Patent Hunters

Lauren Cohen Umit G. Gurun S. Katie Moon Paula Suh*

May 9, 2024

Abstract

Analyzing millions of patents granted by the USPTO between 1970 and 2020, we find a pattern where specific patents only rise to prominence after considerable time has passed. Amongst these late-blooming influential patents, we show that there are key players (patent hunters) who consistently identify and develop them. Although initially overlooked, these late-bloomer patents have significantly more influence on average than early-recognized patents and open significantly broader new markets and innovative spaces. For instance, they are associated with a 15.6% ($t = 29.1$) increase in patenting in the late-bloomer's technology space. Patent hunters, as early detectors and adopters of these late-blooming patents, are also associated with significant positive rents. Their adoption of these overlooked patents is associated with a 22% rise in sales growth ($t = 6.55$), a 3% increase in Tobin's Q ($t = 3.77$), and a 4.8% increase in new product offerings ($t = 2.25$). We instrument for patent hunting and find similarly large positive impacts on firm value. Interestingly, these rents associated with patent hunting on average exceed those of the original patent creators themselves. Patents hunted tend to be closer to the core technology of the hunters, more peripheral to the writers, and to be in less competitive spaces. Lastly, patent hunting appears to be a persistent firm characteristic and to have an inventor-level component as well.

Keywords: Innovation chain, innovation rent, patent citations, patent evolution, technological impact, technological trajectories, commercialization.

JEL Classification: O31, O33, L1

*Cohen is at Harvard Business School and NBER, Baker Library 279, Soldiers Field, Boston, MA, 02163 [lcohen@hbs.edu]. Gurun is at the University of Texas at Dallas, 800 W Campbell Rd., SM41, Richardson, TX 75080, [umit.gurun@utdallas.edu]. Moon is at the University of Colorado at Boulder, Leeds School of Business, Campus Box 419, Boulder, CO 80309, USA, [katie.moon@colorado.edu]. Suh is at the University of Georgia, Terry College of Business, 620 S. Lumpkin St., Athens, GA, 30602, [paula.suh@uga.edu].

1 Introduction

Not all ideas that eventually are successful are recognized immediately. Indeed, eventual positive realizations of innovation take many divergent paths to reach that success point. In this paper, we explore millions of patents to identify ideas that catch on late. We show that while these ideas are equally as valuable as ideas that catch on early, the rents along the value chain are shared quite differently. Namely, we provide the first large-sample evidence that there are critical agents in the innovation chain who actively search out ("hunt") these neglected and overlooked ideas and use them as critical inputs in their innovation and commercialization process. We show that the agents are unique and non-substitutable players in the innovation chain. Moreover, the rents to "patent hunting" are substantial – often the most sizable portion of the entire innovation chain. This patent hunting role – and the technology, physical capital, and human capital needed to implement it – thus appears to represent an important component of many innovation chains and, thus, consideration of agents entering across innovation stages.

In order to explore these rich components of the innovation chain, we examine the past nearly 50 years of patenting in the United States to identify those patents that eventually do catch on and become influential patents. While many of those are identified early on as influential patents that other innovators build upon, a sizable portion ends up being "late-bloomer" patents. These are patents that end up being influential but are not recognized as so until much later in their life. As all patents are, upon approval, publicly available for other innovators to read, build upon, and cite, one could imagine that conditioning on patents that end up as influential, those that are passed over initially are of lower ending value on average. However, we find that this is not the case. These late-bloomer patents, in contrast, when compared with patents that bloom earlier appear just as valuable on average, and by some markers have even more impact.

To better understand our approach, consider the example of GPU (Graphical Processing Units) technology. In 1991, US Patent #5,025,407 was granted to Texas Instruments, Inc.

Texas Instruments (TI) is a publicly traded technology firm based in Dallas, TX specializing in semiconductors and other circuitry technology. This patent was in the technology classes of both G06F (Electric digital data processing) and G06T (Image data processing), as shown in Figure A1. TI's core technology class was H01L (Semiconductor devices). This patent's technology proximity to TI's core technology was then calculated to be 0.13 (with smaller values meaning more distant; larger values meaning closer to "core").

In its early years, the patent garnered comparatively little attention. In fact, it was in 2006, the fifteenth year following its approval, that it saw a large and record spike in citations and innovation build-out.[1] And this was largely driven by a single firm: Nvidia, Corp. Nvidia (also a publicly traded technology company) was founded in the early 1990s, and is based in Santa Clara, CA. It was founded specifically focusing on the promise of graphics technology, with its technology proximity (0.32) much closer to the TI patent than the inventing firm itself. Nvidia used this patent to continue to develop and build out its product portfolio, which both contributed to – and was positively buoyed by – the positive demand trends of gaming (especially mobile), and GPU-reliant AI, machine learning, and crypto-asset demands. Subsequent to this period, prominent innovators like Apple, Inc. entered the fray, focusing particularly on the computationally intensive requirements that emerged in the years leading up to and throughout the early 2020s.

Interestingly, the patent inventor – Texas Instruments – did not end up taking further part in building out GPUs in earnest, nor did it continue to take and build on this critical patent. Instead, other key players, such as Nvidia, years later took the patent, built upon it, and developed and commercialized an industry of products in the space.

We demonstrate that instances like this are far from rare, occurring regularly and extensively across the entire spectrum of patenting and innovation.[2] In particular, patent hunters (*e.g.,* Nvidia) routinely are critical players in the development and commercialization of

[1] See panel (b) of Appendix Figure A1 for the patent's citations over time.

[2] In Appendix Figure A2 and Table **??**, we catalog a selection of other notable technologies—ranging from battery packs to carbon refrigerators, liquid dishwashing detergents, or control logic interfaces for embedding microprocessors in gate arrays—that display analogous trends.

late-blooming high-impact patents across their innovation chain. Moreover, we find that patent-hunting is associated with sizable rents across a number of dimensions. For instance, patent hunters' sales growth increases by a large and significant percentage following their discovery and incorporation of the hunted patent. In particular, sales growth increases by 22.04% ($t = 6.55$) on average. Moreover, Tobin's Q also increases by a significant 2.86% ($t = 3.77$), coupled with a significant increase in the real quantity of new products developed by the patent hunter – a 4.82% ($t = 2.25$) expansion on average. These differences are taken relative to a rich set of "counter-factual" agents within the innovation space – namely, the writers of the forgotten patents (*e.g.,* Texas Instruments), those innovators that build on the more easily identifiable "early blooming" patents, and the patent hunting firms *themselves* pre-hunted patent.

We next explore what types of firms are involved. The writers of the initially overlooked, late-bloomer patents tend to be older, larger, value firms such as General Electric, Eastman Kodak, and Xerox. In contrast, patent hunters tend to be smaller, consumer-focused, growth firms such as Sandisk, Broadcom, and Tivo Corp. Moreover, patent hunting appears to be a persistent firm characteristic. The same firms continue hunting over time, and the rents to their hunting appear to have a learning component, with successively hunted patents increasing in benefits accrued to the hunter over time.

Next, we explore the nature of the patents that become the target of hunting. One might expect that even if the patents are overlooked by other agents, the patenting firm itself should be well aware of the patent and could build upon it. Thus, one might expect there to be some reason as to why the original patent writer does not further develop the patent and its technology area, and somehow "allows" the patent to be hunted and developed by an outside firm. We find a number of systematic markers of these patents consistent with this notion. First, and in line with the TI-Nvidia example above, we find that patents that are hunted by outside firms are on average significantly more distant from the patent writers' core technology areas than their average patent, and conversely significantly closer to the technology focus area of the respective patent hunter. Second, we find that these patents that are later hunted are in technology spaces that are less competitive at their time of patenting. This is consistent with the writers expecting less time pressure to develop the patent immediately, and so developing other closer and more competitive spaces first.

Once developed, the late-blooming, hunted patents also take a diverging path from all other influential patents. We define influential ("superstar") patents as a non-parametric measure of patent success of those patents that receive the 95th percentile or above vintage- and technology class-adjusted citations. We do this as past literature has shown this right-tail parameterization to more closely capture commercializability of a patent, and more highly correlate with patent value, given the highly skewed distribution of patent citations (Trajtenberg, 1990; Sampat and Ziedonis, 2004). For all other superstar patents, the technology distance of citing patents from the original patent increases over time, suggesting technology moving past and away from the original idea. In contrast, for initially overlooked patents, patents and innovators continue to build on the patent. Once hunted and "bloomed," these patents define new spaces with significantly more patenting and innovation happening directly around them – specifically, a 15.6% ($t = 29.19$) increase in new patents granted in the late-bloomer patents' technology space.

Moreover, we then drill down to the individual inventor level to explore inventor-level components in the patent-hunting process. We find two aspects consistent with their being inventor-level determinants of the process and its dynamics. First, we look at firms that end up hiring the inventor of the original patent that they "hunted." In the TI-Nvidia example above, this would be Nvidia hiring one of David Gulley or Jerry Van Aken from Texas Instruments (the two listed patent inventors on the patent) following their detection of the patent. We find that in these cases, the benefits that accrue to the patent hunting firm (sales growth and Tobin's Q) more than double. Of course, it could be that hiring the original inventors is part of a broader scope of investments in the new technology space (and so the real effects we are measuring are not due solely to bringing in the original inventors of the focal patent itself), but at the least, these inventor acquisitions along with the hunted patent provide a signal that results in an economically and statistically reliably greater value gained from patent hunting. Second, we find that inventors themselves who engage in patent hunting (so the inventor at Nvidia who cited the TI patent on a graphic patent), are significantly more likely to continue to patent hunt at future firms they work for (*e.g.,* when moving to Tesla). In particular, they are 7% ($t = 11.23$) more likely to continue to be hunting inventors at subsequent firms. Much like the above, it is difficult to disentangle whether this is the inventor herself, or simply the inventor's selection of firms that subsequently also patent hunt, but irrespective, it suggests that there is an inventor-level component of the hunting.

Stepping back, given the large, positive rents that are associated with patent hunting – larger, in fact, than those that accrue even to the original patent writer itself – one might wonder why anyone would choose to be a patent writer of the original patent at all, instead of specializing in patent hunting. First, as we mention above, there are a number of moderating effects of patent hunting. The rents to patent hunting are attenuated when there are too many same technology class focal patents to search from, along with when there are too many same technology class patent hunters that already exist. Both of these are consistent with increased search costs (lower equilibrium rents) impacting the rents to patent hunting and moreover suggest that the patent hunting mechanism might follow a search cost model. Motivated by this, we develop a simple search model framework in Section 3 to frame thinking around this. Second, there are certainly firm-specific characteristics that cause cross-sectional and time-series heterogeneity in the relative costs and benefits of patent hunting. While labor may appear somewhat mobile (so that patent writing firms like TI could simply attract researchers from Nvidia with sufficiently high wages, benefits, etc.) there are many non-transferrable characteristics such as location, agglomeration, intangible capital (for instance, brand) complementarities or other firm-specific components that make patent hunting uniquely and privately valuable for certain firms, and unprofitable for others.

In that sense, we do not believe any firm is necessarily solving along the innovation chain sub-optimally, or making a mistake by being a writer or patent hunter. That said, we are the first paper to provide large sample evidence on the rents to this activity and portion of the value chain, and given the sizable and repeatable nature of these pay-offs, it might be worth bringing to this innovation conversation an assessment of investment opportunities throughout the chain, when available.

The remainder of the paper proceeds as follows. Section 2 provides a literature review. Section 3 develops a simple framework for an innovator thinking through investment in the write or hunt process. Section 4 describes the data and sample selection, while Section 5 explores the dynamics of influential early- and late-bloomer patents. Section 6 presents the main results on patent hunters and the benefits that accrue to them. Section 7 explores the incentives and characteristics of the firms and inventors involved in patent hunting. Section 8 concludes.

2 Literature Review

Our paper is mainly related to the literature studying the path of knowledge production, technological innovation, and their impacts on economic growth. Along these lines, Weitzman (1998) presents a model to analyze the determinants of long-term growth by considering a production function of new knowledge that uses new configurations of old knowledge as an input. This paper emphasizes the role of building upon existing ideas to create new knowledge. Likewise, analyzing citation patterns over 18 million scientific papers, Uzzi, Mukherjee, Stringer, and Jones (2013) show the most influential research tends to be rooted in conventional combinations of existing knowledge but also includes elements of unusual combinations. Similarly, Escolar, Hiraoka, Igami, and Ozcan (2023) focus on technological trajectories and the reuse of knowledge in subsequent inventions, emphasizing the significance of combining dissimilar technological components with strong scientific content to shape trajectories with high technological impact. In a similar dynamic setting, past work such as Gal-Or (1987) and Chamley and Gale (1994) have modeled Stackelberg games, in which in the traditional leader-follower set-ups there are distinct advantages and certain agents who benefit from a second-mover advantage and entry strategy. Glode and Ordonez (2023) then take the tack of modeling surplus-creating and -appropriating activities of firms, relating this decision to industry-wide technological shocks that change firms' incentives to invest in the latter versus the prior.

In another related study, Pezzoni, Veugelers, and Visentin (2022) investigate the reuse of novel technologies in subsequent inventions and explore the various factors influencing this process. Using European patent data spanning from 1985 to 2015, they identify new combinations of existing technological components, marking the inception of technological trajectories. Instead of relying on citations to trace these trajectories, the authors identify a novel technology as the first occurrence of a particular combination of IPC classes in a patent, akin to the approach taken in Strumsky and Lobo (2015) and Verhoeven, Bakker, and Veugelers (2016). Pezzoni et al. (2022) also observe that technological trajectories tend to follow an S-shaped curve, with variations in their take-off time and maximum technological impact. Specifically, they noted that complex technologies, involving dissimilar components with strong scientific content, often have a longer take-off time but result in higher impact. Conversely, simpler technologies with familiar components tend to have a shorter take-off time but yield lower impact. We add richness to this not only in exploring firms, even down to individual inventors, who specialize in critical components of this process, but the expansive set of value implications of identifying and building on a patent within its technological trajectory.

Our paper also aligns with research exploring the ramifications of technology spillovers on both economic growth and technological advancement. For example, Bloom, Schankerman, and Van Reenen (2013) delve into the pivotal role played by technology spillovers in stimulating economic growth, underscoring the significance of incentivizing research and development (R&D) and considering the scale of firms involved. Furthermore, Kelly, Papanikolaou, Seru, and Taddy (2021) employ textual analysis of patent documents to gauge technological innovation. They identify patents that stand apart from previous work but still have relevance to subsequent innovations. Notably, they classify these distinctive patents as breakthrough innovations, investigate the domains in which these breakthroughs occur, and establish connections between these breakthrough innovations and the overall total factor productivity. Collectively, these studies offer valuable insights into the intricate interplay between knowledge creation, technological progress, and economic growth. Our contribution to this body of research lies in the identification of specific technology groups where spillovers may require time to materialize, with the velocity of these spillovers contingent on the presence of patent hunters.

Finally, our paper also contributes to the ongoing discourse surrounding the merits and demerits of a patent system where inventors publicly unveil their innovations in exchange for patent protection, and the potential ramifications of such a system on future innovations. One perspective posits that the imperative for patent disclosure might dissuade individual inventors, potentially eroding the incentives inherent in the patent system. Conversely, an opposing viewpoint contends that patent disclosure serves as a mechanism to stimulate fresh ideas and foster innovation (as discussed by Williams (2017)). On the empirical side, Furman, Nagler, and Watzinger (2021) delve into the role of information disclosure via patents, revealing that increased accessibility to technical knowledge significantly bolsters local patenting and business establishment. Their findings underscore the role of patent disclosure in advancing cumulative innovation. In contrast, Kim and Valentine (2021) report that firms compelled to disclose their innovations more promptly by the American Inventor's Protection Act (AIPA) of 1999 reduce their R&D investments and generate fewer innovations (see also Graham and Hegde (2015) and Hegde, Herkenhoff, and Zhu (2015)). Our paper adds to this literature by showing the unique merits of a patent disclosure system where patent hunters detect and build on initially neglected ideas to create new knowledge.

In the case of scientific papers, Garfield (1970) proposes that the use of science citation indices not only prevents inadvertent neglect of useful work but also reduces duplicative effort in research and publication. Subsequently, Garfield (1980, 1989a,b, 1990) provide concrete examples of articles that experienced delayed recognition within the scholarly community. Expanding on this notion, Glänzel, Schlemmer, and Thijs (2003), through an extensive literature survey, offer an estimate of the prevalence of delayed recognition and explore common characteristics shared by papers that receive belated recognition. Van Raan (2004) proposes a framework for measuring delayed recognition, often referred to as *sleeping beauties*, along three dimensions: (i) the length of sleep, signifying the duration of the "sleeping period" (ii) the depth of sleep, denoting the average number of citations during the sleeping period and (iii) awake intensity, indicating the number of citations accumulated after the sleeping period.

In a similar vein, Ke, Ferrara, Radicchi, and Flammini (2015) reexamine the concept of delayed recognition, introducing parameter-free methods to identify papers that might escape detection by the methods proposed by Van Raan (2004). Van Raan and Winnink (2019) document instances in medical research where publications went unnoticed for several years after their initial release, only to suddenly garner citation attention subsequently. Moreover, Van Raan (2017) and Hou and Yang (2019) extend the analysis of delayed recognition from scientific papers to patents, exploring various evolutionary trajectories of patents in this context. Our paper builds upon the concept of delayed recognition in the realm of patents, specifically mapping the benefits and costs experienced by innovators whose patents gain recognition after a significant delay. It then moves a step beyond to explore important agents in the recognition process, and the value that accrues to these agents who identify these late-blooming patents.

3 Theoretical framework

In this section, we provide a simple framework for the research development decisions of a firm: whether to write or hunt for a patent. Time is discrete, and the horizon is infinite. At each time $t > 0$, a firm can either write a new patent itself or hunt for a patent that has been produced before by others. If the firm writes the patent itself, the patent will help produce a good at zero marginal cost in each period t. Consumers value this good as $v > 0$, and the price is $p = \alpha v$, where $\alpha \in (0,\ 1)$ denotes the surplus the firm can capture in the market by commercializing the patent.

Alternatively, the company explores the patent landscape by incurring a cost, c, to pinpoint a patent that potentially holds greater commercialization value than the patents it could develop internally, *i.e.*, $V - L - A \geq v$. Here, V represents the value of the goods produced by the new patent. L is the cost of acquiring a technology license from the original innovator, and A is the expenses associated with adapting the product for successful commercialization. It is important to note that an innovator is more likely to demand a higher licensing fee when the original technology closely aligns with its core technologies. The likelihood of discovering this second category of patent is $\lambda \in (0,\ 1)$. Upon identifying such a patent, the company transitions from producing the previous generation of goods to creating novel products. In each time period, there exists a probability $\delta \in (0,\ 1)$ that the scenario concludes, bringing the game to an end.

In light of these fundamental parameters, we provide a comparative analysis of a firm's strategic choices, delineated by two distinct avenues:

a) Firm valuation in the absence of patent space exploration:

The cumulative summation over time t of $(1-\delta)^t(1-\lambda)^t\alpha v$ gives us $\frac{\alpha v(1-\delta-\lambda+\lambda\delta)}{\delta+\lambda-\lambda\delta}$.

b) Firm valuation upon discovery of a "Hunt" at time t:

$$\sum_{t=1}^{\infty}(1-\delta)^t((1-\lambda)^t\alpha(v-c)+(1-\lambda)^{t-1}\lambda H)),$$

where $H = \alpha(V - L - A)(1 - \delta)\delta$ is the perpetuity value of the newly hunted technology.

