Financial Innovation in the 21st Century: Evidence from U.S. Patenting[1]

Josh Lerner Amit Seru Nicholas Short Yuan Sun

August 7, 2020

Patents provide a window into the evolution of financial innovation over the past two decades. We first highlight the growth and importance of these awards in the 21st century. The dramatic surge in financial patenting has been driven largely by information technology and other non-financial firms, while banks and other financial institutions have narrowed the scope of their innovations. These shifts have been an accompanied by a decline in the links between patent awards and academic knowledge, and shifts in the geographic location of innovation, especially within incumbent firms.

[1] jlerner@hbs.edu; aseru@stanford.edu; nshort@g.harvard.edu; ysun@hbs.edu. All authors are affiliated with Harvard University except for Seru, who is at Stanford University. Lerner and Seru are affiliates of the National Bureau of Economic Research. Harvard Business School's Division of Research and Doctoral Programs provided financial support for this project. We thank Howell Jackson, Adam Jaffe, Andy Toole, and seminar participants at Stanford Law School for helpful comments. Rick Townsend kindly allowed us to use the patent-venture capital mapping. Patrick Clapp, Amin Gulamali, Stephen Moon, Mahlon Reihman, Kathleen Ryan, Rhys Sevier, and James Zeitler provided excellent research assistance. Lerner has received compensation for consulting with financial institutions. All errors and omissions are our own.

1. Introduction

The extent and consequences of innovation in the finance industry has been a subject of intense skepticism since the global financial crisis (GFC). These doubts were underscored by Philippon's (2015) evidence that the cost of financial intermediation had been constant over the past 130 years, despite the dramatic increase in assets under management and advances in information technology. The extremely low levels of reported R&D and modest measured productivity gains in this sector reinforce suspicions that social welfare-enhancing innovation in this sector have been rare.[2]

Reflecting these doubts, recent academic studies have focused on the development and diffusion of what Krugman (2007) terms the "alphabet soup of C.D.O.'s and S.I.V.'s, R.M.B.S. and A.B.C.P." A rich empirical literature has examined the frequently deleterious consequences of the innovative financial securities in the run-up to the GFC.[3] Similarly, theoretical work has highlighted concerns about financial innovation and its impact: important papers include Biais, Rochet, and Woolley (2015), Caballero and Simsek (2013), and Thakor (2012), as well as the prescient arguments of Rajan (2006). In one such paper, Gennaioli, Shleifer, and Vishny (2012) adapt the standard depiction of financial innovators catering to investors' desire for certain cash flow patterns (Allen and Gale, 1994; Ross, 1976). In their variant, financial intermediaries meet investor tastes by engineering securities perceived to be safe but actually riven with hidden risks. When these weaknesses subsequently become apparent, the valuations of the novel securities drop sharply and social welfare suffers.

At the same time, practitioner accounts suggest that the rise of smart mobile devices, cloud-based computing, and artificial intelligence have had important implications for finance. Various applications of these technologies (from mobile payments to blockchain) have been the focus of intensive investments by major financial institutions, large technology companies, and start-ups. Many venture capital groups have adopted a strategy of financing companies that seek to supplant incumbents in established industries, and finance is no exception. The rapid growth in venture capital activity in the financial sector during the 2010s can be gleaned from AngelList's database of young firms that are currently or were formerly seeking financing from angels or venture investors. In the database as of January 2020, 1043 ventures are associated with the "payment" keyword, 888 with "investment," 545 with "blockchain," and 335 with "trading."

Despite the intense interest in financial innovations and its consequences, we know remarkably little about where or by whom these new products and services are developed. This paper seeks to address this gap, focusing on finance patents filed between 2000 and 2018 and awarded by

---

[2] For instance, in 2016, the finance and insurance sector spent 0.17% of total revenue on R&D, as opposed to 13.5% for pharmaceuticals, 10.7% for computers and electronic products, and 3.4% for manufacturing as whole (based on calculations by Kung (2020), which in turn used, among other sources, https://ncses.nsf.gov/pubs/nsf19318/). Industry productivity for the financial sector in major Western nations is compiled in OECD (2019). Both productivity and R&D measures in service sectors may be problematic for reasons analyzed in Baily and Zitzewitz (2001) and National Research Council (2005).

[3] A selected list of papers would include Chernenko and Sunderam (2014), Fostel and Geanakoplos (2012), Keys et al. (2010), and Simsek (2013).

February 2019. These awards allow us to examine in what areas and by what organizations financial innovation was pursued, and how these trends evolved over time. The data's richness also permits us to explore the drivers of shifts in new product development, given the widespread consensus that innovation responds to shifting demand and regulatory conditions (Acemoglu and Linn, 2004; Finkelstein, 2007). As discussed in Section 2, while not all financial innovations receive patents, and the legal treatment of these awards has shifted over time, patents provide a valuable window into the nature of financial innovation.

We start by highlighting three key facts that seem to contradict the consensus summarized in the first two paragraphs of this paper (the details of the data set construction are described below):

- *Financial innovation is substantial.* As Figure 1 depicts, the volume of financial patent awards (the red line) in the U.S. has surged, from very low levels in the early 2000s to between 0.4% and 1.1% of all grants subsequently. While this level was still modest compared to the finance and insurance's share of GDP (7.6% in the third quarter of 2019[4]), financial patents in the period under study totaled 24 thousand awards.
- *Financial innovation encompasses many areas in addition to novel security design.* As Panel A of Figure 2 depicts, the bulk of the awards were not in areas related to commercial or investment banking, insurance, or wealth management. Rather, they were dominated by payments and various supporting back-office technologies.
- *Financial innovators extend far beyond commercial and investment banks.* As Panel B suggests, banks and other financial institutions represented a modest share of the awards during this period, with information technology (IT) companies dominating.

Section 2 discusses patents as a measure of financial innovation and addresses concerns that finance patents might be inconsequential. We describe the construction of the dataset in Section 3.

We then turn to exploring the dynamics of financial innovation in the 21st century. We highlight that the identity of financial innovators has changed. Section 4 explores the trends in financial patenting through a decomposition analysis across patent types, assignee characteristics, and time. Both the subject matter of financial patents and the mixture of firms pursuing these innovations has changed sharply. Most dramatically, the surge in financial patenting highlighted in Figure 1 was driven by U.S. information technology firms and those in other industries outside of finance, as well as by patents on topics other than banking.

A further decomposition indicates that banks and payment firms have focused more intensely on their core areas, while IT firms and other financial firms continued to patent widely in finance. Several explanations can be offered for this seeming failure on the part of financial institutions to expand the subject matter in which they innovate. The need to support financial incumbents' existing lines-of-business may make it optimal for these firms to focus on core areas. Alternatively, banks and other financial institutions may have wanted to shift their areas of focus, but their limited organizational and technological skill-sets or pressures from financial regulators may have deterred their ability to do so.

[4] https://fred.stlouisfed.org/series/VAPGDPFI.

We then highlight two additional changes in the financial innovation landscape. The first is the apparent "fishing out" of the relevant academic knowledge base, which appears to have most adversely affected innovation in the banking sector. Section 5 documents these dynamics. Over the sample period, citations in finance patents were associated with more impactful patents, an effect that held for academic citations in general, as well as those to articles in business, economics, and finance journals specifically. The relationship between citations and patent impact was stronger historically for banks and other financial institutions than for information technology firms.

Over time, however, the relationship changed. The number of citations in patents to academic papers fell. This shift had the most dramatic consequences for banks, which had the strongest relations between patent value and academic prior art. Moreover, the academic work being cited became progressively older. These findings were consistent with the depictions by Bloom et al. (2020) and Jones (2005) of the "burden of knowledge," whereby relevant knowledge becomes harder to come by, and the rate of innovation suffers as a result.

Section 6 highlights a second parallel shift, a shift in the geography of financial innovation. In particular, we document shifts across metropolitan areas in financial innovation, most dramatically the rise of the greater San Francisco region and the decline of the New York area.

We show that this pattern is largely due not to the exit and entry of financial innovators, but rather to the shifts in the locus of innovation by firms that continue to innovate. Not all firms are equally likely to switch their hub of innovative activity: the shifts seem to be concentrated among IT firms, those firms under the active control of a venture capitalist, and those located in historical hubs of venture capital financing of financing activity other than New York and San Francisco. We posit that this result may reflect the lower costs that these firms face when they seek to shift their innovative activities.

In the final section, we argue that these findings suggest two broad conclusions. First, financial innovation is far more complex and richer than has been depicted in the academic literature to date. The extent to which finance patenting has been increasingly dominated by firms outside the finance industry is striking. So is the importance of payments technologies, as well as back-office functions such as security and communications.

Second, the results illustrate the stresses that traditional financial institutions have faced in terms of maintaining pace in innovation. This declining share of financial institutions in patents granted largely reflects choices that the firms made, doubtless driven by existing product and service offerings and regulatory concerns. But the results also hint at additional factors that may have exacerbated these declining share of innovation by banks, especially the decreased relevance of contemporaneous academic discoveries. These patterns were consistent with the argument of Philippon (2019) regarding the impediments to innovation by incumbent banks, and the potential for breakthroughs by new fintech entrants.

While these analyses cannot ultimately address the questions regarding social welfare delineated in the opening paragraphs, they suggest that the nature of financial innovation is richer and more

ambiguous than the finance literature to date has suggested. In the final section, we also discuss some of the opportunities for future research.[5]

2. Patents as Indicators of Financial Innovation

2.1 A Historical Perspective

The financial services industry has historically differed from the bulk of manufacturing industries with regard to the ability of innovators to appropriate their discoveries. There has long been ambiguity about the patentability of financial discoveries in the United States. At least since a 1908 court decision established a "business methods exception" to patentability,[6] many judges and lawyers have presumed that business methods were not patentable subject matter. While the USPTO issued patents on financial and other business methods during the twentieth century, many observers questioned their validity. Another concern limiting patenting was that it was very difficult for firms to detect infringement of their pricing and trading-related patents. (Of course, the same considerations also affected the decision to file process patents in other industries.)

Consequently, awardees were reluctant to incur the time and expense to file for awards. Instead, new product ideas diffused rapidly across competitors (Tufano, 1989). But a theoretical analysis by Herrera and Schroth (2011) argued that even when inventions could not be patented, investment banks had considerable incentives to develop new products.

As a result, patents traditionally only provided a limited guide to innovative activity in finance, in contrast to other fields (Griliches, 1990). This disparity was highlighted in Lerner (2002), who documented that between 1971 and 2000, only 445 financial patents were issued by the USPTO. These represented less than 0.02% of all awards during this period. A disproportionate share of these awards were made to individual inventors. Academic research, while highly relevant to many of these patents, was rarely cited or identified by the patent examiners.

Attitudes toward business method patents changed with the July 1998 appellate decision in *State Street Bank and Trust v. Signature Financial Group*. This case originated with a software program used to determine the value of mutual funds, on which Signature had obtained a patent in 1993. State Street Bank sued to have the patent invalidated on grounds that it covered a business method and hence was not patentable. While State Street's argument prevailed in the district court, the Court of Appeals for the Federal Circuit (the central appellate court for patent cases, also known as the CAFC) reversed the finding. The court affirmed the patentability of financial software that

---

[5] This paper relates to a number of earlier works on the empirics of financial innovation, reviewed in Frame and White (2004) and Lerner and Tufano (2011). The most closely related pieces are Lerner (2002), discussed in Section 2, and Lerner (2006), which examines *Wall Street Journal* stories on financial innovation between 1990 and 2002. Another related paper is Beck et al. (2016), which relates data on banks' R&D expenditures from an OECD innovation survey across 32 countries to measures of growth and bank fragility. This paper focuses on innovations developed since 2000, when the patent record provides a far richer depiction of financial innovation than any previously employed metrics.

[6] *Hotel Security Checking Co. v. Lorraine Co.*, 160 F. 467 (2d Cir. 1908).

valued mutual funds since it produced a "useful, concrete, and tangible result."[7] The Supreme Court declined to hear State Street's appeal in January 1999.

State Street thus established that business methods were statutory subject matter on an equal playing field with more traditional technologies. Numerous trade press articles interpreted the case as unambiguously establishing the patentability of business methods. While this decision was refined in important subsequent rulings such as *Bilski v. Kappos* and *Alice Corp. v. CLS Bank* (discussed in Appendix A), it nonetheless represented a sharp discontinuity.

Conversations with practitioners actively involved in prosecuting finance patents suggest that the historical differential between patenting in finance and in other technological domains has narrowed considerably in recent decades. In addition to the greater (though not iron-clad) confidence in the enforceability of finance patents, two factors have contributed to this change in practice. One reason is that greater regulatory disclosures and more public scrutiny has made it hard to keep discoveries secret. In these settings, the disclosure associated with patent awards may be less problematic. A second reason has the emergence of fintech firms that are not vertically integrated. Since these new firms cannot capture the returns from their inventions directly, they regularly file financial patents. These filings have in turn spurred many incumbents who did not traditionally patent to protect their innovations.

2.2. *Empirical Evidence on Financial Patent Quality*

The qualitative discussion above suggested that the mapping between financial innovations and patenting has become closer. These arguments are borne out in three preliminary empirical analyses.

The first analysis looks at the extent to which patent awards were scrutinized by USPTO. As noted above, Lerner (2002) suggested that the pre-*State Street* awards were subject to ineffective reviews. To examine the quality of review in the 21st century, we create a sample of U.S. patents filed after 2000 and awarded through February 2019 whose original applications were published by the U.S. We compare the crucial independent claims in the applications and awards, and determine the extent to which they number and length of these claims were modified during the review process, following the methodology of Marco et al. (2019).[8]

---

[7]In particular, the court held "... that the transformation of data, representing discrete dollar amounts, by a machine through a series of mathematical calculations into a final share price, constitutes a practical application of a mathematical algorithm, formula, or calculation, because it produces 'a useful, concrete and tangible result'–a final share price momentarily fixed for recording and reporting purposes and even accepted and relied upon by regulatory authorities and in subsequent trades." See *State Street Bank and Trust v. Signature Financial Group*, 149 F.3d 1368 (Fed. Cir. 1998).

[8] An independent claim "is a standalone claim that contains all the limitations necessary to define an invention" (https://www.uspto.gov/sites/default/files/documents/Website%20PDF%20-%20Invention%20Con%202017%20Claim%20Drafting%20Workshop%20-%20OPLA.pdf) and as such, are the most important such rights granted. Not all patents have published applications: for instance, those applications only filed in the U.S. are often not published

Table 1 presents a comparison of 2.6 million non-finance patents and almost 16 thousand finance ones. The analysis shows that finance patents are more likely to have the number of independent claims reduced than non-finance patents (by one-half, rather than one-third, of an independent claim) and to have the shortest independent claim lengthened (by 84 words, as opposed to 49).[9] Both of these results are consistent with more intensive scrutiny of finance patents during the past two decades.

The differentials over time are illustrated graphically in Panels A and B of Figure 3. In both cases, since mid-2000s, finance patents have been associated with greater revisions of awards.

The second analysis looks at the relative impact of patent awards. Table 1 examines all finance and non-finance patents filed between 2000 and 2018, and awarded by February 2019, using two leading measures of patent impact. The first of these were the subsequent citations (through October 2019) that the patent garnered. Because the propensity to cite patents varied across technologies and over time, we normalized the citations by the mean number received by other patents in that four-digit CPC class and awarded in the same quarter. The second impact measure was the Kogan et al. (2007) estimate of patent value, based on market reactions to the award grants. The latter measure can only be calculated for publically traded firms, of course.

Using the citation measure, finance patents were on average 25% more impactful than the typical award. Using Kogan et al. (2017) market values, the finance patents were four-and-a-half times more valuable. Again, this result is inconsistent with thee awards being trivial discoveries devoid of economic value.

Finally, we look at whom is filing the finance patents. We examine the identity of the assignees of all patents applied for between 2000 and 2018 and awarded by February 2019. We use the classification of assignees provided by the USPTO, and assume that all unassigned patents are awarded to individuals.

Table 3 shows that 8.6% of finance patents since 2000 were assigned to individuals. This share can be contrasted with the 27% in the pre-*State Street* sample of Lerner (2002). (While the overall share of patents granted to individuals has dropped over time, the differential between finance and other patents in the individual share of awards is less than 1% in patents filed in 2000 and later. This gap was 10% in the earlier period.) Since many of the most problematic patents in the earlier era were those of individual inventors, this result is again consistent with the suggestion that patent awards filed in recent decades may provide a valuable window into changing trends in financial innovation more broadly.

(https://www.uspto.gov/web/offices/pac/mpep/s1122.html#d0e120159). We determine the count of length of independent claims in issued patents using the Patentsview database. Due to the difficulty in obtaining the claim text in application publications, we only use the applications analyzed by Marco et al. (2019) and archived at https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-claims-research-dataset.

[9] In patent claims, patentees generally strive to have the broadest claims, i.e., those with the fewest limitations. An increase in claim length is thus often associated with a narrowing of claim breadth.

3. Construction of a Financial Patent Dataset

3.1 Identification of Financial Patents

The first step in the construction of our dataset was to develop an approach for identifying a "financial patent." Social scientists have generally relied on three types of information when classifying patents: the patent's technological classification code, the firm to which it was initially assigned (usually the inventor's employer), and/or keywords from some subset of the patent text, such as the title or abstract.

Each approach has advantages and disadvantages. Classification codes, for example, are created to help patent examiners identify prior art and often evolve in a somewhat piecemeal fashion. As a result, the codes do not necessarily map into broad technological categories like "finance." For example, while most finance patents are classified under the current system within G06Q 40 (Finance; Insurance; Tax strategies; Processing of corporate or income taxes), a substantial number of blockchain and cryptocurrency patents are classified within H04L 09 (Cryptographic mechanisms or cryptographic arrangements for secret or secure communications).

Another problem with identifying financial patents by classification code is that the U.S. changed from the U.S. Patent Classification (USPC) to Combined Patent Classification (CPC) scheme in January 2013, during our main period of interest. The USPTO offers a concordance between CPC and USPC codes. However, this crosswalk is based on an unpublished statistical association between the old and new codes. As a result, CPC codes for patents issued before January 2013 are essentially imputed and may contain inaccuracies. Moreover, the USPTO stopped using USPC codes in 2015, so the use of those codes would limit our study and exclude recent technologies like blockchain.

