chicago Press, 287–343.

Growiec, J., P. McAdam, and J. Mućk (2022): "Are Ideas Really Getting Harder To Find? R&D Capital and the Idea Production Function," .

Guvenen, F., R. J. Mataloni, D. G. Rassier, and K. J. Ruhl (2022): "Offshore Profit Shifting and Aggregate Measurement: Balance of Payments, Foreign Investment, Productivity, and the Labor Share," American Economic Review, 112, 1848–1884.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2005): "Market Value and Patent Citations," RAND Journal of Economics, 36, 16–38.

Hall, B. H., J. Mairesse, and P. Mohnen (1998): "Measuring the Returns to R&D," in Handbook in Economics, Elsevier B.V., vol. 2, chap. 24, 1034–1082.

Haltiwanger, J. C., R. S. Jarmin, and J. Miranda (2013): "Who Creates Jobs? Small vs. Large vs. Young," Review of Economics and Statistics, 95, 347–361.

Hausman, J., B. H. Hall, and Z. Griliches (1984): "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," Econometrica, 52, 909–938.

Howell, S. T. (2017): "Financing Innovation: Evidence from R&D Grants," American Economic Review, 107, 1136–1164.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," Quarterly Journal of Economics, 124, 1403–1448.

Jaffe, A. B. and G. de Rassenfosse (2017): "Patent Citation Data in Social Science Research: Overview and Best Practices," Journal of the Association for Information Science and Technology, 68, 1360–1374.

Jarmin, R. S. and J. Miranda (2002): "The Longitudinal Business Database," CES Working Paper CES-WP-02-17.

Jones, B. F. (2009): "The Burden of Knowledge and the Death of the Renaissance Man: Is Innovation Getting Harder?" Review of Economic Studies, 76, 283–317.

Jones, C. I. (1995): "R&D-Based Models of Economic Growth," Journal of Political Economy, 103, 759–784.

——— (2019): "Paul Romer: Ideas, Nonrivalry, and Endogenous Growth," Scandinavian Journal of Economics, 121, 859–883.

Kamal, F. (2023): "A Portrait of US Factoryless Goods Producers," in Measuring the Global Economy, ed. by N. Ahmad, B. Moulton, J. D. Richardson, and P. V. D. Ven, University of Chicago Press.

Kamal, F., J. McCloskey, and W. Ouyang (2022): "Multinational Firms in the U.S. Economy: Insights from Newly Integrated Microdata," Working Paper 22-39, Center for Economic Studies.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021): "Measuring Technological Innovation over the Long Run," American Economic Review: Insights, 3, 303–20.

Kerr, W. R. and S. Fu (2008): "The survey of industrial R&D—patent database link project," The Journal of Technology Transfer, 33, 173–186.

Kline, P., N. Petkova, H. Williams, and O. Zidar (2019): "Who Profits from Patents? Rent-Sharing at Innovative Firms," The Quarterly Journal of Economics, 132, 1343.1404.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017): "Technological Innovation, Resource Allocation, and Growth," The Quarterly Journal of Economics, 132, 665–712.

Koh, P.-S. and D. M. Reeb (2015): "Missing R&D," Journal of Accounting and Economics, 60, 73–94.

König, M., K. Storesletten, Z. Song, and F. Zilibotti (2022): "From Imitation to Innovation: Where Is All That Chinese R&D Going?" Econometrica, 90, 1615–1654.

Kortum, S. (1993): "Equilibrium R&D and the Decline in the Patent-R&D Ratio: U.S. Evidence," American Economic Review Papers and Proceedings, 83, 450–457.

Kortum, S. S. (1997): "Research, Patenting, and Technological Change," Econometrica, 65, 1389–1419.

Lanjouw, J. O. and M. Schankerman (2004): "Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators," Economic Journal, 114, 441–465.

Meyers, K. R. and L. Lanahan (2022): "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy," American Economic Review, 112, 2393–2423.

Peters, B., M. J. Roberts, V. A. Vuong, and H. Fryges (2017): "Estimating Dynamic R&D Choice: An Analysis of Costs and Long-Run Benefits," RAND Journal of Economics, 48, 400–437.

Philippon, T. (2022): "Additive Growth," NBER Working Paper 29950, National Bureau of Economic Research.

Ribeiro, B. (2025): "Growth with New and Old Technologies," Working paper, Princeton University.

Romer, P. M. (1990): "Endogenous Technological Change," Journal of Political Economy, 98, S71–S102.

Schlingemann, F. P. and R. M. Stulz (2022): "Have Exchange-Listed Firms Become Less Important for the Economy?" Journal of Financial Economics, 143, 927–958.

Teng, S. (2025): "Innovation Through Recombination," mimeograph, Yale University.

Online Appendix for

Growth is Getting Harder to Find, Not Ideas*

Teresa C. Fort† Nathan Goldschlag‡ Jack Liang§ Peter K. Schott¶ Nikolas Zolas‖

A Matching Patents to Census Business Data

This appendix describes how we match patent assignees from the United States Patent and Trademark Office (USPTO) data to firms in the Census Business microdata files from 1977 to 2021. Throughout this appendix, "assignee" refers to a firm in the USPTO data, "firmid" corresponds to firms in the Census data, and "patent-assignee record" refers to a patent number-assignee sequence pair. A patent document can have multiple assignees, which are distinguished by assignee sequence.

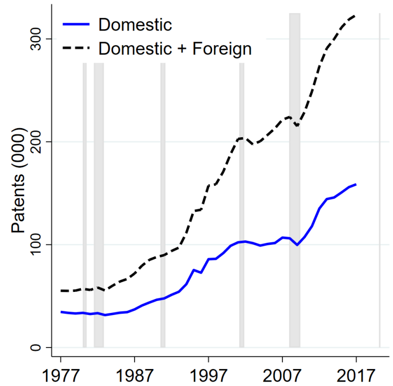
A.1 Data Preparation and Cleaning

We first compile and clean the USPTO patent data and the combined County Business Patterns and Business Register (CBPBR) microdata files. For patents, we combine extracts from Google patent XML files with information from PatentsView. PatentsView is a patent data visualization and analysis platform that transforms patent documents stored in XML into a relational database with consistent identifiers of the same assignees. The PatentsView data construction implements clustering algorithms to group patent-assignee records with similar name and geography information, which improve our ability to match assignees to unique firms in the Census data. We augment the patent-assignee record geography with location information on inventors, attaching each unique geography among a patent's inventors to the patent-assignee record. This addition allows us to search within the inventors' geography for a business name match for the patent-assignee record, in addition to the assignee's geography listed on the patent. Figure A1 reports the relative prevalence of foreign and domestic granted patents by application year.

*Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7083300: CBDRB-FY25-CES002-001)

†Tuck School at Dartmouth, US Census Bureau, NBER, and CEPR.
‡US Census Bureau.
§Temple University.
¶Yale School of Management, CEPR, and NBER.
‖Department of State

Figure A1: Domestic versus Foreign Patents by Application Year



Source: PV and authors' calculations. *Notes:* Figure reports total (black dashed line) and domestic (blue solid line) patents in the PV dataset over our sample period. Domestic patents are defined as having at least one assignee with a US location.

We obtain establishment business name and geography information (physical and mailing) from the CBPBR files, using re-timed files when available.[1] We use the name1 and name2 fields from the CBPBR, along with the concatenation of the two. We match by establishment identifiers to the Longitudinal Business Database (LBD) to attach firm identifiers (lbdfid and firmid) to the CBPBR records. Since our goal is a firm-level match, we keep only unique firm identifier, name, and geography combinations. It is important to note that firmids in the Census data are not inherently longitudinally consistent. Indeed, they break by construction whenever a firm transitions from single to multi-unit status (or vice-versa).

We clean the CBPBR and assignee name and address fields using identical algorithms. We apply SAS Data Quality (DQ) name and geography standardization macros, along with a customized suite of "find and replace" string commands that remove or shorten words or phrases within the name field.[2] After standardizing and cleaning the name and geography information, we generate DQMATCH codes for fuzzy matching.[3] We also create "partial name" variables, one containing the first eight characters, another containing the first two words of the business name, one that removes all spaces from the string, and a "stop word" name field, which removes common words and phrases from the name field.[4]

[1]Re-timed files exist in intercensal years as they incorporate establishment births and deaths that are imputed from EC data.

[2]For example, we change several variations of "COOPERATIVES", "COOPERATIVE", and "CO OPERATIVE" to "COOP".

[3]DQMATCH codes use proprietary algorithms to generate hashcodes that group together similar strings. Two records will be given the same hashcode if their string characteristics are sufficiently similar. The SAS DQMATCH function allows for varying levels of "fuzz" in how distant strings are allowed to be before being assigned different hashcodes.

[4]For example, in the "stop word" name we might remove the word "INDUSTRIES" from the name field.

A.2 Name and Geography Matching

We organize the name and geography matching into blocks and passes. A match pass identifies records in the CBPBR and USPTO data that agree on a given set of criteria. A block consists of multiple match passes with broadly similar match criteria. We remove patent-assignee records that receive at least one match in a given block before proceeding to the next block. Organizing the match passes this way allows us to: (1) distinguish between different tiers of match quality across match blocks, and (2) account for ambiguity in which match pass is of higher quality within match blocks. We match in the following four blocks, using the various passes noted in each one:

Block 1. **Full patent-assignee record name and 2 geographic elements**: We use up to 5 cleaned variants of the patent-assignee record name and 3 variants of its address (city, state, and county) for a total of 16 passes.