The net benefit of following a search strategy to find a hunt is then given by the difference between these values:

$$\alpha(V-L-A)(1-\delta)\lambda - \ c\delta(1-\lambda).$$

This expression is increasing in

(i) λ – As the probability of hunt increases, the avidity of the pursuit escalates;

(ii) α – The magnitude of commercialization potential amplifies the incentive for hunting.

Contrarily, the expression wanes with

(i) c – A diminished cost attached to patent exploration kindles a more fervent pursuit;

(ii) L and A – The appeal of these endeavors increases when licensing fees L and adaptation costs A are low.

4 Data and sample selection

Our data come from various data sources. U.S. patent data are obtained from Thomson Innovation, which covers all patents granted in the U.S. between 1835 and 2020 and contains information on backward and forward citations. We merge this patent data with PatentsView data, which provides information on assignees, claims, inventors, and examiners of the patents granted since 1976. Our main analyses focus on U.S. public firms with Compustat financial data. To identify patents belonging to the U.S. public firms, we match patent numbers to CRSP permno using Kogan, Papanikolaou, Seru, and Stoffman (2017) data. Lastly, the data on new products of U.S. public firms are obtained from Mukherjee, Thornquist, and Žaldokas (2022). Our final data for the analyses consists of all public U.S. firm patents between 1976 and 2020. See Appendix Table A1 for more details on our sample selection procedure.

5 Superstar and late-bloomer patents

We define a patent as a superstar patent if the cumulative citation of the patent within its CPC class and grant year cohort is in the top 95th percentile at any point in time during the patent term of 20 years since its grant.[3] To make a fair comparison of citations across patents granted in different years, we only consider citations during the term of patenting, *i.e.,* the first 20 years, and compute the cumulative citation percentiles within the same grant year and CPC class. By doing so, we restrict our sample to patents granted between 1976 and 1999, so that the last cohort of patents granted in 1999 has full 20-year citation data ending in 2020. We remove self-citations from the citation count as we aim to capture the use of the patents by external users.

By definition, superstar patents are significantly more successful patents than non-superstar patents based on the number of forward citations. We illustrate the stark differences in citations between superstar patents and non-superstar patents in Figure 1. Panels (a) and (b) compare citation counts of superstar patents and non-superstar patents over the patent age. Panel (a) plots the average number of citations for each group using different scales for the visual clarity. We note in panel (a) that citations of both superstar and non-superstar patents grow rapidly in the initial five years of patent life. The average number of citations for non-superstar patents reaches its peak at the age of five and remains relatively constant thereafter. In contrast, the average number of citations for superstar patents continues to

[3] Our approach using percentile distributions of cumulative citations is well accepted in the literature. The right-tail parametrization has been shown to highly correlate with the patent value and commercializability of a patent (Trajtenberg, 1990; Sampat and Ziedonis, 2004). Also, see Brogaard, Engelberg, and Van Wesep (2018) as an example of academic citations.

grow, reaching its peak much later at the patent age of 16. Superstar patents also receive a significantly larger number of citations at every point in the patent's age (*e.g.,* 3 vs. 0.5 at the age of five). Consistent with this observation, Panel (b) shows that the cumulative number of citations for superstar patents grows at a much faster rate compared to that of non-superstars. In fact, by the patent age of 20, superstar patents have accumulated more than five times the number of non-superstar patents' cumulative citations.

[Insert Figure 1 Here]

We then explore different paths to success by examining the time it takes to become a superstar patent. Panel (c) plots the distribution of the years it takes for a patent to be recognized as a superstar patent.[4] The average (median) time until a superstar patent reaches the top 5% of the cumulative citation distribution within the same grant year and CPC class is five (three) years. We use the 90th percentile cutoff (14 years) in this distribution to further characterize patents that take a significantly longer time to become a superstar patent and classify them as "late-bloomer patents." We call the remaining superstar patents "early-bloomer patents."

Figure 2 contrasts the divergent paths to the success point of the early- and late-bloomer patents more precisely. Late-bloomer patents receive a small number of citations near the grant year but accumulate a large number of citations later in their life. A rapidly growing cumulative citation around the patent age of 10 in Panel (a) confirms this point. The box plot in Panel (b) clearly shows this convexity in cumulative citations over the patent age. In contrast, early-bloomer patents start to accumulate citations early on as shown in Panel (c), and the box plot in Panel (d) shows the concavity of cumulative citations over the patent age.

[Insert Figure 2 Here]

[4]We denote the year when a given patent becomes a superstar patent "superstar-year." We use this variable extensively in our subsequent analyses when we analyze the timing of the patent-hunting benefits.

Next, we examine the differences in patent characteristics at the time of patent grant. We present the summary statistics that compare different groups of patents in Table 1. Panel A presents the descriptive statistics of superstar and non-superstar patents and the differences between them. We find some economically and statistically significant differences between superstar and non-superstar patents. Superstar patents tend to be broader and manifested in a larger number of CPC class categories and patent claims. They are more likely to be assigned to public corporations, make more backward citations, and experience positive market responses on the grant date as measured by the KPSS (Kogan et al., 2017) patent value metric. When we compare early-bloomer patents to late-bloomer patents in Panel B, we find that the economic magnitudes of the differences at the time of the patent grant are barely meaningful despite the statistical significance. In particular, the difference in KPSS patent values between early-bloomer and late-bloomer patents is both statistically and economically indistinguishable from zero.

[Insert Table 1 Here]

The results in Panel B particularly suggest that it is difficult to predict different paths to success based on the patent characteristics at the time of patenting and that the stock market reactions are also futile in identifying the initially neglected patents that eventually become a great success. These results offer a valuable insight that the path to eventual innovation success is determined by the external users that make use of (*i.e.,* cite) these patents. This motivates us to further examine those users and compare citing patents' characteristics between early-bloomer and late-bloomer patents.

Panel C presents the descriptive statistics of citing patents. Compared to the focal patents in Panels A and B, we first find that citing patents appear to be less successful than focal patents in terms of the number of forward citations that they receive. In contrast, the citing patents make a substantially broader search of patents, *e.g.,* significantly more backward citations. It is possible that this result is partially driven by the fact that those citing patents relative to focal patents are granted later in time, which increases the size of the entire patent pool for backward citations. However, we find that within citing patents, those that cite late-bloomer patents particularly make a larger number of backward citations in comparison to those that cite early-bloomer patents. In our analyses later, we further focus on this aspect (*i.e.,* the breadth of users) of late-bloomer patents to investigate the distinctive process of once-neglected inventions developing into one of the most successful innovations.

6 Results

6.1 Late-bloomer Patent Writers and Users

We begin our analysis by contrasting firms that write late-bloomer patents and firms that use those late-bloomer patents. Because firms can both write and cite late-bloomer patents, we define users of late-bloomer patents to be more exclusive as firms that have never written a late-bloomer patent during the sample period while they cite at least one. Writers of late-bloomer patents exclusively for this analysis only are defined as firms that have at least one late-bloomer patent during the sample period regardless of whether they have ever cited late-bloomer patents of other firms. The first thing we note from the list of writers and users in the sample is that the writers appear to be much older (*e.g.,* U.S. Surgical Corp, Johnson & Johnson, and IBM) than the users, *i.e.,* log(age) of writers is 2.17 whereas that of users is 1.56. Since the age gap can drive large differences in many financial variables mechanically regardless of the writer or user identity, we report summary statistics using a more refined age-matched sample. Specifically, the users in the refined sample are the five nearest neighbors of each writer. Panel A of Table 2 presents the summary statistics.

[Insert Table 2 Here]

We first examine various patenting characteristics. We find that the average number of granted patents in a year is significantly larger for writers than users (30 vs. 3). Furthermore, the average number of patents in a year that eventually become superstar patents is tenfold greater for writers than users (5 vs. 0.5). By construction, users have no late-bloomer patents while writers generate 0.6 late-bloomer patents per year, and 13% of their patents are identified as late-bloomer patents. Regarding patent impacts, we find that late-bloomer writers receive 78 citations (64 after netting self-citations) per year while users receive 5 (4 external) citations per year. Considering the fact that writers file significantly more patents per year, we normalize the citation counts per year by the total number of patents per year. We still find that writers' normalized citation counts are greater, *i.e.,* 2.7 for writers vs. 1.6 for users, and the difference is statistically significant. All these results collectively suggest that late-bloomer writers are better at patenting and innovation, both in quantity and quality. Despite this conclusion, we find an interesting point regarding patent claims and commercialization. We note that the average number of claims is larger for users than writers (17.4 vs. 16.6), and the difference is statistically significant. The larger number of patent claims implies a broader patent applicability, and, hence, a higher chance of bringing the invention to markets as a new product. For a proxy for patent commercialization, we use the total number of new products (Mukherjee et al., 2022) per year divided by the total number of new patents per year. Based on this measure, we find that the commercialization rate is significantly higher for users relative to writers (26% vs. 18%). Both results are consistent with the interpretation that late-bloomer users particularly stand out as market players that are relatively more capable of commercializing new inventions.

We then compare financial variables between writers and users. We confirm that the average age of firms is 5 years in both groups and no longer shows the big age gap after the nearest neighbor matching. However, even after the age matching, we find that the firm size, measured by all aspects including book assets, market assets, and total revenues, is significantly larger for writers than users. Consistent with this result, we find that CAPX investment rates are higher for writers than for users and that writers are more likely to be mature firms that pay dividends to shareholders. We also note that writers invest more in R&D than users with the investment rate difference at around 1.6%. However, we highlight that users also invest a fair amount in R&D, indicating they are still innovative firms but differ in the way of doing innovation. We do not find any other differences in the remaining financial variables in the table including profitability, leverage, and advertisement. Lastly, we use the 2002 Input-Output Accounts from the Bureau of Economic Analysis (BEA) and create a variable for industry consumer dependence. We define an industry as more consumer-dependent if the industry's production percentage for "Personal consumption expenditures" in the BEA Input-Output Accounts is in the top tercile. We find that user firms are significantly more likely to be in consumer-dependent industries.

In sum, writers of late-bloomer patents are big value firms with a bigger stock of patents and citations and greater R&D spending. Writers appear to have enough resources given their size and investment scales but do not commercialize every good innovation they generate. Conversely, users of late-bloomer patents are relatively smaller in size but generate a lot more new products per patent, and they are more likely to serve consumers directly with their products.

6.2 Persistence in Being Late-bloomer Writers or Users

We note that our analysis in Panel A of Table 2 does not allow for switching between the two groups of writers and users by defining writers and users only cross-sectionally for the entire sample period. In this section, we relax our classification scheme and allow firms to switch between the writer and user groups. Panel B of Table 2 presents a transition matrix where we examine the likelihood of firms changing their writer or user identity in the next period. We use the sample of public firms that have ever written or cited a late-bloomer patent during our sample period. We define four sets of write/user status. Strict Writer is a firm that produces a late-bloomer patent based on its grant year but does not cite any late-bloomer patent in that year. Flexible Writer/User is a firm that produces a late-bloomer patent based on its grant year and also cites late-bloomer patents in that year. Strict User is a firm that cites late-bloomer patents in a given year but does not produce any late-bloomer patents in that year. Lastly, Idle indicates a firm that neither produces a late-bloomer patent nor cites late-bloomer patents in a given year. We present both the number and percentage of observations for each category.

First, we find that the number of Strict Writers is substantially small in any given year in Row (a). At a given year, only 5.7% (861 out of 14,946) firm-years are considered as Strict Writer state. The likelihood that a Strict Writer produces another late-bloomer patent in the next period (*i.e.,* stays as a Strict Writer or becomes a Flexible Writer/User) is approximately 30% (=13.12% + 17.19%). A Strict Writer can also become a Strict User or do nothing in the next period with the probabilities of 21% and 48%, respectively. Row (b) shows that the number of Flexible Writers/Users is three times larger than that of Strict Writers. A Flexible Writer/User's transition likelihood to a Strict Writer is 1.65% while 61% of the Flexible Writers/Users stay as a Flexible Writer in the next period. Jointly, any late-bloomer writers including both Strict and Flexible Writers are more likely to be late-bloomer writers in the next period again with a likelihood of 47%. In Row (c), we find that there exist 4.6 times more Strict Users than Strict Writers. Patent hunting appears to be a very persistent firm characteristic as 51% of the current Strict Users continue to be Strict Users in the next year.

Overall, the transition matrix results in Panel B of Table 2, collectively with the descriptive statistics in Panel A of the same table, suggest that late-bloomer writers and users are originally different in terms of their innovation styles. These results do not appear to support the idea that young firms start as users when resources are constrained but evolve to eventually become writers when the abundance of resources can sustain more innovative activities. Thus far, existing studies in the literature mainly focus on patent writers as a whole and their innovation outcomes, but we uniquely highlight in this paper the important role of patent users (*i.e.,* hunters) who exhibit a separating equilibrium from writers.

6.3 Benefits of Patent Hunting

6.3.1 Firm-level Analyses

As discussed in the previous section, late-bloomer patent writers and users have distinct innovation styles and firm characteristics. In this section, we examine the incentives of patent hunting relative to originating completely new ideas. Before going deeper to compare patent writers and users at the patent level, we first consider users' firm-level analyses to gain a clear insight into the firm benefits of patent hunting. We estimate the following regression model:

$$Y_{j\{t,t+4\}} = a + b_1 log(1 + LBhunting)_{jt} + b_2 log(1 + EBhunting)_{jt} + \gamma_j + \eta_t, \qquad (1)$$

where j is user firm and t is year. $log(1 + LBhunting)_{jt}$ is the log of one plus the total number of late-bloomer patents that user firm j cites in year t. To compare the benefits of citing late-bloomer patents to those of citing early-bloomer patents as the "counter-factual" patents with a similar impact, we also include $log(1 + EBhunting)_{jt}$ in the model. $log(1 + EBhunting)_{jt}$ is the log of one plus the total number of early-bloomer patents that user firm j cites in year t. The dependent variable for firm benefits is sales growth or firm value as measured by Tobin's Q in the subsequent five years. $Sales\ growth_{j\{t,t+4\}}$ is computed as $(sales_{j\{t+4\}}/sales_{j\{t\}}) - 1$, and $Avg\ Tobin's\ Q_{j\{t,t+4\}}$ is the five-year firm average of Tobin's Q. The regression is at the user firm-year level and includes firm fixed effects (γ_j) and year fixed effects (η_t). Standard errors are clustered at the firm level. Table 3 presents the results from the estimation of Eq.(1).

[Insert Table 3 Here]

We find in Columns 1 and 2 that the number of citing late-bloomer patents is associated with higher future sales growth and firm value. The coefficients translate into a 6.4% increase in sales growth and a 2.2% increase in Tobin's Q by doubling the number of late-bloomer patent citations. In contrast, we find coefficients on $log(1 + EBhunting)$ in both columns are significantly negative. While it is highly probable that a firm cites both late-bloomer and early-bloomer patents in a year or even within a single patent, the incremental benefits of using early-bloomer patents do not seem as significant as those of late-bloomer patents.

In Columns 3 and 4, we examine the differential benefits of citing late-bloomer patents early vs. later by replacing $log(1 + LBhunting)$ with $log(1 + earlyLBhunting)$ and $log(1 + laterLBhunting)$. $earlyLBhunting$ is the number of late-bloomer patents a user cites, where the citation was one of the first three on a given late-bloomer patent since its grants and prior to its superstar-year. $laterLBhunting$ is the number of the rest of late-bloomer patents that are not early-hunting. We find that the benefits of early hunting are 1.1 times and 1.5 times larger for sales growth and firm value, respectively, than those of later hunting. These findings are consistent with the interpretation that earlier users that possibly discover those neglected inventions reap greater benefits relative to followers.[5]

Overall, the findings from our user firm-level analyses in this section show that the rents to patent hunting are substantial in terms of future firm growth and value. The rents appear to be exclusive to hunting late-bloomer patents and do not extend to hunting early-bloomer patents.

6.3.2 Patent-level Analyses

We now attempt to refine our understanding of the benefits associated with patent hunting through detailed patent-level analyses. In the following regression analyses, we compare patent writers and users across three types of focal patents: late-bloomer, early-bloomer, and non-superstar patents. We consider a difference-in-differences analysis that exploits the

[5]In Appendix Table A2, we consider analogous tests for patent writers. In the first two columns, we examine the association between writer firms' financial benefits and the total number of late bloomers produced in a year. In the last two columns, we present the corresponding results of users as in the first two columns of Table 3 for a comparison purpose. The coefficient estimates in all columns are standardized for ease of comparison. In Columns 1 and 2, we find writing late-bloomer patents is positively and significantly related to writers' future firm value while there is no relation to sales growth. However, such benefits are smaller than those of users shown in Columns 3 and 4, and the firm value increase is only about one-third of writing early-bloomer patents based on the coefficients on $log(1 + EBwriting)$ in Column 2.

differences in patent writers and users, particularly around each patent's peak impact year as measured by the number of cumulative forward citations. Specifically, we estimate the following model:

$$Y_{ijpt} = a + b_1 user_{ijp} + b_2 ssyear^{post}_{pt} + b_3 user_{ijp} \times ssyear^{post}_{pt} + \gamma_{ij} + \eta_t, \tag{2}$$

where j is user firm, p is focal patent created by writer firm i, and t is year. $user_{ijp}$ is an indicator that is one for the user firm j of the patent p created by firm i and zero for writer firm i of the same patent. We consider writer-user pairs within the 20-year period since the focal patent's grant. The dependent variable for firm benefits is sales growth or firm value as measured by Tobin's Q. For a superstar patent, $ssyear^{post}_{pt}$ is one if t is after the year when the patent p becomes a superstar patent (*i.e.,* reaches the top 5% of the cumulative citation distribution within the same grant-year and CPC class) and zero otherwise. For a non-superstar patent, $ssyear^{post}_{pt}$ is one if t is after the peak cumulative forward citation year and zero otherwise. The regression is at the focal patent-firm-year level and includes firm pair fixed effects (γ_{ij}) and year fixed effects (η_t). Standard errors are clustered at the focal patent level. Table 4 presents the results from the estimation of Eq.(2).

[Insert Table 4 Here]

Columns 1 and 2 present results for the late-bloomer patents. We find that late-bloomer users, on average, have larger sales growth and firm value relative to writers. These results are consistent with our findings in Panel A of Table 2 that late-bloomer writers are bigger and older value firms relative to late-bloomer users. The main variable of interest in this analysis is the interaction term between $user$ and $ssyear^{post}$. We find that the coefficient estimates for the interaction term in both columns are positive and statistically significant at the 1% level. The coefficients translate into a 22.4% increase in sales growth and a 2.9% increase in Tobin's Q for users, compared to their pre-superstar-year mean. These results indicate that the incremental benefits in sales and firm value after the superstar-year are significantly larger for late-bloomer users than those for late-bloomer writers.