Alternatively, we can identify financial firms using published lists of fintech firms, such as the Forbes 100, the KPMG 50, or the CB Insights Fintech 250, and assume that the patents held by these firms are all financial patents. For firms in the start-up phase this may be reasonable, but as firms grow larger and potentially expand into multiple lines of business, it no longer make senses to assume that all of their issued patents are in finance. For example, a subsidiary of payments firm Square, Weebly, holds several patents. But Weebly is a website builder, rather than a financial company, and thus the bulk of their awards are associated with web site design and manipulation. Thus, it would be incorrect to assume that patents held by Square and its subsidiaries are financial patents. A similar issue surfaces when considering patents owned by established financial institutions. Thus, this approach might bias the sample of financial patents in unpredictable ways.

Finally, we can use Google BigQuery to execute SQL queries for certain keywords across the corpus of all published U.S. patent documents, using the IFI Claims patent data. We thus can generate a suitable set of keywords predictive of "financial" status—for example, some form of the word "finance"—and search for those keywords across all patents. As above, the main challenge here is to identify a suitable set of keywords without arbitrarily picking words that might bias the sample towards specific examples of financial innovation (like cryptocurrency) known to the researcher. Another challenge is to identify words that have high specificity and will not pick

up too much noise (e.g., patents that use some form of the word "finance" but are not financial patents).

Of course, one can also use any combination of the sets of financial patents produced from each of these three techniques, like (A U B) ∩ C (Hall and MacGarvie, 2010). However, without extensive auditing, we cannot easily identify the best combination of techniques, nor can we evaluate how well these various combinations eliminate or reduce inherent bias in the merged dataset.

We broke with prior literature by employing supervised machine learning (ML) techniques to develop an algorithm for appropriately classifying patents as "financial" (treatment) or "not financial" (control), based on each patent's features. As with any standard supervised machine learning, one must first choose a way to label the training set of patents. Based on our survey of existing classification techniques above, we elected to use CPC codes, under the belief that the codes would allow us to label a large sample of financial patents with relatively high accuracy. We chose CPC over USPC codes to enable future work and comparisons (as patents today and in the future are only classified using the CPC scheme). We experimented with various feature sets—the patent text, inventors, assignees, and the CPC codes of backward citations—before settling on the patent text and inventor names as the two feature sets which produce, in combination, the highest and most balanced levels of accuracy.

To determine which CPC codes might allow us to label a set of financial patents, we first looked at the USPTO's concordance file for the financial patent classes analyzed in Lerner (2002) (former USPC class 705, subclasses 35-38). We determined that CPC groups G06Q 20 and G06Q 40 broadly capture what we consider to be financial patents. Patents in G06Q 20 involve significant data processing operations and generally relate to payment architectures, schemes, or protocols, while those in G06Q 40 generally relate to finance, insurance, tax strategies, and the processing of corporate or income taxes. Patents with a primary CPC code (note the USPTO typically places patents into one primary and multiple secondary categories) in these two groups constituted our treatment set (set A).

Within subclass G06Q, we excluded groups 10 and 30, as those groups cover data processing systems or methods specially adapted to administrative or managerial purposes (group 10) and electronic commerce (group 30), categories that are not financial in our view. We also excluded group 50 and all subsequent groups, as they either do not involve significant data processing steps or cover technologies specially adapted for certain non-financial industries or technologies outside of our view of finance (e.g. business processing using cryptography). Patents with a primary CPC subclass in G06Q but not in groups 20 or 40 constituted our control set (set B).

Next, we merged our treatment set and control set, then bifurcated the data into a training set with 70% of the data and a testing set with 30% of data. Then we applied natural language processing techniques to each patent's text and the inventor names. When we first experimented with this approach, we used patent titles and abstracts for the patent text, but neither of these textual sources

produced models with suitable accuracy.[10] Our initial model runs produced high sensitivity (also called the true positive rate, the proportion of actual positives correctly identified as such) of about 98 percent. But the specificity (the true negative rate, the proportion of actual negatives that are correctly identified as such) was very poor: about 30 percent. We therefore elected to use each patent's entire written description, as the much richer set of language features obtained from the written descriptions produced much better results. With the entire written description as features, we obtained 91 percent sensitivity and 85 percent specificity.

Figure A-1 in the Appendix depicts how we applied the standard supervised machine learning process to predict financial patents.

We then repeated a similar natural language processing procedure for other features of interest, in addition to the written text. We also generated feature sets of the prior art cited in each patent, the names of the firms to which the patent was initially assigned, and the names of the inventors. When we applied each model to the test data, we found that the text model was the most accurate, followed by the inventor model. The prior art and assignee models could not improve accuracy beyond what could be achieved with the text and inventor models. Compared to the text-only model, the text-inventor model slightly decreased sensitivity from 91.3 to 89.9 percent (a drop of 1.4 percentage points), but significantly improved specificity from 85.3 to 90.0 percent (an increase of 4.7 percentage points). As a result, our new model generated false positives and false negatives at about a similar rate. This low rate (10 percent) was a tremendous improvement compared to our initial model.[11] The structure of our model is presented in Figure A-2.

We then deployed the model to capture financial patents outside G06Q by applying it to other "supplemental" classifications where some financial patents might reside. After analyzing all patents that had "any" (but not a "primary") classification in G06Q groups 20 or 40, we found that nearly 80% of those patents had a "primary" subclass in nine other categories that we had not considered (G06F, G06K, G07C, G07F, G07G, H04L, H04M, H04N, and H04W). There were 12,010 such patents. Our next step was therefore to generate text and inventor feature sets for these patents, and apply our text-inventor model to that data to predict which could be financial. This process identified 6,777 of those patents as financial. The final data set of financial patents thus consisted of 17,511 patents with a primary CPC group in G06Q 20 or 40 plus an additional 6,777

[10] Intermediate steps included the removal of extra blank spaces, the converting of accented characters to ASCII characters, the removal of non-English characters, the removal of stop words, the stemming of each word, and the lowercasing of the text. (Stop words are very common words such as "we" or "are," which do not provide necessary differentiable information for machine learning classifiers.)

[11] Our initial strategy was to adopt a stacking technique, an ensemble learning method that has the potential to improve further the classification accuracy but requires the combination of multiple classification models via a meta-classifier. After experimenting with different types of stacking architecture, we settled on the use of a Naive Bayes model for the patent description text, and a Logistic Regression model for the inventor names (Jurafsky and Martin, 2019, chapters 4-5). A concise "sum up" text-inventor model was adopted, in which a patent was predicted to be financial if either the text model or the inventor model makes such a prediction.

patents in the nine subclasses listed above that were predicted by the model to be financial, for a total of 24,288 patents.

To verify the quality of the ML model, we audited the results. Appendix B describes the auditing process in more detail.

3.2 *Joining with Other Data Sets*

After generating a list of financial patents and auditing the results of our ML models, we then obtained additional information about the financial patents and the firms to which our financial patents were assigned.

The first step in our process was to obtain additional patent-level data on financial patents from Derwent. Such information included the publication date, inventor names, assignee names, and abstract. We noticed one discrepancy in the assignee field when comparing the Derwent data and the IFI Claims patent data (accessed through Google BigQuery), but determined that the discrepancy could be readily addressed after auditing (see Appendix B).

We also obtained from Patentsview the patent assignee type (corporation, government, or individual, divided by domestic or foreign),[12] the number of forward citations, and the geographical location of the first-named inventor. We included all patents applied for between January 2000 and December 2018 and issued by February 2019, and citations through October 8, 2019.[13]

We then matched firms listed as assignees on financial patents to Capital IQ firms, so that we could obtain firm-level data about the assignees of financial patents. First, we matched a subset of our financial patents with the GVKEY of the first assignee using the Global Corporate Patent Dataset (GCPD) developed by researchers at the University of Virginia (Bena et al., 2017). This allowed us to match 12,351 patents to a Compustat GVKEY, which can be easily linked to the associated Capital IQ identifier because both Compustat and CapitalIQ are Standard & Poor's databases. Then, after removing inventor-assignees, we used a Levenshtein distance-based fuzzy name matching technique to match the remainder of assignee names with 12 million firm names in the Capital IQ database.[14]

After examining the data, we determined that a matching score of 0.95 or higher was sufficiently accurate that the match could be accepted without further scrutiny. This yielded an additional 6,237

---

[12] Between 7% and 8% of the patents in the financial patent and overall samples had no assignee type in Patentsview. We audited 2% of the financial patents with a missing assignee type, and discovered that 99% of these were assigned to individuals (also known as inventor-assignees). In the analyses below, we treat all patents with a missing assignee type as assigned to an individual.
[13] We also used Patentsview data to assign payments to primary CPC classes in some ambiguous cases where patents had more than one "primary" CPC code in the IFI data.
[14] We divided the Capital IQ database into three subsets, with four million company names in each subset, to execute the fuzzy name-matching algorithm in parallel and save computing time, and to get multiple optimum matches within each subset.

patents matched to Capital IQ firms. Similarly, we found that matches with scores below 0.8 were so poor that they should be rejected outright. For the 1,940 potential matches with scores between 0.80 and 0.95, we had a research assistant examine the potential matches, ultimately identifying an additional 818 patents with good assignee matches. This yielded Capital IQ identifiers for 19,406 patents, or 80% of the sample. We used the Capital IQ identifier to join to our financial patent database a host of detailed financial information about each firm in the year of the patent application in addition to its industry, employment, and whether it was publicly traded at the time. We used the Refintiv VentureXpert database to determine whether the firms were actively venture-backed at the time of the patent filing, following the methodology in Akcigit et al. (2020).

The industry groups that we focused on, and the associated GICS codes, were as follows:

- Banks cover large and geographically diverse institutions, as well as regional ones, with significant business activity in retail banking and small and medium corporate lending. This category also includes thrifts and mortgage finance firms providing mortgage and mortgage related services, and diversified financial services firms (GICS 401010, 401020, and 402010).
- Other finance includes providers of consumer services like personal credit and lease financing (GICS 402020), capital markets including asset management, and financial exchanges for securities, commodities, and derivatives (GICS 402030), and insurance (GICS 403010).
- Payments firms are classified under Data Processing and Outsourced Services (GICS 45102020).
- Information technology firms cover a wide variety of computer hardware and software developers, as well as technology consulting firms (GICS 45 outside of payments).
- All other.

We thus constructed a database containing, for each financial patent in our list, Derwent patent data, Patentsview patent data, and financial data from Capital IQ data (for each assignee that could be matched). Figure A-3 depicts the process we used in this step.

We then used similar techniques to match assignee names with the names of:

- SIFIs,[15]
- Patent Assertion Entities (PAEs or "patent trolls"), which were identified in the Stanford NPE Litigation Database (Miller et al., 2018), and
- Major fintech firms included in the Forbes 100, the KPMG 50, and the CB Insights Fintech 250 lists.[16]

---

[15] Data on SIFIs was taken from https://www.fsb.org/work-of-the-fsb/policy-development/addressing-sifis/global-systemically-important-financial-institutions-g-sifis/. We focused on the initial SFIs designated in November 2011.

[16] When matching the patentees against names of known fintech firms, we had a very low match rate. We manually inspected 20 fintech firms that had not matched to the patent data. We confirmed that 19 out of 20 of these entities did not have any patents issued to them. This reinforced our prior belief that using fintech firm names is not an ideal way for identifying financial patents.

We also matched the list of patentees against a list of patents assigned to academic institutions over this period that we constructed. We compiled this list by identifying all assignees containing the word "university," as well as those on the various annual lists of the most active academic patentees compiled by the Association of University Technology Managers (which allowed us to capture entities as the Massachusetts Institute of Technology and the Wisconsin Alumni Research Foundation).

We matched all patents to the database of citations to academic articles compiled by Marx and Fuegi (2019). This database contained all academic citations contained within patent documents (whether on the front page or in the text), as well as information about the subject matter of the articles, and the name and impact factor of the journals in which the articles appeared. We downloaded these data for all U.S. patents applied for between 2000 and 2018, and awarded by February 2019.

As a last step, we associated financial patents with particular functions in financial services, which we refer to as patent type or subject matter. The patent classification scheme was insufficient here, as many categories did not map readily to particular subject matters. Instead, we created a set of keywords (listed in Table A-1 in the Appendix) associated with accounting, commercial banking, communications, cryptocurrency, currency, insurance, investment banking, payments, real estate, retail banking and wealth management. We based these keywords on a review of the patent abstracts, finance glossaries, and industry knowledge. Some keywords were associated with a single patent type; others with multiple ones. Accordingly, for each patent that fell into more than one category, we assigned it a fractional share to all of the associated types.

We adopted four progressively wider searches to identify these keywords. First, we just examined the patent abstracts. For the patents with no matches, we examined the first 100 words of the background section of the patent. For firms with no matches, we examined the entirety of the background section. For the remaining firms without matches, we examined the entirety of the patent text. Tables A-2 and A-3 summarize the matching process. For the 345 patents without a match, we read the patents. For the 33 patents that could not be classified even after manual examination, we excluded them from our dataset. Hence the final dataset contains 24,255 (24,288-33) patents. For the purposes of the analyses below, we consolidated the patent types into banking (encompassing commercial, investment, and retail), payments, and all others.

Figure A-4 presents an overview of the financial dataset construction procedure.

4. Shifts in Financial Patenting

In this section, we examine the changes in financial patenting since 2000 in a decomposition analysis. While there was a dramatic increase in financial patenting of all types, these years also saw a substantial shift in the nature of the innovators. In particular, awards to U.S. information technology and other non-financial assignees have surged. We have also seen a shift in patent subject matter to outside of banking.

Before we turn to this analysis, we can illustrate the churn qualitatively. While the ranks of top patenting firms overall have remained largely constant over the 21st century (with companies like IBM, Canon, Hitachi, and Samsung dominating the compilations year after year), there is considerable volatility in the financial patentees.

Panels B and C of Table 4 show the largest changes in of patent assignees during the period between 2000 and 2004 on the one hand and 2015 and 2018 on the other. (Panel A presents the top total patentees during this period.) The table indicates that the share of innovation fell most sharply for unassigned patents (again, most ones filed by individuals), computer hardware firms (e.g., Fujitsu Hitachi, HP, and Diebold Nixdorf), legacy software and service providers (e.g., First Data, Oracle. and IBM), and investment banks (Goldman and JP Morgan). Meanwhile, the most rapid growth was from commercial banks (Bank of America and Wells Fargo), insurers (State Farm, Allstate, The Hartford, and USAA), and payments firms, whether incumbents or entrants (Capital One, PayPal, Square, and Visa).

We then undertook a decomposition of patenting trends. To do so, we create 456 cells, one for each award year, for each of the three broad patent types (banking, payments, and other), for each broad assignee industry (banking, other finance, payments, and IT plus all other), and for U.S. and foreign inventors. We estimated ordinary least squares (OLS) regressions, where the dependent variable in each cell was the number of patent awards. The independent variables are fixed effects for the award year, patent type, assignee industry, and inventor location, as well as interactions between year and the three other sets of controls. This analysis can help us better understand what is behind the surge of patenting, though it cannot explain what factors led to a surge in a specific category.

All the sets of explanatory variables jointly have significant explanatory power. The joint significance tests are presented in Table A-4. Figure 4 presents the year interaction effects, in each case with 2001 normalized as zero:

- Panel A shows the sharp increase in the number of patents per year across all cells. To calibrate the rise in the year fixed effects from 0 to about 200 patents per cell, the mean cell has 53.2 patent awards.
- Panel B displays the sharp decline in patenting by banks and other financial institutions relative to IT and other firms, a decline that started at the beginning of the sample, accelerated after the GFC, and only began recovering in the mid-2010s. Payments firms, after mirroring the decline of banks, experienced a somewhat more rapid recovery of the 2010s.
- Panel C shows a steady decline in the share of patenting in banking relative to payments and all other subject matters.
- Panel D shows a strong trend towards increasing patenting by domestic assignees, at least up until the mid-2010s. This pattern is consistent with the strong domestic bias in finance patents shown in Table 3.

This analysis also lends itself to a classic difference-in-differences analysis. To examine the changes in Panels B through D in this manner, we substituted for the year dummies an indicator variable for whether the observation was from 2009 or after. The interactions between the indicator

for an assignee in the banking industry and a post-GFC observation was significantly negative (coefficient of -125.4, with a p-value of 0.000), as was that for an assignee in another finance industry and a post-GFC observation (-104.7 and 0.000) and similarly for payment firms (-116.7 and 0.000). The interaction between patents with a subject matter in payments and the post-GFC dummy was insignificant, but that between banking type awards and the post-GFC indicator was significant at the 5% level (-25.2 and 0.031). The interaction between domestic patentees and the post-GFC indicator was significantly positive (79.2 and 0.000).

While this analysis suggests that the years after the GFC saw more patenting by firms outside of finance, and outside of the banking subject matter, it does not explore the interactions between these two factors.

To explore this phenomenon at a deeper level, we repeated the analysis, now with the addition of an interaction between the award year, assignee industry, and dummies denoting whether the patent by a bank patenting an invention in banking or a payments firm patenting in payments. (In addition, we added controls for the interactions between assignee industry and patent type.)

Figure 5 graphically depicts the interactions. Both banks and payments firms become progressively more likely (relative to other firms) to patent in their core areas over time. Thus, banks actually *increased* their share of patenting in banking, controlling for the overall decline for patenting activity by this type of firm and in this subject matter. The null hypothesis that the three-way interaction terms were equal to zero was rejected at the 1% confidence level. In short, innovation is becoming more specialized over time: banks have not responded to the apparent decline in innovative potential in banking by moving their innovative efforts into other areas.

Taken together, the analysis suggests that the financial institutions' share of financial innovation fell sharply over time, in part due to their failure to expand their range of innovative activities. The increased focus on banking patents by these firms may have reflected the fact that, far more so than IT and other companies, they had existing businesses that had continuing needs for ongoing innovation. Another possibility is that these firms wanted to expand into other areas of financial innovation, but found it difficult to do so. Two possible constraints may have been regulatory pressures or their lack of ability to innovate in these new technologies.