Block 2. **Full name and 1 geographic element**: We use the same name variables as above along with 1 of the 3 geographic variables listed above for a total of 16 passes.

Block 3. **Full name but no geographic elements**: We use 4 cleaned variants of the firm name for a total of 4 passes.

Block 4. **Incomplete name and no geographic elements**: We use 2 incomplete name variables for a total of 2 passes.

We retain information on each match's block and pass, and also compute the string lengths and a series of string comparators between the two name fields, including: Levenstein distance (`lev`), Jaro-Winkler (`jw`), SAS compare (`cmp`), and generalized edit distance (`ged`). We use these string comparator measures, along with the match pass, in subsequent post-match cleaning algorithms to identify the best match for each patent-assignee record.

A.3 Raw Match Cleaning and Longitudinal Imputation

The first two rows in Table A1 show that the raw matches assign a Census firmid to 68 percent of the patent-assignee records, though only 48 percent match to a unique Census firmid. These aggregates include foreign assignees with no US address information. Among patent-assignee records with a city-state combination in the United States, which we label US-assignees, 92 percent match to Census firmids, with 74 percent receiving a unique match.

Table A1: Patent-Assignee Observation Match Rates

% of Patent Assignees	US	Foreign	All
Raw Match	92.43	42.94	67.96
Unique Raw Match	74.39	20.93	47.96
Match in Crosswalk	92.23	57.75	75.18
Total Matched Patents	3,336,000	3,263,000	6,599,000

Source: BDSPF-Long, PatentsView. *Notes:* Table shows the percent of US (has US location) and Foreign (no US location) patent-assignee records from 1977 to 2021 with at least one raw match ("Raw Match"), with a unique raw match ("Unique Raw Match"), and with a unique match in the final crosswalk ("Match in Crosswalk").

The goal of the match cleaning steps is to create an unbalanced assignee-grant year panel with a single matched firmid for each observation. We focus on grant year since assignee information is formalized at the time of granting. To this end, we: (1) remove matches outside of +/- 2 years from grant year or to firms not payroll active in the grant year or the year prior, (2) select matches based on upon string comparators, (3) select most frequently matched firmid to an assignee within a given patent record, (4) select most frequently matched firmid within a pooled set of years (+/- 2 years), and (5) select most frequently matched firmid within a given assignee-grant year. Matches for assignee-grant year combinations with multiple matched firmids after these cleaning steps are removed.

A novel contribution of our matching algorithms is to use the sequence of matches for a given assignee over time to increase final match rates. For instance, if an assignee is missing a match in its first year in the panel, but has a match in its second year, we push the firmid matched in the second year "backwards" to cover earlier patents. We use a similar approach to "fill holes" in an assignee's match time series: if an assignee matches to a given firmid in $t-1$ and $t+1$, we impute that firmid as a match in t. In cases where the lead and lagged matched firmids are different, we select the closest (in time) or the lag if they are equally close. Finally, after filling gaps using leading and lagging matches, we assign missing matches to the modal firmid across years for a given assignee.

After the longitudinal processing we remove matches to firmids that do not have positive payroll in $t-1$ and/or t ("payroll denom positive") in the LBD within +/- 2 years of the grant year. Since a matched firmid may not appear in the LBD in the grant year (e.g. if it is not payroll positive in that year or the prior year), we include lbd_yr in the crosswalk, which denotes the year of LBD to which that the patent-assignee record can be matched. All assignee-grant year to firmid matches in the final crosswalk are unique.

The bottom row of Table A1 shows that this cleaning results in 92% of US patent-assignee records and 58% of foreign patent-assignee records matching to a unique firmid. Since only those foreign assignees with an employer footprint in the US can match to the CBPBR, it is unclear what the maximum share of matched foreign assignees could be. These percentages translate to 3.336 million matched US patent-assignee records and 6.6 million total matches.

Table A2 shows the counts and percentages of these raw and final matches by block. The last two columns show the share of raw matches that are unique, as well as the "conversion rate" of raw matches into final matches in a given block. Conversion rates are less than 100 percent because we cannot assign all raw patent-assignee observation matches to a unique firmid.

Roughly 31% of raw matches come from Block 1, our highest quality set of match passes, but most (55%) are made in Block 3, where we match on name only. Although Block 1 comprises only 31% of raw, over 41% of our final patent-assignee record matches originate from Block 1. This high share arises because the raw matches are higher quality, with fewer false positives, and thus more likely to be unique. Indeed, the unique rate for Block 1 matches is almost 84%, which in turn makes conversion to a unique match more likely. We convert over 98% of patent-assignee records with Block 1 matches, versus 88% for Block 4 matches.

The final row of Table A2 (Block "None") highlights a key contribution of our new crosswalk. About 32% of patent-assignee records in the final crosswalk did not receive a raw match. These final matches occur because we exploit the full time-series information for an assignee across all its patents.

Table A2: Raw Match Block Distribution

	Raw Match		Final Match		Unique	Conversion
Block	Count	Percent	Count	Percent	Rate	Rate
1	4,786,000	31.04	2,728,000	41.33	83.87	98.24
2	364,000	2.36	182,000	2.76	75.27	95.60
3	8,551,000	55.45	1,386,000	21.00	48.41	87.09
4	1,720,000	11.15	190,000	2.88	36.05	87.89
None	0	0.00	2,114,000	32.03		34.67
Total	15,421,000	100.00	6,600,000	100.00	47.95	75.17

Source: BDSPF-Long, PatentsView. *Notes:* Table shows the count and distribution of raw matches ("Raw Match") by match block. A single patent-assignee record may have multiple raw matches. The "Final Match" columns present the count and percent of patents by the type of raw match used in the crosswalk. "Unique Rate" shows the percent of patents with a given raw match type that had a unique raw match. "Conversion Rate" shows the percent of patents with a raw match that are also in the final crosswalk (e.g., the raw match was "converted" into a final match). Block 1 matches use name, city, and state. Block 2 matches use name and city. Block 3 matches use name only. Block 4 matches use an incomplete name match. The positive conversion rate and final matches for patent-assignee records with no raw match are due to assignee matches and longitudinal imputations.

Panel A of Table A3 shows the count of firmids per assignee in the raw and final matches. While the raw matches include instances in which the same assignee maps to multiple firmids in a year, the final match has a unique assignee-to-firmid mapping within each year by construction. An assignee may change firmids over time because the Census firmid breaks. Such breaks occur for various reasons, including mergers and acquisitions. While such variation occurs in the final match, the cleaned final match has substantially lower firmid counts per assignee than the raw match. Over 60% of assignees do not match to any firmid, though the last column of Table A1 indicates that these firms account for just under 25% of total patent-assignee records.

Panel B of Table A3 reports the percent of assignees and assignee-years for which there is a change, from year-to-year, in the matched firmid. For the raw match, we define a "break" if the *set* of firmids is different between $t-1$ and t. In the raw match, 8.2% of assignees had at least one break, or change in firmid, across their time series. By contrast, only 5% of assignees in the final match change firmid over the period. These breaks correspond to 7.9% of year-to-year firmid breaks the raw match, compared to just 3% in the final crosswalk. The cleaning algorithms thus reduce the probability that a firmid changes from year-to-year for a given assignee by more than 50%.

Table A3: Match Ambiguity and Time Series Breaks

Panel A: Ambiguity				
	Raw Match		Final Match	
Firmids per Assignee	Count	Percent	Count	Percent
0	353,000	61.81	374,000	65.41
1	145,000	25.39	169,000	29.55
2	30,500	5.34	21,500	3.76
3	11,500	2.01	5,000	0.87
4	6,100	1.07	1,400	0.24
5-9	11,500	2.01	900	0.16
10+	13,500	2.36	20	0.00
Panel B: Breaks				
Assignees with Breaks	47,000	8.23	28,500	4.99
Assignee-Year Breaks	117,000	7.92	45,000	3.04

Source: BDSPF-Long, PatentsView. *Notes:* Panel A shows the count and distribution of patent assignees (`pv_asgid`) that have a given number of matched `firmids` across all of their patents in the raw match ("Raw Match") and in the final crosswalk ("Final Match"). Panel B shows the count of patent assignees and patent-assignee year combinations that experience a change in their matched firmid between years in the raw match ("Raw Match") and in the final crosswalk ("Final Match").

Table A4 reports the count and percent of matched patent-assignee records by their longitudinal imputation status. About 7% of matches are identified using longitudinal imputes, which account for about 5.8% of all records. The majority of longitudinal-impute cases are "Lagging Gaps", where we push matches forward in time to unmatched assignee-grant years. The next most frequent types of longitudinal imputations are Leading Gaps (pushing matches backwards in time) and complex interior matches (filling gaps with different matches on either end).