We do not find similar patterns on early-bloomer patents in Columns 3 and 4. There are negative benefits in sales and Tobin's Q for the users of early-bloomer patents relative to writers after the superstar-year although the effects are close to zero in magnitude. For non-superstar patents in Columns 5 and 6, we similarly find no user benefits relative to writers after the superstar-years. The interaction term coefficient estimate is statistically insignificant for the sales growth and significantly negative for Tobin's Q.

In Appendix Table A3, we examine patent hunting benefits based on alternative regression specifications for robustness. The dependent variable is the difference in the outcome variable between firms using and writing a focal patent. In Panel A, we use early-bloomer patents as a benchmark group and find consistent results with Table 4 that the user benefits are significantly larger for hunting late-bloomer patents relative to hunting early-bloomer patents post-superstar year. In Panel B, we use a different benchmark group that comprises non-superstar patents and find that the positive late-bloomer hunting benefits are robust in comparison to non-superstar patents. In Appendix Table A4, we further examine whether the user benefits are prevalent among any superstar patents given their extraordinary success. To examine this possibility, we alternatively compare early-bloomer patents and non-superstar patents (Panel A) and superstar patents and non-superstar patents (Panel B). We find that the user benefits of hunting early-bloomer patents or superstar patents relative to those of non-superstar patents are statistically indistinguishable from zero. The results in Appendix Tables A3 and A4 strongly support our conclusion that there are unique user benefits in hunting late-bloomer patents.

We also graphically illustrate the patent hunting benefits in Figure 3. We plot sales and firm values of early-bloomer and late-bloomer users. The observation window is [-10, +10] around the focal patent's superstar-year, and we consider financial data in the grant year of each user patent (*i.e.,* citing year).[6] In Panel (a), we find that sales start similarly for both

[6] Appendix Figure A3 presents alternative figures using financial data cumulatively since the citing year. We find qualitatively similar results.

early-bloomer and late-bloomer users but diverge significantly three years before the focal patent's superstar-year. The gap does not revert in 10 years after the superstar-year. In Panel (b), we present an analogous figure for early-bloomer and late-bloomer writers. We find that the early-bloomer writers' sales grow much faster than those of late-bloomer writers, and the early-bloomer writers' sales growth notably increases right before their patents' superstar-years. In contrast, we do not observe any notable changes in sales growth for late-bloomer writers around their patents' superstar-years. We also point out an important finding that late-bloomer writers' sales also keep increasing over time. This implies that writers still retain some benefits by producing late-bloomer patents despite the fact that their users enjoy more of those benefits.

[Insert Figure 3 Here]

In Panel (c), we compare the firm value of early-bloomer and late-bloomer users and find consistent, and even stronger, results with Panel (a). In Panel (d), we compare the firm values of early-bloomer and late-bloomer writers. Consistent with the sales growth result in Panel (b), we find that late-bloomer writers' firm value increases after their patents' superstar-years. This indicates that producing late-bloomer patents creates firm value for late-bloomer writers. Interestingly, the firm value of early-bloomer writers increases before their patents' superstar-years but drops right after the superstar-years. This result conveys the interpretation that markets for early-bloomer patents might already exist before their impact peaks, and, thus, the utilization of those patents drops after the peak.

Collectively, the findings from our patent-level analyses in Table 4 and Figure 3 show that the rents to patent hunting are significant and a sizable portion of the entire innovation chain as the user benefits, on average, exceed the original patent writers' benefits.

6.4 Late-bloomer Patents and Creation of New Markets

6.4.1 Patent Applicability

In this section, we focus on where the benefits to late-bloomer users come from. To understand the mechanism for the patent-hunting benefits, we first delve into how late-bloomer patents are used by examining user patents' technologies. We consider the technology proximity between a focal patent and its citing patent. For a technology proximity measure of two patents, we use the cosine similarity between the CPC class vectors of the two patents.[7]

In Figure 4, we present how technology proximity between a patent and its citing patents, on average, changes over time. Panel (a) shows the changes in the technology proximity between an average non-superstar focal patent and its citing patents over the focal patent's life. As the figure clearly shows, the focal patent is initially cited by patents that are very close in technology but, over time, gets cited by technologically more distant patents. The average proximity between a focal patent and its citing patents is 0.67 right after the focal patent's grant but it drops by 25% to 0.5 when the focal patent's term ends in 20 years. This implies that innovation is extensively used by more focused users initially but applies more broadly to innovation in relatively far fields over time (Kuhn, Younge, and Marco, 2020).

[Insert Figure 4 Here]

In Panel (b) when we consider superstar patents, we find stark differences in the patterns between early bloomers and late bloomers. While early-bloomer patents follow the same trend that we find in Panel (a) for non-superstar patents, late-bloomer patents deviate from the normal trend. First, technology proximity between a late-bloomer patent and its users starts significantly lower at around 0.55 compared to 0.6 for early-bloomer patents (0.67 for normal patents), and this gap does not converge afterward. Second, the technology

[7] We also consider an alternative technology-proximity measure for two patents using the distance between their CPC classes (class-to-class proximity). The distance between CPC classes is computed using the vector of how many patents with other CPC classes cite the patents in a given CPC class. We find results are robust to using this alternative proximity measure.

proximity of late-bloomer patents drops in the beginning similar to any other patents but the decline discontinues later in time. Late-bloomer patents' technology proximity to their users notably stabilizes around the superstar-year and remains flat at approximately 0.46 on average thereafter. To assist in interpretation, we show the figure in Panel (b) with different presentation methods in Panels (c) and (d). In Panel (c), we index the level of proximity by setting the average proximity in the superstar-year as 100 and show the extent of deviation from the superstar-year proximity level. In Panel (d), we consider a detrended technology proximity between a patent and its citing patents which is the difference between the actual technology proximity value and the predicted value from its regression on the time trend based on the data of pre-superstar years. Both figures support our interpretation that late-bloomer patents' technology proximity to their users increases significantly after their superstar-years relative to their own predicted levels or compared to those of early-bloomer patents.

We reinforce this interpretation with an additional analysis in Table 5. We predict that when late-bloomer patents create their own fields (markets), the number of subsequent patents in those newly created fields will increase significantly. We test this prediction using the number of subsequent patents in a focal patent's CPC classes (Column 1) and the major overlapping CPC class pair among citing patents of the focal patent (Column 2). We consider only superstar patents and regress one of the measures of the subsequent patent count on an indicator for late-bloomer patents. If a given CPC class or a given major overlapping CPC class pair among citing patents is shared by both early-bloomer and late-bloomer patents, we drop the CPC class or the major overlapping CPC class pair of citing patents.

[Insert Table 5 Here]

We find in the first column that the number of subsequent patents in the same CPC class of an early bloomer declines significantly after its superstar-year based on the coefficients for standalone $ssyear^{post}$. This result is consistent with the interpretation that the technology field of early bloomer patents generally peaks before their superstar-years. In contrast, late-bloomer patents' superstar-years are associated with a 16% increase in the number of subsequent patents in the same CPC class. We note that a new field of technology can be created from the late-bloomer patent and does not have to be in the same CPC class as that of the late-bloomer patent. Therefore, in the second column, we use an alternative definition of a new field by considering exhaustive pairs of all CPC classes reported by the citing patents of a focal patent and identifying the most frequent pair. Then, we count the number of subsequent patents whose CPC classes include the most frequent pair of user CPC classes. With this alternative measure, we find similar results that overlapping technology fields among the patents citing late-bloomers show a significant increase in the number of subsequent patents after the late-bloomer patents' superstar-years. The effect with this alternative measure is estimated at around 4%.

6.4.2 Commercialization

Next, we investigate new market creation by late-bloomer patent users. We use a measure of new product launches from Mukherjee et al. (2022) based on the search of media articles mentioning new product introductions. We set the number of new products to zero when there is no media announcement of an important product launch. We then estimate Eq.(1) and Eq.(2) by replacing the dependent variable with the log of one plus the total number of new products in a given year. We present the results in Table 6. Column 1 shows the results from estimating Eq.(1). Columns 2 and 3 compare late-bloomer and early-bloomer patents, and Column 4 shows results for non-superstar patents.

[Insert Table 6 Here]

First, we find from the firm-level analysis in Column 1 that patent hunting is significantly and positively associated with new product launches in the subsequent five years since hunting. The coefficient on $log(1 + LBhunting)$ translates into a 2.2% increase in the number of new product launches when doubling the number of late-bloomer patent citations. We also find that the number of early-bloomer citing is related to an increase in new product launches during the next five years, but the effect is one-half compared to the effect for late-bloomer citing.

Results reported in Column 2 show that late-bloomer users have more new product launches than late-bloomer writers in general and that their superstar-years are particularly associated with about 5 percentage-point incremental increase in new product launches. However, we do not observe that late-bloomer writers show similar increases in new products after their superstar-years based on the significantly negative coefficient estimate for standalone $ssyear^{post}$. Column 3 shows the results for early-bloomer patents. The coefficient for standalone $ssyear^{post}$ is significantly negative indicating that early-bloomer writers also face a decrease in the number of new products after their superstar-years. However, the interaction term between $user$ and $ssyear^{post}$ is statistically insignificant and close to zero. Lastly, when we examine non-superstar patents in Column 4, users have significantly fewer new product launches than writers regardless of the superstar-year timing.

We note that the benefits of commercialization as measured by new product launches generally accrue to the writer of a patent based on the results for non-superstar patents. Importantly, the results in Table 6 provide strong evidence that late-bloomer patents are distinguished by their users (not writers) that appear to reap greater benefits from developing new markets. Such user benefits from creating new markets likely explain the greater sales growth and firm value for late-bloomer users relative to writers.

7 Mechanism and Equilibrium

7.1 Late-bloomer Writers' Incentives

Thus far, our results suggest that late-bloomer patents are unique in that they provide greater benefits to users than writers and that the benefits are likely associated with the creation of new markets and commercialization. If so, one might argue that only patent hunters should exist in an equilibrium. This argument also implies that late-bloomer writers have no incentives to create a patent that benefits other firms more. In this section, we investigate possible reasons why late-bloomer writers still produce late-bloomer patents.

First, we examine the characteristics of late-bloomer patents within their writers' patent portfolios. To do so, we regress an indicator variable for a late-bloomer patent on several patent and firm characteristics at the time of patent grant. For those characteristics, we particularly focus on the following three constraints under which writers may optimally neglect their ideas and patents: (i) capacity constraints, (ii) competitive threat, and (iii) financial constraints. For measures of capacity constraints, we consider *tech-class weight* and *tech-class dist to core*. *tech-class weight* is the fraction of the writer's patents in a specific CPC class of a given patent over the entire sample period. This measure captures the importance of a particular technology class to the patent's writer. *tech-class dist to core* is the class-to-class proximity between the CPC class of a given patent and the core CPC class of its writer. The core CPC class is the CPC class with the highest *tech-class weight* within a firm. The class-to-class proximity measure is the distance between two CPC classes computed using the vector of how many patents with other CPC classes cite the patents in a given CPC class. For a measure of competitive threat, we use *log(competing patent stock)* which is the log of the total number of patents from U.S. public firms with the same CPC class up to the grant year of a given patent. This measure captures how many other players exist in the same technology space at the time a given patent is produced. Lastly, for financial constraints, we consider the KZ index (Kaplan and Zingales, 1997), firm size and age following Hadlock and Pierce (2010), and the textual analysis-based financial constraints measure by Linn and Weagley (2021).[8] In addition to these constraint measures, we control for firm size, age, profitability, CAPX investment, and R&D investment. The regression is at the patent level (one observation per patent) by taking the averages of relevant variables

[8]Hoberg and Maksimovic (2015) are the first attempt to develop the textual analysis for measuring financial constraints. However, we utilize Linn and Weagley (2021) due to our extended sample period. Linn and Weagley (2021) estimate a mapping between firm-level accounting variables and financial constraints using random decision forests, a machine learning algorithm, for the period from 1972 to 2021.

when a patent has multiple CPC classes and using the grant-year data when control variables are time-varying. The regression includes both the writer and grant year fixed effects. Table 7 presents the results.

[Insert Table 7 Here]

In Columns 1 and 2, we use *tech-class weight* as a measure for intellectual capacity constraints. In Columns 3 and 4, we use *tech-class dist to core* alternatively. Columns 1 and 3 (2 and 4) consider the KZ index (the Linn and Weagley (2021) measure) for firm financial constraints in a given year. We also show the coefficient estimates for firm size and age in all columns as measures for financial constraints following Hadlock and Pierce (2010). Throughout the columns, we find that a patent is more likely to become a late-bloomer when the idea of the patent is not in the primary technology space of the writer or is more distant from the core technology space. Also, when there are fewer competing patents in the same technology space in the past (*i.e.,* the idea of a patent is relatively new), the patent is more likely to become a late bloomer. These findings suggest that writers have lower incentives to rapidly commercialize a new idea when the development cost is high—due to increased learning or opportunity costs—or when they face a minimal competitive threat. Interestingly, our analysis indicates that financial constraints do not significantly influence the timing of late-bloomer patent commercialization. This is evidenced by the negative coefficient estimates for both financial constraint measures, particularly the significantly negative coefficient for the KZ index (or the positive coefficient for firm size). These results are in line with the notion that firms producing late-bloomer patents possess sufficient financial flexibility to invest in innovations that do not demand immediate attention. Alternatively, it is possible that late-bloomer patents represent ideas that their writers have attempted to commercialize but ultimately did not succeed.

The insights gleaned from our analysis in this section provide innovative perspectives on the deliberate sidelining of certain ideas by writers of late-bloomer patents. The example of Texas Instrument and Nvidia, as discussed in the introduction, aptly encapsulates these nuanced insights.

7.2 Moderating Factors of Patent Hunting Benefits

In this section, we explore aspects of the patent hunting process that make patent hunting uniquely valuable - or less valuable - in certain states. We focus specifically on two forces: 1.) how the complexity of the patent itself - and thus likely the cost of processing and integrating it into own-innovation - impacts patent hunting benefits; and 2.) how competition intensity for a given patent impacts the end-benefits to a hunter. Given that complexity is likely unknown ex-ante (before searching), one might expect that ex-post complexity could decrease the end total benefits through the costs of processing and integration. Similarly, as competition for a given technology space is somewhat uncertain ex-ante, if ex-post competition is more intense (i.e., there are more firms vying to integrate this idea), this might be expected to decrease the total size of the pie to any given hunter.

In Table 8, we explore both of these dimensions. The sample for this analysis consists of all public firms that use late-bloomer patents. The dependent variables are late-bloomer users' next five-year sales growth, computed as $(sales_{t+4}/sales_t) - 1$ since the citation of the late-bloomer patents, and next five-year average Tobin's Q.

[Insert Table 8 Here]

In Columns 1 and 2, we consider search costs using the Gunning fog index as a proxy for the complexity of patent text. The index quantifies document readability, with higher values indicating greater complexity. Consistent with the above dynamics, the results in Columns 1 and 2 suggest that the benefits of using late-bloomer patents decline with the complexity of patent. This effect is particularly pronounced for firm value, and negative in point-estimate, but not statistically significant, for sales growth.

Columns 3 and 4 explore competition dynamics using the number of competing patent hunters in the same technology class. In both columns, patent hunter benefits decrease significantly - in both firm value and sales growth - when competition is more intensive to exploit similar technologies.

Overall, the results in this section suggest that the benefits to patent hunting do vary. In particular consistent with the search cost model framework in Section 3, higher costs of patent hunting reduce the net benefits associated with the search strategy of hunting.

7.3 Skills and Deliberateness of Patent Hunting

In this section, we investigate the deliberate nature of patent hunting. One might be concerned that patent hunters such as Nvidia possess inherent capabilities and are thus predisposed to pioneering new markets with exceptional innovations even without citing late-bloomer patents. If this is the case, citing late-bloomer patents could be a mere formality for patent hunters, representing a courteous acknowledgment of earlier work. It is also plausible that the recognition of earlier work is not voluntary but rather compelled by examiners.

We address these concerns based on the following three approaches. First, we examine the learning component of patent hunting. We expect that patent hunting is indeed a skill, not a mere formality, if more experienced patent hunting leads to greater rents. Second, we examine the technology distance between patent hunters' own patents and the late-bloomer patents that they cite. We expect that patent hunting is more technologically meaningful if the distance to cited late-bloomer patents is smaller than the distance to other cited patents. Third, we examine in-text mentions of late-bloomer patents and examiner-added citations. We expect that late-bloomer patents are deliberately studied by inventors if they are mentioned in the text part of patents and not added by examiners. Table 9 presents the results of these analyses.

[Insert Table 9 Here]

In Panel A, we compare the financial benefits of patent hunting by experienced and less-experienced patent hunters. We consider the specifications in the first two columns of Table 4 for subgroups of experienced users (Columns 1 and 2) and less-experienced users (Columns 3 and 4). Experienced users are the firm-years in the top 10% distribution in the firm-level average of the late-bloomer fraction among all cited patents in the past 5 years from a citing year. The rest of the firm-years are less experienced users. We find, based on the coefficient estimates on the interaction term between $user$ and $ssyear^{post}$, that financial benefits are significantly greater for experienced users in both sales growth and firm value. For experienced users, sales growth benefits are four times larger, and Tobin's Q benefits are significantly positive while they are negative for less-experienced users. These results support our interpretation that patent hunting is firm skills that have a learning component.

In Panel B, we analyze the technological proximity between patent hunters' own patents and cited late bloomers compared to other patents. If citing a late-bloomer patent lacks intentional consideration, we expect to observe no significant difference in the distance between hunters' own patents and late-bloomer patents compared to other patents cited. However, Column 1 shows that the distance to cited late bloomers is notably closer than to other cited patents. We find that late bloomers are much closer to hunters' own patents than non-superstar patents in Column 3, albeit slightly farther than cited early bloomers in Column 2, with the difference being significantly smaller than that in Column 3.

Lastly, in Panel C, we examine (i) the likelihood of being mentioned in the text part of citing patents, (ii) the number of text mentions, (iii) the sentiment of neighboring words around text mentions, and (iv) the likelihood of being cited by examiners. We compare these statistics of cited patents across late-bloomer, non-late-bloomer, and early-bloomer groups. In Column 4, we find that cited late-bloomers, when compared to cited non-superstar patents, are more likely to be mentioned in the text part of hunters' own patents, have more positive sentiment when mentioned in the text part, and are less likely to be cited by examiners. In Column 5, however, cited early-bloomer patents are more likely to be mentioned in the text part than cited late-bloomer patents while the sentiment around the text mentions of early-bloomer patents is more negative compared to that of late-bloomer patents.

In Figure A4, we explore further dimensions of cited late-bloomer patents including their overall impact on the broader innovation landscape after being cited and the impact of patent hunters' own patents after citing them. Panel (a) plots the growth in forward citations that late bloomers, early bloomers, and non-superstar patents receive after they are cited by patent hunters. If patent hunters' citation of a late bloomer merely served to acknowledge an earlier work, we would not observe notably distinct subsequent impacts on the entire innovation landscape compared to other cited patents. The figure shows that late bloomers experience a 20% increase in forward citations on average during the five years after patent hunters cite them. In stark contrast, early bloomers exhibit a decline in citation growth, with non-superstar patents showing an even more pronounced negative trajectory. These findings suggest that citing late bloomers carries distinctive importance and should not be equated with citing other patents. Panel (b) plots the subsequent citations of patent hunters' own patents, distinguishing between their organic patents and those hunting late bloomers. Organic patents are defined as patents with greater backward self-citations. The figure shows that patents hunting late bloomers in their backward citations receive markedly more forward citations over the following ten years since their grant. These findings imply that patents hunting late bloomers exhibit significantly greater influence across the wider innovation landscape than patent hunters' own work without late bloomers.