5. Trends in Academic Ties

In this section, we examine a second dynamic: the change in the utilization of academic knowledge. Using citations in patents to academic prior art, we show that there traditionally was a strong association between connections to academic knowledge and patent impact, particularly in patents assigned to banks. But in recent years, the rate of citation to academic research has fallen, and the papers being referenced are older.

Table 5 presents a first look at the journals most frequently cited in finance patents. Aside from one anomalous case (discussed in the note to the table), the journals were well known ones that fell into three reasonable categories: journals devoted to computer technologies, academic finance journals, and practitioner-oriented finance publications.

Table 6 examines the overall use of academic citations in finance patents. We compared the finance patents to two broader populations: the entire population of patents applied for and awarded over the same period, and those in "academic-heavy" patent classes. To determine the academic-heavy classes, we used first identified patents assigned to academic institutions using the procedure outlined in Section 3.2. We then extracted the four-digit CPC subclasses in which these patents most frequently have a primary assignment. We designate the 53 top classes (all those with 500 or more patent awards by academic institutions in the sample period) as academic heavy.

The table makes clear that in general, finance patents cited less academic work than other patents. The disparity between the finance and the academic-heavy awards were particularly striking. When we look at the citations to articles in business, economics, and finance, and even more so top finance and top practitioner finance journals,[17] a very different picture emerges: the financial patents make significantly more such citations. Moreover, the finance patents cite significantly fresher prior art: that is, the mean lag between the article publication and patent application was nearly a year shorter for finance patents.

Table 7 looks in more depth at which financial patents cite academic prior art. We estimated OLS regressions, with each observation consisting of a finance patent. The dependent variables were the number of academic citations in these patents, the number of citations to business, economics, and finance journals, the number of citations to Top 3 finance journals, and the mean age of the citations in each patent, defined as above.[18] All regressions included controls for patent type, inventor location, and assignee characteristics (age of firm, revenue, and status as an academic institution, other non-corporate entity, publicly traded firm, or SIFI).

In Panel A, the key independent variables were the assignee industry and the application year. (IT and other was the omitted assignee type.) The analysis suggests that financial institutions were considerably more likely to cite academic work, no matter how measured. (The coefficients were substantial relative to the means reported in the first column of Table 6.) Finance patents by IT and other firms were higher than payments when overall academic citations were used, but otherwise close to payments. The propensity to cite academic work fell over time—the time trend is significantly negative. When we rerun the analysis, assigning the application year to one of four categories (2000-04, 2005-09, 2010-14, and 2015-15), the results show a similar downward trend. Turning to citation age, citations were increasingly made to older articles over time. Payment firms had the youngest citations.

Panel B used similar specifications, but the key independent variables were the interaction between assignee industry and time period. Here, several patterns stood out:

[17] We identify the "Top 3" finance journals (the *Journal of Finance*, *Journal of Financial Economics*, and *Review of Financial Studies*), from numerous efforts to rate journals in the literature, such as Chan, Chang, and Chang (2013).

[18] We did not report the results using citations to journals with above median impact factors or to business/economics/finance journals with above median impact factors, but the results were similar.

- The propensity to cite academic work dropped at the 1% significance level between the 2000-04 and 2010-18 periods for banking, payments, and IT and other firms. Other finance was the only exception.
- All firms cited more stale academic knowledge over time, that is, older papers. There was no significant differences across the classes of firms in the age of the citations or the change over time.
- Table A-5 presents some supplemental analyses. It highlights how many of these patterns are driven by the patenting practices of U.S. corporations, the most frequently represented assignees.
- When we looked at these patterns in unreported analyses assigning the application year to one of four categories (2000-04, 2005-09, 2010-14, and 2015-15) and interacting these dummies with industry, the results showed a similar downward trend. The decline in banking was particularly striking.

Thus, it appears that these institutions are accessing academic work less than they did in the past, with other finance firms being the only exception. Whether this reflected the shifts in the supply of relevant academic knowledge or the ability of firms to absorb this knowledge (Cohen and Levinthal, 1990) was not obvious from this analysis, but below we find some clues to this question.

It is natural to wonder how consequential these shifts in academic citations are. While academic citations have been shown to be linked to patent impact in other fields (Watzinger and Schnitzer, 2019; Poege et al., 2019), to what extent is this knowledge relevant for financial patents?

In Table 8, we took a first look at this question by comparing the impact of finance patents with and without academic citations. As in Table 2, we used two metrics of patent value: citation weights and Kogan et al. (2017) patent values.

A striking association between more academic citations and greater patent impact appeared. Using citation weights, there was a statistically significant relationship for all academic cites, high impact academic citations, business/economic finance citations, and high-impact ones. The only exception was citations to Top 3 journal, where the results were directionally similar, but insignificant. The results using Kogan et al. values were similar directionally, and consistently statistically significant. Moreover, the results were large in economic magnitude: for instance, a financial patent without an academic citation was subsequently 8% more cited than a typical patent in its subclass; the differential for ones with such citations was six times larger.

Table 9 looks at the relationship between patent impact and academic citations in a regression analysis. We again estimated OLS regressions, using each patent as an observation. The dependent variable was the normalized citations and the Kogan et al. (2017) value. The key independent variable was the number of academic citations, either alone (Panel A), interacted with the industry of the assignee (Panel B), or interacted with the time period of the patent application (Panel C).[19]

[19] We do not report the results using citations to journals with above median impact factors or to business/economics/finance journal with above median impact factors, but the results are similar. Results are similar using application year.

We controlled for patent type, inventor location, and assignee characteristics (age of firm, revenue, and status as an academic institution, other non-corporate entity, publicly traded firm, or SIFI).

Three conclusions stood out:

- First, Panel A suggested a strong relationship between academic citations and patent impact. A positive and statistically significant relationship appeared in all cases except for that between "Top 3" finance journal citations and normalized citations.
- Second, the association between academic citations and patent impact in Panel B was consistently greater for banking firms (and in regressions using citation weights, other finance was also very large). In all cases, an F-test rejected the equality of the coefficients for the interaction terms between citations and the dummies for the different types of firms at the 0.1% confidence level, with the exception of column 6, where the null hypothesis is rejected at about the 1% confidence level.
- Finally, the association over time between academic citations and patent impact seems to be highly mixed. While it is generally increasing for the citation weights, it is declining for the Kogan at al. weights.

Overall, the analysis suggested a second substantial change. Academic knowledge appeared to be an important driver of financial patent impact, but it appears that this this relation had abated. This likely affected the banking industry most deleteriously: it was the primary beneficiary of this relationship, because they had a greater relationship between academic citations and patent impact. Moreover, financial patents in general were citing increasing older academic knowledge. This suggests that there had been a "drying up" of new knowledge relevant to innovators.

6. The Changing Geography of Innovation

In this section, we focus on a third shift: the changing geography of financial innovation. Focusing on the United States (which as shown above, was the primary and increasingly important locus of financial innovation), we document two distinct effects. First, the locus of innovation has dramatically shifted to the San Jose-San Francisco metropolitan areas, largely at the expense of the New York-Newark one. While part of this was due to the entry and exit of firms, it was primarily driven by the shifts in the locus of innovation within incumbent firms. IT firms, venture backed firms, those in regions with historical venture activity financing outside of San Francisco, were particularly likely to shift their innovative hubs.

6.1. Summarizing the Shifts

In order to undertake both analyses, we needed to map each patent to a combined statistical area (CSA). To do this, we used the state and county Federal Information Processing Standard (FIPS) code of the first-named inventor, also provided by Patentsview, and a crosswalk, compiled by the U.S. Bureau of the Census, between county-level FIPS codes and CSA codes as of mid-2013.[20] Patents whose first assignee was outside the U.S. or in the U.S. but not in a CSA (collectively between 20% and 25% of all patents in each year) are not included in the CSA-level analyses.

[20] https://www.nber.org/cbsa-csa-fips-county-crosswalk/List1.xls.

Table 10 shows the share of patenting by CSA for the ten CSAs with the highest total patent counts applied for between 2000 and 2018. The table tabulates for three periods these patents as a share of those assigned to U.S. CSAs, using simple patent counts, citation weights, and Kogan et al. (2017) weights.

The table shows that financial patenting has been progressively more concentrated over time, particularly when weighted. Using Kogan et al. weights, the top ten CSAs represent 80% of the CSA-based financial patenting in 2012-18. The rise of patenting in the San Jose-San Francisco CSA drives much of the increase in concentration. The decline in the importance of New York and the rise of Charlotte (which passed New York using weighted patents by the 2015-18 period) are also evident.

In Panels A through C in Table A-6, we assemble a variety of patenting measures for these three CSAs. The "tale of three cities" paints a sharp set of contrasts:

- San Jose-San Francisco-Oakland saw dramatic growth, whether measured using raw or weighted patenting. This was driven by mid-sized firms (i.e., those where the firm's revenue in the application year was more than $100 million but less than $10 billion), rather than by small and large ones. The patenting activity was driven by firms in the IT and other category, but especially by payments firms.
- New York-Newark, by way of contrast, saw a sharp decline in patenting. This was driven by a decline in patenting by large firms, especially SIFIs. Meanwhile, innovation by small firms increased sharply, reflecting the rise of the active fintech cluster there. Firms in the IT and other category saw the fastest growth.
- Charlotte-Concord saw rapid growth, particularly using Kogan weighting. This growth was driven by patenting by large firms and SIFIs in banking. A closer look at the data shows that this change was largely driven by Bank of America, which not only consolidated its patenting activity in Charlotte (in the 2000-04 period, the largest CSA for patent applications by the bank was New York, with 26% of the total; in 2015-18, Charlotte-Concord represented 66% of its awards), but also greatly accelerated its innovative activities.

Table A-7 provides another view of the overall patterns, focusing on activity across U.S. Census regions. More details on the construction of the CSA data set are in Appendix C.

6.2. Decomposing the Geographic Changes

We then sought to understand these changes in more detail. In particular, we explore what drives these shifts in patenting. The results highlight the importance of shifts by existing firms in their innovative activities.

Table 11 undertakes an initial decomposition of firms. Panel A divides them into three categories:

- Exiting innovators, who filed an (ultimately successful) financial patent in 2000-04, but not in 2015-18;

- Entrant innovators, who filed an (ultimately successful) financial patent in 2015-18, but not in 2000-18, and;
- Continuing innovators, who filed an (ultimately successful) financial patent in 2000-04and in 2015-18.

(Note we do not include firms that did not patent in 2000-04 and 2015-18, but just in intermediate years.)

For the third category, we also break out firms that shifted their modal CSA for patenting between these two periods. This latter category of location-switching continuers is relative few in number (26 firms), but very significant when patents are tabulated. In fact, these firms represent 55% of the awards by continuing innovators, and 41% of the awards across all three categories.

Panel B looks at the 26 location-switching continuers in more depth. The analysis highlights that nine of the firms (representing 9293 patents in total) moved their modal location from New York-Newark; no other CSA is close in losses. Meanwhile, San Jose-San Francisco was the destination of choice for four of the switchers, representing 5562 patents. These results suggest the importance of location-switching continuers in the analysis below.

Table 12 presents another way to dramatize the impact of shifts in innovative location by continuers. The table presents counterfactual calculations of patenting shares in 2015-18, under two assumptions; (a) that the 26 continuing financial innovators that shifted their modal location retained the same geographic distribution of patenting that they had in 2000-04, and (b) that all 130 continuers retained the same geographic distribution of patenting that they had in 2000-04. Focusing again on the three cities highlighted above, had all continuers maintained the innovative locations they had in 2000-04, the decline in patenting in New York and the rise in Sam Francisco each would have been half as large. The dramatic growth in Charlotte would not have happened at all, because (as discussed above) it was largely driven by a single switcher.

The last analysis in Table 13 looks at which continuing financial innovators are switchers in a probit analysis. We use all 130 continuing innovators as observations. A firm is defined as a switcher if its modal location for innovation changes from 2000-04 to 2015-18. The results suggest:

- IT and other firms are consistently the most likely firms to switch.
- Venture-backed firms, as well as those that are publicly traded, are more likely to switch.
- Firms in regions with extensive venture capital financing of financial firms in 2000-04, with the exception of New York and San Francisco, are more likely to switch their modal innovative location.

These results seem very consistent with the finding of Moretti (2019) of the importance of location to innovative efficiency. It appears that firms actively shift their location, in hopes of pursuit of innovative advantages. These shifts have had profound impacts on the location of financial innovation. The greater willingness of IT firms and venture-backed ones may reflect the lower costs associated with moving innovative activities for these firms.

7. Conclusion

In this paper, we explored the evolution of financial innovation by examining U.S. patents applied for between 2000 and 2018. We highlighted three fundamental shifts:

- The first is the entry of U.S. IT and payments firms, and the associated reduction in innovation by banks and other financial institutions. Banks have not responded by the decline of innovation in their core area by moving their innovative focus: in fact, they have become more focused on banking innovations.
- The second is the reduction in importance of academic knowledge that was traditionally incorporated in these patents. This trend has affected banks most adversely, given the historically strong relationship here between academic citations and patent importance.
- Finally, the geography of financial innovation is being reshaped. This shift was largely driven by continuing innovators moving their locus of innovative activity, rather than the entry or exit of financial innovators. IT firms, venture-backed firms, and those based in venture hubs outside of New York and San Francisco appear to be the most mobile.

We conclude with two observations. The first is the difference between the focus of academic studies of the financial innovation discussed in the introduction and the patterns documented here. The literature on financial innovations has largely highlighted new financial instruments created by banks and capital market firms, as well as cryptocurrencies. While these areas are doubtless important, the extent to which innovation is occurring in areas like payments, and has been driven by firms outside the traditional definition of financial institutions, has received little attention in the literature.

A second observation relates to the pressure that financial institutions have felt in regard to innovation. The declining share of financial institutions in patents granted and the continuing focus on banking technologies likely reflects (at least in part) optimization decisions based on existing product lines. But other shifts may be beyond their control, such as the decreased value-relevance of academic research.

Of course, there are many areas for future exploration. Foremost of these is assessing the social impact of these discoveries. As Lerner and Tufano (2011) highlighted, the assessment the social impact of financial innovations is particularly subtle: unlike a new chemotherapy or solar panel, these discoveries can have dramatically different impacts over time as they diffuse and the behavior of consumers and issuers changes. Some of the conceptual approaches highlighted in papers such as Budish, Roin, and Williams (2016) may represent a way forward.

References

.

Acemoglu, Daron, and Joshua Linn. 2004. "Market Size and Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics* 119 (3): 1049–1090.

Akcigit, Ufuk, Sina T. Ates, Josh Lerner, Richard R. Townsend, and Yulia Zhestkova, 2020, "Fencing Off Silicon Valley: Cross-Border Venture Capital and Technology Spillovers." Unpublished working paper.

Allen, Franklin, and Douglas Gale. 1994. "Limited Market Participation and Volatility of Asset Prices." *American Economic Review* 84 (4): 933–955.

Baily, Martin N., and Eric Zitzewitz. 2001. "Service Sector Productivity Comparisons: Lessons for Measurement." In *New Developments in Productivity Analysis*. Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, editors. Chicago: University of Chicago Press, pp. 419-464.

Beck, Thorsten, Tao Chen, Chen Lin, and Frank M. Song. 2016. "Financial Innovation: The Bright and Dark Sides." *Journal of Banking and Finance* 72 (C): 28–51.

Bena, Jan, Miguel A. Ferreira, Pedro Matos, and Pedro Pires. 2017. "Are Foreign Investors Locusts? The Long-Term Effects of Foreign Institutional Ownership." *Journal of Financial Economics* 126 (1): 122–146.

Biais, Bruno, Jean-Charles Rochet, and Paul Woolley. 2015. "Dynamics of Innovation and Risk." *Review of Financial Studies* 28 (5): 1353–1380.

Bloom, Nicholas, Charles Jones, John Van Reenen, and Michael Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review*, forthcoming.

Budish, Eric, Benjamin N. Roin, and Heidi Williams, 2016. "Patents and Research Investments: Assessing the Empirical Evidence." *American Economic Review Papers and Proceedings* 106 (5): 183-187.

Caballero, Ricardo J., and Alp Simsek. 2013. "Fire Sales in a Model of Complexity." *Journal of Finance* 68 (6): 2549–2587.

Chan, Kam C., Chih-Hsiang Chang, and Yuanchen Chang. 2013. "Ranking of Finance Journals: Some Google Scholar Citation Perspectives." *Journal of Empirical Finance*. 21: 241-250

Chernenko, Sergey, and Adi Sunderam. 2014. "Frictions in Shadow Banking: Evidence from the Lending Behavior of Money Market Mutual Funds." *Review of Financial Studies* 27 (6): 1717–1750.

Cohen, Wesley M., and Daniel A. Levinthal. 1990. "Absorptive Capacity: A New Perspective on Learning and Innovation." *Administrative Science Quarterly* 35 (1): 128-152.

Finkelstein, Amy. 2007. "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare." *Quarterly Journal of Economics* 122 (1): 1–37.

Fostel, Ana, and John Geanakoplos. 2012. "Why Does Bad News Increase Volatility and Decrease Leverage?" *Journal of Economic Theory* 147 (2): 501–525.

Frame, W. Scott, and Lawrence J. White. 2004. "Empirical Studies of Financial Innovation: Lots of Talk, Little Action?" *Journal of Economic Literature* 42 (1): 116-144.

Gennaioli, Nicola, Andrei Shleifer, and Robert Vishny. 2012. "Neglected Risks, Financial Innovation, and Financial Fragility." *Journal of Financial Economics* 104 (3): 452–468.

Griliches, Zvi. 1990. "Patent Statistics as Economic Indicators: A Survey." *Journal of Economic Literature* 28 (4): 1661–1707.

Hall, Bronwyn H., and Megan MacGarvie. 2010. "The Private Value of Software Patents." *Research Policy* 39 (7): 994–1009.

Herrera, Helios, and Enrique Schroth. 2011. "Advantageous Innovation and Imitation in the Underwriting Market for Corporate Securities." *Journal of Banking and Finance* 35 (5): 1097–1113.