Table A4: Longitudinal Imputations

Longitudinal Imputation	Count	Percent
Unmatched	1,396,000	21.16
None	4,827,000	73.16
Leading Gap	110,000	1.67
Lagging Gap	133,000	2.02
Simple Interior	49,500	0.75
Complex Interior	82,500	1.25
Total	6,598,000	100.00

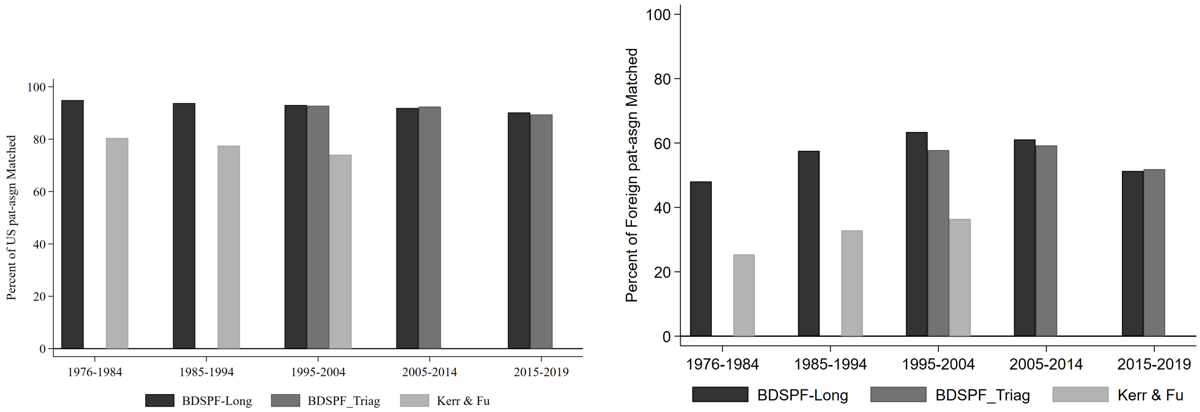
Source: BDSPF-Long, PatentsView. *Notes:* Table shows the count of patent-assignee records (`patnum`, `assg_seq`) by the type of longitudinal imputation they received, if any.

A.4 Match Quality Analysis

We assess our match rates and quality by comparing them to those in the Kerr and Fu (2008) (Kerr & Fu) and BDSPF-Triangulation (Graham et al., 2018; Dreisigmeyer et al., 2018) crosswalks. Figure A2 shows that the match rates in the BDSPF-Long (our match) are higher than those in than Kerr & Fu, particularly for foreign patents, and similar to those in the BDS-Triangulation bridge. Most notably, the BDSPF-Long covers a longer time period than its predecessors. While it is difficult to interpret patent-assignee record match rates because the true subset of employer businesses among assignees is unknown, the match rate for US patent-assignee records falls from 95% to 90% between the early 1980s and the late 2010s. We investigate these declines below.

To assess the quality of our matches we compare them to the very high quality BDSPF-Triangulation model "A1" matches. These matches are ones for which the BDSPF-Triangulation algorithms "fully triangulate" patent-inventor and patent-assignee records, both being corroborated by employment records in the Longitudinal Employer-Household Dynamics (LEHD) data. There are approximately 1.4 million "A1" matches. Among the BDSPF-Triangulation "A1" matches, we compute the number of "correct" BDSPF-Long matches (those for which

Figure A2: Match Rate Time Series



(a) US Patent Assignees

(b) Foreign Patent Assignee

Source: BDSPF-Long, PatentsView. *Notes:* Figure shows the percent of patent-assignee records, granted within a given time window, that received a match in each crosswalk. BDSPF-Long is the crosswalk introduced in this paper. BDSPF-Triag is the LEHD-based crosswalk described in Graham et al. (2018) and Dreisigmeyer et al. (2018). Kerr & Fu is the crosswalk developed by Kerr and Fu (2008).

both crosswalks agree), the number of incorrect BDSPF-Long matches (those for which the crosswalks disagree) and the number of records for which BDSPF-Long is unable to identify a match.

We measure precision of the BDSPF-Long matches as the number of correct matches divided by the sum of correct and incorrect matches. This metric captures the share of matches we make that are correct. We compute recall as the count of correct and incorrect matches divided by the count of correct, incorrect, and missing matches. This metric captures the share of records for which we make a match, correct or not. Table A5 shows that we match 93% of patent-assignee records in this set correctly. Precision of matches for US patent-assignee records is significantly higher than for foreign patent-assignee records (94% vs. 81%). Recall shows an even bigger gap between US and foreign records (98% vs 80%). Among the US patent-assignee records in the A1 set, our matches tend to be high quality.

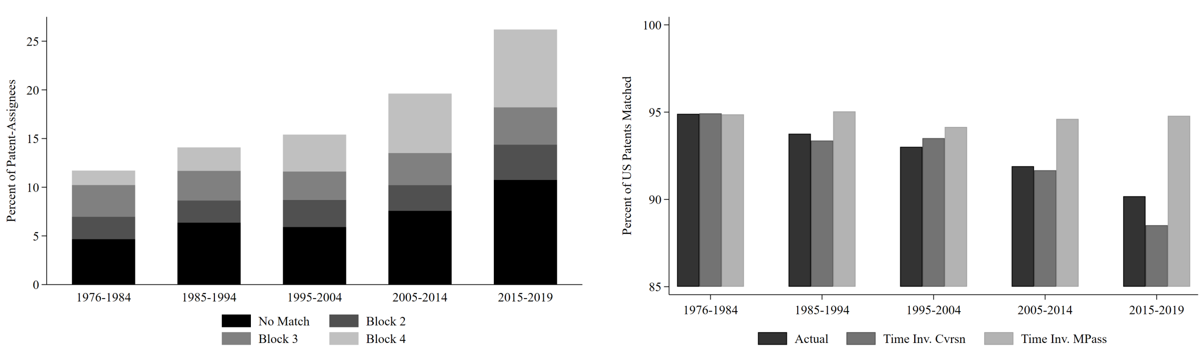
Table A5: BDSPF-Long Match Precision & Recall

	US	Foreign	All
Precision	93.60	80.98	93.08
Recall	97.94	80.44	97.23
Total Patent-Assignee Records	1,350,000	57,500	1,407,500

Source: BDSPF-Long, PatentsView. *Notes:* Table shows the precision and recall of BDSPF-Long matches using the "A1" model (e.g., highest quality) matches from the BDSPF-Triangulation crosswalk as a truth set. Precision is the count of records with agreement between BDSPF-Long and BDSPF-Triag divided by the count of all matched records. Recall is the count of records with agreement between BDSPF-Long and BDSPF-Triag divided by the sum of all matched and unmatched records.

We investigate the declining match rates by assessing relative changes in raw match rates. Panel A of Figure A3 depicts the share of patent-assignee record raw matches by block, with the highest quality Block 1 matches as the excluded share. Block 1 matches starkly decline (88% to 74%), which is evident in the rising shares accounted for the other blocks and the "No Matches" This decline is driven by rising shares of "No Matches" (5% to 11%) and Block 4 matches (which are the noisiest and hardest to convert).

Figure A3: Raw Match and Counterfactual Match Rates



(a) Raw Match Status

(b) Counterfactual Match Rates

Source: BDSPF-Long, PatentsView. *Notes:* Panel A shows the distribution of patent-assignee records by the type of raw match, if any. The excluded category is records with a Block 1 match. If the Block 1 percent were included all groups would sum to 100 within a given time period. Panel B shows actual and simulated match rates for the BDSPF-Long crosswalk over time. Time-invariant conversion ('Time Inv. Cvcrsn") shows simulated match rates using actual match compositions in each period, but holding the rate at which matches from each block are converted into final matches constant at 1976–1984 values. Time-invariant match pass ("Time Inv. Mpass") shows what match rates would be using actual conversion rates, but holding the raw match composition each period equal to the 1976–1984 period values.

Panel B of Figure A3 illustrates how the changing raw match composition maps to declining match rates. We present two counterfactual match rates in which either "conversion rates" are held constant at initial levels, or the match-pass composition is held constant. The first set of bars shows the actual US-based patent-assignee record match rate, as depicted in the left panel of Figure A2. The second bar in Panel B of Figure A3 shows what match rates would have been if the conversion rates by match pass were held at their initial levels and match-pass shares evolved as they do in the underlying data. The third bar in Panel B of Figure A3 shows what match rates would have been if the match-pass composition was held at its initial level and conversion rates varied by year as they do in the underlying data.

A3 shows that the decline in match rates would be exacerbated if conversion rates were held constant at their initial levels. In contrast, if the match-pass composition had remained constant, match rates for US-based patent-assignee records would have stayed unchanged at about 94%. This exercise suggests that the composition of raw matches explains the decline in match rates. Over time, more raw matches are made in Block 4 or not at all, putting downward pressure on match rates. Although we have no information on the unmatched firms, the fact that they can only be matched via noisy name matching, or cannot be matched at all, suggests they are smaller with fewer establishments in the CBPBR. They may also be non-employers, which will not appear in the CBPBR since it is built from administrative tax records of employer establishments.

B Merging Patent Data into Compustat

To match the USTPO patents to Compustat, we start with an extract of CRSP/Compustat Merged (Fundamentals Annual) downloaded from Wharton Research Data Services (WRDS) on July 9, 2024.[5] Firms in Compustat are identified by two variables, *gvkey* and *permno*. The former is a unique number assigned to each firm, while *permno* uniquely identifiers the share class of a publicly traded security. Though infrequent, a *gvkey* can be associated with more than one *permno* if the firm has more than one share class.

As with the LBD, we match patents to firms in Compustat by application year, in this case by combining the patent-to-*permno* mapping in KPSS with the patent-to-*gvkey* mapping of Dyevre and Seager (2024) (hereafter DS).[6] As in our LBD match, though our matching

[5] We choose consolidation level C, Industry Format INDL and FS, Data Format STD, Population Source D, Currency USD, Company Status Active and Inactive for years 1950 to 2024. Though both Compustat and our patent datasets extend beyond 2016, our sample period ends in that year due to the lag between the application and patent years noted earlier.