Taken together, the findings in Table 4 and Figure A4 provide strong evidence that patent hunters intentionally incorporate late-bloomer patents, resulting in a substantially greater influence across the broader innovation landscape and such deliberate patent hunting skills improve with experience. These results contradict the notion that citing late-bloomer patents is merely a procedural courtesy for patent hunters to acknowledge prior work.

7.4 Hunter Firms or Hunter Inventors

We recognize that inventors can play important roles in patent hunting. In this section, we investigate the role of inventors. In particular, we examine whether hunting benefits will be larger when the inventors of a late-bloomer patent join a firm that uses the late-bloomer patent. This test aims to answer if patent hunting is initiated by inventors who move from one firm to the other. We also examine whether inventors who once used late-bloomer patents in the past are more likely to use late-bloomer patents in the future regardless of who they work for. If this is the case, patent hunting can also be inventor traits in addition to firm attributes.

We present in Table 10 the results for the test of whether hunting benefits will be larger when the user firm hires the original inventors of the late-bloomer patents. We extend the regression specification in Eq.(2) to include a triple interaction term with an indicator for inventor moves. We only focus on late-bloomer patents as in Columns 1 and 2 of Table 4. The variable, *inventor move*, is one if the actual inventor of a given patent joins the firm that cites the patent and zero otherwise. On average, inventors of 3.4% of late-bloomer patents move to user firms of their late-bloomer patents in the future. In both Columns 1 and 2, we find that the user benefits in sales and firm value after the superstar-year of a late-bloomer patent significantly intensify when the user firm hires the inventors of the late-bloomer patent. When inventors are not shared by writers and users, the user benefits are still economically and statistically significant. However, when inventors are shared, the benefits are 3-4 times larger. Additionally, in Columns (3) and (4), we find that the incremental benefits are even larger when we use the sub-sample of inventors that move before their patents become superstar patents, allowing the new firm to reap all the benefits from the late-bloomer patents.

[Insert Table 10 Here]

While the results thus far support that there are patent-hunting firms and the benefits of late-bloomer patents primarily accrue to them, it could be the case that our results are driven by inventors who actively engage in patent hunting. Thus, we turn our focus to individual inventors and examine whether an inventor who patent hunts in the current employer is also more likely to patent hunt in her subsequent employers. If we find a positive association between individual inventors' hunting behaviors in current and next firms that the inventors work for, that will support the conclusion that patent hunting is partly driven by individual inventors as well. We explore this possibility in Table 11.

[Insert Table 11 Here]

In the first two columns, we only consider an inventor's subsequent employer. In the last two columns, we take the average of up to three subsequent employers. The dependent variable in Columns 1 and 3 is an indicator for whether the inventor is involved in any patent hunting during the subsequent employment, while the dependent variable in Columns 2 and 4 is the average number of late-bloomer patents that the inventor cites during the subsequent employments. We note that we take into account the inventor's general citing behavior by controlling for the use of superstar patents that include both early-bloomer and late-bloomer patents in addition to the individual characteristics including gender, total number of invented patents, and total number of employments during the sample period, and the current employer's characteristics. The variables of interest are the indicator for whether the inventor's current employment is involved with any patent hunting and the total number of late-bloomer patents that the inventor's current employment has used.

Throughout all four columns, we find that patent hunting of an inventor in her current employment is associated with patent hunting of the inventor in subsequent employments positively and significantly. The estimated association is at around 7% for the extensive margin and 14-16% for the intensive margin. We also find that when the inventor cites superstar patents with the current employer, the inventor is also more likely to cite late-bloomer patents with subsequent employers. However, this association is much weaker than the association with the late-bloomer use at one-fourth in the magnitude in Columns 1 and 3, for example.

Overall, these results are consistent with two different interpretations. One is that patent hunting is individual inventors' skills, and the other is that patent hunting firms are more likely to hire inventors from another patent hunting firm. Although we cannot confidently determine which of these two explanations dominates the other, the results in the inventor-level analyses highlight the important role of individual inventors who can help identify late-bloomer patents or add patent-hunting skills to their firms.

7.5 Identification Strategy for Causal Interpretation

A central finding thus far is that late-bloomer patent hunting appears to promote sales growth of user firms by facilitating successful commercialization of inventions in new markets. However, one might be concerned that this could be driven instead by some unobserved firm characteristic affecting both sales growth and patent hunting behavior, or even perhaps that expected future growth in sales or firm value (that on average materialized) prompted and allowed slack to hunt for neglected ideas. To mitigate such endogeneity concerns, we further consider an identification strategy in this section based on a novel instrumental variable (IV) approach that exploits our findings on hunting inventors in Section 7.4.

The results in Section 7.4 suggest that late-bloomer hunting is, in part, an inventor trait that persists across different employers. We utilize patent-hunting inventor moves forced by firm bankruptcies as our instrument. However, as firm-level bankruptcies of neighboring firms might have an impact on firms' sales through other channels (e.g., increasing the competitive advantage of firms now facing reduced local competition), we construct a very fine specification on which to instrument patent-hunting variation. In particular, we only compare firms that both have nearby firms going bankrupt. We then use the relative level of patent hunting intensity at the respective bankrupting firms as the varying treatment instrument for differential shock to patent hunter supply of the pair of treated neighboring firms. This then orthogonalizes against the bankrupt neighbor exclusion restriction concern, and solely focuses on patent hunting intensity differences for identification.

In this IV analysis, we thus only consider firms in our sample located within a 100-mile radius of bankrupt neighbors and examine inventor spillovers from the bankrupt neighbors to nearby firms. The idea behind this approach is that when a neighboring firm goes bankrupt, inventors from the bankrupt firm are highly likely to join nearby firms and continue hunting late-bloomer patents at the new firms if they have previously engaged in such activities as shown in Table 11.[9] Specifically, our instrument is the hunting intensity of the bankrupt neighbor *prior to* bankruptcy. We expect a positive correlation between the bankrupt neighbor's hunting intensity and the subsequent hunting activities of nearby firms, given that inventors from bankrupt firms often move to nearby firms. Conversely, we do not expect the bankrupt neighbor's hunting intensity to be correlated with the *future* sales growth of nearby firms, except through the effects of inventor moves.[10] We present our IV analysis results in Table 12.

[Insert Table 12 Here]

We measure the bankrupt neighbor's hunting intensity in two different ways. The first two columns use the average of inventor-level hunting intensities within firm as the instrument, while the last two columns use the firm-level fraction of hunting inventors as the instrument. The bankrupt neighbor's hunting intensity is calculated using the past three-year moving average considering that inventors may not file patents every year. We then estimate the following first-stage regression for each IV and present the results in the odd-numbered columns:

$$log(1 + LB\ hunting)_{jkt} = a + b_1\ bankrupt\ neighbor\ hunting\ intensity_{jk\{t-3\}} + \gamma_j + \eta_t, \quad (3)$$

[9]In support of this, we find that approximately 20% of inventor moves occur within a 100-mile radius, both unconditionally and in association with bankruptcies.

[10]Again, by restricting the sample for this analysis to only those firms with bankrupt neighbors, we already alleviate concerns that firms experiencing peer bankruptcies nearby would have local competitive advantages, such as increased sales, compared to firms without bankrupt neighbors.

where j is user firm, k is the bankrupt neighbor, and t is year. To account for the fact that it takes time for the moving inventor to develop new patents at the new firm, we lag the instruments by three years ($t-3$). The dependent variable is $log(1 + LB\ hunting)$ of firms near bankrupt neighbors. The regressions are at the user firm-year level consistent with Table 3 and include firm fixed effects (γ_i) and year fixed effects (η_t). Standard errors are clustered by the bankruptcy-spillover-area (within 100 miles from the bankrupt neighbor) and year level to account for the correlation within the area and year. We find a strong IV effect on nearby firms' late-bloomer hunting with statistical significance at the 1%-level in the first stage. The F-statistics exceed the threshold of 10 suggested by Stock and Yogo (2005), indicating a strong instrument.

We then estimate the following second-stage regression in the even-numbered columns:

$$Sales\ growth_{j\{t,t+4\}} = a + b_1 log(1 + \widehat{LBhunting})_{jt} + \gamma_j + \eta_t. \tag{4}$$

The coefficient estimates for the instrumented late-bloomer hunting are all positive and statistically significant. The results in the second-stage estimations imply that an increase in late-bloomer hunting, driven solely by the hunting-inventor spillovers from bankrupt neighbors, has a significant positive impact on nearby firms' future sales growth. Our IV analysis in this section reinforces our main finding by showing that the user benefits of hunting late-bloomer patents are indeed causal.

8 Conclusion

We use the universe of patents granted over the past five decades to provide new insight into the fundamental chain of experimentation, search, and implementation that underlies the innovation process. Namely, we document large sample evidence of the importance of patent hunters – agents in the later stages of the innovation chain that search out, develop, and commercialize overlooked patents – in the eventual life-cycle of influential patents. We show that amongst all influential patents, a sizable portion is characterized by these "late-blooming" patents on which patent hunters play a role. These late-blooming patents, even though initially overlooked, on average are more influential than early-blooming patents, and open up significant new markets.

Patent hunters amass significant rents from detecting neglected patents – in terms of sales growth, Tobin's Q (market value), and new products. The patents they search out tend to be closer to their core (and more peripheral to the writers), along with being in – at that moment they are patented – less competitive idea and innovation spaces. Patent hunting is persistent at the firm level and appears to have a learning component, as the rents increase with successive patents hunted within firm. It also appears to have inventor-level components, as hunted patents are more valuable when tied with inventors, along with patent-hunting inventors continuing across workplaces.

This patent-hunting process also appears to have spillovers for the system in terms of creating more attention, innovation, and new product development in the hunted patents' idea spaces. Taken together, the results represent a new understanding of the latter stages of the innovation process – an area that is less well-understood and has received relatively less attention. Future research should continue to explore these dynamics, including other important agents and dynamics that underlie ultimate successful (and unsuccessful) realizations following the initial idea and patenting stages of the innovation chain.

References

Bloom, Nicholas, Mark Schankerman, and John Van Reenen, 2013, Identifying technology spillovers and product market rivalry, *Econometrica* 81, 1347–1393.

Brogaard, Jonathan, Joseph Engelberg, and Edward Van Wesep, 2018, Do economists swing for the fences after tenure?, *Journal of Economic Perspectives* 32, 179–194.

Chamley, Christophe, and Douglas Gale, 1994, Information revelation and strategic delay in a model of investment, *Econometrica: Journal of the Econometric Society* 1065–1085.

Escolar, Emerson G, Yasuaki Hiraoka, Mitsuru Igami, and Yasin Ozcan, 2023, Mapping firms' locations in technological space: A topological analysis of patent statistics, *Research Policy* 52, 104821.

Furman, Jeffrey L., Markus Nagler, and Martin Watzinger, 2021, Disclosure and subsequent innovation: Evidence from the patent depository library program, *American Economic Journal: Economic Policy* 13, 239–270.

Gal-Or, Esther, 1987, First mover disadvantages with private information, *The Review of Economic Studies* 54, 279–292.

Garfield, Eugene, 1970, Would Mendel's work have been ignored if the Science Citation Index was available 100 years ago?, *Essays of an Information Scientist* 1, 69–70.

Garfield, Eugene, 1980, Premature discovery or delayed recognition–why?, *Essays of an Information Scientist* 4, 488–493.

Garfield, Eugene, 1989a, Delayed recognition in scientific discovery-citation frequency-analysis aids the search for case-histories, *Current Contents* 23, 3–9.

Garfield, Eugene, 1989b, More delayed recognition. Part 1. Examples from the genetics of color blindness, the entropy of short-term memory, phosphoinositides, and polymer rheology, *Current Contents* 38, 3–8.

Garfield, Eugene, 1990, More delayed recognition. Part 2. From Inhibin to Scanning electron microscopy, *Current Contents* 9, 3–9.

Glänzel, Wolfgang, Balázs Schlemmer, and Bart Thijs, 2003, Better late than never? On the chance to become highly cited only beyond the standard bibliometric time horizon, *Scientometrics* 58, 571–586.

Glode, Vincent, and Guillermo Ordonez, 2023, Technological progress and rent seeking, *The University of Pennsylvania Working Paper Series* .

Graham, Stuart, and Deepak Hegde, 2015, Disclosing patents' secrets, *Science* 347, 236–237.

Hadlock, Charles J, and Joshua R Pierce, 2010, New evidence on measuring financial constraints: Moving beyond the kz index, *The review of financial studies* 23, 1909–1940.

Hegde, Deepak, Kyle Herkenhoff, and Chenqi Zhu, 2015, Patent publication and innovation, *National Bureau of Economic Research* .

Hoberg, Gerard, and Vojislav Maksimovic, 2015, Redefining financial constraints: A text-based analysis, *The Review of Financial Studies* 28, 1312–1352.

Hou, Jianhua, and Xiucai Yang, 2019, Patent sleeping beauties: Evolutionary trajectories and identification methods, *Scientometrics* 120, 187–215.

Kaplan, Steven N, and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints?, *Quarterly journal of economics* 112, 169–215.

Ke, Qing, Emilio Ferrara, Filippo Radicchi, and Alessandro Flammini, 2015, Defining and identifying sleeping beauties in science, *Proceedings of the National Academy of Sciences* 112, 187–215.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, 2021, Measuring technological innovation over the long run, *AER: Insights* 3, 303–320.

Kim, Jinhwan, and Kristen Valentine, 2021, The innovation consequences of mandatory patent disclosures, *Journal of Accounting and Economics* 71, 101381.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *Quarterly Journal of Economics* 132, 665–712.

Kuhn, Jeffrey, Kenneth Younge, and Alan Marco, 2020, Patent citations reexamined, *RAND Journal of Economics* 51, 109–132.

Linn, Matthew, and Daniel Weagley, 2021, Uncovering financial constraints, *Journal of Financial and Quantitative Analysis* 1–36.

Mukherjee, Abhiroop, Tomas Thornquist, and Alminas Žaldokas, 2022, New products, *Working paper* .

Pezzoni, Michele, Reinhilde Veugelers, and Fabiana Visentin, 2022, How fast is this novel technology going to be a hit? antecedents predicting follow-on inventions, *Research Policy* 51, 104454.

Sampat, Bhaven N, and Arvids A Ziedonis, 2004, Patent citations and the economic value of patents: A preliminary assessment, in *Handbook of quantitative science and technology research: the use of publication and patent statistics in studies of S&T Systems*, 277–298 (Springer).

Stock, James, and Motohiro Yogo, 2005, Testing for Weak Instrument in Linear IV Regression. In D. W. K. Andrews and J.H. Stock (eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg* pp. 80–108. Cambridge, UK: Cambridge University Press.

Strumsky, Deborah, and José Lobo, 2015, Identifying the sources of technological novelty in the process of invention, *Research Policy* 44, 1445–1461.

Trajtenberg, Manuel, 1990, A penny for your quotes: Patent citations and the value of innovations, *RAND Journal of Economics* 172–187.

Uzzi, Brian, Satyam Mukherjee, Michael Stringer, and Ben Jones, 2013, Atypical combinations and scientific impact, *Science* 342, 468–472.

Van Raan, Anthony F. J., 2004, Sleeping beauties in science, *Scientometrics* 59, 467–472.

Van Raan, Anthony FJ, 2017, Sleeping beauties cited in patents: Is there also a dormitory of inventions?, *Scientometrics* 110, 1123–1156.

Van Raan, Anthony FJ, and Jos J Winnink, 2019, The occurrence of 'sleeping beauty' publications in medical research: Their scientific impact and technological relevance, *PLoS One* 14, e0223373.

Verhoeven, Dennis, Jurriën Bakker, and Reinhilde Veugelers, 2016, Measuring technological novelty with patent-based indicators, *Research policy* 45, 707–723.

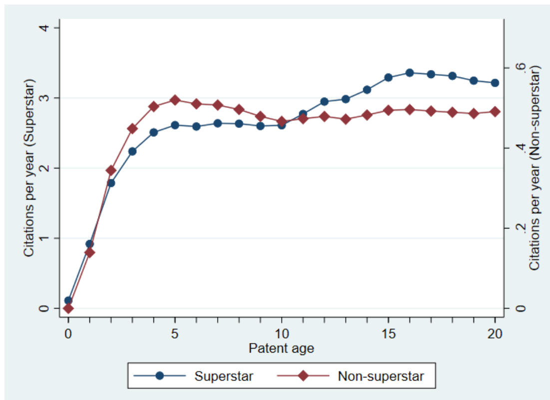
Weitzman, Martin L, 1998, Recombinant growth, *The Quarterly Journal of Economics* 113, 331–360.

Whited, Toni M, and Guojun Wu, 2006, Financial constraints risk, *Review of Financial Studies* 19, 531–559.

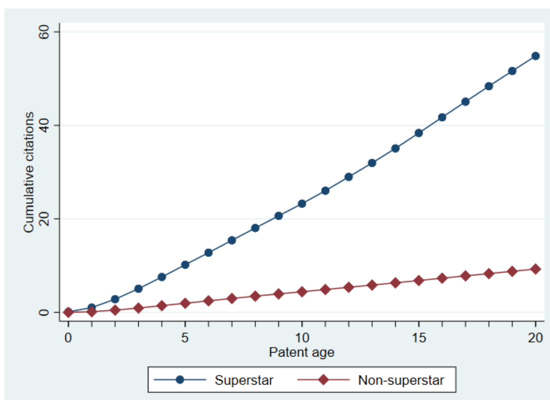
Williams, Heidi L, 2017, How do patents affect research investments?, *Annual Review of Economics* 9, 441–469.

Figure 1: Superstar patent forward citations over time

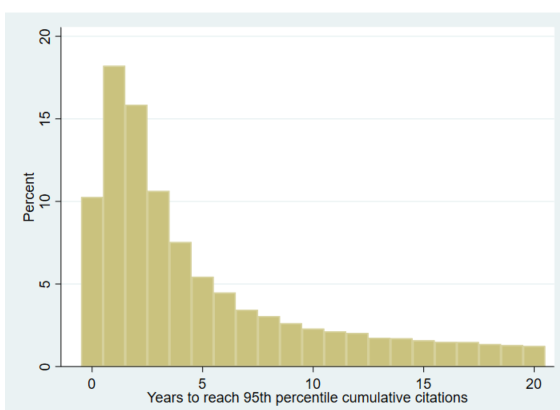
The sample consists of all patents granted between 1976 and 1999. A patent is classified as a superstar patent if the cumulative citations within CPC class and grant year cohort are in the top 95th percentile at any point in time during the patent term of 20 years since its grant. Panel (a) plots the average number of citations each year excluding self-citations. Panel (b) plots the cumulative number of citations over the patent age. Panel (c) shows the distribution of time to become a superstar patent.