Hunter, III, Starling D. 2004. "Have Business Method Patents Gotten a Bum Rap? Some Empirical Evidence." *Journal of Information Technology Theory and Applications* 6 (1): 1–24.

Japanese Patent Office. 2019. "Recent Trends in Business-Related Inventions." https://www.jpo.go.jp/e/system/patent/gaiyo/recent_trends_biz_inv.html

Jones, Benjamin. 2009. "The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder?" *Review of Economic Studies.* 76(1): 283-317.

Jurafsky, Daniel, and James H. Martin. 2019. *Speech and Language Processing*, draft of 3rd edition, https://web.stanford.edu/~jurafsky/slp3/.

Keys, Benjamin J., Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 2010. "Did Securitization Lead to Lax Screening? Evidence from Subprime Loans." *Quarterly Journal of Economics* 125 (1): 307–362.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth." *Quarterly Journal of Economics* 132 (2): 665–712.

Krugman, Paul. 2007. "Innovating Our Way to Financial Crisis." *New York Times*, December 3, 2007, https://www.nytimes.com/2007/12/03/opinion/03krugman.html.

Kung, Edward, 2020, "Innovation and Entrepreneurship in Housing and Real Estate," in *The Role of Innovation and Entrepreneurship in Economic Growth*, Aaron Chatterji, Josh Lerner, Scott Stern, and Michael J. Andrews, editors, University of Chicago Press for the National Bureau of Economic Research, forthcoming.

Lerner, Josh. 2002. "Where Does State Street Lead? A First Look at Finance Patents, 1971 to 2000." *Journal of Finance* 57 (2): 901–930.

Lerner, Josh. 2006. "The New New Financial Thing: The Origins of Financial Innovation." *Journal of Financial Economics* 79 (2): 223–255.

Lerner, Josh, and Peter Tufano. 2011. "The Consequences of Financial Innovation: A Counterfactual Research Agenda." *Annual Review of Financial Economics* 3: 41–85.

Marco, Alan C., Joshua D. Sarnoff, and Charles A.W. deGrazia. 2020. "Patent Claims and Patent Scope." *Research Policy* 48, forthcoming.

Marx, Matt and Aaron Fuegi. 2019. "Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles." Boston University Questrom School of Business Research Paper No. 3331686, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3331686.

Miller, Shawn P., Ashwin Aravind, Bethany Bengfort, Clarisse De La Cerda, Matteo Dragoni, Kevin Gibson, Amit Itai, Charles Johnson, Deepa Kannappan, Emily Kehoe, Katherine Mladinich, Roberto Pinho, John Polansky, and Brian Weissenberg. 2018. "Who's Suing Us? Decoding Patent Plaintiffs since 2000 with the Stanford NPE Litigation Dataset." *Stanford Technology Law Review.* 21 (2): 235–275.

Moretti, Enrico. 2019. "The Effect of High-Tech Clusters on the Productivity of Top Inventors." National Bureau of Economic Research Working Paper No. 26270, https://eml.berkeley.edu//~moretti/clusters.pdf.

National Research Council. 2005. *Measuring Research and Development Expenditures in the U.S. Economy*. National Academies Press, Washington.

Philippon, Thomas. 2015. "Has the US Finance Industry Become Less Efficient? On the Theory and Measurement of Financial Intermediation." *American Economic Review* 105 (4): 1408-1438.

Philippon, Thomas. 2019. "The FinTech Opportunity." In *The Disruptive Impact of FinTech on Retirement Systems.* Julie Agnew and Olivia S. Mitchell, editors. New York: Oxford University Press, pp. 190-217.

Poege, Felix, Dietmar Harhoff, Fabian Gaessler, and Stefano Baruffaldi. 2019. "Science Quality and the Value of Inventions." *Science Advances* 5 (12): https://advances.sciencemag.org/content/5/12/eaay7323.

Rajan, Raghuram G. 2006. "Has Finance Made the World Riskier?" *European Financial Management* 12 (4): 499-533.

Ross, Stephen A. 1976. "Options and Efficiency." *Quarterly Journal of Economics* 90 (1): 75–89.

Simsek, Alp. 2013. "Speculation and Risk Sharing with New Financial Assets." *Quarterly Journal of Economics* 128 (3): 1365–1396.

Thakor, Anjan V. 2012. "Incentives to Innovate and Financial Crises." *Journal of Financial Economics* 103 (1): 130–148.

Tufano, Peter. 1989. "Financial Innovation and First Mover Advantages." *Journal of Financial Economics* 25 (2): 213–240.

Watzinger, Martin, and Monika Schnitzer. 2019. "Standing on the Shoulders of Science," Centre for Economic Policy Research Discussion Paper No. DP13766, https://ssrn.com/abstract=3401853.

Figure 1: Financial patents and applications as a share of total U.S. patenting. The red line shows the number of financial patents granted annually by the total number of patents granted that, for patents applied from January 2000 to December 2018, and issued by February 2019. The blue line shows the number of financial patents applied for annually divided by the total number of patents applied for.

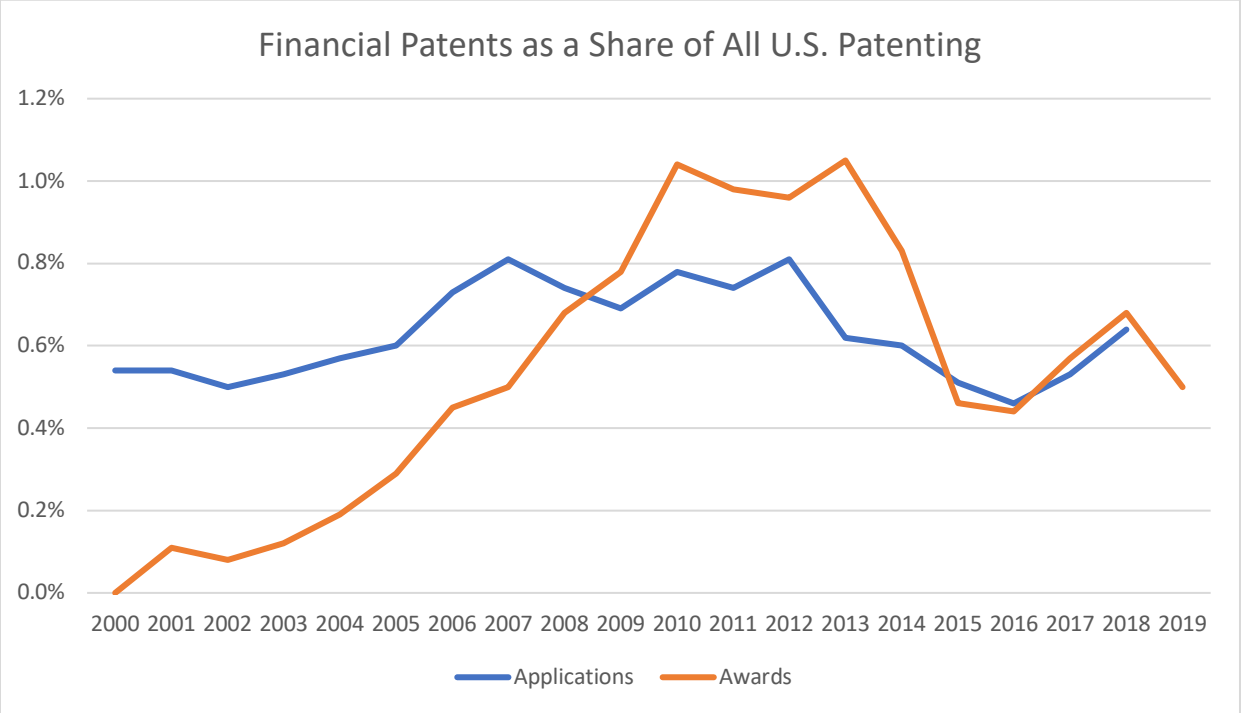
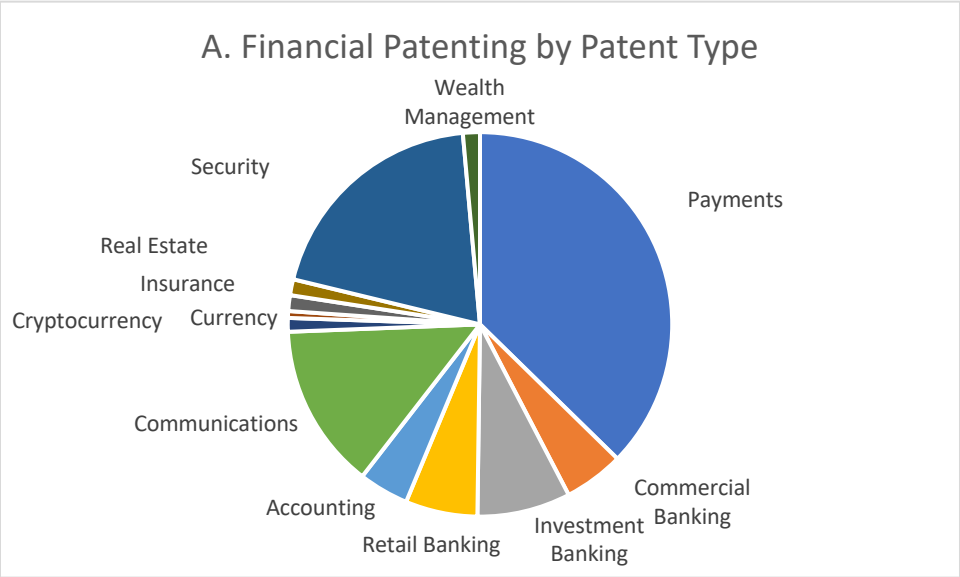


Figure 2. Composition of financial patents. The table present the breakdown of patent type (Panel A) and assignee industry (Panel B) for patent applied for between 2000 and 2018 and awarded by February 2019. The tabulation in Panel B excludes patents assigned to governments or individuals (as well as unassigned patents, which are overwhelmingly individually assigned).



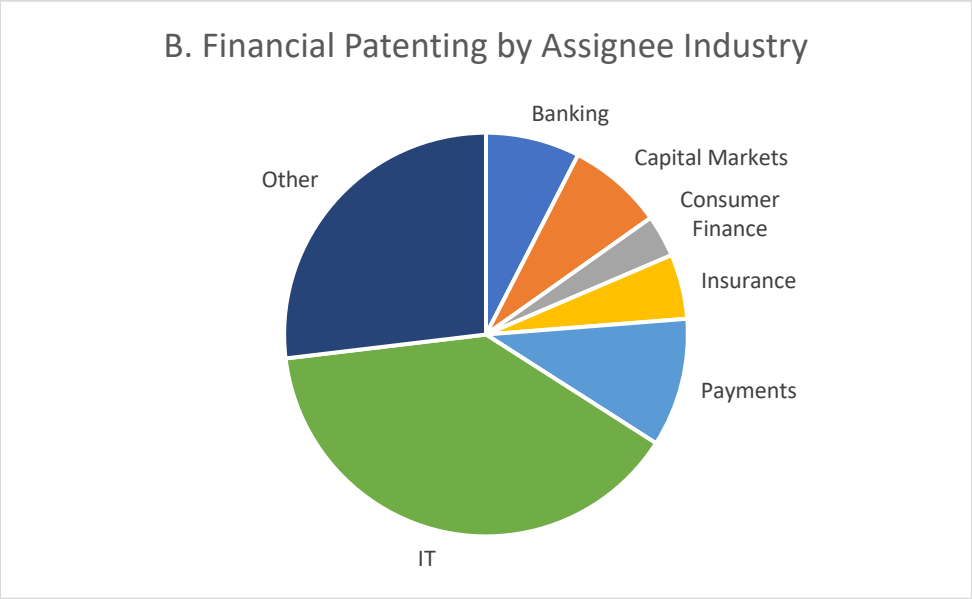
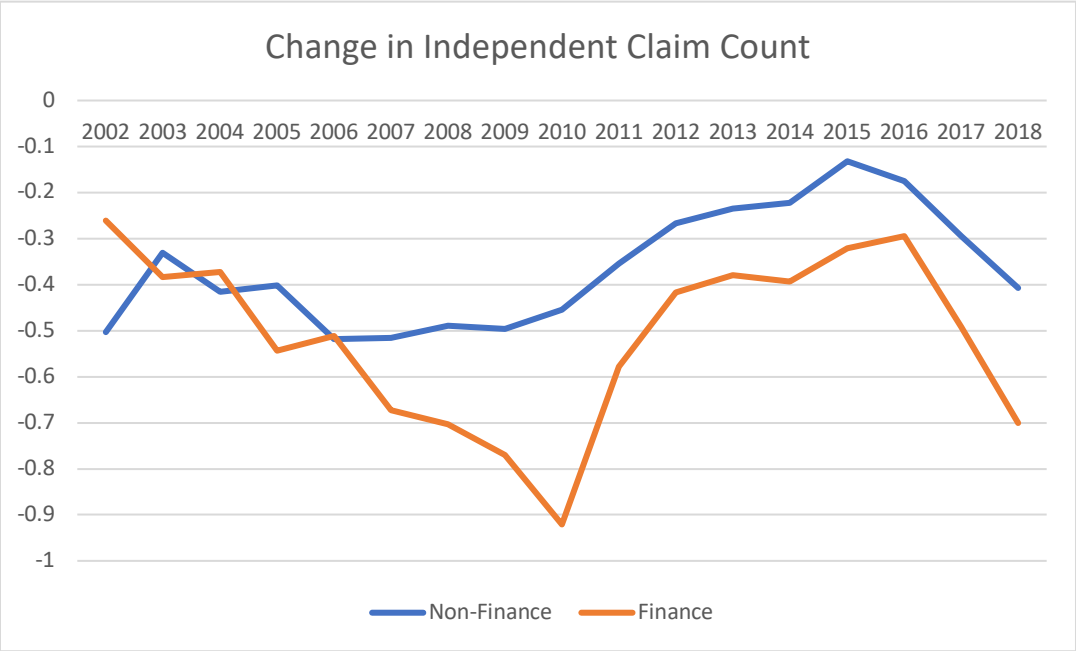


Figure 3. The extent of patent change between application publication and award, over time. Panel A reports the change in the number of independent claims at the time of application publication and award and the length of the shortest independent claim at these two points, for finance and other patents. The mean values are presented by year of award.

Panel A.



Panel B.

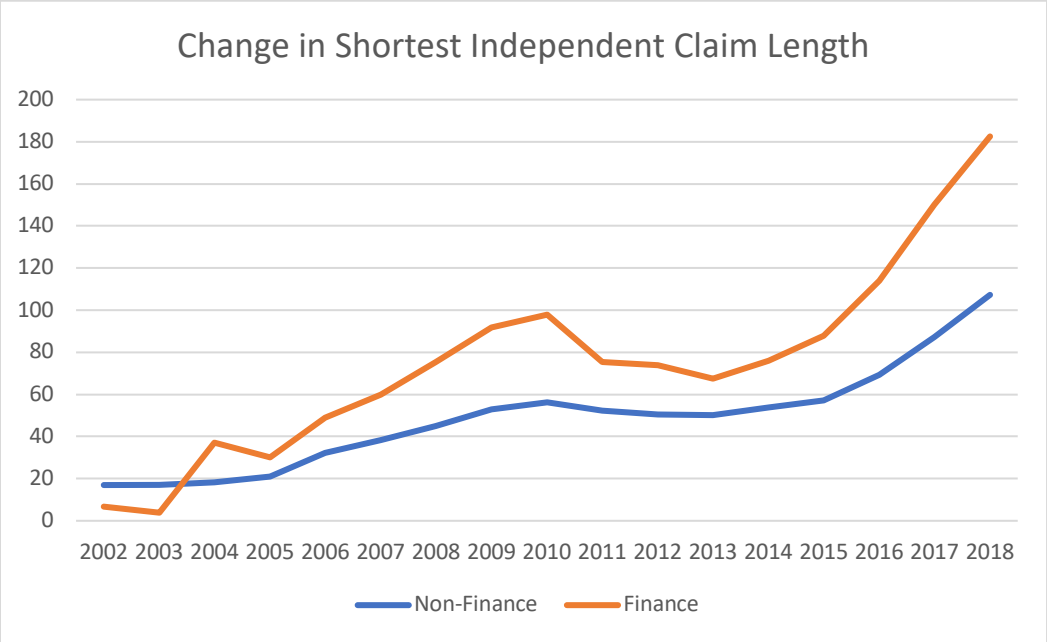
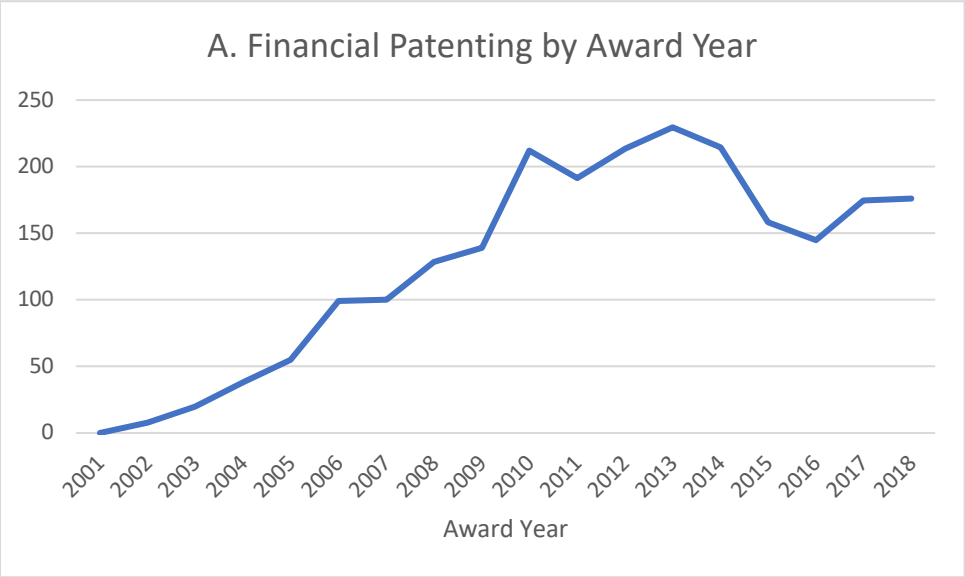