[6] KPSS consider patents from 1926 to 2022 while DS focus on 1950 to 2020. The KPSS mapping is available on Github here. The DS concordance is available on Github here. We consider these mappings versus others in the literature because they are the most recent and because they extend to our sample

algorithm matches by grant year, we associate patents to firms in their application year, since this timing most closely approximates the state of the firm in the year the innovation occurred. Our use of two separate mappings provides a cross check on each mapping's assignment of patents to firms, as well as their patent coverage. To make the DS mapping comparable to that of KPSS, we use the *gvkey* to *permno* concordance implicit in our downloaded Compustat file to assign one or more *permno*s to each *gvkey* in the DS mapping. (This concordance is simply all unique *gvkey-permno* pairs in the Compustat file.) We then merge the KPSS mapping into the amended DS mapping by patent and *permno*. This merge identifies patent-firm pairs that are in both mappings, as well as patent-firm pairs that appear in only one of the two mappings.[7] Our matching of patents to *permno* represents the union of these mappings.

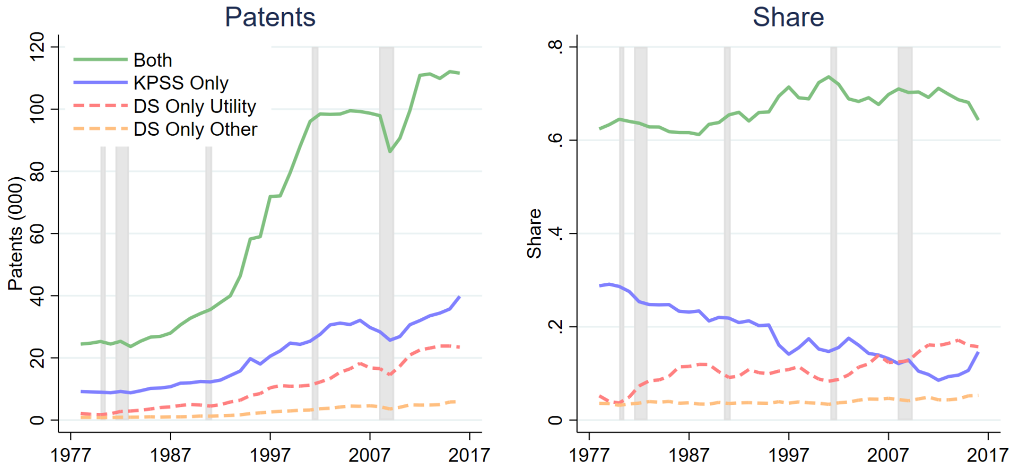
The stacked-line scatter plot in the left panel of Figure B1 reports the cumulative number of patents matched to Compustat by either KPSS or DS in each application year, with the top line showing the total number of patents in the union of the two mappings, referred to as the "total" number of patents for the remainder of this section. As indicated in the figure, about a third of total patents appear only in one of the two datasets. One reason for this discrepancy is that DS consider all patents while KPSS (as well as KPST) focus on utility patents.[8] As indicated in the right panel of the figure, non-utility patents included only in DS (represented by the lowest dashed line) account for about 5 percent of total patents across application years. The gap between the first and second dashed lines captures utility patents contained in DS but not KPSS. These account for 5 to 15 percent of all patents across years. The distance between the highest dashed and lowest solid line picks out utility patents that are in KPSS but not DS. These account for a declining share over time, from about 30 to about 10 percent. Finally, the gap between the two solid lines captures the share of patents in both mappings, which represent about two-thirds of patents across the years.

period of interest.

[7]Reassuringly, across the 3,664,570 patents in the combined mappings, we find only 77,672 (about 2 percent) for which KPSS and DS differ in their assignments of a *permno*. In these cases, we default to the KPSS assignment.

[8]Utility patents account for the vast majority of patents issued by the USPTO and are for products, processes or machines that are new or improved. Other types of patents include "design" patents, which are drawings of a design with only minimal associated text, and "plant" patents for discovered or created plants.

Figure B1: Patents in the KPSS vs DS Mappings to Compustat



Source: PV, KPSS, DS, and authors' calculations. *Notes:* Figure is a stacked-line scatter plot that reports number (left panel) or share (right panel) of patents matched to Compustat by KPSS and DS over our sample period, by application year. The two dashed lines represent patents utility and other patents appearing only in DS. The two solid lines capture utility patents appearing only in KPSS and in both KPSS and DS.

B.1 Match-Rate Comparisons

Table B1 presents the match rates of patents to the LBD and Compustat panels.

Table B1: LBD vs Compustat Match Rates

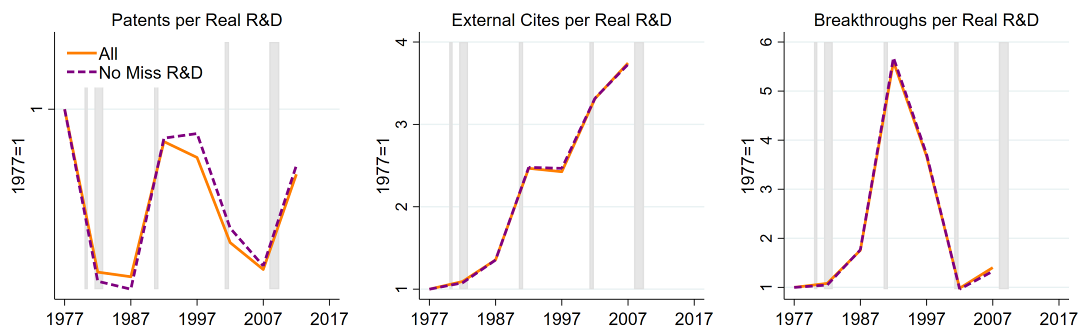
	LBD Panel		Compustat Panel	
semidecade	Patents	Breakthroughs	Patents	Breakthroughs
1977	89	91	63	73
1982	88	87	60	63
1987	88	87	54	60
1992	87	88	56	67
1997	87	86	58	74
2002	86	86	60	75
2007	86	87	57	72
2012	85		55	
Average	87	87	57	70

Source: PV, LBD, KPST, BDSPF-Long, DS, RADS, Compustat and authors' calculations. *Notes:* Figure reports the match rates for patents and breakthroughs for the LBD and Compustat panels. These match rates correspond to the totals reported in Figure 2.

C Missing R&D

Koh and Reeb (2015) report substantial missing R&D cells in Compustat. In this section, we examine these missing cells in our sample. Among the sample to firms with patents from 1977 to 2022, the share of firm-year observations with missing R&D is 19 percent (11216/58346). Figure C1 shows that the patterns displayed in Figure 3 of the main text is robust to excluding firms with missing R&D.

Figure C1: Compustat "Simple" research productivity Robustness to Firms with Missing R&D



Source: PV, LBD, KPST, RADS, BDSPF-Long, DS, Compustat and authors' calculations. *Notes:* Figure reports patents, breakthroughs, and external citations per real R&D expenditure for the firms tracked by the Compustat panel. Ratios are averages across 5-year semidecades from 1977 to 2012, and these averages are indexed to 1 in the first, 1977 to 1981 semidecade. Patents are assigned to firm by application year. Figure compares the trends for our baseline results in the main text to those that exclude firms with any missing R&D expenditures in Compustat. In both cases, results are restricted to firms in our patent elasticity regressions. The orange lines in this figure are identical to the orange lines in Figure 3.

D Census R&D Surveys

The Census Bureau conducts R&D surveys, collectively referred to as RADS, with available data from 1972 to 2021. The surveys are conducted in collaboration with the National Center for Science and Engineering Statistics (NCSES) and used to publish aggregate US R&D statistics.[9]

The RADS comprise 4 different surveys that have changed over time. The Survey of Industrial Research & Development (SIRD) was conducted from 1953 to 2007, followed by the Business Research & Development and Innovation Survey (BRDIS) from 2008 to 2016,

[9]These totals are publicly available at https://ncses.nsf.gov/data-collections/national-patterns/2021#data.

the Business Research & Development Survey (BRDS) from 2017 to 2018, and the Business Enterprise Research & Development (BERD) survey from 2018 to the present. These surveys provide detailed information on firms' basic and applied R&D expenditures, as well as their funding source, *i.e.*, federal versus private firm.

The SIRD and BRDIS covered firms with five or more employees, while the BERD covered businesses with 10 or more employees. The BERD excludes companies that performed or funded less than $50k of R&D . While such firms were included in the past surveys, Census estimates suggest that such firms accounted for small portions of aggregate R&D. Starting in 2017, the Annual Business Survey (ABS) collected R&D activity for firms with one to nine employees. The ABS also collects information on business innovation activities.

The R&D surveys are collected by Employer Identification Number (EIN). We use the BR to collapse EINs to the firm level and develop code to the match the data to our LBD panel. In practice, the enterprise identifiers in the RADs are not always EINs. We cycle over various Census identifiers and also match to a window of years to improve matching. This procedure is especially important in certain years around transitions in the identifiers used in the underlying Census data.

The survey re-designs were driven by changes in how R&D is performed and funded, and included changes to both the information collected and sampling frames. From 1957 to 1992, the SIRD focuses exclusively on manufacturing firms known to have conducted R&D in the previous 5 years. In 1992, sampling expanded substantially to include firms with unknown R&D expenditures. These "uncertainty" cases comprise the majority of the current sampling in terms of firms. The sampling frame is partitioned into three groups: companies known to perform R&D, companies that report two consecutive years of R&D expenditure; and companies whose R&D activity is unknown. To be included in the sample, a firm must meet a minimum size threshold (5 or more employees for the SIRD and 10 or more for the subsequent surveys). Firms that are in the top 50 of their state by payroll, or have an R&D lab (NAICS 5417) are included in the sample with certainty. Firms that report two consecutive years of zero R&D expenditures are dropped from future surveys. Surveyed firms are sent a "long" form if they are known R&D performers, and a "short" form otherwise. The "short" firm is designed to establish whether the firm engages in R&D, and thus does not contain many of the detailed questions appearing on the long form, which often requires different parts of a firm to weigh in on responses. Beginning with the 2017 cycle, all companies received the same questionnaire, BRD-1, which combines the standard form BRDI-1 and the abbreviated form BRDI-1(S) used in prior cycles.