(a) Average citations



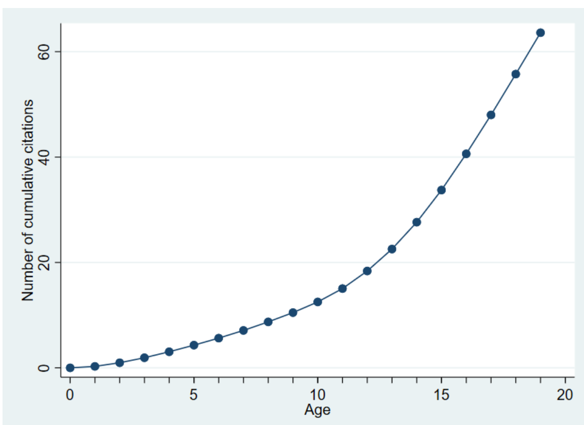
(b) Cumulative citations



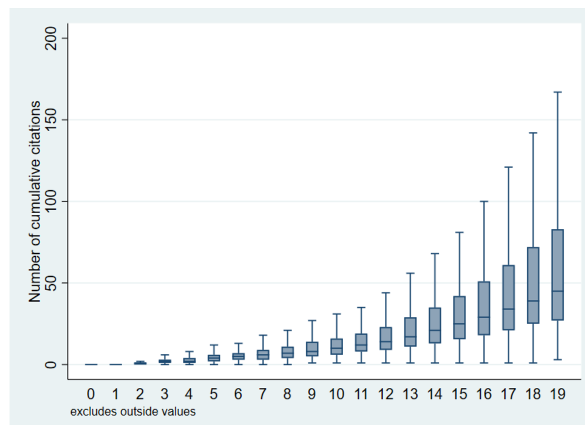
(c) Time to become superstar-patent

Figure 2: Late-bloomer vs. early-bloomer cumulative citations

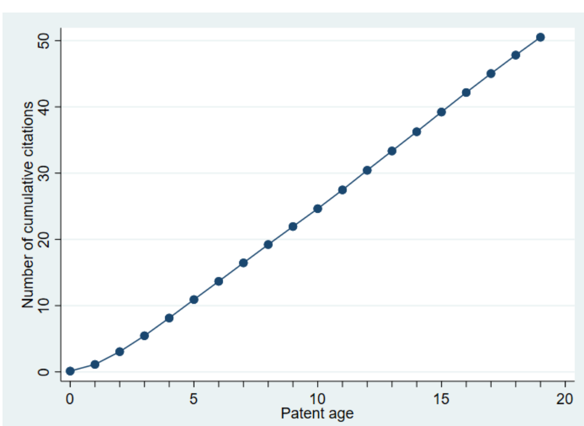
The sample consists of superstar patents granted between 1976 and 1999. The figures on the left panel (a and c) plot the average cumulative number of citations by patent age. The figures on the right panel (b and d) present the box plots of the number of cumulative citations, where the mid-line and upper/lower hinge represent the median and the interquartile range between the 25th and 75th percentiles, respectively. A superstar patent is a patent that has ever reached the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. A late-bloomer patent is a patent that takes an excessively long time period before it becomes a superstar patent. We use the 90th percentile point in the time-to-superstar distribution (14 years) to define the excessively long time period. An early bloomer is a superstar patent that is not classified as a late-bloomer patent.



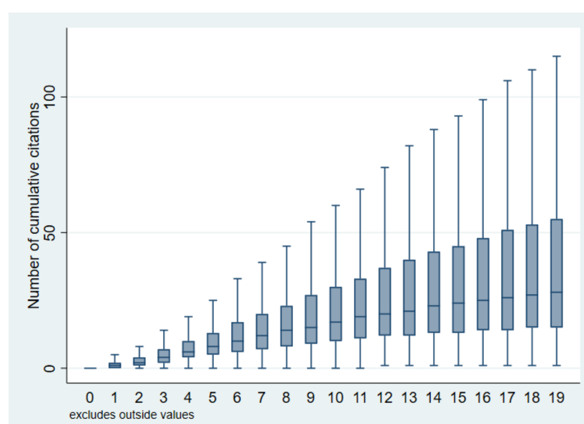
(a) Late-bloomer patents



(b) Late-bloomer patents



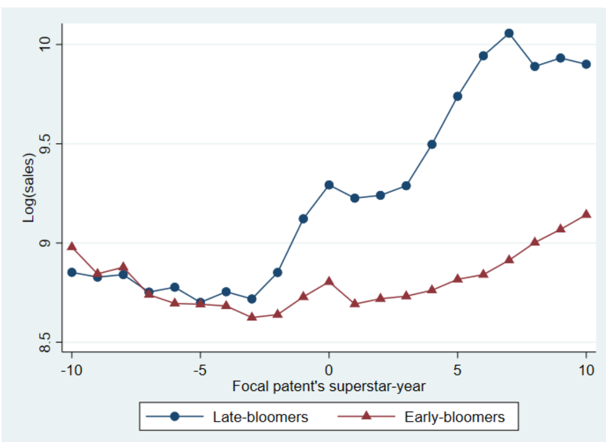
(c) Early-bloomer patents



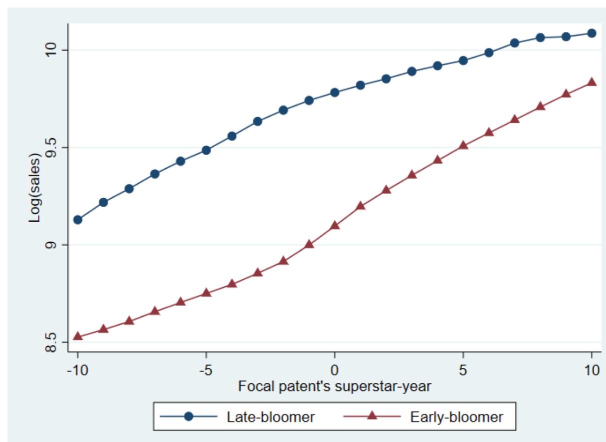
(d) Early-bloomer patents

Figure 3: Financial outcomes of late-bloomer writers and users

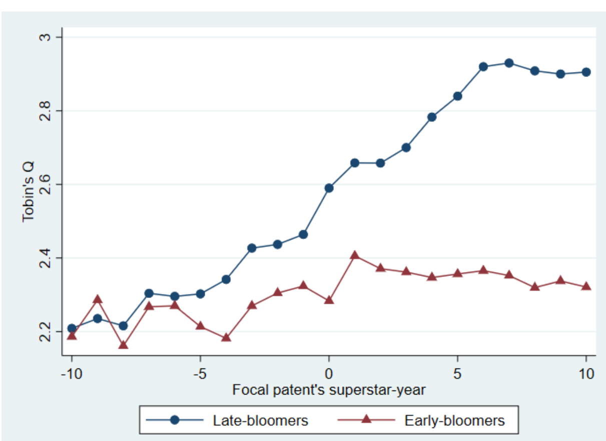
The figures present sales and Tobin's Q of late-bloomer patent writers and users around the time when the late-bloomer patent becomes a superstar patent (*i.e.,* superstar-year: its cumulative citations reach the top 95th percentile within CPC class and grant-year cohort). The x-axis represents the year relative to the superstar-year of each late-bloomer patent and is zero at the superstar-year. A superstar patent is a patent that has ever reached the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. A late-bloomer patent is a patent that takes an excessively long time period before it becomes a superstar patent. We use the 90th percentile point in the time-to-superstar distribution (14 years) to define the excessively long time period. An early bloomer is a superstar patent that is not classified as a late-bloomer patent. Late-bloomer patent writers are the firms that have at least one late-bloomer patent during the sample period. Late-bloomer patent users are the firms that have never written a late-bloomer patent but have cited at least one late-bloomer patent.



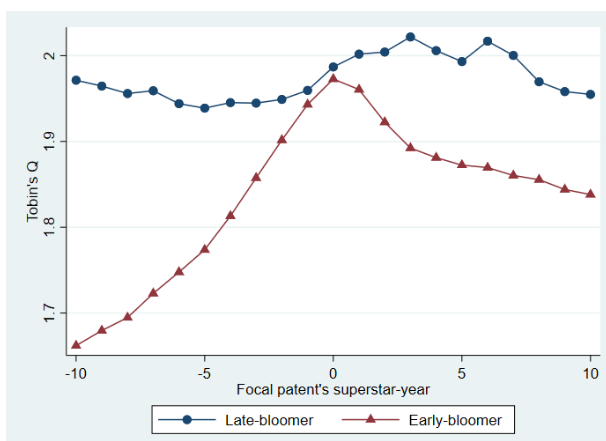
(a) User sales



(b) Writer sales



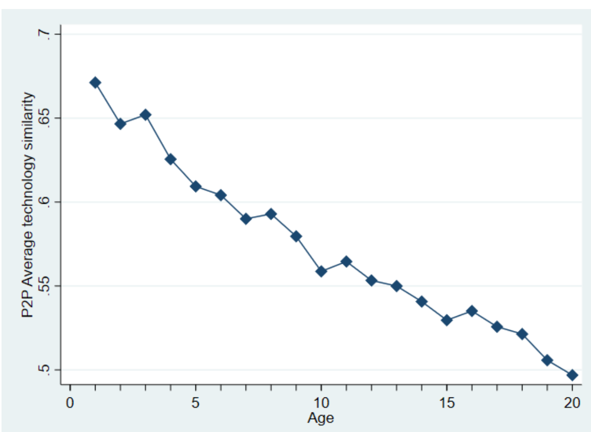
(c) User Tobin's Q



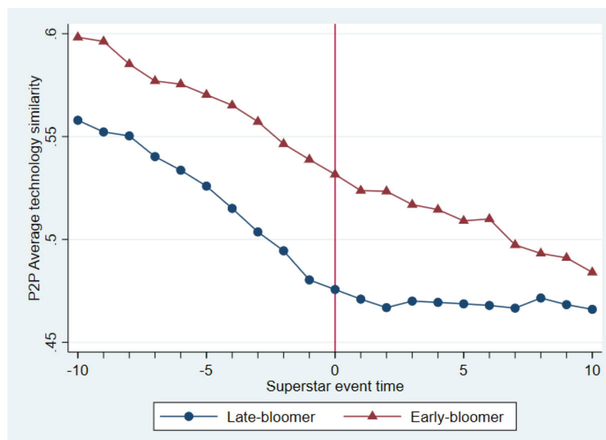
(d) Writer Tobin's Q

Figure 4: Technology proximity of citing patents

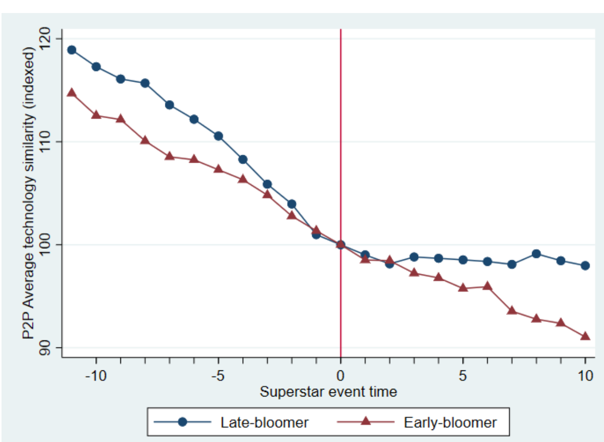
The figures present technology proximity between a pair of cited and citing patents. The sample consists of all cited-citing patent pairs. The cited patents are patents granted between 1976 and 1999 including both non-superstar and superstar patents. A superstar patent is a patent that has ever reached the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. Non-superstar patents in the sample comprise 100,000 randomly selected patents for a comparable sample size with superstar patents. A late-bloomer patent is a patent that takes an excessively long time period before it becomes a superstar patent. We use the 90th percentile point in the time-to-superstar distribution (14 years) to define the excessively long time period. An early bloomer is a superstar patent that is not classified as a late-bloomer patent. Technology proximity is the cosine similarity between two patents using their section-class-subclass level CPC classifications. Panel (a) plots the technology proximity between non-superstar patents and their citing patents over the focal patent age. Panel (b) plots the technology proximity between superstar patents and their citing patents by late-bloomers and early-bloomers around the superstar event time. Superstar event time is 0 when a patent becomes a *superstar patent* for the first time (*i.e.*, superstar-year). In Panel (c), the technology proximity values are indexed to superstar-year (*i.e.*, zero for the superstar event time). In Panel (d), the technology proximity is detrended from the pre-superstar year trend.



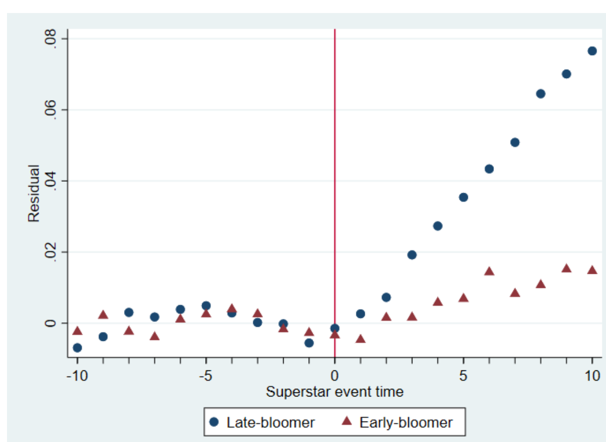
(a) Non-superstar patents



(b) Superstar patents



(c) Indexed technology proximity



(d) Detrended technology proximity

Table 1: Summary statistics

The table presents summary statistics of patent characteristics. Panel A compares non-superstar patents and superstar patents using the sample of all USPTO patents granted between 1976 and 1999. Panel B compares early-bloomer patents and late-bloomer patents using the sample of only superstar patents. A superstar patent is a patent that has ever reached the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. A late-bloomer patent is a patent that takes an excessively long time period before it becomes a superstar patent. We use the 90th percentile point in the time-to-superstar distribution (14 years) to define the excessively long time period. An early bloomer is a superstar patent that is not classified as a late-bloomer patent. Panel C compares citing patents of early-bloomer patents and late-bloomer patents. See Appendix A for other variable definitions in detail. ***p<0.01, **p<0.05, *p<0.1.

Panel A: Superstars vs. non-superstars

	Superstar patents			Non-superstar patents			
	mean	p 50	sd	mean	p50	sd	difference
issue year	1989.90	1991.00	6.84	1989.61	1991.00	6.95	-0.29***
cum. citations at age 5	10.17	8.00	10.00	1.96	1.00	2.39	-8.20***
cum. citations at age 10	23.12	16.00	24.40	4.43	3.00	5.07	-18.69***
cum. citations at age 15	38.04	23.00	47.16	6.84	4.00	8.23	-31.20***
cum. citations at age 20	54.26	30.00	76.22	9.31	6.00	12.17	-44.95***
count class	2.06	2.00	1.29	1.83	2.00	1.07	-0.23***
count claims	15.85	12.00	13.98	12.27	10	10.17	-3.58***
avg. claim word count	77.70	62.75	56.40	76.45	61.44	57.10	-1.25***
two-examiners	0.42	0.00	0.49	0.39	0	0.49	-0.03***
backward citation	12.13	8.00	15.82	9.39	7	10.52	-2.74***
individual inventor	0.02	0.00	0.15	0.02	0	0.15	0.00***
public	0.46	0.00	0.50	0.39	0	0.49	-0.07***
KPSS value	11.28	3.90	31.45	9.08	3.26	23.56	-2.19***
Unique number of patents		213,772			1,499,277		

Panel B: Late-bloomers vs. early-bloomers

	Late-bloomers			Early-bloomers			
	mean	p 50	sd	mean	p50	sd	difference
issue year	1989.44	1991.00	7.06	1989.95	1991.00	6.82	0.51***
cum. citations at age 5	3.80	3.00	3.32	10.88	8.00	10.24	7.08***
cum. citations at age 10	11.29	9.00	8.33	24.44	17.00	25.23	13.15***
cum. citations at age 15	31.19	23.00	24.52	38.80	23.00	48.97	7.61***
cum. citations at age 20	69.80	49.00	64.73	52.53	28.00	77.20	-17.27***
count class	2.17	2.00	1.39	2.05	2.00	1.28	-0.12***
count claims	16.02	13.00	14.13	15.83	12.00	13.96	-0.19**
avg. claim word count	74.01	60.12	53.84	78.11	63.00	56.66	4.09***
two examiners	0.41	0.00	0.49	0.42	0.00	0.49	0.01**
backward citation	12.41	8.00	17.64	12.09	8.00	15.61	-0.32***
individual inventor	0.02	0.00	0.14	0.02	0.00	0.15	0.00*
public	0.46	0.00	0.50	0.46	0.00	0.50	0.01**
KPSS value	11.15	4.31	29.61	11.29	3.85	31.65	0.14
Unique number of patents		21,960			191,812		

Panel C: Citing patents of late-bloomers vs. and early-bloomers

	Late-bloomer citing patents			Early-bloomer citing patents			
	mean	p50	sd	mean	p50	sd	difference
issue year	2006.42	2008.00	9.20	2003.70	2004.00	10.35	-2.72***
cum. citations at age 5	5.05	2.00	11.71	4.04	2.00	8.47	-1.01***
cum. citations at age 10	15.39	7.00	29.49	11.05	5.00	21.08	-4.34***
cum. citations at age 15	27.43	12.00	51.16	18.50	8.00	36.63	-8.94***
cum. citations at age 20	35.99	16.00	69.11	23.21	10.00	47.37	-12.77***
count class	2.22	2.00	1.54	2.00	2.00	1.31	-0.22***
count claims	19.56	17.00	15.54	17.26	15.00	13.43	-2.30***
avg. claim word count	64.91	53.45	109.04	70.34	57.30	73.48	5.43***
two-examiners	0.37	0.00	0.48	0.39	0.00	0.49	0.02***
backward citation	96.89	31.00	195.04	43.96	16.00	113.79	-52.94***
individual inventor	0.01	0.00	0.08	0.01	0.00	0.08	0.00**
public	0.40	0.00	0.49	0.41	0.00	0.49	0.01***
KPSS value	16.34	5.81	42.47	13.65	4.51	38.77	-2.69***
Observations		790,936			2,797,100		

Table 2: Late-bloomer writers vs. users and their persistence

Panel A compares firm-level patenting and financial characteristics of late-bloomer writers and users. Writers and users are matched by firm age. Late-bloomer writers are the firms that have produced at least one late-bloomer patent during the sample period. Late-bloomer users are the firms that have never produced a late-bloomer patent but cited at least one late-bloomer patent during the sample period. The sample consists of 3,097 firms in total with 1,892 late-bloomer writers and 1,205 late-bloomer users that are mutually exclusive. We use five nearest neighbor users for each late-bloomer writer based on the firm age. ATE stands for the average treatment effect, and SE stands for the standard error. Panel B presents the transition matrix of patent hunter statuses from year t to year $t+1$. Strict Writer is one for a firm that produces a late-bloomer patent in a given year but does not cite any late-bloomer patent in that year, and zero otherwise. Flexible Writer/User is one for a firm that produces a late-bloomer patent in a given year and also cites late-bloomer patents in that year, and zero otherwise. Strict User is one for a firm that cites late-bloomer patents in a given year but does not produce any late-bloomer patents in that year, and zero otherwise. Idle is one for a firm that neither produces a late-bloomer patent nor cites late-bloomer patents in a given year. In each status (a) to (d), the value is the number of observations with its corresponding percentage in the brackets. See Appendix A for other variable definitions in detail. ***p<0.01, **p<0.05, *p<0.1.