Figure 4: Decomposition of the rise of financial patenting. The charts depict the results of a regression analysis of the rise of patenting, where the dependent variable is the number of financial patents awarded by each year-assignee firm industry-patent type-inventor location cell. The charts depict the annual fixed effects (Panel A) and the interactions between year with assignee industry (Panel B; relative to "IT and Other Industries"), patent type (Panel C, relative to "Other Types"), and inventor location (Panel D, relative to "Non-U.S. Inventors").



r45re

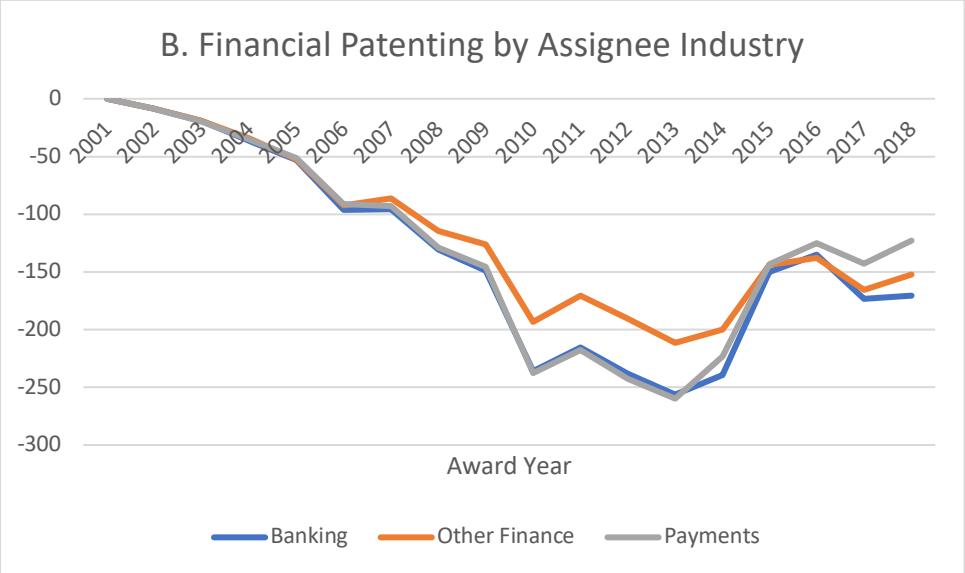
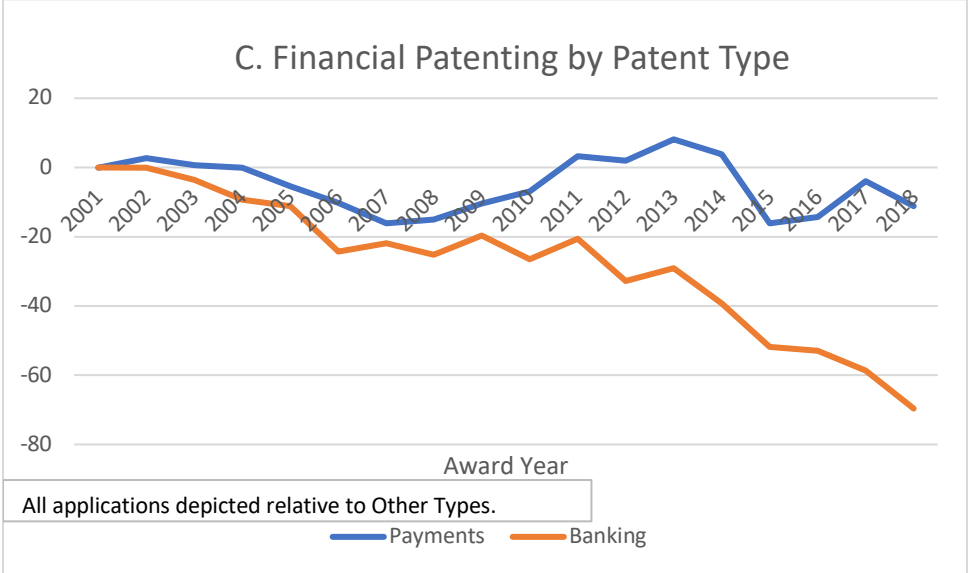


Figure 4 (continued).



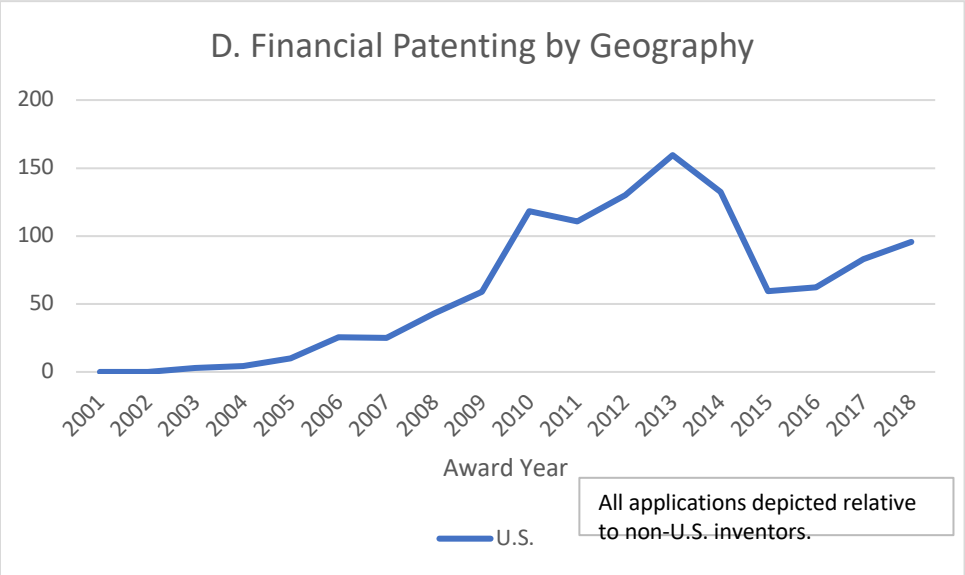


Figure 5: Decomposition of the rise of financial patenting. The charts depict the results of a regression analysis of the rise of patenting, where the dependent variable is the number of financial patents awarded by each year-assignee firm industry-patent type-inventor location cell. The chart depicts the annual fixed effects of the interactions between year with assignee industry and patent type (relative to the year 2000, "IT and Other Industries," "Other Types," and other interactions).

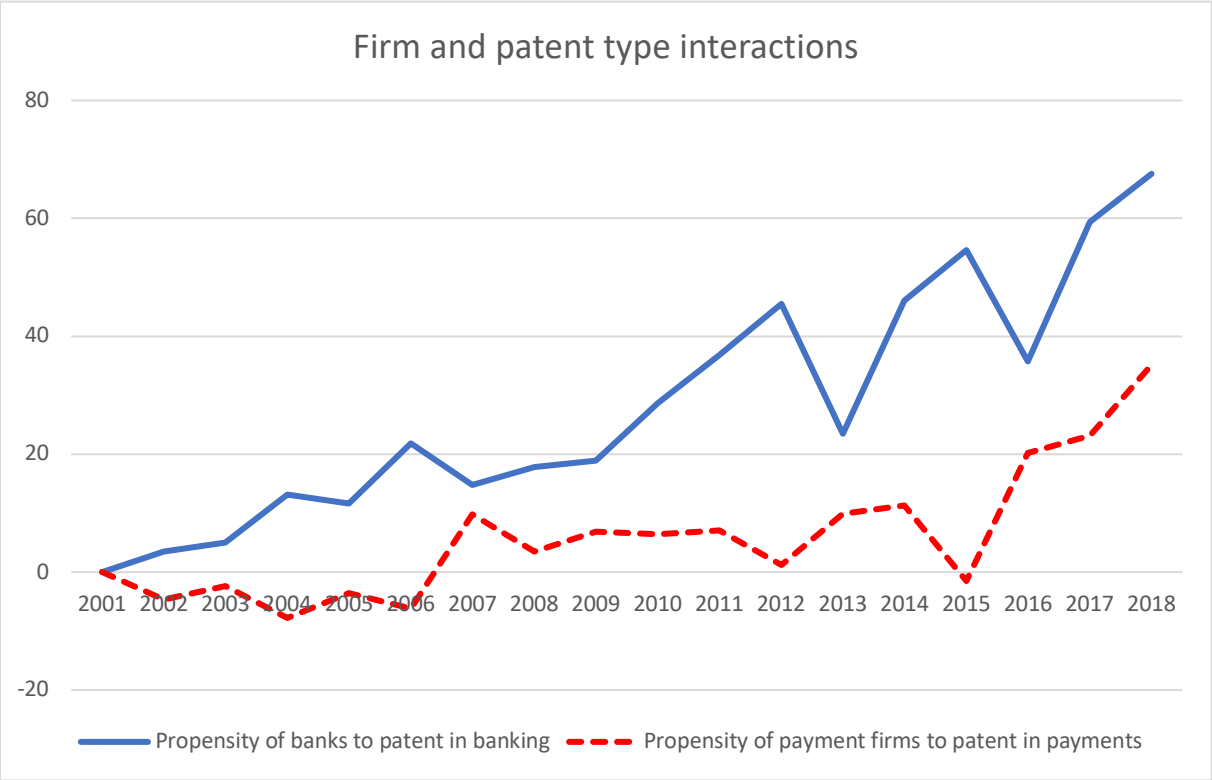


Table 1. The extent of patent change between application publication and award. The table reports the number of independent claims at the time of application publication and award, the length of the shortest independent claim at these two points, and the change in these measures for finance and other patents. The sample consists of all patents applied for between 2000 and 2014, and issued by February 2019. It reports as well the significance of t-tests of the equality of these measures for finance and non-finance patents. * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Finance Patents	Non-Finance Patents
Application publication		
Count of independent claims	3.60	3.00***
Length of shortest independent claim	117.60	111.52**
Patent		
Count of independent claims	3.07	2.66***
Length of shortest independent claim	201.18	160.55***
Change, count of independent claims	-0.53	-0.33***
Change, length of shortest independent claim	83.58	49.04***
Count of patents	15,922	2,600,032

Table 2. Impact of finance patents and all patents, by assignee type. The table presents the citation weights and Kogan et al. (2007) weights for finance patents and all other patents applied for between 2000 and 2018 and awarded by February 2019, computed using patent counts and citations weights.

	Mean Citation Weights	*Mean Kogan Et Al. Weight*	*Number of patents*
Finance Patents	1.25	53.61	24,255
All Other Patents	1.00	11.81	3,781.439
p Value, t-test	0.000	0.000	

Table 3. The assignees of financial and non-financial patents. The sample consists of finance and non-finance patents applied for between 2000 and 2018 and awarded by February 2019. We compare the finance and finance patents in a t-test: * denotes rejection of the null hypothesis of no difference in the means at the 10% level; ** at the 5% level; and *** at the 1% level.

	Finance Patents	Non-Finance Patents
Assignee Type:		
U.S. corporation	81.29%	46.50%***
Foreign corporation	17.40%	50.21%***
Individual	8.65%	7.79%***
U.S. government	0.08%	0.36%***
Foreign government	0.01%	0.09%***
U.S. university	0.21%	1.45%***
Foreign university	0.07%	0.75%***
Share active VC backed	4.02%	2.22%***
Share VC-backed, U.S. inventors only	4.99%	4.43%***

Table 4. The assignees of financial patents. Panel A presents the most frequent assignees of finance patents applied for between 2000 and 2018 and awarded by February 2019. Panel B and C presents the most sharply declining (growing) financial patent assignees. These are identified by comparing the share of financial patents in the sample applied for between 2000 and 2004 and the share applied for between 2015 and 2018.

Panel A: Most frequent assignees.

	Number of patents
Bank of America Corporation	652
Trading Technologies International	645
Visa Inc.	608
Diebold Nixdorf, Inc.	597
International Business Machines Corporation	589
Mastercard Inc.	418
JPMorgan Chase & Co.	407
American Express Company	404
United Services Automobile Association	351
Intuit	310

Panel B: Most rapidly declining finance patent assignees.

	Change in share
Unassigned	-6.1%
First Data Corporation	-2.4%
Goldman Sachs Group, Inc.	-1.5%
JPMorgan Chase & Co.	-1.4%
Fujitsu Limited	-1.3%
Hitachi, Ltd.	-1.3%
HP Inc.	-1.2%
International Business Machines Corporation	-1.2%
Oracle Corporation	-1.0%
Sony Corporation	-1.0%
Diebold Nixdorf, Inc.	-1.0%

Panel C: Most rapidly growing finance patentee assignees.

	Change in share
Bank of America Corporation	+6.1%
Square, Inc.	+4.3%
State Farm Mutual Automobile Insurance Company	+3.8%
Mastercard Inc.	+3.3%
PayPal Holdings, Inc.	+3.1%
Visa Inc.	+2.7%
Capital One Services, LLC	+2.2%
The Allstate Corporation	+1.5%

The Hartford Financial Services Group, Inc.	+1.1%
Wells Fargo & Company	+0.9%
United Services Automobile Association	+0.8%

Table 5. Most frequently cited academic journals in finance patents. The table present the ten journals most frequently cited in finance patents applied for between 2000 and 2018 and awarded by February 2019.

Journal Name	*Number of Citations*
Communications of the ACM	949
Journal of Finance	633
Journal of Animal Science	499
Financial Analysts Journal	363
IEEE Computer	301
Journal of Portfolio Management	279
ABA Banking Journal	242
Computers & Security	216
Journal of Financial Economics	191
Management Science	190

Notes: In addition to the journal citations, there are 257 citations to working papers archived at www.ssrn.org. The prominent role of the *Journal of Animal Sciences* reflects the presence of one dozen patents that are continuations (or continuations-in-part) of a single application originally filed by Micro Beef Technologies, relating to an accounting system for cattle farms. Each of the patents cites an (almost identical) list of approximately 40 papers from the *Journal of Animal Science*.

Table 6. Number of academic citations in finance patents and all patents. The table presents the mean number of citations to academic output, the number in publications with an above-median impact factor, the number in publications of various types (all business and economics journals, all business and economic journals with an above-median impact factor, and "Top 3" finance journals), and the lag between article publication and patent application filing. The totals are reports for finance patents, all patents, and all patents in the 53 four-digit CPC patent classes in which universities most frequently file patents. All analyses use patents applied for between 2000 and 2018 and awarded by February 2019. * denotes statistical significance of the differences in t-tests at the 10% level; ** at the 5% level; and *** at the 1% level.

	Financial Patents	*All Other Patents*	*All Other Patents in Academic Classes*
Total Citations	2.45	6.17***	10.36***
Total Citations to High-Impact Factor Journals	0.07	1.38***	2.53***
Total Citations to Business/Economics Journals	0.54	0.02***	0.02***
Total Citations to High-Impact Business/Economics Journals	0.07	0.00***	0.00***
Total Citations to Top 3 Finance Journals	0.04	0.00***	0.00***
Article-Patent Application Lag (years)	9.38	10.50***	10.02***
Number of Observations	24,255	3,781,439	1,823.420

Table 7. Regression analysis of academic citations and patent characteristics. The sample consists of finance patents applied for between 2000 and 2018 and awarded by February 2019. The dependent variables are the number of academic citations in these patents, the number of citations to business, economics, and finance journals, the number to Top 3 finance journals, and the mean age of the citations in each patent (years between the article publication and patent application date). In Panel A, the key independent variables are assignee industry and time period (payments is the omitted assignee type, and applications from 2010 to 2018 is the omitted category); in Panel B, the key independent variables are the interactions between assignee industry and time period (payment applications from 2010 to 2018 is the omitted category). All regressions control for the patent subject, inventor location, and assignee characteristics. Robust standard errors are in italics; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Academic Citations		*Bus/Econ/Fin Citations*		*Top 3 Citations*		*Citation Age*
Panel A: Trends							
Banking assignee	1.81		0.57		0.04		1.95
	*0.17****		*0.09****		*0.02***		*0.63****
Other finance assignee	2.07		0.77		0.06		1.77
	*0.13****		*0.06****		*0.02****		*0.42****
IT/other	1.51		0.11		-0.004		0.90
	*0.14****		*0.03****		*0.002**		*0.39***
Application year	-0.08		-0.02		-0.002		0.45
	*0.02****		*0.003****		*0.001****		*0.03****
Panel B: Interactions							
Banking * Appl. year	-0.10		-0.04		-0.004		0.56
	*0.01****		*0.01****		*0.001****		*0.05****
Other finance * Appl. year	0.01		0.02		0.002		0.54
	0.02		*0.01****		*0.001*		*0.03****
Payments * Appl. year	-0.19		-0.04		-0.003		0.34
	*0.02****		*0.003****		*0.001****		*0.05****
IT/other * Appl. year	-0.09		-0.03		-0.003		0.40
	*0.02****		*0.003****		*0.001****		*0.03****

Table 8. The relationship between patent impact and academic citations. The sample consists of in finance patents applied for between 2000 and 2018 and awarded by February 2019. The table reports the mean citation weight and the Kogan et al. [2007] weight, for patents that do and do not cite any academic output, publications with an above-median impact factor, and publications of various types (all business and economics journals, all business and economic journals with an above-median impact factor, and "Top 3" finance journals). * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Mean, Weighted Citations			Mean, Kogan Et Al. Value	
	Yes	*No*		*Yes*	*No*
Academic Citation(s)?	1.52	1.08***		59.4	50.1***
Citation(s) to High-Impact Factor Journals?	1.74	1.23***		67.1	53.0***
Citation(s) to Business/Economics Journals?	1.38	1.22***		85.3	48.0***
Citation(s) to High-Impact Business/Economics Journals?	1.52	1.27**		96.6	52.2***
Citation(s) to Top 3 Finance Journals?	1.30	1.25		184.2	52.1***