E Research-Input Stocks

We transform these research input flows into stocks using the perpetual inventory method, assuming a discount rate of 15 percent, consistent with prior literature (Griliches, 1998). We compute the payroll-based research inputs only for the years in which the firm has the relevant establishment(s), *i.e.*, we do not impute initial stock for firms with the relevant establishments in the first year of the LBD (1976), nor do we include any "residual" stock that depreciates if the firm closes the relevant establishments. We make this assumption based on the premise that the establishment 'houses' a particular type of knowledge or expertise via its workers, which persists only while that establishment is alive and has such an industry code. When constructing research-input stocks based on the RADS, we assume flows are zero during any gaps in firms' participation in the survey and set the resulting stocks to missing in those years. We only use stocks constructed from unimputed flows, namely, we drop stocks after gaps in firms' participation in the RADS.[10] We follow a similar approach to compute SG&A and R&D real expenditure stocks for Compustat firms, though we treat missing flows as true zeros for this subset of publicly traded firms.

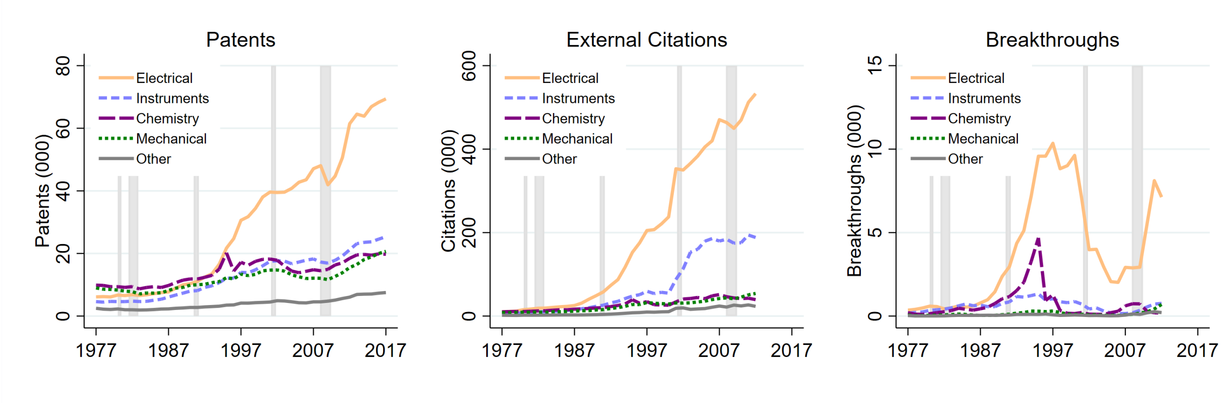
F Patent Activity Over time

Figure F1 presents a breakdown of US patenting activity by World Intellectual Property Organization (WIPO) categories, as recorded in PV.[11] The increase in patenting shown in Figure A1 is largely driven by innovation in Electrical Engineering, which includes patents related to computing, telecommunications, information technology, and semiconductors. This growth occurs in two bursts starting in the late 1980s and after the Great Recession. Electrical Engineering patents also account for the greatest share of external citations and breakthroughs. There is less agreement among across these measures for the second largest contributor to patenting activity. Indeed, while Instrument patents are responsible for a surging share of external citations in the latter part of the sample period, there is no commensurate jump in Instrument breakthroughs. By contrast, patents in Chemistry, which includes pharmaceuticals and biotech, account for a relatively large share of breakthroughs *vis-à-vis* citations.

[10] In undisclosed results, we have computed two variants of research-input stocks. The first, applicable only to RADS R&D expenditure, imputes gaps in firms' survey responses with the average flow of the two endpoints before computing stocks. The second computes an initial stock for firms present in the first year of the LBD equal to $\bar{x}_f/(d+g)$ where $\bar{x}_f$ is the firm's average observed flow, and d and g are the depreciation and long-run flow growth rates, which we set to 15 and 2 percent, respectively.

[11] While Figure F1 plots just the first WIPO category listed for each patent in PV, we find similar patterns when patents are assigned to all categories listed for them.

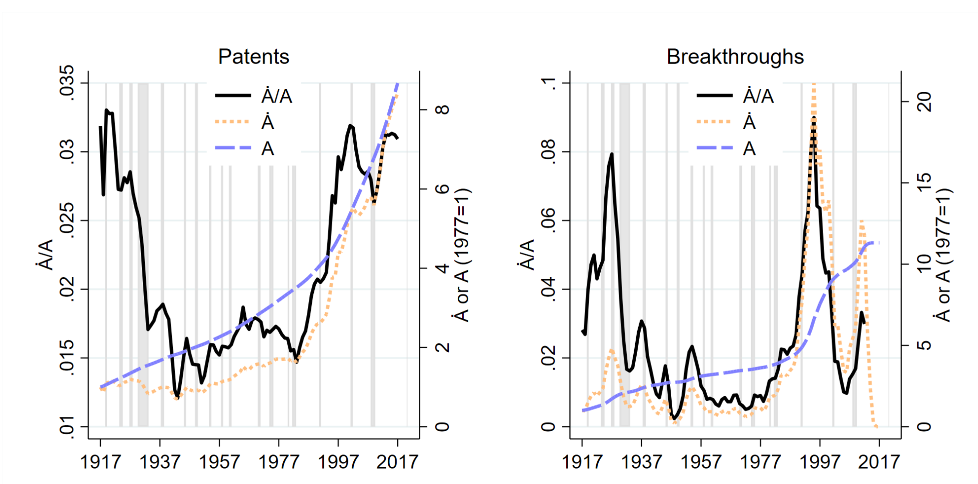
Figure F1: Patenting Activity by WIPO Code



Source: PV, KPST and authors' calculations. *Notes:* Figure provides a breakdown of US domestic patents, external citations, and breakthroughs by WIPO category and application year.

Figure F2 depicts aggregate stocks (A), flows ($\dot{A}$), and growth rates ($\dot{A}/A$) for granted patents and breakthroughs. We initialize the stocks in 1850. Note that growth rates always decline in the initial gears after initialization of the stock construction. This figure shows the cyclical nature of patent growth rates over a longer time period than the 40 years studied in the the paper.

Figure F2: $\dot{A}/A$, $\dot{A}$ and A for Patents and Breakthroughs, 1917 to 2017



Sources: PV and authors' calculations. *Notes:* Figure plots the flow ($\dot{A}$), stock (A), and growth rate ($\dot{A}/A$) of granted patents and breakthroughs by application year. Stocks are computed as cumulative breakthroughs or patents starting in 1850. Vertical gray bars indicate recessions. External citations are excluded due to data limitations.

G Additional Patent-by-Firm Statistics

Table G1 presents the number of patents per firm and firms in our samples in each semidecade.

Table G1: Patents Per Firm and Firm Counts, by Decade

	Patents Per Firm						Firm Counts					
	LBD Panel				Compustat Panel		LBD Panel				Compustat Panel	
semidecade	All	5455	5417	RADS	R&D	SG&A	All	5455	5417	RADS	R&D	SG&A
1977	.22	1	4.3	4.3	.4	1.2	138	23	4	5	14	4
1982	.17	.7	3.2	4	.3	.8	172	31	5	5	18	6
1987	.18	.7	2.7	4.7	.2	.9	207	42	8	5	21	9
1992	.23	.8	2.7	5.4	.3	1.7	235	51	10	7	26	10
1997	.31	1	2.7	8.4	.4	2.4	261	63	14	6	25	10
2002	.32	1	2.5	6.5	.7	1.7	269	71	17	10	20	8
2007	.34	.9	2.3	4.9	1	1.5	269	76	20	14	18	7
2012	.48	1.3	2.9	5.4	1.6	2.1	262	76	20	16	17	6

Source: PV, LBD, KPSS, KPST, BDSPF-Long, DS, RADS, Compustat and authors' calculations. *Notes:* Left panel reports average annual patents per firm by semidecade; right panel reports firm-year counts (in thousands) by semidecade. Both panels cover four groups of firms: those that patent at least once in our sample, and the subsets of those firms that have a Professional Services and Management (NAICS 54-55) establishment, an R&D lab, or report R&D expenditure in the Census RADS survey.

Table G2 presents share of total patents by sector, size, and age bins. Table G3 presents firm counts and employment by sector, size, and age bins using the publicly available BDS data.