Panel A: Cross-sectional firm characteristics

	Late-bloomer writers	Late-bloomer Users	ATE	SE
	(1)	(2)	(3)	(4)
log(age) without matching	2.176	1.557	0.619***	0.030
no. patents per year	29.71	2.840	26.87***	3.234
no. superstars per year	4.739	0.512	4.227***	0.485
no. late bloomers per year	0.620	0	0.620***	0.0409
no. external cites per year	63.91	4.135	59.78***	6.330
no. external cites/no. patents	2.450	1.517	0.933***	0.112
no. claims/no. patents	16.58	17.37	-0.795**	0.346
no. new products/no. patents	0.181	0.256	-0.0749***	0.0271
log(asset)	5.212	4.665	0.546***	0.0827
log(sale)	4.930	4.383	0.547***	0.0968
log(age)	1.799	1.798	0.000116	0.000170
tobinq	2.495	2.523	-0.0281	0.0698
salegr	0.167	0.157	0.00998	0.0102
roa	-0.0686	-0.0651	-0.00343	0.0110
leverage	0.192	0.196	-0.00421	0.00666
ppe_asset	0.479	0.476	0.00335	0.0118
rnd_sale	0.507	0.437	0.0697	0.0559
capx_sale	0.187	0.163	0.0238	0.0174
adv_sale	0.0113	0.0111	0.000140	0.00110
d_dv	0.425	0.384	0.0414***	0.0143
consumer dependent	0.231	0.256	-0.0250**	0.0116

Panel B: Persistence of patent hunting

	status at $t+1$				
	strict writer	flexible writer/user	strict user	idle	total
status at t	(1)	(2)	(3)	(4)	
(a) strict writer	113 [13.12]	148 [17.19]	184 [21.37]	416 [48.32]	861 [100]
(b) flexible writer/user	46 [1.65]	1,709 [61.12]	788 [28.18]	253 [9.05]	2,796 [100]
(c) strict user	118 [2.97]	832 [20.94]	2,019 [50.82]	1,004 [25.27]	3,973 [100]
(d) idle	444 [6.07]	379 [5.18]	1,308 [17.88]	5,185 [70.87]	7,316 [100]
total	721 [4.82]	3,068 [20.53]	4,299 [28.76]	6,858 [45.89]	14,946 [100]

Table 3: Benefits from patent hunting (firm level)

The table presents results from the regressions that examine financial benefits to patent users. The observations are at the firm-year level for the period of 1976 to 2020. *Sales growth* is the five-year sales growth, computed as $(sales_{t+4}/sales_t) - 1$, and $Avg\ Tobin's\ Q$ is the five-year average of Tobin's Q over t to $t+4$. *LBhunting* and *EBhunting* are the numbers of hunted late-bloomer and early-bloomer patents in a given year, respectively. *earlyLBhunting* is the number of hunted late-bloomer patents completed in the late-bloomers' first three hunting patents since their grants and prior to their superstar-year. A superstar-year is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. *laterLBhunting* is the number of the rest hunted late-bloomer patents that are not early hunted. The regressions include the firm fixed effects and year fixed effects. Standard errors in parentheses are clustered at the firm level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Sales growth (1)	Avg Tobin's Q (2)	Sales growth (3)	Avg Tobin's Q (4)
log(1+LBhunting)	0.0831***	0.0658***		
	(0.0275)	(0.0168)		
log(1+EBhunting)	-0.0962***	-0.0217**	-0.0955***	-0.0250***
	(0.0181)	(0.00866)	(0.0180)	(0.00852)
log(1+earlyLBhunting)			0.0749**	0.0848***
			(0.0305)	(0.0238)
log(1+laterLBhunting)			0.0658**	0.0576***
			(0.0290)	(0.0179)
log(asset)	-0.757***	-0.309***	-0.757***	-0.309***
	(0.0366)	(0.0139)	(0.0366)	(0.0139)
log(age)	-0.443***	-0.225***	-0.443***	-0.226***
	(0.0350)	(0.0155)	(0.0351)	(0.0155)
roa	-1.837***	-0.251***	-1.837***	-0.251***
	(0.149)	(0.0435)	(0.149)	(0.0436)
leverage	-0.940***	0.00403	-0.939***	0.00523
	(0.139)	(0.0557)	(0.139)	(0.0557)
Mean	0.901	2.080	0.901	2.080
$H_0: LB = EB$ (p-value)	0.000	0.000		
$H_0: earlyLB = laterLB$ (p-value)			0.847	0.417
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	75589	98776	75589	98776
Adjusted R^2	0.350	0.719	0.350	0.719

Table 4: Benefits from patent hunting (patent level)

The table presents results from the difference-in-differences regressions that examine financial benefits to patent users. The observations are at the focal patent-user firm-year level for the period of 1976 to 2020. The samples in Columns 1 and 2, 3 and 4, and 5 and 6 consist of late-bloomer, early-bloomer, and non-superstar patents and their citing patents, respectively. *user* is an indicator variable that is one if the firm cites a focal patent and zero otherwise. *ssyear* is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. For a non-superstar patent, *ssyear* is the year with the greatest number of citations during 20 years since its grant. ssyear_{post} is an indicator variable equal to one if the year is after *ssyear* of a given patent and zero otherwise. The non-superstar patents comprise 100,000 randomly selected patents for a comparable sample size with superstar patents. Appendix Tables A3 and A4 use alternative comparison groups for the robustness tests. The regressions include the cited patent fixed effects and year fixed effects. Standard errors in parentheses are clustered at the patent-by-year level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Late-bloomers		Early-bloomers		Non-superstars	
	Sales growth	Tobin's Q	Sales growth	Tobin's Q	Sales growth	Tobin's Q
	(1)	(2)	(3)	(4)	(5)	(6)
user × ssyear_{post}	0.00730***	0.0179**	0.000159	-0.0193***	0.00146	-0.0482***
	(0.00125)	(0.00816)	(0.000570)	(0.00363)	(0.000922)	(0.00542)
user	0.0143***	0.0644***	0.00732***	0.0191***	0.0146***	0.0308***
	(0.000923)	(0.00525)	(0.000511)	(0.00327)	(0.000810)	(0.00468)
ssyear_{post}	-0.00146	-0.000481	-0.00408***	0.0157***	-0.000298	0.0128***
	(0.00145)	(0.00625)	(0.000573)	(0.00301)	(0.000856)	(0.00384)
Mean	0.047	2.075	0.057	2.020	0.046	2.022
Control variables	Y	Y	Y	Y	Y	Y
Cited patent FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Observations	1523717	1534074	11017213	11086629	759482	763142
Adjusted R^2	0.226	0.386	0.287	0.452	0.261	0.490

Table 5: Late-bloomer patents' blooming and subsequent patents

The table presents results from the regressions that examine the effect of superstar patents on the number of subsequent patents in the same tech-class group of the focal patents (Column 1) and in the same tech-class group of citing patents of the focal patents (Column 2). The observations are at the tech-class group and year level. For the tech-class groups in Column 1, we use a given focal patent' reported CPC technology classes. We drop a tech-class group if both late-bloomer patents and early-bloomer patents are assigned to the group. For the tech-class groups in Column 2, we consider exhaustive pairs of all CPC technology classes reported by citing patents of a given focal patent. Then, we count the number of subsequent patents whose CPC classes include the most frequent pair of citing patents' CPC classes. *ssyear* is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. ssyear_{post} is an indicator variable equal to one if the year is after *ssyear* of a given patent and zero otherwise. The regressions include the focal patent fixed effects. Standard errors in parentheses are clustered at the tech-class group level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Log(Patent counts in tech-class groups)	
	Focal CPC	Citing CPC
	(1)	(2)
latebloomer × ssyear_{post}	0.145***	0.0398**
	(0.00495)	(0.0198)
ssyear_{post}	-0.147***	-0.00416
	(0.00219)	(0.00623)
log(totalpat)	0.0189***	0.0377**
	(0.00204)	(0.0150)
Control variables	Y	Y
Focal patent FE	Y	Y
Observations	696851	1274268
Adjusted R^2	0.410	0.458

Table 6: Commercialization of technology

The table presents results from the regressions that examine the commercialization of technology by patent users. The observations are at the firm-year level in column 1 and focal patent-user firm-year level in columns 2, 3, and 4 for the period of 1990 to 2015. The sample period is shorter due to the availability of the New Product data. *New product count* is the total number of new products from Mukherjee et al. (2022). *Avg log*$(1 + \textit{new products})$ is the five-year average number of new products over t to $t+4$. *LBhunting* and *EBhunting* are the numbers of hunted late-bloomer and early-bloomer patents in a given year, respectively. The sample in columns 2, 3, and 4 consists of late-bloomer, early-bloomer, and non-superstar patents and their citing patents, respectively. The non-superstar patents comprise 100,000 randomly selected patents for a comparable sample size with superstar patents. *User* is an indicator variable that is one if the firm cites a focal patent and zero otherwise. *ssyear* is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. For a non-superstar patent, *ssyear* is the year with the greatest number of citations during 20 years since its grant. ssyear_{post} is an indicator variable equal to one if the year is after *ssyear* of a given patent and zero otherwise. The regression in column 1 includes the firm fixed effects and year fixed effects, and the regressions in columns 2, 3, and 4 include cited patent fixed effects and year fixed effects. Standard errors are clustered at the firm level in column 1 and focal patent-by-year level in columns 2, 3, and 4. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Avg log(1+new products)	Log(1+new products)		
	Late-bloomers	Late-bloomers	Early-bloomers	Non-superstars
	(1)	(2)	(3)	(4)
log(1+LBhunting)	0.0313***			
	(0.0105)			
log(1+EBhunting)	0.0124***			
	(0.00463)			
user × ssyear_{post}		0.0472***	0.000301	-0.0286***
		(0.00842)	(0.00386)	(0.00570)
user		0.0359***	0.00362	-0.0456***
		(0.00569)	(0.00346)	(0.00495)
ssyear_{post}		-0.0404***	-0.0106***	0.0196***
		(0.00877)	(0.00391)	(0.00513)
Mean	0.356	1.119	0.898	1.080
Control variables	Y	Y	Y	Y
Firm FE	Y	N	N	N
Cited patent FE	N	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	66636	1201198	10045142	640206
Adjusted R^2	0.706	0.626	0.898	0.669

Table 7: Constraints for late-bloomer patent writers

The table presents results from the regressions that examine potential constraints for late-bloomer writers. The sample consists of superstar patents only, and the analysis compares late-bloomer and early-bloomer patents. Observations are at the patent level (one observation per patent) by taking the averages of relevant variables when a patent has multiple CPC tech classes. We consider (i) intellectual capacity constraints measured by *tech-class weight* or *tech class dist to core*, (ii) competitive threat measured by *log(competing patent stock)*, and (iii) financial constraints measured by *fin_const (KZ)* or *equity(debt)_const (LW)*. *tech-class weight* is the fraction of the patents in the CPC tech class of a given patent in all patents of its assignee over the entire sample period. *tech class dist to core* is the class-to-class proximity between the CPC tech class of a given patent and the core CPC tech class of its assignee. *log(competing patent stock)* is the log of the number of all patents from U.S. public firms with the same CPC tech class up to the grant year of a given patent. The regressions include the writer firm fixed effects. Standard errors in parentheses are clustered at the firm and grant year levels. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Late-bloomers			
	(1)	(2)	(3)	(4)
tech-class weight	-0.0432***	-0.0455**		
	(0.0145)	(0.0167)		
tech-class dist to core			0.0143**	0.0142**
			(0.00534)	(0.00551)
ln(competing patent stock)	-0.00556***	-0.00534***	-0.00561***	-0.00543***
	(0.00134)	(0.00138)	(0.00131)	(0.00135)
fin_const (KZ)	-0.00701**		-0.00695**	
	(0.00296)		(0.00293)	
equity_const (LW)		-0.00913		-0.00936
		(0.00776)		(0.00776)
debt_const (LW)		-0.00177		-0.00165
		(0.00770)		(0.00770)
log_asset	0.00782*	0.00618	0.00769	0.00604
	(0.00455)	(0.00464)	(0.00454)	(0.00464)
log_age	0.000186	0.00214	0.000170	0.00221
	(0.00575)	(0.00769)	(0.00577)	(0.00768)
Control variables	Y	Y	Y	Y
Writer FE	Y	Y	Y	Y
Grant year FE	Y	Y	Y	Y
Observations	94889	86801	94889	86801
Adjusted R^2	0.033	0.033	0.033	0.033

Table 8: Costs of patent hunting

The table presents results from the cross-sectional regressions that examine the effects of costs of hunting on user benefits. The sample consists of all late-bloomer citing firms. The observations are at the focal patent-user firm-year level for the period of 1976 to 2020. *Sales growth* is the five-year sales growth, computed as $(sales_{t+4}/sales_t) - 1$. *Avg Tobin's Q* is the five-year average of Tobin's Q. *complexity* is the Gunning fog index of the text of each cited late-bloomer patent (focal patent). *competition* is the average number of competitors with overlapping CPC tech classes among all users that cite the focal patent. The regressions include the focal-patent fixed effects and year fixed effects. Standard errors in parentheses are clustered at the focal-patent level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Sales growth (1)	Avg Tobin's Q (2)	Sales growth (3)	Avg Tobin's Q (4)
complexity	-0.000675	-0.0191***		
	(0.00174)	(0.00387)		
competition			-0.00251***	-0.00795***
			(0.000925)	(0.00209)
log_asset	-0.0531***	-0.155***	-0.0480***	-0.143***
	(0.00305)	(0.00602)	(0.00287)	(0.00566)
log_age	-0.126***	0.0156*	-0.141***	-0.0194**
	(0.00537)	(0.00937)	(0.00528)	(0.00904)
leverage_b	-0.231***	-1.316***	-0.291***	-1.309***
	(0.0275)	(0.0566)	(0.0266)	(0.0558)
Mean	0.316	2.267	0.316	2.267
Focal patent FE	N	N	N	N
Focal patent class FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	95841	116692	108953	135942
Adjusted R^2	0.117	0.220	0.117	0.212

Table 9: Skills and deliberateness of patent hunting

The table examines skills and deliberateness of patent hunting. Panel A presents results from the regressions that examine financial benefits to patent users by the users' hunting experience. We consider the specifications in the first two columns of Table 4 for subgroups of experienced users (Columns 1 and 2) and less-experienced users (Columns 3 and 4). *Experienced* users are the top 10% firm-years in the firm-level average of the late-bloomer fraction among all cited patents in the past 5 years from a citing year. The rest of the firm-years are *less experienced* users. The regressions include the cited patent fixed effects and year fixed effects. Standard errors are clustered at the focal patent-by-year level. Panel B examines the within citing patent comparison of technology proximity between the user's citing patent and the cited LB patents with those between all the other cited patents (EB and non-superstar). Panel C compares the following statistics for cited patents across late-bloomer, non-late-bloomer, and early-bloomer groups: (i) the likelihood of being mentioned in the text part of citing patents, (ii) the number of text mentions, (iii) the sentiment of neighboring words around the text mentions, and (iv) the likelihood of being cited by examiners. The measures (i), (ii), and (iii) are from Moon, Suh, and Zhou (2024). ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

Panel A: Experienced hunting

	Experienced		Less experienced	
	Sales growth	Tobin's Q	Sales growth	Tobin's Q
	(1)	(2)	(3)	(4)
user × ssyear$_{post}$	0.0160***	0.0993***	0.00405***	-0.0384***
	(0.00404)	(0.0282)	(0.00123)	(0.00734)
user	0.0184***	0.312***	0.0134***	0.0406***
	(0.00365)	(0.0243)	(0.000910)	(0.00483)
ssyear$_{post}$	-0.0162***	-0.0586***	0.00113	0.0203***
	(0.00367)	(0.0189)	(0.00143)	(0.00596)
Control variables	Y	Y	Y	Y
Cited patent FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	241937	247010	1281769	1287061
Adjusted R^2	0.222	0.410	0.241	0.424

Panel B: Proximity between hunters' and hunted patents

	Technology Proximity		
	(1)	(2)	(3)
LB	0.00795***	-0.00282*	0.0253***
	(0.00192)	(0.00159)	(0.00293)
patage_f	-0.00213***	-0.00205***	-0.00146***
	(0.0000930)	(0.000161)	(0.0000897)
Comparison group	EB, non-superstar	EB only	non-superstar only
Citing patent FE	Y	Y	Y
Tech class FE	Y	Y	Y
Observations	2938358	1812704	1546350
Adjusted R^2	0.523	0.563	0.508

Panel C: Deliberate hunting

	Late-bloomers	Non-late-bloomers	Early-bloomers	(1)-(2)	(1)-(3)
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$(in-text cited)	0.0552	0.0446	0.0733	0.0106***	-0.0180***
No.(in-text mentions)	0.0676	0.0558	0.0912	0.0117***	-0.0236***
Sent(in-text mentions)	0.373	0.329	0.345	0.0438***	0.0282***
$\mathbb{1}$(examiner cited)	0.178	0.362	0.285	-0.184***	-0.107***

Table 10: Benefits from patent hunting with inventor moves

The table presents results from the difference-in-differences regressions that examine the incremental financial benefits to patent users that hire the late-bloomer patent inventors. The observations are at the firm and year level for the period of 1976 to 2020. *inventor move* is an indicator variable that is one if the late-bloomer inventor moves to the user firm and zero otherwise. *user* is an indicator variable that is one if the firm cites a focal patent and zero otherwise. *ssyear* is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. For a non-superstar patent, *ssyear* is the year with the greatest number of citations during 20 years since its grant. ssyear_{post} is an indicator variable equal to one if the year is after *ssyear* of a given patent and zero otherwise. The regressions include the cited patent fixed effects and year fixed effects. Standard errors in parentheses are clustered at the focal patent-by-year level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	All inventor moves		Moves before ssyear	
	Sales growth	Tobin's Q	Sales growth	Tobin's Q
	(1)	(2)	(3)	(4)
inventor move × user × ssyear_{post}	0.0371***	0.115***	0.0473***	0.241***
	(0.00567)	(0.0388)	(0.00658)	(0.0471)
user × ssyear_{post}	0.00505***	0.0135*	0.00580***	0.0139
	(0.00126)	(0.00814)	(0.00153)	(0.00923)
ssyear_{post}	-0.000185	0.0000323	-0.000185	-0.000779
	(0.00146)	(0.00628)	(0.00146)	(0.00629)
inventor move × ssyear_{post}	-0.0193***	-0.0581***	-0.0263***	-0.0864***
	(0.00340)	(0.0187)	(0.00388)	(0.0224)
inventor move × user	-0.00167	0.0670***	-0.00194	0.0522*
	(0.00401)	(0.0243)	(0.00460)	(0.0277)
inventor move	0.00961***	0.0164	0.0110***	0.0257*
	(0.00232)	(0.0126)	(0.00261)	(0.0146)
user	0.0137***	0.0632***	0.0137***	0.0627***
	(0.000934)	(0.00525)	(0.000934)	(0.00525)
Control variables	Y	Y	Y	Y
Cited patent FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	1501145	1510968	1485859	1495647
Adjusted R^2	0.228	0.397	0.228	0.398