Table 9. Regression analysis of financial patent impact and academic citations. The sample consists of in finance patents applied for between 2000 and 2018 and awarded by February 2019. The dependent variables are the citation weight for each patent and the Kogan et al. [2007] weight. The key independent variables in Panel A is the number of academic citations in these patents; in Panel B, the interaction between the number of academic citations and the assignee industry; and in Panel C, the interaction between the number of academic citations and the time period. The first two regressions use the total number of citations to academic products; the second two, the number of citations to business, economics, and finance journals; and the final two, the number to Top 3 finance journals. All regressions include controls for the assignee type, patent subject, inventor location, application time period, and applicant characteristics. Robust standard errors are in italics; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Using All Academic Citations			Using Bus/Economics Citations			Using Top 3 Journal Citations	
	Citation weight	*Kogan weight*		*Citation weight*	*Kogan weight*		*Citation weight*	*Kogan weight*
Panel A: Basic								
Citations	0.05	0.80		0.07	2.39		0.0003	41.34
	*0.01****	*0.17****		*0.02****	*0.68****		*0.03*	*16.17***
*Panel B: Cites * Industry*								
Citations * Banking	0.09	11.08		0.09	20.99		0.14	124.93
	*0.03****	*2.17****		*0.05**	*5.05****		*0.25*	*46.53****
Citations * Other Finance	0.25	0.52		0.15	-0.56		-0.01	-7.11
	*0.05****	*0.37*		*0.04****	*0.89*		*0.02*	*5.03*
Citations * Payments	0.10	1.39		-0.06	3.67		-0.73	
	*0.04****	*1.23*		*0.05*	*5.39*		*0.12****	
Citations * IT	0.02	0.001		-0.01	-0.01		0.15	-16.28
	*0.01****	*0.08*		*0.02*	*0.42*		*0.30*	*9.68*
*Panel C: Cites * Time Period*								
Citations * 2000-04 Apply	0.02	1.27		0.05	4.04		0.06	71.72
	*0.004****	*0.51***		*0.03*	*1.59***		*0.08*	*42.01**
Citations * 2005-09 Apply	0.02	0.84		-0.01	3.27		-0.05	61.43
	*0.01****	*0.28****		*0.01*	*1.41***		*0.09*	*27.56***
Citations * 2010-14 Apply	0.09	0.41		0.04	0.30		-0.05	-3.17
	*0.02****	*0.14****		*0.02**	*0.55*		*0.02*	*8.21*
Citations * 2015-18 Apply	0.59	2.07		4.20	5.54			
	*0.17****	*1.02***		*0.83***	*6.55*			

Table 10. Financing patenting by U.S. urban area over time. The table presents the share of patenting by CSA for the ten CSAs with the most financial patents overall. All analyses use patents applied for between 2000 and 2018 and awarded by February 2019. The table presents financial patents as a share of all patents, computed using patent counts, citation weights, and Kogan et al. (2017) weights. We assign patents based on the location of the first inventor.

	Patent Count					Citation Weighted					Kogan et al. Weighted			
	2000-04	2005-09	2010-14	2015-18		2000-04	2005-09	2010-14	2015-18		2000-04	2005-09	2010-14	2015-18
San Jose-San Francisco-Oakland	8.5%	10.7%	15.7%	18.3%		11.5%	16.2%	21.3%	21.5%		8.4%	14.8%	25.0%	25.6%
New York-Newark	13.4%	11.6%	9.5%	5.7%		14.6%	7.8%	6.4%	5.7%		34.6%	19.8%	14.4%	5.7%
Chicago-Naperville	3.4%	6.2%	7.5%	3.9%		5.6%	5.8%	7.3%	3.0%		2.9%	4.5%	4.4%	4.4%
Washington-Baltimore-Arlington	4.0%	3.4%	3.2%	4.0%		4.7%	6.0%	3.3%	2.2%		3.1%	2.6%	1.4%	4.1%
Los Angeles-Long Beach	2.4%	2.1%	2.8%	1.8%		3.1%	2.8%	5.0%	3.7%		0.3%	0.9%	0.7%	0.9%
Cleveland-Akron-Canton	2.4%	2.8%	2.7%	1.7%		1.3%	1.8%	2.3%	0.7%		0.6%	0.5%	0.3%	0.3%
Atlanta-Athens-Clarke County	2.0%	2.6%	2.0%	2.8%		2.5%	3.7%	1.8%	1.3%		0.7%	1.4%	1.1%	2.1%
Seattle-Tacoma	1.9%	2.5%	2.3%	1.8%		2.0%	2.5%	2.5%	1.7%		1.8%	1.7%	2.4%	2.8%
Charlotte-Concord	0.3%	1.7%	2.3%	4.2%		0.4%	1.5%	3.2%	1.6%		0.4%	11.0%	8.7%	13.7%
Denver-Aurora	2.2%	2.0%	2.1%	1.3%		1.9%	1.4%	1.2%	0.5%		2.7%	1.2%	1.3%	0.6%

Table 11. Movement of financial patentees. Panel A reports the number of firms and the number of total patents awarded to these firms, divided into those that filed a successful financial patent application in 2000-04 but not 2015-18, those that did so in 2015-18 but not 2000-04, those that did so In both periods, and the subset that moved their modal location of patenting between these two periods. In Panel B, for the switchers only, the three most common departure and destination CSAs are also reported. We assign patents based on the location of the first inventor.

Panel A: Breakdown of firms and associated patents

	Firms	*Total patents*
Firms that patented in 2000-04, but not in 2015-18	792	3876
Firms that patented in 2015-18, but not in 2000-04	306	1895
Firms that patented in 2000-04 and in 2015-18	130	16539
Of these, firms that shifted modal CSA	26	9137

Panel B: Departure and arrival city of switchers

	Firms	*Total patents*
Three most frequently departed 2000-04 CSAs:		
New York-Newark, NY-NJ-CT-PA	9	8283
Denver-Aurora, CO	1	297
San Jose-San Francisco-Oakland, CA	3	188
Three most frequently arrived 2015-18 CSAs:		
San Jose-San Francisco-Oakland, CA	4	5562
Charlotte-Concord, NC-SC	1	652
Rochester-Austin, MN	1	589

Table 12. Counterfactuals regarding the movement of financial patentees. The first two columns present the unweighted share of patenting by CSA for the ten CSAs with the most financial patents overall for 2000-04 and 20015-18 (from Table S-10). The third column presents the distribution of patenting in 2015-18 had the 29 firms that switched their modal location of financial patenting between 2000-04 and 2015-18 retained the distribution as it was in 2000-04. The fourth column presents the distribution of patenting in 2015-18 had all 130 firms that patented in 2000-04 and 2015-18 retained the distribution as it was in 20000-04. All analyses use patents applied for between 2000 and 2018 and awarded by February 2019.

	2000-04		*2015-18*		
	Actual		*Actual*	*If switchers stayed in place*	*If all continuers stayed in place*
San Jose-San Francisco-Oakland, CA	8.5%		18.3%	18.1%	13.5%
New York-Newark, NY-NJ-CT-PA	13.4%		5.7%	7.8%	8.4%
Chicago-Naperville, IL-IN-WI	3.4%		3.9%	3.9%	2.9%
Washington-Baltimore-Arlington, DC-MD-VA-WV-PA	4.0%		4.0%	4.1%	4.2%
Cleveland-Akron-Canton, OH	2.4%		1.7%	1.6%	2.4%
Los Angeles-Long Beach, CA	2.4%		1.8%	1.9%	1.6%
Atlanta--Athens-Clarke County--Sandy Springs, GA	2.0%		2.8%	2.6%	2.4%
Seattle-Tacoma, WA	1.9%		1.8%	1.6%	1.9%
Denver-Aurora, CO	2.2%		1.3%	2.0%	2.4%
Charlotte-Concord, NC-SC	0.3%		4.2%	0.2%	0.2%

Table 13. Determinants of the movement of financial patentees. The sample consists of 130 firms that filed financial patents in 2000-04 and 2015-18. The dependent variable is a dummy indicating f the firm shifted its modal CSA for patenting. The independent variables include dummies for firm industry (payments is the omitted category), whether the firm is venture-backed or publicly traded (all of the time of the first patent filing in the 2000-04 period), and whether its modal patenting location in 2000-04 were New York or San Jose CSAs, as well as the volume of finance venture capital investments in 2000 in the modal CSA. The observations are weighted by the number of patents filed in 2000-04. Robust standard errors are in italics; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Did firm switch CSAs?		
Is firm a bank?	-1.33	-0.92	-0.46
	*0.40****	*0.45***	*0.98*
Is firm other financial services?	0.0001	0.50	0.98
	0.28	*0.27**	*0.30****
Is firm IT or other?	0.88	0.99	1.41
	*0.19****	*0.16****	*0.20****
Is firm venture-backed?	2.18	2.00	2.30
	*0.51****	*0.50****	*0.54****
Is firm publicly traded?	1.93	1.48	1.48
	*0.09****	*0.11****	*0.12****
Is modal patent in 2000-04 in NY CSA?		-0.53	-0.88
		*0.09****	*0.10****
Is modal patent in 2000-04 in SJ/SF CSA?		0.92	-1.21
		*0.008****	*0.26****
2000 Finance VC investments in modal CSA			2.04
			*0.25****
Number of observations	130	130	130
Weighted observations	3484	3484	3484
p-Value, χ^2-test	0.000	0.000	0.000
Pseudo R^2	0.31	0.39	0.40

Appendix A: Major Judicial Decisions and Policy Changes Post-State Street that Affected Financial Patenting

Two important Supreme Court decisions revisited the validity of business method patents during the period studied in this paper (2000-2019). First, in *Bilski v. Kappos*, the Supreme Court affirmed a CAFC decision rejecting the patentability of a method for hedging against price risk in commodities trading but also rejected a *per se* exclusion against patenting business methods.[21] The decision also rejected the judicial standard by which the CAFC had assessed the patentability of business method patents, which injected uncertainty into questions about the validity of such patents.[22]

Next, in June 2014, the Supreme Court ruled in *Alice Corp. v. CLS Bank* that Alice's patent for a computerized trading program that mitigated settlement risk and facilitated the exchange of financial obligations was invalid. The Court found the patent to be merely an abstract idea and thus ineligible for patent protection.[23] At the same time, the Court again made no categorical rejection of business methods or software, *Alice* amplified concerns over the extent of financial-related software patentability.

Patent law changes in 2011 also affected financial patenting. Specifically, the Leahy-Smith America Invents Act (P.L. 112-29) added a new method of post-grant review for "covered business methods" (CBMs), a provision which took was in effect between 2012 and 202. This legislation was motivated by critics of the financial patents, summarized in Hunter (2004, Table 1), who questioned (a) the capabilities of the USPTO to process applications, (b) the validity of such patents in terms of obviousness and novelty, and (c) its overall impact on innovation and competition.

In this context, a CBM is essentially a financial patent.[24] The provision was meant to stifle litigation over questionable patents by enabling alleged infringers being sued in district court to challenge patent validity in a less expensive forum with a faster timeline, before a board perceived as being harsher on questions of patentability. Practitioners suggest that while current attitudes

---

[21]"Section 101 similarly precludes a reading of the term 'process' that would categorically exclude business methods." See *Bilski v. Kappos*, 561 U.S. 593 (2010).

[22]The *en banc* CAFC rejected its prior test for determining whether a claimed invention was a patentable "process" under Patent Act, 35 U. S. C. §101—i.e., whether the invention produced a "useful, concrete, and tangible result," as delineated in *State Street*—holding instead that a claimed process is patent eligible "if: (1) it is tied to a particular machine or apparatus, or (2) it transforms a particular article into a different state or thing." See *In re Bilski*, 545 F.3d 943, 88 U.S.P.Q.2d 1385 (Fed. Cir. 2008).

[23]In particular, the Supreme Court held that "an instruction to apply the abstract idea of intermediated settlement using some unspecified, generic computer is not 'enough' to transform the abstract idea into a patent-eligible invention." See *Alice Corp. v. CLS Bank Int'l* 573 U.S. 208 (2014).

[24] A covered business method patent is defined as "a patent that claims a method or corresponding apparatus for performing data processing or other operations used in the practice, administration, or management of a financial product or service...." 37 *C.F.R.* 42.301(a).

towards granting finance patents are quite permissive within the USPTO, the Federal Circuit is taking a harder line on the validity of finance patents in their rulings.

The ambiguities associated with finance patents in the U.S. have also manifested elsewhere. European patent law explicitly excludes methods of doing business and finance from patent protection. But given the complexity of the definitions, some finance patents appear to have made it past these categorical exclusions. Meanwhile, Japan has shifted from one of the most skeptical patent offices regarding business methods to a much more permissive one: its rejection rate for these patents, of which finance constitutes a considerable number, fell from 92% in 2000 to 34% in 2012 through 2014 (Japanese Patent Office, 2019).

Appendix A: Financial Database Validation Analyses

Auditing the Sorting between Finance and Non-Finance Patents

Within our initial sample, there were 66,534 patents assigned to CPC subclasses G06Q. Of these, 17,511 were assigned to CPC groups G06Q 20 or 40, and the remaining 47,023 to other groups. These patents were divided with random assignment, with 70% (45,174) of the patents as the training data, and 30% (19,360) patents as the testing data.

As is routine with machine learning models, after we estimated the model with the training data, we tested its accuracy using the testing data: that is, we used the testing data to quantify the extent to which the model successfully distinguished between patents that were actually in CPC groups G06Q 20 and 40 and those that were not. Our chosen model operated with about 90 percent sensitivity and specificity: that is, the true positive and true negative rates were both quite high.

Even so, the test set contained 1,426 patents (out of 14,106) that were not actually in CPC groups G06Q 20 and 40 that were predicted to be financial (false positives), and 526 patents in CPC groups G06Q 20 and 40 (out of 5,253) that were predicted to be non-financial (false negatives). (See the schematic below.) To determine whether these inaccuracies represented the performance limits of our model or suggested some noise in the primary CPC codes we used to classify patents, we had a research assistant audit a 10% random sample from each group of misclassifications (false positives and false negatives). He read the title and abstract (and more text if needed) and determined whether the patent is financial or not based on these descriptions.

		Predicted		
		Negative	Positive	Total
Actual	Negative	True Negative (12,680)	False Positive (1,426)	Actual Negative (14,106)
	Positive	False Negative (526)	True Positive (4,727)	Actual Positive (5,253)

The research assistant found that 61 out of 143 (43 percent) allegedly false positives were actually financial patents, and that 39 out of 53 (74 percent) allegedly false negatives were actually not financial patents. In other words, of the patents not included in CPC groups G06Q 20 and 40 but predicted to be financial, 43% percent turned out to actually be financial upon an examination of the patent text itself. Similarly, of the patents included in CPC groups G06Q 20 and 40 but predicted to be not financial, 74% turned out to be not financial. These results broadly suggest some error in the labelling for "marginal" patents——those patents for which a judgment call is difficult.

These results raised the concern that the initial labelling of patents in the training and test sets based on CPC codes could be erroneous. To satisfy ourselves that this was not the case, and that the large inaccuracies only affected approximately 10 percent of the data (the "marginal" patents),

we had the same research assistant do a similar audit for the "true positives" and "true negatives": those patents that the model correctly predicted were or were note in CPC groups 20 and 40. He found that 231 out of 254 (91%) true positives (patents with CPC codes in G06Q 20 or 40 and predicted to be financial by the model) were actually financial patents. He also found that only 4 out of 95 (96%) true negatives (patents not in G06Q 20 or 40 and predicted to be "not fintech" by the model) are financial in nature. These accuracy levels were much higher than the 43 and 74 percent accuracies found in samples of false positives and negatives, and suggested that the low levels of accuracy in those samples stemmed from the difficulty of determining whether the patent is financial or not, rather than from any major flaw in the CPC classifications.

We then used the model to identify financial patents with a primary subclass or group outside of G06Q, where we believed (after analyzing other common CPC codes for known financial patents) finance patents could be located. We did not generate a "test set" to evaluate the performance of our model when deployed to patents with a primary CPC subclass outside of G06Q. Instead, we had a research assistant audit small samples of patents that were predicted to be financial or not financial when we deployed the model on these "supplemental" subclasses. He found that 23 out of 67 (34%) patents identified as financial were actually financial, and that 51 out of 53 (96%) identified as not financial were actually not financial. For these patents, our model appeared to have high sensitivity but relatively poor specificity, a common problem.

This was expected because we did not include any financial patents with a primary CPC subclass outside of G06Q in the treatment group when we built and tested the machine learning model. Hence just like many other in many tests and applications, it is easier to precisely eliminate negative cases than identify positive ones. As a result, our list of financial patents should be considered a broad and perhaps over-inclusive sample of true financial patents.

Assessing an Alternative Assignment Method

We also explored whether an alternative approach using patents assigned to fintech firms would have generated better results. Using the lists mentioned above, we had a research assistant manually search Google patents to identify the standardized assignee names of known fintech firms in the underlying IFI Claims patent data. Through these searches, and additional web searches and examinations of patent filings, our assistant was able to identify common spellings of each firm and some of its publicly known subsidiaries.

Using this list of standardized firm names, we identified 1,065 patents assigned to known fintech firms. We found that only 32 percent of these patents ended up on our final list of financial patents using the methodology described above. Another research assistant audited a random sample of 101 of the patents assigned to known fintech firms that did not end up on our list. He found that only six of these patents were indeed financial. These results confirmed our belief that using firm names to label financial patents would not be appropriate in this context.

Issues with Proper Assignment Names

After pulling over patent-level data from Derwent, we noticed that Derwent often carried the inventor or applicant over into the assignee field in many instances in which it was not appropriate

to do so (i.e., when the inventors were not assignees in the raw USPTO data from IFI). We therefore audited a two percent sample of the financial patents with multiple assignees (a sample of 150 patents) by having research assistants categorize the nature of the discrepancies between Derwent data and raw patent data. We found that in most instances (136 out of 150), the data either agreed (and contained only inventors or corporate entities as assignees) or the data disagreed but Derwent simply appended the inventor names onto a list of true corporate assignees. In some instances (13 out of 150), the raw data contained no assignee but the Derwent data listed all the inventors, a result which is consistent with the pre-2012 rule vesting ownership in inventors in the absence of a written assignment (see MPEP Section 301, 37 C.F.R. 3.1(I)). Reflecting these findings, we purged all inventor names from the assignee field except when the only assignees were the inventors. In one instance (0.7 percent of the sample), in actuality the patent listed both the inventor and corporate entities as assignees. In this instance, our process caused a discrepancy by purging the individual inventor from the list of assignees. These incorrect corrections affected only a very small portion of the data set.