Table G2: Breakdown of Patents, Citations and Breakthroughs by Decade (LBD Panel)

	Patents					External Citations					Breakthroughs				
Sector	1970	1980	1990	2000	2010	1970	1980	1990	2000	2010	1970	1980	1990	2000	2010
30 Manufacturing	78	72	61	42	30	77	69	54	33	23	73	58	42	16	12
42 Wholesale	3	5	10	17	15	4	5	12	18	18	6	7	16	22	15
51 Information	3	4	3	8	12	4	5	5	13	15	7	8	9	26	37
54 Professional	2	4	8	14	19	3	5	10	16	18	3	9	11	16	19
55 Management	5	8	9	9	11	4	7	10	9	12	3	5	12	13	10
Other	8	8	9	10	14	8	9	9	12	13	8	13	11	7	8
Size															
1–19	7	7	9	9	10	8	8	10	10	11	6	7	6	6	5
20–99	7	8	8	8	8	7	9	9	9	9	5	9	6	5	5
100–499	7	7	8	8	8	6	8	8	8	8	4	7	7	5	4
500–2499	8	8	10	11	10	8	8	10	10	10	6	7	8	8	8
2500–9999	13	11	15	15	16	13	12	16	15	14	10	10	16	12	11
10,000+	58	58	49	49	48	58	55	47	48	49	70	60	57	65	67
Age															
0–5 years	2	6	8	8	7	2	9	11	10	10	2	9	8	8	5
6–10 years	0	2	7	6	6	0	3	8	8	7	0	4	8	7	5
11–15 years	0	0	6	6	6	0	0	8	7	7	0	0	9	8	14
16–20 years	0	0	4	6	7	0	0	5	7	8	0	0	7	7	10
21–25 years	0	0	1	6	6	0	0	1	6	6	0	0	1	5	6
26+ years	0	0	0	4	13	0	0	0	5	13	0	0	0	3	12
Left Censored	98	91	74	64	55	98	87	67	57	49	98	87	68	62	48

Source: PV, LBD, KPSS, KPST, BDSPF-Long, DS, RADS, Compustat and authors' calculations. Figure provides a breakdown, in percent, of patent grants, external citations and breakthroughs by LBD firms' major 2-digit NAICS sector, size, and age. Firms in the LBD cannot be observed prior to 1977. As a result, the majority of firm observations have left-censored age in the early years of the panel. Cells representing less than 0.5 percent are coded as 0.

Table G3: Firm Counts and Employment by Sector, Size and Age (LBD Panel)

	Firms					Employment				
Sector	1970	1980	1990	2000	2010	1970	1980	1990	2000	2010
30 Manufacturing	7	7	6	5	5	26	21	17	12	10
42 Wholesale	7	7	6	6	5	5	5	5	5	5
51 Information	1	1	1	1	1	3	3	3	3	3
54 Professional	7	9	11	13	13	7	8	9	9	10
Other	78	76	75	75	75	59	62	67	71	73
Size										
1–19	90	89	89	88	88	22	22	21	19	18
20–99	9	9	9	10	9	18	18	18	17	17
100–499	1	1	2	2	2	13	14	14	14	14
500–9999	0	0	0	0	0	21	22	23	23	24
10,000+	0	0	0	0	0	25	24	24	27	28
Age										
0–5 years	27	44	41	38	33	8	16	15	13	11
6–10 years	0	10	19	18	18	0	6	11	9	8
11–20 years	0	1	17	23	22	0	1	13	17	15
21+ years	0	0	0	10	18	0	0	1	11	22
Left Censored	73	45	22	13	9	92	77	60	50	45

Source: PV, LBD, KPSS, KPST, BDSPF-Long, DS, RADS, Compustat and authors' calculations. Figure provides a breakdown, in percent, of patent grants, external citations and breakthroughs by LBD firms' major 2-digit NAICS sector, size, and age. Firms in the LBD cannot be observed prior to 1977. As a result, the majority of firm observations have left-censored age in the early years of the panel. Cells representing less than 0.5 percent are coded as 0.

H Patent Elasticity Estimates

Tables H1 and H2 present the estimates for the LBD and Compustat samples.

Table H1: Patent Elasticity Estimates for Various Research-Input Stocks

	Patent Grants				External Citations				Breakthrough Patents			
	R&D	Payroll in Estabs			R&D	Payroll in			R&D	Payroll in		
	Exp.	54-55	R&D Labs	All	Exp.	54-55	R&D Labs	All	Exp.	54-55	R&D Labs	All
j =1982 - 1986	-.809***	-.7315***	-.6818**	-.5662***	-.4648**	-.1968	-.05548	.1565	-.5139	.3311	1.59*	.6816***
	(.1694)	(.1542)	(.3062)	(.1242)	(.2025)	(.2391)	(.4342)	(.1606)	(.5671)	(.375)	(.8326)	(.244)
j =1987 - 1991	-1.245***	-.7096***	-.01644	-.5524***	-.7196**	.3964	.8656*	.5743***	-1.085*	.7766	2.679***	.7816***
	(.3027)	(.2526)	(.4322)	(.1505)	(.2918)	(.3475)	(.5032)	(.1973)	(.6463)	(.5056)	(.8862)	(.2847)
j =1992 - 1996	-1.278***	-.1765	-.03553	-.3856*	-.1619	1.573**	1.533***	1.124***	-.6897	1.863***	3.372***	1.06***
	(.4056)	(.4292)	(.4152)	(.2137)	(.6327)	(.6271)	(.5427)	(.3133)	(.7518)	(.6919)	(.9173)	(.3674)
j =1997 - 2001	-1.674**	-.03857	.3345	-.5837*	-.5353	1.777**	1.796**	1.054***	-1.844	.1976	2.107	-1.304**
	(.6919)	(.6315)	(.6228)	(.3072)	(1.026)	(.7852)	(.7777)	(.409)	(1.681)	(.9362)	(1.371)	(.5675)
j =2002 - 2006	-2.155***	-.6836	.2423	-1.29***	-.06887	1.793***	2.798***	.8046**	-3.32*	-2.08***	1.278	-4.124***
	(.7324)	(.5793)	(.6442)	(.3164)	(.914)	(.594)	(.8266)	(.3634)	(1.788)	(.7181)	(1.388)	(.6774)
j =2007 - 2011	-2.732***	-1.133	-.3119	-1.851***	-.2668	1.419**	2.658***	.1257	-3.185	-1.247	1.227	-4.661***
	(.9899)	(.7061)	(.8473)	(.3898)	(1.039)	(.6232)	(.9613)	(.4049)	(1.989)	(.8403)	(1.299)	(.7021)
j =2012 - 2016	-3.333***	-1.403*	-.6483	-2.222***								
	(1.14)	(.8116)	(.9981)	(.458)								
Stock ×												
j =1977 - 1981	.265***	.2601***	.08763**	.4649***	.3146***	.2688***	.1509***	.4135***	.1609	.2101***	.1426**	.4722***
	(.07356)	(.0394)	(.0403)	(.0343)	(.08659)	(.05491)	(.04793)	(.03673)	(.1022)	(.06015)	(.06791)	(.05152)
j =1982 - 1986	.2992***	.2914***	.1288***	.4754***	.3444***	.2848***	.1677***	.4011***	.1957*	.1935***	.04061	.426***
	(.07489)	(.04116)	(.0432)	(.03509)	(.08651)	(.05516)	(.06077)	(.03707)	(.1159)	(.06822)	(.08643)	(.05425)
j =1987 - 1991	.3347***	.297***	.0953**	.4812***	.3797***	.2655***	.1301**	.3953***	.2662**	.199***	.00967	.4509***
	(.07514)	(.03794)	(.0462)	(.03481)	(.07868)	(.04208)	(.05411)	(.03456)	(.1034)	(.04952)	(.06636)	(.05196)
j =1992 - 1996	.3516***	.2738***	.1138***	.4852***	.3812***	.2239***	.1257***	.3987***	.3054***	.1939***	.03422	.4962***
	(.05513)	(.0331)	(.03572)	(.03019)	(.05845)	(.03929)	(.04463)	(.03301)	(.09233)	(.05055)	(.0647)	(.05207)
j =1997 - 2001	.3941***	.2792***	.1064**	.5116***	.4233***	.2293***	.1299***	.4193***	.3633***	.2859***	.105	.6173***
	(.0479)	(.04111)	(.04491)	(.02832)	(.06334)	(.04426)	(.04836)	(.03351)	(.1365)	(.06253)	(.09883)	(.05743)
j =2002 - 2006	.4256***	.3252***	.1152**	.5555***	.4152***	.256***	.0829**	.4572***	.3738***	.3576***	.07774	.7098***
	(.0497)	(.03727)	(.04534)	(.03025)	(.06615)	(.03849)	(.03803)	(.03764)	(.1438)	(.06274)	(.08684)	(.07728)
j =2007 - 2011	.4592***	.3556***	.1563***	.5895***	.4291***	.2892***	.09792**	.5038***	.3847**	.3277***	.09796	.7588***
	(.0467)	(.04342)	(.05882)	(.02989)	(.07253)	(.04763)	(.04255)	(.04157)	(.1569)	(.07312)	(.07319)	(.07642)
j =2012 - 2016	.5063***	.3864***	.1981***	.6241***								
	(.05322)	(.05082)	(.07185)	(.03277)								
Constant	1.187	.9297**	3.74***	-2.699***	1.643	1.873***	3.946***	-.8275*	.5904	-.4515	.8871	-4.734***
	(.935)	(.4708)	(.4534)	(.4262)	(1.117)	(.6634)	(.5988)	(.46)	(1.37)	(.7382)	(.8203)	(.6879)
R^2	.8916	.8894	.9126	.8519	.9088	.9053	.924	.8835	.8310	.8157	.8433	.7876
Observations (000s)	68	430	96.5	1813	50	310	70.5	1290	23.5	78	24.5	186

Source: LBD, PV, RADs, KPST and authors' calculations. *Notes:* Table reports estimates of equation (8) for patent grants, external citations, and breakthroughs via PPML. Column headers denote patent measure and research stock used: 'R&D' Exp is R&D expenditure, '54-55' is Professional Services and Management, 'R&D Labs' is NAICS 5417, and 'All' denotes all establishments. First set of coefficients reports estimates of the semidecade fixed effects ($\hat{\gamma}_j$) displayed in Figure 7. Second set of estimates reports the time-varying patent elasticities ($\hat{\eta}_j$) displayed in Figure 6. All specifications include firm fixed effects. Standard errors are clustered by firm.