Table 11: Patent hunting inventors

The table presents results from the regressions that examine whether patent hunting is a firm-specific or an inventor-specific skill. The observations are at the level of inventors whose patents are granted during the period from 1976 to 1999. The sample consists of a universe of patent inventors and their public firm employers. We only consider inventors who change jobs during the sample period. We drop the years where an inventor works for more than one employer at the same time. The dependent variable is either an indicator for whether the inventor uses late-bloomer patents at least once in the next employment ($\mathbb{1}$(late-bloomers)) or the number of late-bloomer patents that the inventor cites in the next employment (no.(late-bloomers)). Columns 1 and 2 consider only the subsequent employer, and Columns 3 and 4 consider the averages of up to three subsequent employers. $\mathbb{1}$(late-bloomers)) and no.(late-bloomers) on the right-hand side of regressions are the indicators for whether a given inventor cites a late-bloomer patent in the current employment and the number of total late-bloomer patents that the inventor cites in the current employment, respectively. $\mathbb{1}$(superstars) and no.(superstars) are the indicators for whether a given inventor cites a superstar patent in the current employment and the number of total superstar patents that the inventor cites in the current employment, respectively. The regressions control for a given inventor's gender, the total number of patents that the inventor writes, the total number of firms that the inventor works for during the sample period, and the current employer's financial characteristics. The regressions include the current employer-fixed effects and the inventor's work-start year fixed effects. Standard errors in parentheses are clustered at the inventor level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Next firm		Next three firms	
	$\mathbb{1}$(late-bloomers)	no.(late-bloomers)	$\mathbb{1}$(late-bloomers)	no.(late-bloomers)
	(1)	(2)	(3)	(4)
$\mathbb{1}$(late-bloomers)	0.0681***		0.0738***	
	(0.00606)		(0.00580)	
$\mathbb{1}$(superstars)	0.0189***		0.0181***	
	(0.00453)		(0.00432)	
no.(Late-bloomers)		0.136***		0.155***
		(0.0224)		(0.0230)
no.(superstars)		0.00672***		0.00776***
		(0.00212)		(0.00217)
inv_gender	-0.00696	0.0130	-0.0107	0.00915
	(0.00844)	(0.0156)	(0.00903)	(0.0172)
inv_npat	0.00236***	0.00131***	0.00257***	0.00137***
	(0.000291)	(0.000230)	(0.000298)	(0.000249)
inv_nfirms	-0.0178***	-0.0108***	-0.0154***	-0.00475*
	(0.00212)	(0.00226)	(0.00223)	(0.00273)
Control variables	Y	Y	Y	Y
Current employment FE	Y	Y	Y	Y
Work start year FE	Y	Y	Y	Y
Observations	51544	51544	51544	51544
Adjusted R-squared	0.053	0.062	0.062	0.062

Table 12: Benefits from patent hunting - instrumental variable regressions

The table presents results from the instrumental variable regressions that examine sales growth benefits to patent users to mitigate potential endogeneity issues. The observations are at the firm-year level for the period of 1987 to 2020. The sample period in this analysis is shorter because our bankruptcy data from Audit Analytics starts in 1987. We use the hunting intensity of bankrupt neighbors within a 100-mile radius as an instrument for nearby firms' late-bloomer hunting activities. Columns (1) and (2) use the average of inventor-level hunting intensities (the fraction of hunting patents) within the bankrupt neighboring firm as the instrument. Columns (3) and (4) use the firm-level hunting intensity (the fraction of hunting inventors) at the bankrupt neighboring firm as the instrument. The bankrupt neighbor's hunting intensity is calculated using the past three-year moving average considering that inventors may not file patents every year. Columns (1) and (3) are the first-stage results from the regressions of late-bloomer hunting on the instrument. To account for the fact that it takes time for the moving inventor to develop new patents at the new firm, we lag the instruments by three years. We use Cragg-Donald Wald F-stat for the weak instrument test and Kleibergen-Paap rk statistic for underidentification test. Columns (2) and (4) are the second-stage results from the regressions of sales growth on instrumented late-bloomer hunting. *Sales growth* is the five-year sales growth, computed as $(sales_{t+4}/sales_t) - 1$. *LB hunting* is the numbers of hunted late-bloomer patents in a given year, respectively. The regressions include the firm fixed effects and year fixed effects. Standard errors in parentheses are clustered at the bankrupt firm neighbor by year level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Inventor-level hunting intensity		Firm-level hunting intensity	
	Log(1+LB hunting)	Sales growth	Log(1+LB hunting)	Sales growth
	(1)	(2)	(3)	(4)
bankrupt neighbor hunting intensity	0.454***		0.236***	
	(0.0982)		(0.0539)	
instrumented log(1+LB hunting)		2.475***		1.555*
		(0.955)		(0.888)
log(asset)	0.117***	-0.865***	0.117***	-0.757***
	(0.00695)	(0.119)	(0.00695)	(0.112)
log(age)	0.0859***	-0.168*	0.0855***	-0.0882
	(0.0110)	(0.0965)	(0.0110)	(0.0912)
roa	-0.0807***	-1.592***	-0.0809***	-1.665***
	(0.0167)	(0.190)	(0.0167)	(0.187)
leverage	-0.0320	-0.736***	-0.0320	-0.762***
	(0.0266)	(0.180)	(0.0266)	(0.172)
First-stage F-stat	21.41		19.16	
Weak instrument test	21.96		18.47	
Underidentification test	19.71		16.11	
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	25874	25874	25874	25874
Adjusted R-squared	0.776	0.135	0.776	0.253

A Variable Definition

Variable Name	Definition
Superstar patent	A patent that has ever reached the 95th percentile cumulative forward citations (net of self-citations) within the CPC class-grant year cohort.
Late-bloomer patent	A patent that takes more than 14 years (the 90th percentile in time-to-superstar distribution) to become a superstar patent.
Early-bloomer patent	A superstar patent that is not a late-bloomer patent.
Cum. citations	The cumulative number of forward citations net of self-citations.
Count class	The number of unique technology classes.
Count claims	The number of claims.
Avg. claim word count	The average number of words in claims.
Two-examiners	An indicator variable equal to one if a patent was reviewed by two examiners and zero otherwise.
Backward citation	The number of backward citations.
Individual inventor	An indicator variable equal to one if the patent is assigned to an individual and zero otherwise.
Public	An indicator variable equal to one if the patent is assigned to a public firm and zero otherwise.
KPSS value	Kogan et al. (2017) value of patent.
no. patents per year	The number of granted patents of the firm in a year.
no. superstars per year	The number of granted patents of the firm that become superstar patents in a year.
no. latevbloomers per year	The number of granted patents of the firm that become late-bloomer patents in a year.
no. external cites per year	The number of citations received in a year net of self-citations.
no. external cites/no. patents	The total number of citations received (net of self-citations) scaled by the total number of patents.
no. claims/no. patents	The total number of claims scaled by the total number of patents.
no. new products/no. patents	The total number of new product launches scaled by the total number of patents.
log(asset)	The logarithm of total assets.
log(sale)	The logarithm of total assets.
log(age)	The logarithm of firm age.
tobin's Q	The Tobin's Q ratio, calculated as the market value of a company divided by the total assets.
sales growth	Logarithm of the total revenues divided by the previous year's total revenues.
roa	Return on assets, calculated as the net income divided by the total assets.
leverage	The debt-to-assets ratio, calculated as total debt divided by total assets.
ppe_asset	Tangible fixed assets (Property, Plant, and Equipment) scaled by the total assets.
rnd_asset	R&D expense scaled by the total assets.
capx_asset	Capital expenditure scaled by the total assets.
adv_asset	Advertising expense scaled by the total assets.
d_dv	An indicator variable equal to one if the firm pays dividends.
consumer dependent	An indicator for consumer-dependent industries whose production percentage for "Personal consumption expenditures" in the 2002 Input-Output Accounts from the Bureau of Economic Analysis is in the top tercile.
log(totalpat)	The logarithm of the total number of U.S. public firm patents in a given year.
cumulative products	The cumulative number of new product launches since the beginning of the Mukherjee et al. (2022) data set up to $t-1$.
fin const (KZ)	The Kaplan-Zingales index based on the five-factor model in Kaplan and Zingales (1997).
fin const (WW)	The Whited-Wu index from Whited and Wu (2006).
inventor move	An indicator variable that is one if the late-bloomer inventor moves to the user firm and zero otherwise.
ninv	Number of inventors in the citing firm.
inv_gender	An indicator variable equal to one for male inventors and zero for female inventors.
inv_npat	The total number of the patents that the inventor produced during the sample period.
inv_nfirms	The total number of firms that the inventor worked for during the sample period.
dlog(asset)	The difference in the 10-year log(asset) since the citing year between firms citing and writing the focal patent.
dlog(age)	The difference in the 10-year log(age) since the citing year between firms citing and writing the focal patent.
droa	The difference in the 10-year roa since the citing year between firms citing and writing the focal patent.
dleverage	The difference in the 10-year leverage since the citing year between firms citing and writing the focal patent.

Figure A1: Late-Bloomer patent example

The figures show Texas Instrument(TI)'s patent (#5,025,407) and its relation to Nvidia's late-bloomer patent hunting benefits. Panel (a) shows the front page bibliographic data and some exhibits from TI's patent grant. It contains the invention title, assignee names, backward citations, and an abstract. Panel (b) illustrates the forward citations that the patent #5,025,407 has received over the 20-year patent term since its grant year. Each mark represents a forward citation, and the letter inside refers to citing firms' initials. A square and circle each denote a citation made by a writer and a user, respectively. A circle without a letter denotes a citation made specifically by Nvidia, who is a patent hunter. A dotted circle is a non-US assignee. 2006 is the superstar-year, which is the year when the patent's cumulative citations reach the 95th percentile within its cohort of the same CPC class and grant year. Panel (c) presents Nvidia's stock prices over the 2000-2017 period. Panel (d) shows the video game industry revenues by segment over a similar period between 2002 and 2019.

United States Patent [19]
Gulley et al.

[11] **Patent Number: 5,025,407**
[45] **Date of Patent: Jun. 18, 1991**

[54] **GRAPHICS FLOATING POINT COPROCESSOR HAVING MATRIX CAPABILITIES**

[75] Inventors: **David W. Gulley; Jerry R. Van Aken,** both of Sugar Land, Tex.

[73] Assignee: **Texas Instruments Incorporated,** Dallas, Tex.

[21] Appl. No.: **387,459**

[22] Filed: **Jul. 28, 1989**

[51] **Int. Cl.**[5] **G06F 7/52**
[52] **U.S. Cl.** **364/754;** 364/736
[58] **Field of Search** 364/754, 736, 748, 518

[56] **References Cited**

U.S. PATENT DOCUMENTS

3,763,365	10/1973	Seitz	364/754
4,493,048	1/1985	Kung et al.	364/754
4,697,247	9/1987	Grinberg et al.	364/754
4,719,588	1/1988	Tatemichi et al.	364/754
4,878,190	10/1989	Darley et al.	364/752
4,916,651	4/1990	Gill et al.	364/736

OTHER PUBLICATIONS

Mokhoff, N., Graphics Chips Forge High-Res Boards for PCs, Workstations, *Electronic Design,* Mar. 17, 1988, Design Trade-Offs in High-End Graphics Board, *Electronic Design,* Mar. 17, 1988, pp. 77–84.

Foley, J. D., and A. Van Dam, *Fundamentals of Interactive Computer Graphics,* Reading, Mass: Addison-Wesley, 1982, pp. 245–265, 274–279, 297–302.

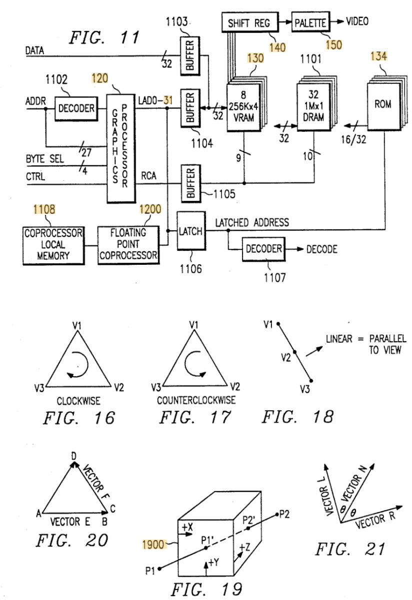
Newman, W. M., and R. F. Sproull, *Principles of Interactive Computer Graphics,* 2nd ed., New York: McGraw-Hill, 1979, pp. 57–60, 333–351, 491–501.

Primary Examiner—Dale M. Shaw
Assistant Examiner—Long T. Nguyen
Attorney, Agent, or Firm—James F. Hollander; James T. Comfort; Melvin Sharp

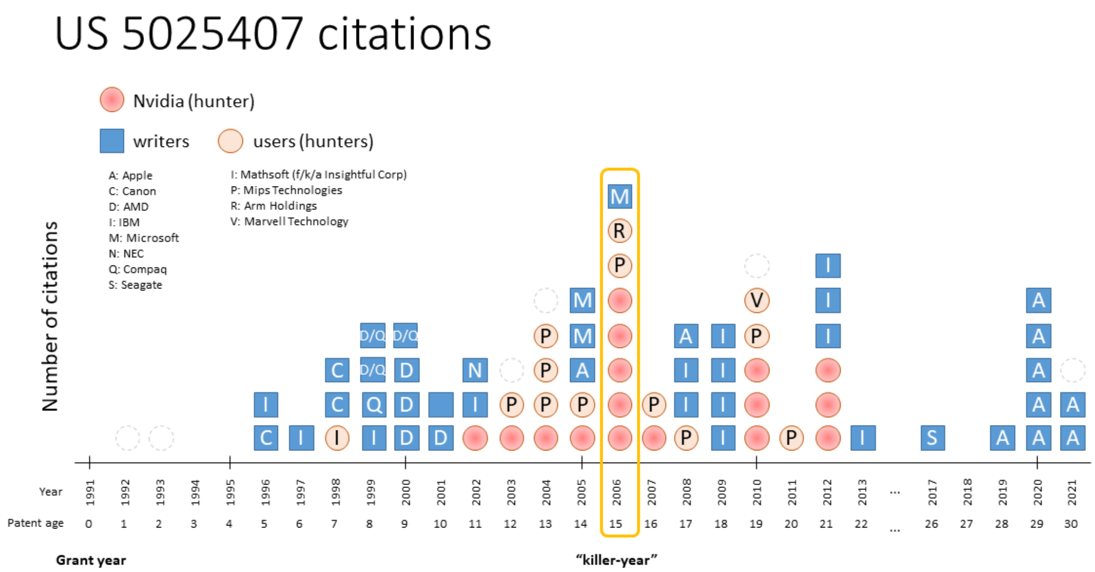
[57] **ABSTRACT**

A graphics coprocessor designed to work in conjunction with a host graphic processor in a graphics system. The coprocessor is adapted to perform arithmetic calculations including matrix calculations. The matrix size is such that the intermediate results require more registers than are practical to include in the coprocessor. This has been solved by arranging for certain selected ones of the intermediate results to continue within the program execution from stage to stage and avoiding intermediate storage.

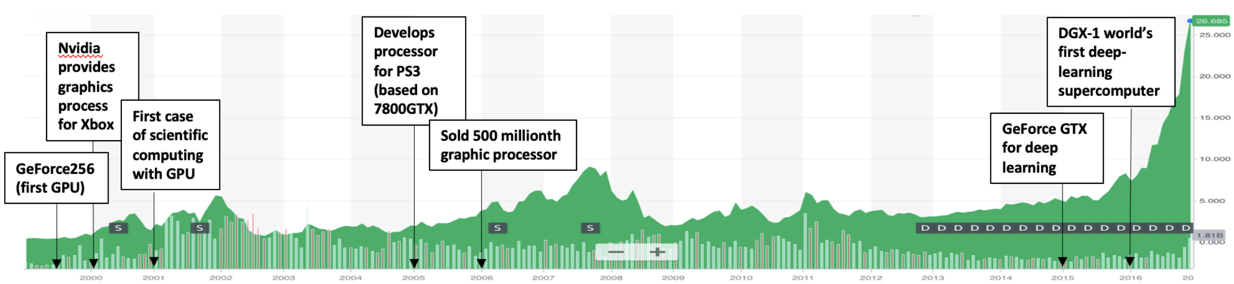
U.S. Patent June 18, 1991 Sheet 4 of 8 **5,025,407**



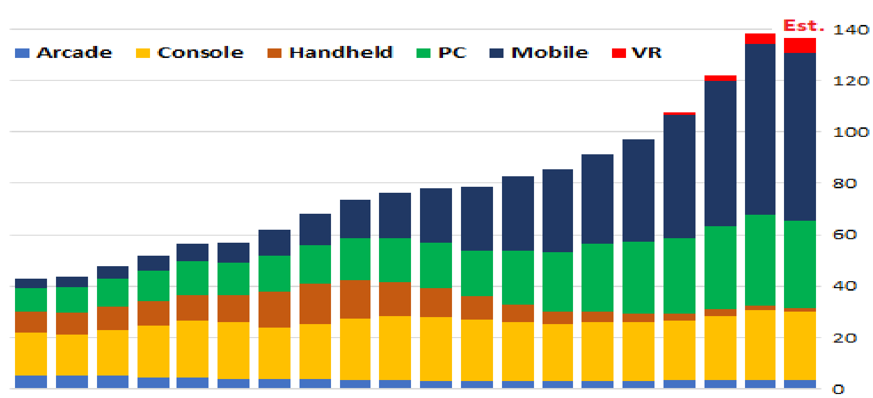
(a) Texas Instrument's Patent #5,025,407



(b) Patents Citing #5,025,407



(c) Nvidia Stock Prices (2000-2017)



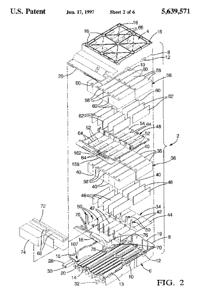
(d) Video Game Industry Revenues ($bn, 2002-2019)

Figure A2: Other hunting examples

Battery pack

DirecTV 1997, No. 5639571
Tesla citing extensively from 2012

A battery pack for easy access to, and uniform cooling/heating of, the individual battery modules thereof. The pack comprises stackable housing parts (i.e., top and bottom) housing multiple tiers of battery modules supported by underlying trays having openings/holes therein aligned with gaps/spaces between adjacent battery modules through which cooling/heating air is uniformly flowed in parallel between the modules from an underlying plenum. The battery modules are compressively immobilized in the housing by resilient foam pads which bear down on the tops of the modules.

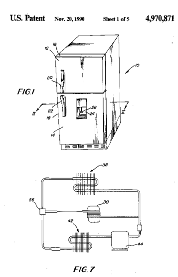


(a) Tesla cites DirecTV

Carbonator refrigeration system

Coca-Cola 1990, No. 4970871
Whirlpool citing extensively from 2000

A carbonator refrigeration system for use in a conventional refrigerator for dispensing a chilled carbonated liquid such as water or a beverage from the front door of the refrigerator. The system includes a compressor, an evaporator, a condenser, a carbonator and a valve member wherein the valve member is responsive to conditions detected within the refrigerator for selectively directing a source of cooling fluid to or away from a heat exchange device provided in connection with the carbonator. The carbonator refrigeration system enables cooling of the carbonator for home dispensing use in a time-share manner with the remaining mechanical refrigeration components.