Capital IQ Identifier Issues

We were concerned that the Capital IQ identifiers used in our financial patent dataset might be associated with subsidiaries rather than the parent companies, despite our efforts to ensure matching to the ultimate parent company. By looking at the list of 2011 Systemically Important Financial Institutions (listed at the last page of https://www.fsb.org/wp-content/uploads/Policy-Measures-to-Address-Systemically-Important-Financial-Institutions.pdf), we identified 1,611 patents with a first assignee among the SIFI list. After auditing this list, we found that 1,563 out of 1,611 SIFI patents (97 percent accuracy) were assigned to the correct parent companies. And if we only looked at the SIFI banks who were awarded more than 20 patents (their granted patents covered 95% of all SIFI patents), the accuracy rate was further increased to 98.7% (1511 out of 1531 patents were correctly assigned).

We identified two reasons for the erroneous matching with subsidiaries instead of parent companies. First, the UVA dataset on which we heavily relied has some errors. For instance, the UVA dataset assigns separate identifiers for "Morgan Stanley Capital International Inc." and "Morgan Stanley," though all patents associated with these companies should be assigned to a single parent company identifier. Second, our fuzzy name matching efforts also had some errors. For example, we matched some patents to the subsidiary "Credit Suisse Securities (USA) LLC" instead of its parent "Credit Suisse."

In total, 5 SIFI patents were not assigned to any identifiers by either UVA dataset or fuzzy name matching method, and 43 SIFI patents were wrongly assigned to the subsidiaries rather than their corporate parents. We did not see any time distribution differences among those problematic patents. In sum, though our analysis of the SIFI patents suggests that there are some errors in our dataset when it comes to matching patents with parent companies, they errors seem to affect only a small percentage of the data and should not affect the analysis materially.

Appendix C: CSA Database Construction

The U.S. Bureau of the Census has used varying definitions for urban areas over time and has periodically redrawn the boundaries of these regions. We attempted to be as consistent as possible in defining geographic regions, subject to the limitations of data availability.

First, we associated each patent to a local geography using the county FIPS of the first inventor, provided by Patentsview. We then matched county FIPS to 2013 CSA regions using Census/NBER crosswalk discussed in the text of the paper. We then aggregated simple and weighted patent counts to the CSA-year level using this mapping. Patents associated with counties outside of the 166 CSAs (we excluded the three CSAs in Puerto Rico) were collectively associated with an aggregate "Not a CSA Region," but observations in this region were dropped from the analysis. The 2013 CSAs include all major finance patenting hubs with the exception of Austin, Texas: the Census Bureau recognized the Austin-Round Rock-Marble Falls, TX CSA in the late 2000s and early 2010s, but then eliminated it after the criteria for selecting CSAs changed.

We similarly obtained from VentureXpert county-by-county data (and the associated FIPS code) on venture capital financing (both for all transactions and for finance transactions) between 2000 and 2018. We computed the number of deals and transaction volume using the 2013 mapping from counties to CSAs.

We then collected additional annual data about each CSA that existed in 2013, including: (1) total population, (2) total number of households, (3) median household income, (4) total adult (aged 25 or older) population, (5) total adult population with an education level of a bachelor's degree or higher, (6) the number of non-employer establishments in finance or insurance (NAICS 52), and (7) the number of employees in finance or insurance.

For census year 2000, data was collected at the county level and aggregated to the CSA-level using the Census/NBER crosswalk. For variables (1)-(2) and (4)-(7), the data were aggregated with simple summations. For median household income, the CSA-level value is a weighted mean of the county median incomes using county households as weights.

For non-decennial census years, these data were not available for the county level in most cases. Variables (1) through (5) above were reported annually for each CSA, however, in the American Community Survey. These data at the CSA level, however, had three limitations:

- The ACS data for 2001-04 (as well as 2000, which we did not use) was removed by the Census Bureau from its online servers due to reliability concerns.
- As noted above, the Census Bureau adds and sometimes removes urban areas from its list of CSAs. The ACS data were reported only for CSAs that were on the Census Bureau list at the time.
- The boundaries of CSAs may change over time.

As a result, for variables (1)-(5), we generally imputed missing values using a simple linear regression based on non-missing data in instances where the variable had two or more

observations. If only one observation of a variable within a CSA was available, we attributed that value to all years in which the variable is missing, making the variable constant over time.

Variables (6)-(7) were taken from the quinquennial economic census from years 2002, 2007, 2012, and 2017. We generally imputed 2000 and 2001 observations in a CSA using the 2002 observation, and the 2018 observation using the 2017 observation. For years 2003-06, 2008-11, and 2013-16, we generally imputed missing values by fitting a linear regression using data from 2002, 2007, 2012, and 2017.

Figure A-1: Financial Patents Supervised Machine Learning Flow Chart. The figure presents how we predict financial patents using supervised machine learning. First, the labeled patents (financial data and non-financial data) are divided into training data (70%) and test data (30%). Then the machine is trained using the training data. Then different ML models are compared and the best model is selected as our prediction model. Finally, the unlabeled supplemental patents are used as the input of the prediction model, and the predicted labels of these patents are obtained.

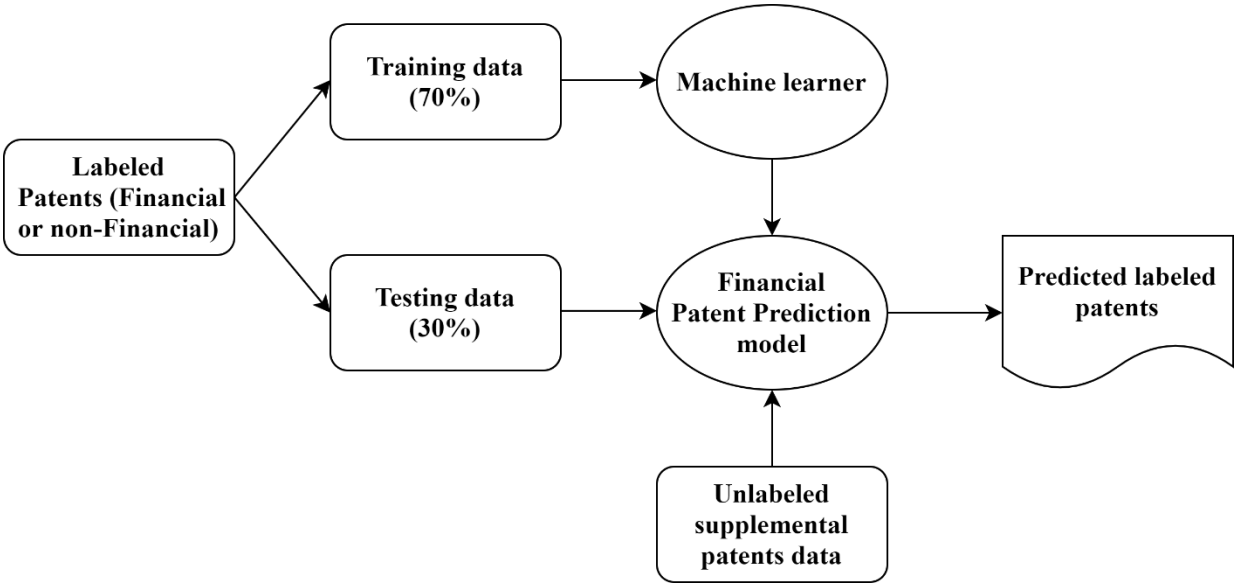


Figure A-2: Financial Patents Machine Learning Model Architecture. The figure presents the structure of our final machine-learning model. Compared to the text-only model, the text-inventor model slightly decreases sensitivity from 91.3 to 89.9 percent (a drop of 1.4 percentage points), but significantly improves specificity from 85.3 to 90.0 percent (an increase of 4.7 percentage points). With about 90 percent sensitivity and specificity, respectively, we consider this model to be reliable and scalable for predictions.

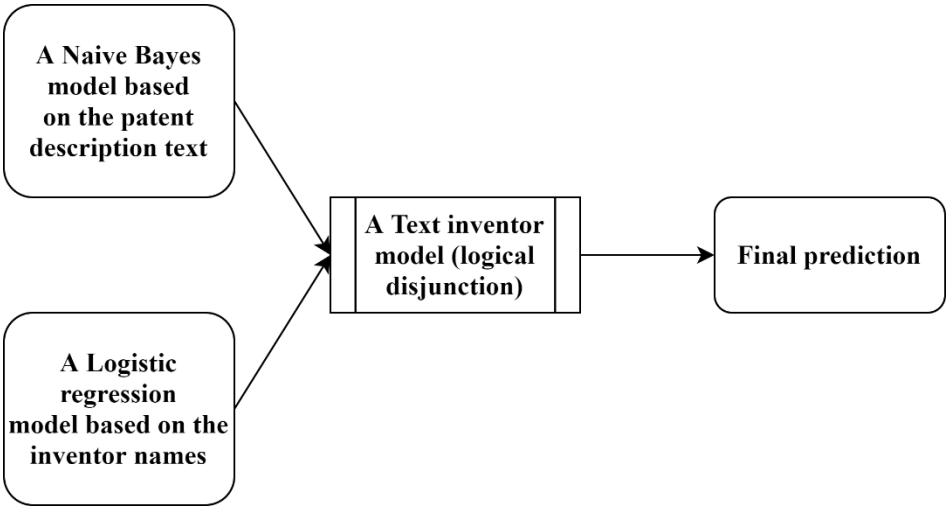


Figure A-3: Fuzzy Name Matching between Assignee Names and Capital IQ Names. This figure presents how we use a Levenshtein distance-based fuzzy name matching techniques to match the remainder of assignee names with 12 million firm names in the Capital IQ database. The Capital IQ database was divided into three3 subsets, with four million company names in each subset. After examining the data, we determine that matches in which the matching score is 0.95 or higher were so accurate that they could be adopted without further scrutiny. Similarly, matches with scores below 0.8 were so poor that they could be rejected outright. For matches with scores between 0.8 and 0.95, the results were inspected to determine which is appropriate. In the last step, the high confidence results are merged.

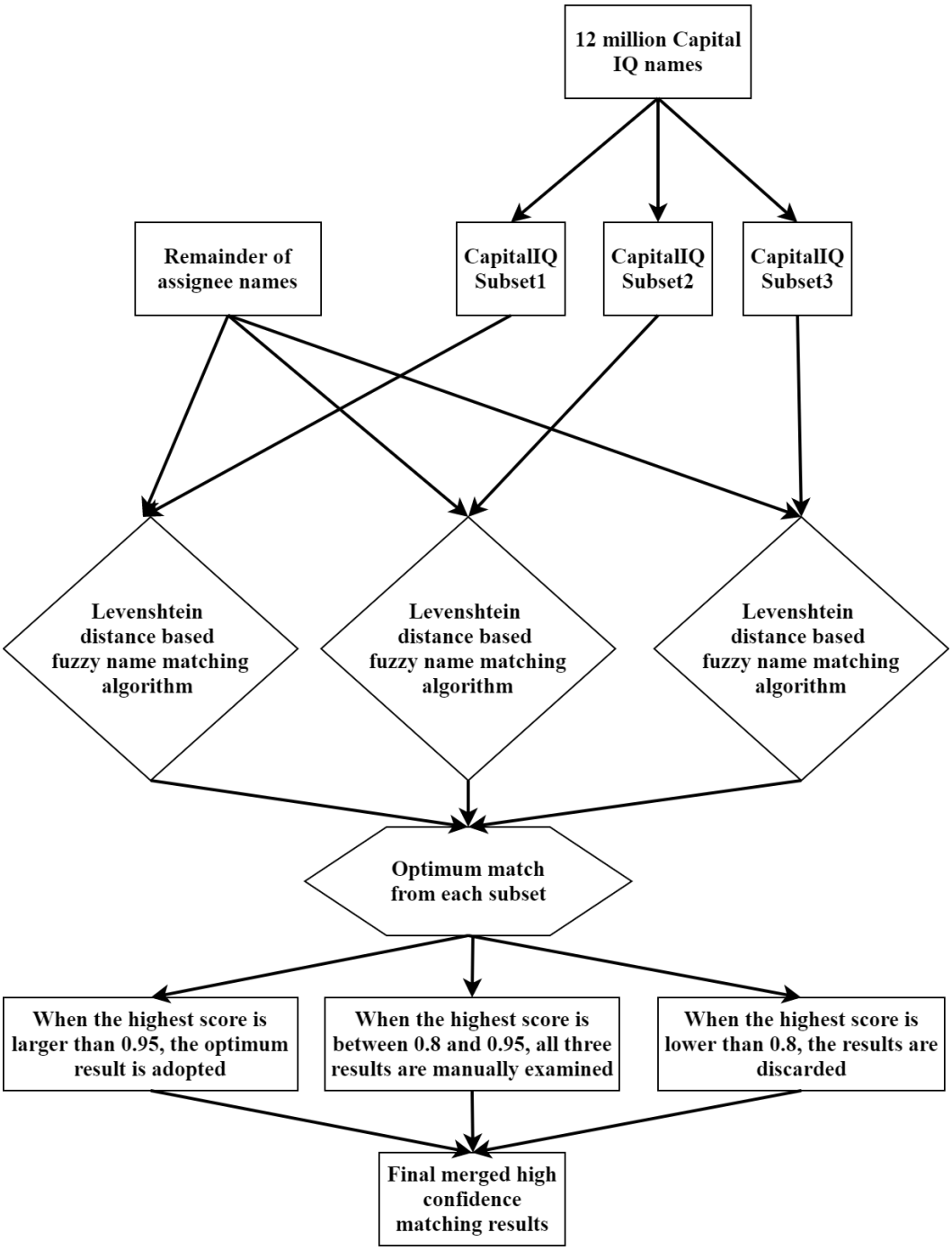


Figure A-4: An Overview of the Financial Dataset Construction Procedure. The first step in our process was to obtain additional patent-level data on financial patents from Derwent. We obtained from Patentsview the patent assignee type and a host of other information. Then the assignee's Capital IQ ID was obtained from either the UVA dataset or fuzzy name matching with Capital IQ company names. The detailed Capital IQ data were merged using a crosswalk file. Finally, we used keywords to describe the patent.

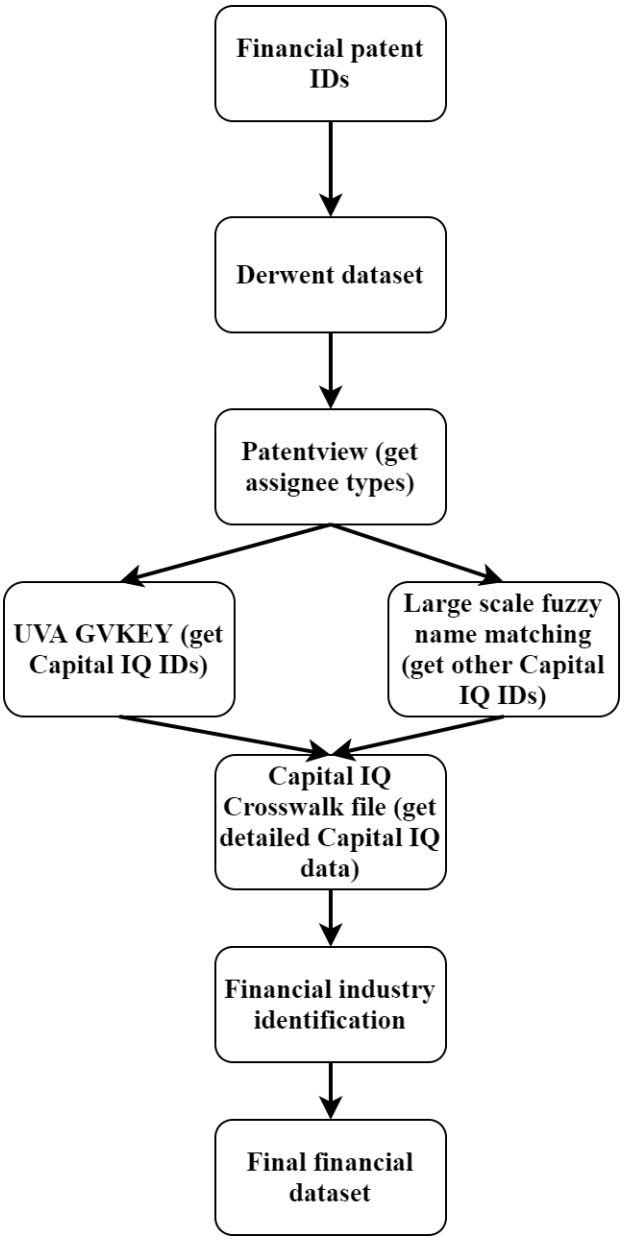


Table A-1: List of keywords.

Accounting	**Consumer Banking**	**Communications**	**Cryptocurrencies**	**Currency**	**Investment Banking**
Accounting	Bridge Finance	Broadcast	Altcoin	Currency Conversion	Asset Analysis
Accounts Payable	Commercial Loan	Broadcasts	Bitcoin	Exchange Rate	Asset Characterization
Accounts Payables	Covenant	Communication	Blockchain	Foreign Exchange	Bid Ask
Accounts Receivable	Debtor Finance	Communications	Cryptocurrency	Forex	Bond
Accounts Receivables	Debtor In Possession	Message	Distributed Ledger	Swap	Call Option
Audit	Default	Messages	Initial Coin Offering		Chinese Wall
Auditor	Event	News Feed	Token		Derivative
Bookkeeper	Indicator Lending Rate	News Feeds			Dummy Order
Budget	Interest Coverage				Dummy Orders
Budgeting	Letter Of Credit				Gilt
Cash Flows	Line Of Credit				Hair Cut

Controller	Material Adverse Change				Hedge Fund
FIFO	Sweep Account				Hidden Liquidity
Financial Controls	Term Loan				Initial Public Offering
First In First Out	Zero Balance Account				Liquidity Pool
Forecasting					Liquidity Provider
Free Cash Flows					Margin
GAAP					Moving Average
Generally Accepted Accounting Principles					Option
Gross Margin					Order Book
Information System					Price Level
Interest Coverage					Price Levels
Inventory					Private Equity

Last In First Out					Put Option
LIFO					Short Selling
Net Present Value					Trading Protocol
Net Working Capital					Trading Protocols
Payables					Valuation
Payback					
Payroll Taxes					
Quick Ratio					
Working Capital					

Table A-1 (continued): List of keywords.