Table H2: Patent Elasticity Estimates (Compustat Panel)

	Patents		Breakthroughs		External Cites	
	SG&A	R&D	SG&A	R&D	SG&A	R&D
j=1982-1986	–0.917***	–0.701***	0.0758	0.0493	–0.277**	–0.262***
	(0.138)	(0.0919)	(0.325)	(0.264)	(0.134)	(0.0900)
j=1987-1991	–1.263***	–0.848***	–0.350	–0.262	–0.0588	–0.0592
	(0.239)	(0.162)	(0.451)	(0.423)	(0.192)	(0.143)
j=1992-1996	–0.572	–0.343	0.780	0.872*	1.155***	0.949***
	(0.362)	(0.211)	(0.539)	(0.509)	(0.260)	(0.172)
j=1997-2001	–0.162	–0.163	–0.151	0.269	1.470***	1.157***
	(0.505)	(0.296)	(0.727)	(0.802)	(0.331)	(0.228)
j=2002-2006	–0.217	–0.0881	–2.153**	–0.919	1.559***	1.471***
	(0.610)	(0.393)	(1.030)	(1.293)	(0.394)	(0.297)
j=2007-2011	–0.417	–0.106	–2.279*	–0.912	1.449***	1.460***
	(0.743)	(0.472)	(1.186)	(1.471)	(0.481)	(0.361)
j=2012-2016	–0.528	–0.146				
	(0.852)	(0.616)				
Stock x j=1977-1981	0.359***	0.370***	0.455***	0.384***	0.401***	0.414***
	(0.0651)	(0.0564)	(0.0862)	(0.0811)	(0.0454)	(0.0407)
Stock x j=1982-1986	0.437***	0.426***	0.437***	0.363***	0.437***	0.448***
	(0.0674)	(0.0566)	(0.0886)	(0.0893)	(0.0448)	(0.0391)
Stock x j=1987-1991	0.492***	0.462***	0.549***	0.487***	0.444***	0.456***
	(0.0712)	(0.0572)	(0.0926)	(0.0915)	(0.0438)	(0.0370)
Stock x j=1992-1996	0.434***	0.418***	0.532***	0.474***	0.354***	0.375***
	(0.0745)	(0.0539)	(0.0914)	(0.0922)	(0.0463)	(0.0368)
Stock x j=1997-2001	0.422***	0.442***	0.611***	0.533***	0.348***	0.384***
	(0.0850)	(0.0601)	(0.105)	(0.125)	(0.0523)	(0.0422)
Stock x j=2002-2006	0.429***	0.428***	0.688***	0.517***	0.391***	0.398***
	(0.0976)	(0.0733)	(0.149)	(0.191)	(0.0573)	(0.0474)
Stock x j=2007-2011	0.443***	0.419***	0.717***	0.545***	0.408***	0.402***
	(0.105)	(0.0780)	(0.157)	(0.203)	(0.0647)	(0.0543)
Stock x j=2012-2016	0.476***	0.451***				
	(0.115)	(0.0963)				
Constant	2.073***	2.718***	–0.612	0.542	2.316***	3.012***
	(0.393)	(0.230)	(0.569)	(0.389)	(0.258)	(0.180)
R^2	0.908	0.904	0.838	0.831	0.907	0.904
Observations	96,233	74,516	36,300	33,322	83,134	64,843

Source: Compustat, PV, KPST and authors' calculations. *Notes:* Table reports estimates of equation (8) via PPML. Column headers denote patent measure and research stock used: 'R&D' Exp is R&D expenditure, 'SG&A' is Sales, General and Administrative.. First set of coefficients reports estimates of the semidecade fixed effects ($\hat{\gamma}_j$) displayed in Figure 7. Second set of estimates reports the time-varying patent elasticities ($\hat{\eta}_j$) displayed in Figure 6. Regressions weighted by the growth rate denominator. Standard errors clustered by firm.

I Growth Regression Estimates

Table I1 presents the coefficient estimates depicted in Figures 11 and 12.

Table I1: Growth Regression Estimates (LBD Panel)

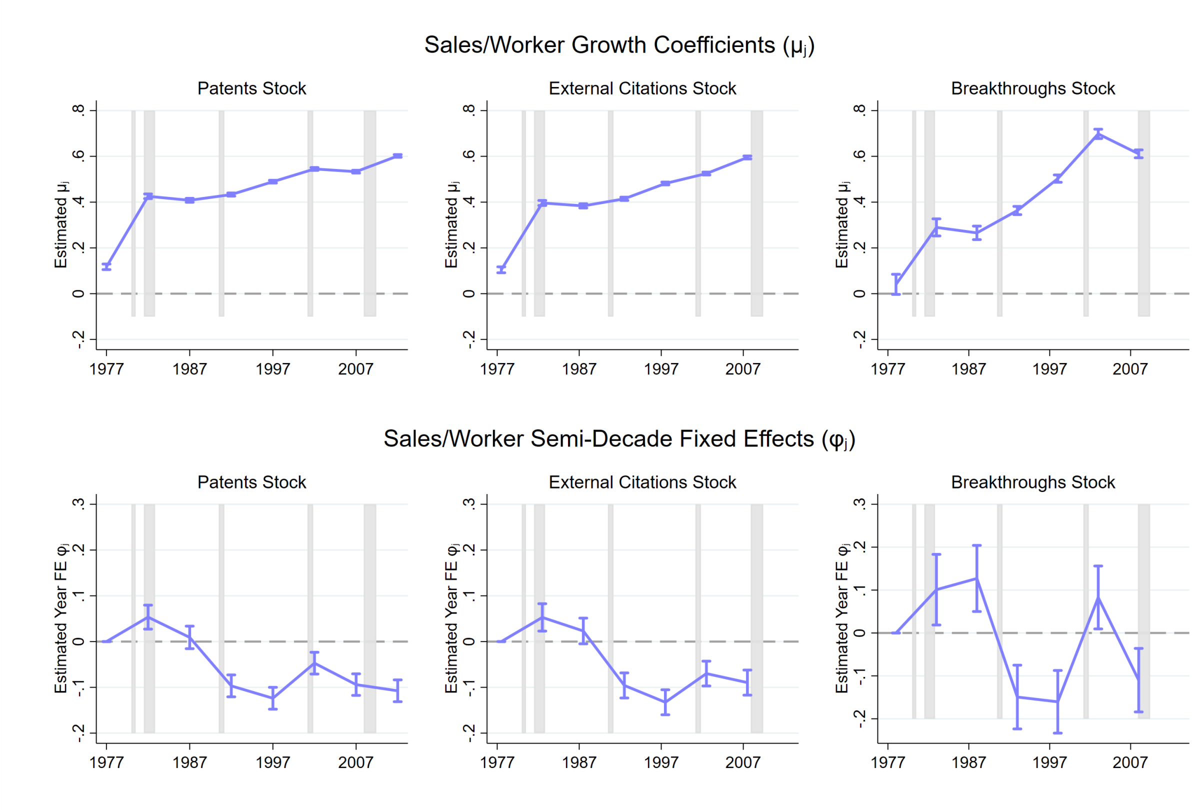
Dependent variable:	Growth(Sales/Worker)			Growth(Sales)		
Idea Stock:	Patents	Breakthroughs	External Cites	Patents	Breakthroughs	External Cites
Growth(Employment)				.8741***	.8665***	.8758***
				(.01007)	(.0185)	(.01223)
j=1982-1987	.2053***	.2284	.1945**	.22***	.216***	.2236***
	(.07421)	(.1573)	(.08953)	(.03907)	(.07659)	(.03979)
j=1987-1992	.4248***	.5583***	.4273***	.3529***	.3744***	.3396***
	(.0746)	(.1551)	(.0883)	(.03788)	(.07219)	(.03966)
j=1992-1997	.203***	.3727**	.2407***	.2289***	.2815***	.2366***
	(.07165)	(.1482)	(.08611)	(.03311)	(.06019)	(.03451)
j=1997-2002	.09297	.2018	.1122	.2094***	.2255***	.2073***
	(.07181)	(.1479)	(.08688)	(.03182)	(.05816)	(.03332)
j=2002-2007	.1953***	.3831***	.2163**	.2498***	.2868***	.2523***
	(.07279)	(.1466)	(.08817)	(.03355)	(.06294)	(.03633)
j=2007-2012	.08921	.2132	.1202	.1471***	.1585***	.1536***
	(.06984)	(.1456)	(.08494)	(.03086)	(.05859)	(.03245)
j=2012-2017	.09846			.152***		
	(.06905)			(.03058)		
Growth(Stock) x j=1977-1982	.1235***	-.02259	.1235***	.01727	-.04346	.03207*
	(.03764)	(.0307)	(.03324)	(.01812)	(.02693)	(.01636)
Growth(Stock) x j=1982-1987	.47***	.2173***	.428***	.08682***	.03265	.0829***
	(.03569)	(.05916)	(.03458)	(.01398)	(.02511)	(.0144)
Growth(Stock) x j=1987-1992	.2903***	.1876***	.2786***	.04779**	.089*	.06735***
	(.03757)	(.06076)	(.03808)	(.02253)	(.04641)	(.02575)
Growth(Stock) x j=1992-1997	.2281***	.1013***	.2058***	.02293*	.003478	.02028
	(.02749)	(.03186)	(.02592)	(.01225)	(.01822)	(.01264)
Growth(Stock) x j=1997-2002	.439***	.3837***	.4069***	.1046***	.1172***	.1024***
	(.02695)	(.0454)	(.02792)	(.02002)	(.04388)	(.02607)
Growth(Stock) x j=2002-2007	.3498***	.4209***	.3327***	.07165***	.09644***	.068***
	(.03288)	(.06487)	(.0308)	(.01536)	(.02941)	(.01471)
Growth(Stock) x j=2007-2012	.3344***	.2722***	.3424***	.04066***	.03589*	.04089***
	(.02959)	(.04741)	(.02935)	(.01402)	(.02131)	(.01349)
Growth(Stock) x j=2012-2017	.365***			.0638***		
	(.02647)			(.01206)		
Constant	-.297***	-.3516***	-.3373***	-.1662***	-.1755***	-.1729***
	(.06239)	(.1274)	(.0746)	(.02668)	(.04926)	(.0276)
R^2	.1949	.1666	.1913	.6920	.6543	.6896
Observations	321,000	33,000	226,000	409,000	42,000	291,000