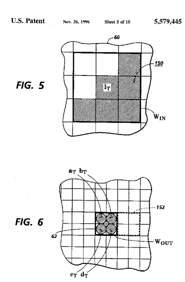


(b) Whirlpool cites Coca-Cola

Image resolution conversion method that employs statistically generated multiple morphological filters

Xerox 1996, No. 5579445
Adobe citing extensively from 2012

A method and apparatus for automating the design of morphological or template-based filters for print quality enhancement. A plurality of different phase, but same resolution, subsampled images are generated from training documents. Statistical data derived therefrom is then employed in an automated process to generate filters. The filters may be used for resolution enhancement and/or conversion of bitmap images. Furthermore, the statistical data is used to produce filters that are intended to not only optimize image structure, but image density as well.

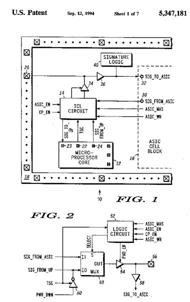


(c) Adobe cites Xerox

Interface control logic for embedding a microprocessor in a gate array

Motorola 1994, No. 5347181
Xilinx citing extensively from 2004

An interface circuit (14) that allows for a flexible three-way interface between a microprocessor (12), an ASIC cell block (16), and the external world has been provided wherein the microprocessor and the ASIC cell block are fabricated within a gate array (10). The interface circuit provides circuitry for each I/O pin (22, 23, 24) of the microprocessor to allow it to readily interface with the customer designed ASIC cell block or external devices via the ASIC I/O pads (20). The interface circuit also allows isolated testing of only the microprocessor, of only the ASIC cell block, or of both the microprocessor and the ASIC cell block.

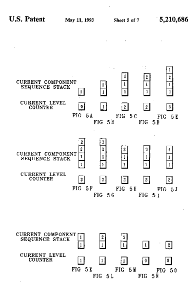


(d) Xilinx cites Motorola

Multilevel bill of material processing

IBM 1993, No. 5210686
Oracle citing extensively from 2002

A method and system for processing a multilevel bill of material contained in a relational database that does not require a pre-established limit on the number of levels that can be processed and minimizes user lock out from the same data. A control table keeps track of each component retrieved at a given level of the bill of material, tagging each table entry with a component item identifier, bill of material level, and component sequence number, which identifies the order in which components are processed at each level. A counter is used to keep track of the next level in the bill of material to be processed and a stack data structure is used to indicate the sequence number of the next component to be processed at a given level.

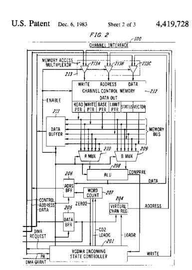


(e) Oracle cites IBM

Channel interface circuit providing virtual channel number translation and direct memory access

AT&T 1983, No. 4419728
Cisco citing extensively from 1996

The subject channel interface circuit functions to provide a high speed interface between a processor and a data link, which link carries data messages having virtual addresses. The message handler is programmable and serves to translate the header portion of the data message from a virtual address into a hardware memory address, which is used to activate a specific location in the processor memory. The data portion of the data message is then directly inputted to this memory location (i.e., DMA) and the appropriate file pointers are reset. When a complete file is received and stored in memory, the message handler generates a processor interrupt.



(f) Cisco cites AT&T

Bleaching composition

Procter and Gamble 1976, No. 3996152
Clorox citing extensively from 1990

Liquid dishwashing detergent containing anionic surfactant…

Procter and Gamble 1985, No. 4492646
Ecolab citing extensively from 2000

(g) Clorox/Ecolab cite P&G

Process for producing laminates of fabric and fluorocarbon copolymer

DuPont 1979, #4165404
3M citing extensively in 1996

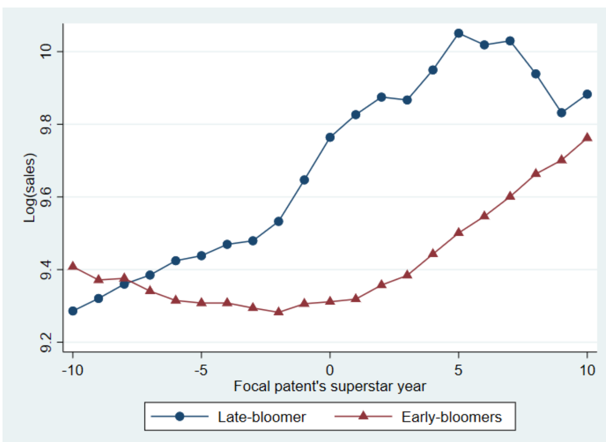
Extended address generating apparatus and method

Unisys 1984 No. 4453212
AMD citing extensively in 1996

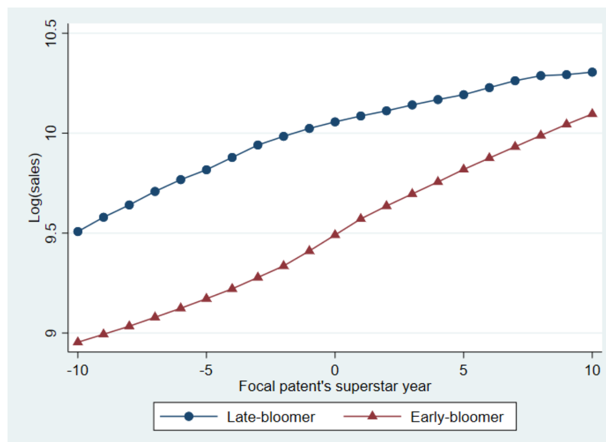
(h) 3M/AMD cite DuPont/Unisys

Figure A3: Financial outcomes of late-bloomer writers and users

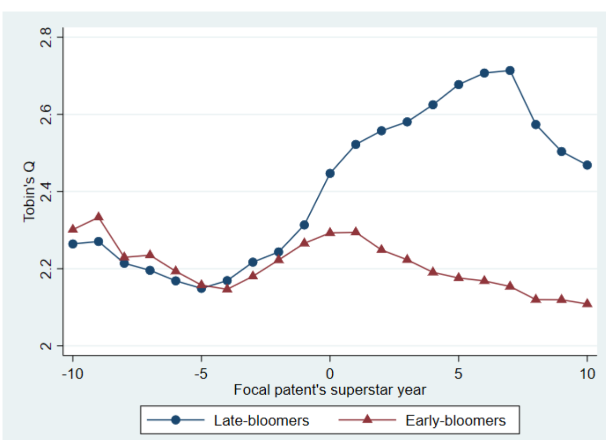
The figures replicates Figure 3 using financial data cumulatively since the citing year. The figures present sales and Tobin's Q of late-bloomer patent writers and users around the time when the late-bloomer patent becomes a superstar patent (*i.e.,* superstar-year: its cumulative citations reach the top 95th percentile within CPC class and grant-year cohort). The x-axis represents the year relative to the superstar-year of each late-bloomer patent and is zero at the superstar-year. A superstar patent is a patent that has ever reached the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. A late-bloomer patent is a patent that takes an excessively long time period before it becomes a superstar patent. We use the 90th percentile point in the time-to-superstar distribution (14 years) to define the excessively long time period. An early bloomer is a superstar patent that is not classified as a late-bloomer patent. Late-bloomer patent writers are the firms that have at least one late-bloomer patent during the sample period. Late-bloomer patent users are the firms that have never written a late-bloomer patent but have cited at least one late-bloomer patent.



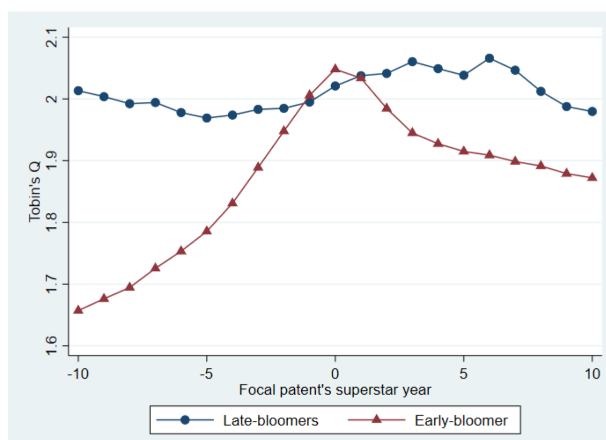
(a) User sales



(b) Writer sales



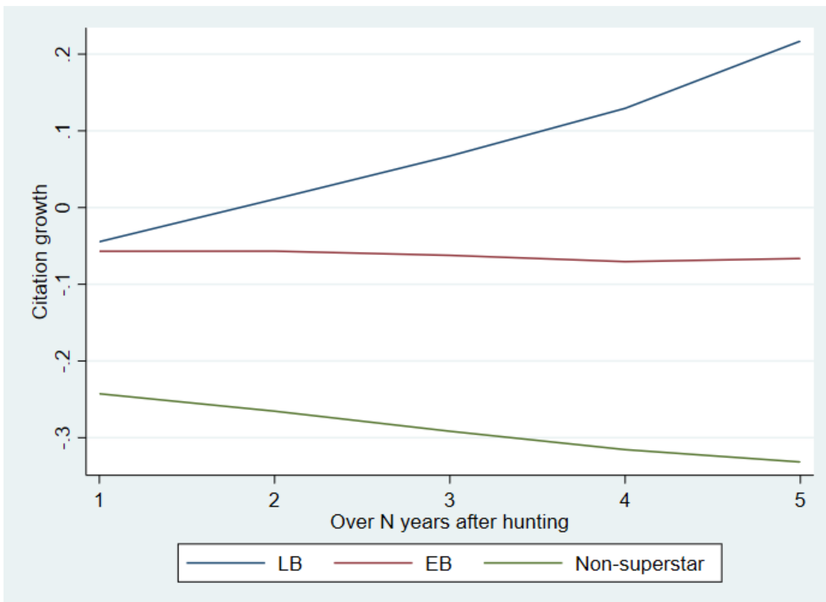
(c) User Tobin's Q



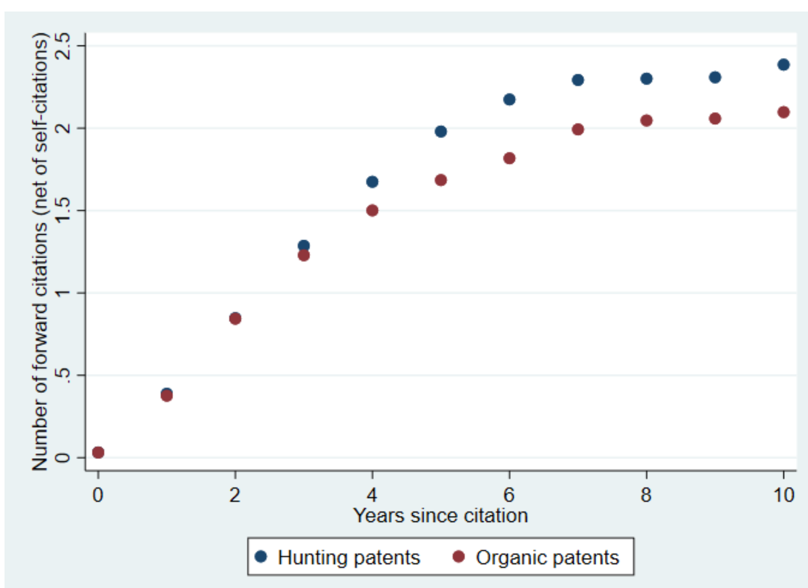
(d) Writer Tobin's Q

Figure A4: Citation growth of hunted patents

The figures show the forward citation growth by different types of user-cited patents and users' organic and hunting patent forward citations. In Panel A, the sample consists of all user-cited patents in our sample between 1976 and 2020. Citation growth is measured over the N years after the user citation. In Panel B, the sample consists of all hunters' patents citing at least one late-bloomer patent. The figure plots the number of forward citations received by hunters' organic and hunting patents over time. A hunter's patent is classified as an organic patent if the number of backward self-citations is larger than the number of late-bloomer citations and a hunting patent otherwise.



(a) Citation growth of hunted patents



(b) Citations of hunters' organic vs. hunting patents

Table A1: Sample selection

Our base patent sample consists of all USPTO patents granted between 1976 and 1999. The table describes our sample-selection procedure and method of classifying patents into superstar patents, late-bloomer patents, and early-bloomer patents with the number of observations in each group.

	Number of patents	Description
Base patent sample	1,712,247	All USPTO patents granted between 1976 and 1999. The sample period starts in 1976 due to the availability of data on patent assignees, inventors, claims, and other information from the PatentsView database. The sample period ends in 1999 as identifying a superstar patent requires 20 years since each patent's grant year. We also exclude approximately 0.45% of the remaining patents from the sample when they have no CPC information.
Superstar patents	213,772	A superstar patent is the patent that has ever reached the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant.
Late-bloomers	21,960	A late-bloomer patent is a patent that takes an excessively long time period before it becomes a superstar patent. We use the 90th percentile point in the time-to-superstar distribution (14 years) to define the excessively long time period.
Early-bloomers	191,812	An early bloomer is a superstar patent that is not classified as a late-bloomer patent.

Table A2: Benefits from patent writing

The table presents results from the regressions that examine financial benefits to patent users. The observations are at the firm-year level for the period of 1976 to 2020. *Sales growth* is the five-year sales growth, computed as $(sales_{t+4}/sales_t) - 1$, and $Avg\ Tobin's\ Q$ is the five-year average of Tobin's Q over t to $t+4$. *LBhunting* and *EBhunting* are the numbers of hunted late-bloomer and early-bloomer patents in a given year, respectively. *log*(*LBwriting*) and *log*(*EBwriting*) are standardized. The regressions include the firm fixed effects and year fixed effects. Standard errors in parentheses are clustered at the firm level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

	Writer benefits		User benefits	
	Sales growth	Avg Tobin's Q	Sales growth	Avg Tobin's Q
	(1)	(2)	(3)	(4)
log(LBwriting)	0.00841	0.0199***		
	(0.0104)	(0.00754)		
log(EBwriting)	-0.00150	0.0707***		
	(0.0211)	(0.0126)		
log(LBhunting)			0.0553***	0.0437***
			(0.0183)	(0.0112)
log(EBhunting)			-0.129***	-0.0292**
			(0.0243)	(0.0116)
log(asset)	-1.036***	-0.286***	-0.757***	-0.309***
	(0.0511)	(0.0173)	(0.0366)	(0.0139)
log(age)	-0.475***	-0.219***	-0.443***	-0.225***
	(0.0511)	(0.0184)	(0.0350)	(0.0155)
roa	-2.072***	-0.415***	-1.837***	-0.251***
	(0.217)	(0.0522)	(0.149)	(0.0435)
leverage	-1.108***	-0.0693	-0.940***	0.00403
	(0.182)	(0.0625)	(0.139)	(0.0557)
Mean	4.172	1.924	0.901	2.080
$H_0 : LB = EB$ (p-value)	0.676	0.000	0.000	0.000
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Observations	45749	58030	75589	98776
Adjusted R^2	0.441	0.815	0.350	0.719

Table A3: User benefits to hunting late-bloomer patents (alternative comparison groups)

The table presents results from the regressions that examine benefits to patent users in general relative to corresponding patent writers. The sample consists of all patents and their citing patents. The observations are at the focal-citing patent pair and year level for the period of 1976 to 2020. The non-superstar patents comprise 100,000 randomly selected patents for a comparable sample size with superstar patents. The dependent variable is the difference in the outcome variable between firms citing and writing a focal patent. We consider the difference-in-differences analyses of the dependent variable between late-bloomer patents and early-bloomer patents (Panel A) and late-bloomer patents and non-superstar patents (Panel B). *ssyear* is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. For a non-superstar patent, *ssyear* is the year with the greatest number of citations during 20 years since its grant. ssyear_{post} is an indicator variable equal to one if the year is after *ssyear* of a given patent and zero otherwise. The regressions include the superstar-year fixed effects. Standard errors in parentheses are clustered at the firm level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

Panel A: Late-bloomer vs. early-bloomer

	Diff(Sales growth)	Diff(Tobin's Q)
	(1)	(2)
latebloomer × ssyear_{post}	0.0130***	0.276***
	(0.00294)	(0.0336)
latebloomer	-0.00158	-0.0629*
	(0.00248)	(0.0380)
ssyear_{post}	0.00673***	-0.0280**
	(0.00105)	(0.0112)
Control variables	Y	Y
Superstar-year FE	Y	Y
Controls	Y	Y
Observations	10428295	10687503
Adjusted R^2	0.051	0.027

Panel B: Late-bloomer vs. non-superstar

	Diff(Sales growth)	Diff(Tobin's Q)
	(1)	(2)
latebloomer × ssyear_{post}	0.0143***	0.210***
	(0.00283)	(0.0299)
latebloomer	0.00444**	0.133***
	(0.00209)	(0.0226)
ssyear_{post}	0.00564***	0.0254*
	(0.00129)	(0.0140)
Control variables	Y	Y
Superstar-year FE	Y	Y
Observations	2115307	2167795
Adjusted R^2	0.039	0.043

Table A4: Is the benefit to hunting prevalent?

The table presents results from the regressions that examine benefits to patent users in general relative to corresponding patent writers. The sample consists of all patents and their citing patents. The observations are at the focal-citing patent pair and year level for the period of 1976 to 2020. The non-superstar patents comprise 100,000 randomly selected patents for a comparable sample size with superstar patents. The dependent variable is the difference in the outcome variable between firms citing and writing a superstar patent. We consider the difference-in-differences analyses of the dependent variable between early-bloomer patents and non-superstar patents (Panel A) and superstar patents and non-superstar patents (Panel B). *ssyear* is the year when a given patent becomes a superstar patent by reaching the 95th percentile of the distribution of cumulative forward citations (net of self-citations) within its cohort of the same CPC class and grant year during 20 years since its grant. For a non-superstar patent, *ssyear* is the year with the greatest number of citations during 20 years since its grant. ssyear_{post} is an indicator variable equal to one if the year is after *ssyear* of a given patent and zero otherwise. The regressions include the superstar-year fixed effects. Standard errors in parentheses are clustered at the firm level. ***p<0.01, **p<0.05, *p<0.1. See Appendix A for variable definitions in detail.

Panel A: Early-bloomer vs. non-superstar

	Diff(Sales growth)	Diff(Tobin's Q)
	(1)	(2)
earlybloomer × ssyear_{post}	-0.0000402	-0.0406***
	(0.00157)	(0.0148)
earlybloomer	0.00661***	0.178***
	(0.00157)	(0.0220)
ssyear_{post}	0.00696***	0.0138
	(0.00124)	(0.0132)
Control variables	Y	Y
Superstar-year FE	Y	Y
Observations	9066974	9283734
Adjusted R^2	0.053	0.021

Panel B: Superstar vs. non-superstar

	Diff(Sales growth)	Diff(Tobin's Q)
	(1)	(2)
superstarpat × ssyear_{post}	0.00247	-0.00813
	(0.00156)	(0.0148)
superstarpat	0.00668***	0.196***
	(0.00144)	(0.0167)
ssyear_{post}	0.00751***	0.0501***
	(0.00126)	(0.0135)
Control variables	Y	Y
Superstar-year FE	Y	Y
Observations	10805288	11069516
Adjusted R^2	0.051	0.026