Insurance	Payments	Real Estate	Retail Banking	Security	Wealth Management
Actuarial	Authorized	Appraisal	ATM	Authentic	Active Management
Auto Insurance	Card Reader	Cap Rate	Automatic Teller Machine	Authenticate	Asset Allocation
Beneficiary	Cash Register	Closing Costs	Availability Policy	Authenticating	Asset Class
Catastrophe Bond	Contactless	Closing Fee	Balance Transfer	Biometric	Back-End Load
Catastrophe Loss	Credit Transaction	Conforming Loan	Certificate Of Deposit	Cipher	Benchmark
Claims Adjustment	Customer	Cumulative Loan To Value	Check	Ciphers	Capital Appreciation
Coinsurance	Debit Transaction	Deed	Checking	Credential	Capital Preservation
Crash	Interbank Fee	Delinquency	Checks	Credentials	Custodian
Disability	Keypad	Dual Agency	Credit Score	Cryptographic	ETF
Driving Behavior	Kiosks	Easement	Direct Deposit	Decipher	Exchange Traded Fund
Driving Environment	Merchant	Eminent Domain	Direct Payroll Deposit	Deciphers	Financial Industry Regulatory Authority

Earned Premium	NFC	Escrow	Interbank Fee	Decrypt	FINRA
Home Insurance	Payment	Eviction	Money Market	Decryption	Front-End Load
Homeowners Insurance	Point Of Sale	Foreclosure	NOW Account	Detection	Index Fund
Indemnity	POS	Home Equity	Online Banking	Encrypt	Individual Retirement Account
Insurance Risk		Home Warranty	Overdraft	Encryption	Mutual Funds
Life Insurance		Jumbo Loan	Passbook	Fraud	Prospectus
Life Settlement		Loan To Value	Savings	Fraudulent	Prospectuses
Long-Term Care		Mortgage	Student Loan	Identifier	Target Date Fund
Malpractice		Non-Conforming Loan	Time Deposit	Identity	Tax Avoidance
Reinsurance		Prepayment	Withdrawal Fee	Public Key	Tax Benefits
Structured Settlement		Real Estate Investment Trust		Secure Key	Tax Cost
Term Insurance		Realtor		Security	Tax Costs

Umbrella Liability		Refinancing		Spoofing	Tax Deduction
Vehicle Damage		REIT		Symmetric Key	Tax Deductions
		Tax Lien		Theft	Wrap Fee
		Title Search		Token	
		Zoning		Verify	

Table A-2: Searching strategy for patent categorization. We search each section of the patent in sequence, for those patents without a keyword match in the earlier sections. We classify the remaining 345 patents without a keyword match through a manual review of the patent text.

	Section of the Patent Examined			
	Abstract	*First 100 Words of Background*	*Entirety of Background Section*	*Entirety of Patent Text*
Patents Searched	24288	5062	2107	1030
Keywords Found:				
0	5062	2107	1030	345
1	9179	1891	321	11
2	6805	866	263	28
3	2606	166	244	70
4	555	30	120	122
5	74	2	64	140
6	6	0	53	146
7	1	0	9	115
8	0	0	3	42
9	0	0	0	8
10	0	0	0	3

Table A-3: Number of keywords found. The table reports the number of cases with zero, one, and more than one keywords, and the mean number of keywords found.

Patent Section Examined:	*Total Search Space*	*# with 0 Keywords*	*# with 1 Keyword*	*# with >1 Keyword*	*Mean Keyword Count for >1 Cases*
Abstract	24288	5062	9179	10047	2.39
First 100 Words of Background	5062	2107	1891	1064	2.22
Entirety of Background Section	2107	1030	321	756	3.26
Entirety of Patent Text	1030	345	11	674	5.30

Table A-4. Decomposition of the rise of financial patenting. The table presents results of a regression analysis of the rise of patenting, where the dependent variable is the number of financial patents awarded in each year-assignee firm industry-patent type-inventor location cell. The table reports the results of F-tests of the joint significance of various sets of independent variables.

Set of Independent Variables	*F-statistic*	*p-Value*
Year Fixed Effects	34.47	0.000
Assignee Industry Fixed Effects	110.82	0.000
Patent Type Fixed Effects	17.00	0.000
Inventor Location Fixed Effect	216.67	0.000
Year * Assignee Industry Fixed Effects	6.43	0.000
Year * Patent Type Fixed Effects	**1.37**	0.081
Year * Inventor Location Fixed Effects	11.45	0.000

Table A-5. Regression analysis of academic citations and patent characteristics. The sample consists of all patents applied for between 2000 and 2018 and awarded by February 2019. The dependent variables are the number of academic citations in these patents, the number of citations to business, economics, and finance journals, the number to Top 3 finance journals, and the mean age of the citations in each patent (years between the article publication and patent application date). In Panel A, the key independent variable is a dummy whether the patent is financial; in Panel B, the key independent variables are dummies whether the patent is financial, the assignee is a U.S. corporation, a foreign corporation, a U.S. university or another type, and the interactions between assignee type and the financial patent dummies (other assignees is the omitted category); and in Panel C, the key independent variables are dummies whether the patent is financial, the assignee is venture backed, and the interactions between the dummies. All regressions control for the time period and inventor location. Robust standard errors are in italics; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Academic Citations		*Bus/Econ/Fin Citations*		*Top 3 Citations*		*Citation Age*
Panel A							
Financial patent	-8.15		0.71		0.07		-0.63
	*0.45****		*0.01****		*0.001****		*0.13****
Panel B							
Financial patent	-1.70		0.47		0.04		-2.63
	1.36		*0.02****		*0.003****		*1.24***
U.S. corporation	5.98		0.04		0.0001		-1.63
	*0.16****		*0.002****		*0.0003*		*0.10****
Foreign corporation	2.93		0.02		-0.0001		-2.42
	*0.23****		*0.0004****		*0.0004*		*0.11****
U.S. university	44.83		0.04		0.0001		-1.36
	*0.31****		*0.005****		*0.0006*		*0.11****
Financial * U.S. corporation	-6.11		0.28		0.04		2.04
	*1.44****		*0.02****		*0.0003****		*1.25*
Financial * Foreign corporation	-3.10		0.05		-0.01		1.44
	2.63		*0.05*		*0.005****		*1.38*
Financial * U.S. university	-36.23		0.37		0.03		2.59
	*7.41****		*0.13****		*0.01****		*1.86*
Panel C							
Financial patent	-8.06		0.80		0.08		-0.36
	*0.52****		*0.01****		*0.0001****		*0.13****
Venture-backed firm	9.82		0.02		-0.0003		0.42
	*0.21****		*0.004****		*0.0004*		*0.05****
Financial * venture-backed	-8.67		-0.02		0.05		-3.78
	*1.95****		*0.03*		*0.004***		*0.50****

Table A-6. Financial patenting in three key regions. Panel A presents the characteristics of patents applied for in each five-year period in San Jose-San Francisco-Oakland CSA; Panel B in the New York-Newark CSA; and Panel C in the South Atlantic and East North Central U.S. Census divisions. The table presents for finance patents applied for between 2000 and 2018 and awarded by February 2019 the share of all finance patents applied for from the region, the share of all finance patents assigned to a CSA, and the share of all finance patents assigned to a firm of a given type. We run a regression using each CSA in each five-year period as an observation, with the patent share in a given five-year period as the dependent variable and independent variables controlling for the CSA, the time tend, the interaction of these two measures, and various demographic characteristics of the CSA in that period. The t-statistic is from the interaction term. All shares are computed using patent counts, citation weights, and Kogan et al. (2017) weights. We assign patents based on the location of the first inventor.

Panel A: San Jose-San Francisco-Oakland, CA CSA

	2000-04	*2005-09*	*2010-14*	*2015-18*	*t-stat*
Unweighted					
Share of all patenting	8.5%	10.7%	15.7%	18.3%	20.37
Share of all CSA patenting	14.2%	16.9%	23.2%	28.0%	22.01
Normalized by CSA patenting of that type					
Small firms	19.5%	18.6%	21.4%	25.0%	4.11
Medium firms	18.2%	28.8%	34.0%	48.6%	16.00
Large firms	10.7%	11.0%	26.0%	22.9%	4.55
SIFIs	3.7%	3.6%	6.2%	6.4%	4.42
Banking industry	4.6%	3.3%	6.3%	6.6%	3.03
Other finance industry	8.1%	4.2%	6.5%	2.9%	-3.67
Payment industry	15.3%	39.0%	58.0%	63.9%	8.02
IT/other industry	16.1%	18.5%	22.3%	23.5%	8.93
Cite weighted					
Share of all patenting	11.5%	16.2%	21.3%	21.5%	5.66
Share of all CSA patenting	16.7%	23.4%	28.4%	29.6%	5.71
Normalized by CSA patenting of that type					
Small firms	21.2%	24.9%	26.9%	20.9%	-0.10
Medium firms	21.3%	45.8%	49.5%	72.4%	12.73
Large firms	10.2%	14.2%	30.6%	11.4%	0.58
SIFIs	6.2%	7.1%	8.4%	15.7%	4.34
Banking industry	5.4%	5.6%	8.9%	14.5%	5.52
Other finance industry	9.4%	5.3%	4.2%	0.0%	-7.97
Payment industry	26.4%	60.5%	72.1%	76.8%	4.81
IT/other industry	17.8%	21.7%	24.0%	33.8%	7.87
Kogan weighted					
Share of all patenting	8.4%	14.8%	25.0%	25.6%	7.41
Share of all CSA patenting	10.7%	18.7%	32.6%	34.4%	8.64
Normalized by CSA patenting of that type					
Small firms	33.9%	42.2%	16.7%	0.0%	-4.07
Medium firms	19.6%	55.5%	38.0%	42.1%	1.15
Large firms	8.1%	6.3%	31.7%	32.6%	6.32
SIFIs	6.4%	5.1%	13.0%	15.1%	6.28
Banking industry	9.2%	5.9%	13.9%	14.9%	3.57
Other finance industry	2.5%	1.9%	3.5%	0.4%	-1.06
Payment industry	11.4%	72.7%	63.4%	64.6%	2.08
IT/other industry	32.6%	35.6%	58.5%	40.4%	1.46

Panel B: New York-Newark CSA

	2000-04	*2005-09*	*2010-14*	*2015-18*	*t-stat*
Unweighted					
Share of all patenting	13.4%	11.6%	9.5%	5.7%	-8.49
Share of all CSA patenting	22.4%	18.4%	14.2%	8.7%	-15.74
Normalized by CSA patenting of that type					
Small firms	14.4%	16.5%	14.3%	25.0%	2.59
Medium firms	15.6%	11.6%	9.8%	6.2%	-14.46
Large firms	32.0%	23.2%	15.6%	5.6%	-33.64
SIFIs	63.3%	33.4%	24.1%	4.0%	-13.42
Banking industry	27.6%	18.0%	12.5%	6.0%	-22.74
Other finance industry	56.1%	46.2%	33.0%	4.4%	-7.71
Payment industry	11.1%	6.7%	6.8%	5.8%	-4.82
IT/other industry	16.7%	13.7%	11.6%	11.4%	-10.89
Cite weighted					
Share of all patenting	14.6%	7.8%	6.4%	5.7%	-5.04
Share of all CSA patenting	21.3%	11.3%	8.5%	7.8%	-5.01
Normalized by CSA patenting of that type					
Small firms	5.0%	6.5%	22.7%	42.5%	6.43
Medium firms	16.2%	6.5%	4.1%	6.0%	-3.11
Large firms	33.1%	12.3%	7.1%	1.7%	-5.86
SIFIs	50.8%	12.4%	7.1%	7.7%	-3.12
Banking industry	34.5%	12.3%	9.4%	14.7%	-2.13
Other finance industry	54.1%	33.8%	9.6%	0.0%	-14.64
Payment industry	17.0%	3.1%	3.2%	5.7%	-2.17
IT/other industry	16.0%	10.2%	9.6%	15.0%	-0.68
Kogan weighted					
Share of all patenting	34.6%	19.8%	14.4%	5.7%	-12.87
Share of all CSA patenting	44.2%	25.0%	18.9%	7.7%	-12.09
Normalized by CSA patenting of that type					
Small firms	28.2%	10.5%	6.6%	0.0%	-7.50
Medium firms	14.9%	12.9%	18.1%	12.0%	-0.90
Large firms	52.0%	29.1%	18.9%	6.5%	-12.70
SIFIs	57.7%	30.7%	24.0%	5.5%	-12.63
Banking industry	34.5%	19.2%	16.3%	6.0%	-11.90
Other finance industry	77.9%	65.8%	52.1%	4.8%	-5.64
Payment industry	16.1%	8.4%	13.2%	7.6%	-3.33
IT/other industry	7.4%	5.3%	5.8%	18.3%	-1.89

Panel C: Charlotte-Concord CSA

	2000-04	*2005-09*	*2010-14*	*2015-18*	*t-stat*
Unweighted					
Share of all patenting	0.3%	1.7%	2.3%	4.2%	13.52
Share of all CSA patenting	0.5%	2.7%	3.3%	6.5%	11.76
Normalized by CSA patenting of that type					
Small firms	0.0%	0.0%	0.0%	0.0%	-0.55
Medium firms	0.0%	0.2%	0.4%	0.3%	0.68
Large firms	0.7%	10.0%	11.0%	16.9%	8.36
SIFIs	2.3%	27.0%	36.1%	54.9%	16.63
Banking industry	3.1%	25.3%	33.1%	52.2%	17.95
Other finance industry	0.4%	1.0%	0.3%	1.0%	-0.75
Payment industry	0.0%	0.0%	0.6%	0.6%	1.54
IT/other industry	0.3%	0.3%	0.3%	0.7%	0.77
Cite weighted					
Share of all patenting	0.4%	1.5%	3.2%	1.6%	1.32
Share of all CSA patenting	0.6%	2.2%	4.3%	2.3%	1.31
Normalized by CSA patenting of that type					
Small firms	0.0%	0.0%	0.0%	0.0%	-0.60
Medium firms	0.0%	0.0%	0.0%	0.0%	0.16
Large firms	1.2%	9.0%	6.9%	4.7%	0.59
SIFIs	3.8%	32.3%	43.7%	63.0%	14.73
Banking industry	0.4%	2.2%	3.7%	2.3%	35.36
Other finance industry	0.8%	0.8%	0.0%	0.0%	-2.42
Payment industry	0.0%	0.0%	0.0%	0.0%	0.24
IT/other industry	0.2%	0.1%	3.2%	0.4%	0.64
Kogan weighted					
Share of all patenting	0.4%	11.0%	8.7%	13.7%	4.15
Share of all CSA patenting	0.5%	13.9%	11.4%	18.3%	4.69
Normalized by CSA patenting of that type					
Small firms	0.0%	0.0%	0.0%	0.0%	-0.25
Medium firms	0.0%	0.1%	3.8%	1.1%	1.33
Large firms	0.7%	18.5%	13.6%	22.8%	4.07
SIFIs	0.9%	22.9%	23.6%	39.5%	8.94
Banking industry	1.3%	26.8%	25.2%	39.2%	6.08
Other finance industry	0.0%	0.1%	0.0%	1.3%	0.32
Payment industry	0.0%	0.0%	3.3%	1.8%	2.31
IT/other industry	0.0%	0.0%	0.1%	0.1%	0.32

Table A-7. Financing patenting by U.S. region over time. The table presents the share of patenting by region for the nine U.S. Census regions. All analyses use patents applied for between 2000 and 2018 and awarded by February 2019. The table presents financial patents as a share of all patents, computed using patent counts, citation weights, and Kogan et al. (2017) weights. We assign patents based on the location of the first inventor.

	Patent Count					Citation Weighted					Kogan et al. Weighted			
	2000-04	2005-09	2010-14	2015-18		2000-04	2005-09	2010-14	2015-18		2000-04	2005-09	2010-14	2015-18
East North Central	8.2%	11.0%	13.0%	10.9%		9.2%	9.2%	13.6%	27.9%		4.7%	6.9%	6.2%	6.1%
East South Central	0.6%	0.6%	0.4%	0.3%		0.5%	0.6%	0.2%	0.2%		0.4%	0.2%	0.3%	0.1%
Middle Atlantic	15.6%	14.9%	12.0%	7.1%		16.4%	12.4%	13.7%	9.3%		42.4%	26.8%	19.3%	7.3%
Mountain	5.9%	5.9%	5.2%	5.3%		7.5%	5.8%	4.0%	2.8%		6.3%	5.6%	3.1%	2.7%
New England	6.4%	5.5%	6.0%	4.5%		6.5%	4.0%	4.0%	2.9%		4.7%	3.2%	3.9%	2.2%
Pacific	16.7%	19.2%	25.5%	26.9%		22.4%	27.4%	34.6%	32.6%		11.3%	19.0%	32.7%	33.5%
South Atlantic	11.2%	12.3%	11.7%	15.2%		12.5%	15.4%	11.8%	6.4%		15.1%	21.1%	16.9%	23.9%
West North Central	3.2%	4.0%	3.3%	3.3%		2.6%	4.0%	3.2%	1.0%		3.6%	4.8%	4.1%	7.2%
West South Central	5.4%	6.6%	4.8%	4.4%		5.6%	9.0%	5.5%	7.4%		5.8%	5.3%	2.8%	4.1%
Outside the US	26.8%	20.0%	18.1%	22.1%		16.8%	12.2%	9.4%	9.5%		5.7%	7.1%	10.7%	12.9%