Source: LBD, PV, KPST and authors' calculations. *Notes:* Table reports estimates of equation (10) via OLS. Column header denotes the measure of output growth and idea stock growth for each regression. First set of coefficients reports estimates of the semidecade fixed effects ($\hat{\phi}_j$) displayed in Figure 7. Second set of estimates corresponds to time-varying coefficients on the Growth in Idea Stock ($\hat{\mu}_j$) displayed in Figure 6. Regressions also include firm size and age bins, whose coefficients are not reported. Regressions weighted by the growth rate denominator. Standard errors clustered by firm.

I.1 Unweighted LBD Growth Regression Estimates

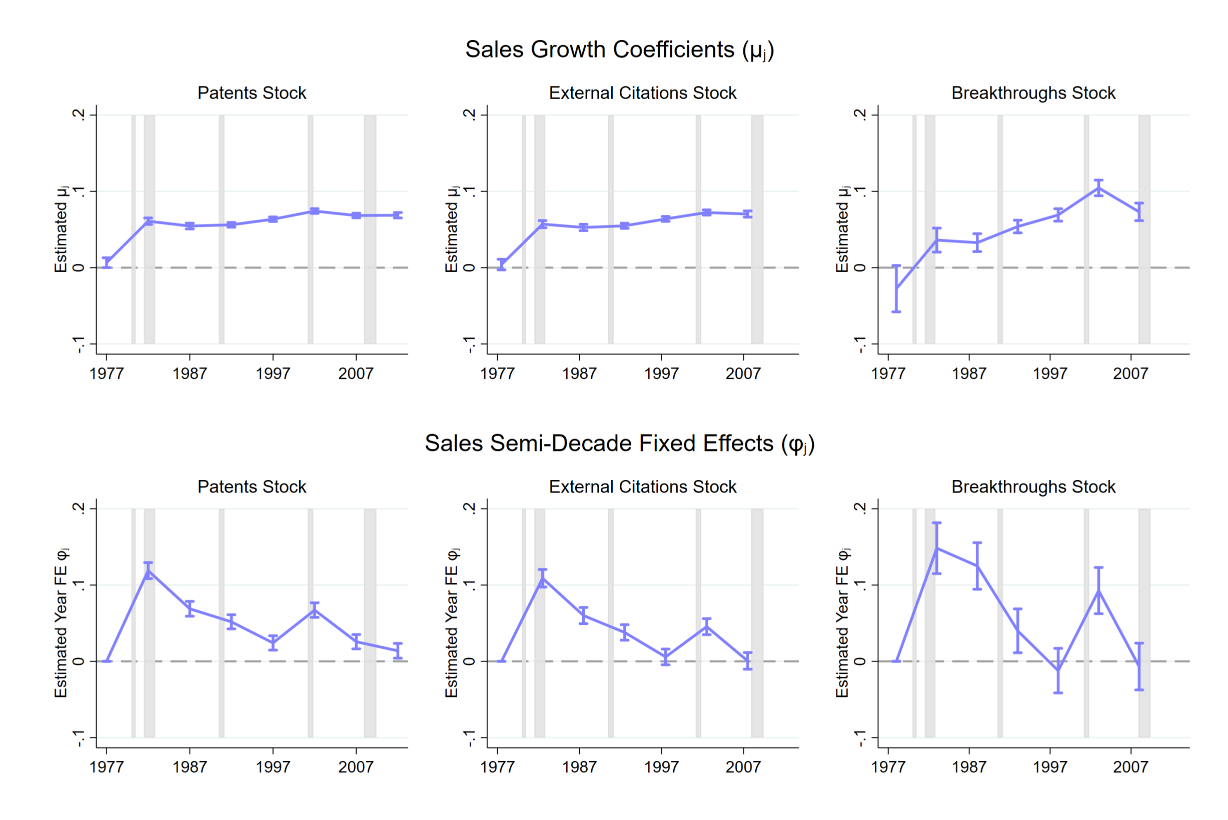
Figures I1 and I2 present unweighted growth regression estimates.

Figure I1: Unweighted Estimated Sales per Worker Growth Coefficients



Source: LBD, EC, PV, KPST, and authors' calculations. *Notes:* Figure reports relationship between LBD firms' sales per worker and idea stock growth between years t and $t+5$. Standard errors clustered by firm with 95% CIs depicted.

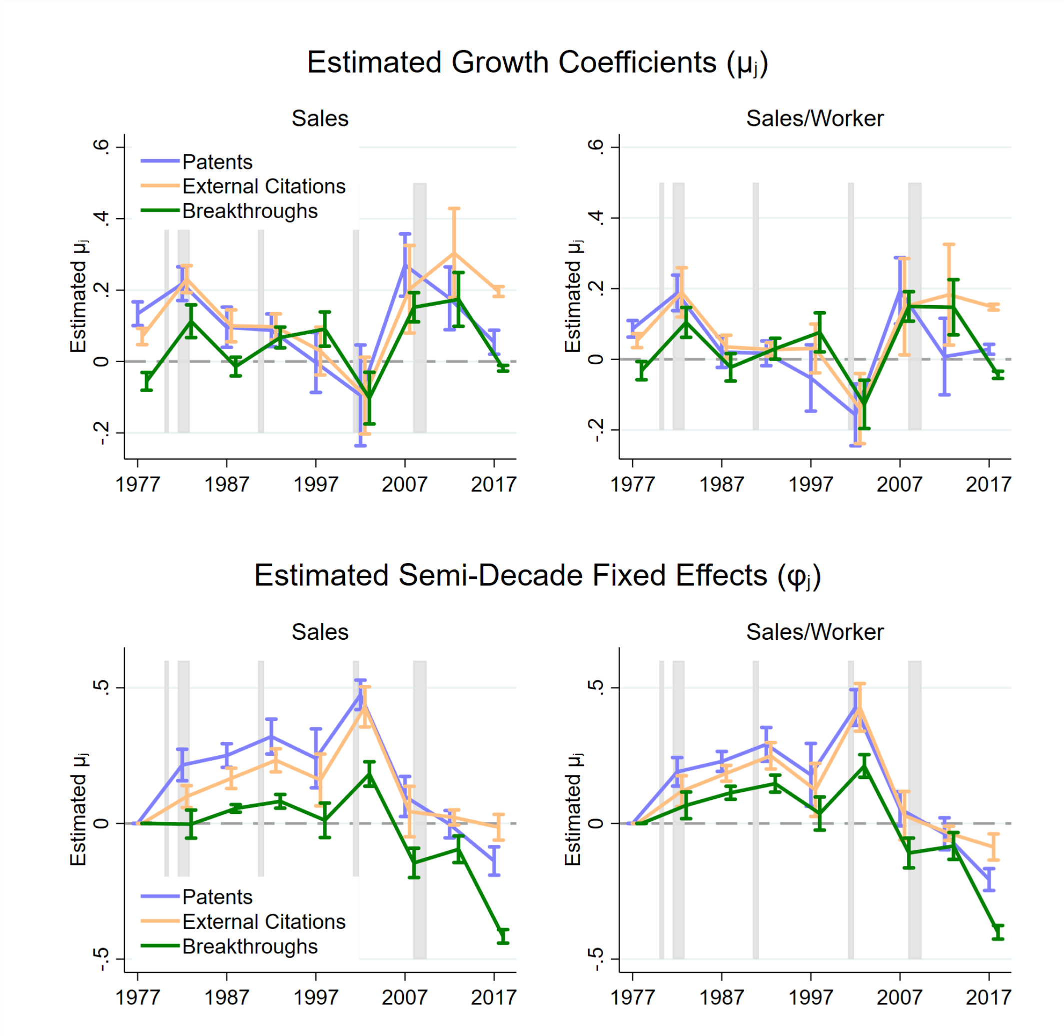
Figure I2: Unweighted Estimated Sales Growth Coefficients



Source: LBD, EC, PV, KPST, and authors' calculations. *Notes:* Figure reports relationship between LBD firms' sales and idea stock growth between years t and t+5. Standard errors clustered by firm with 95% CIs depicted.

I.2 Compustat Estimates

Figure I3: Estimated Growth Coefficients (μ_j) and Time Fixed Effects (ϕ_j) (Compustat Panel)



Source: Compustat, PV, KPST, and authors' calculations. *Notes*: First two panels report relationship between Compustat firms' noted outcome and idea stock growth by semidecade. Second two panels report the growth regression semidecade fixed effects for each outcome. Standard errors are clustered at the firm-level with 95 percent confidence intervals depicted.