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

Several 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 in 2010 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.[1] 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.[2]
- The court's 2012 decision in *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*,[3] while specifically determining that a method of giving a drug to a patient was not patentable subject matter, was seen as weakening the ability to patent abstract subject matter more generally.
- 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.[4] While 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 2020. 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 evaluate such applications, (b) the validity of issued finance patents in terms of obviousness and novelty, and (c) such patents' overall impact on innovation and competition.

In this context, a CBM is essentially a financial patent.[5] The provision was meant to reduce 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 more skeptical on questions of patentability. Practitioners suggest that while current

[1] "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).

[2] The *en banc* CAFC rejected its prior test for determining whether a claimed invention was a patentable "process" under 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).

[3] 566 U.S. 66 (2012).

[4] 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).

[5] 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).

attitudes 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).

References Not Cited in the Paper

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

Appendix B: Two 20th and Two 21st Century Financial Innovations

20th Century Financial Innovations

Automated teller machine

The automated teller machine, or ATM, enables customers of financial institutions to withdraw funds and complete a variety of other financial transactions (e.g., checking balances). The origins of the ATM have been traced back to Luther Simjian's automated deposit machine installed in New York City in 1961, the Bankograph, which accepted but did not disburse funds. The first modern ATM—with a card reader and cash dispensing—was introduced by Barclays Bank in North London in June 1967. The device was developed not by the bank, but an engineering team led by John Shepherd-Barron of the printing firm De La Rue.

Similar devices were introduced in the subsequent weeks and months by a consortium of Swedish banks, Westminster Bank, Chemical Bank, and Lloyds Bank, each using slightly different principles regarding the technology behind the tokens used to access the device (one-time vs. multiple use tokens/cards, the use of Carbon-14 vs. magnetism in the tokens/cards, and the presence of personal identification codes). Many of the initial devices were developed by start-ups, but larger computer manufacturers firms such as Burroughs and IBM soon entered the market. The diffusion of ATMs appears to have peaked about 2013, when there were 3.5 million devices installed worldwide.

Reflecting the strong reliance of hardware manufacturers on formal intellectual property protection, the ATM inventions were frequently patented. Simjian filed for a U.S. patent on the Bankograph in 1960, which was granted in 1963. An early U.K. patent was awarded to Adrian Ashfield for the concept of a card system for ATM users. One of the most important early patent families was for PIN identifier, which was issued between 1966 and 1970 to a group of engineers working at Smiths Group in the U.K. This patent was licensed by many of the subsequent ATM developers, including IBM and NCR.

Sources and References Not Cited in the Paper

"Automated Teller Machine," *Wikipedia,* https://en.wikipedia.org/wiki/Automated_teller_machine.

Batiz-Lazo, Bernardo, and Robert J. Reid. 2008. "Evidence from the Patent Record on the Development of Cash Dispensing Technology," MPRA Paper 9461, University Library of Munich, Germany.

"Depository Machine Combined with Image Recording Mean," U.S. patent no. 3,079,603, *Google Patents,* https://patents.google.com/patent/US3079603A/en.

Harper, Tom R., and Bernardo Batiz-Lazo. 2013. *Cash Box: The Invention and Globalization of the ATM*, Louisville, KY, Networld Media Group.

CICS

IBM's Customer Information Control System (CICS) transaction processing software, a highly centralized system that ran only on IBM mainframes and "greenscreen terminals," was extensively used by banking, securities, and brokerage firms.

The system was originally developed, starting in 1966, at a series of IBM facilities (first in Illinois, then Palo Alto, and finally in a series of overseas development laboratories), with the objective of meeting the information-handling needs of the public utility industry. Soon after the initial product release in 1969, IBM realized that the product would have robust demand from other vendors as well and broadened its marketing.

CICS was distinguished from its predecessors along two crucial dimensions. The first was its ability to process transactions in real time. Previously, most applications used batch processing, where numbers of punched cards would be prepared and loaded together into a computer. The second distinguishing feature was the development of what is today termed "middleware." As IBM describes it, "CICS also provided a collection of standard general-purpose programs, which were delivered as functions that customers could include in their own applications… such as security, recovery and scalability."[6]

CICS became rapidly adopted by the financial sector, including by banks, insurers, and payments firms. *Personal Computing* magazine characterized it "probably the most successful piece of software of all time. … Millions of users unknowingly activate CICS every day, and if it were to disappear, the world economy would grind to a halt."[7]

Like many of the IBM products of that era, the CICS software was initially free to purchasers of IBM computers. IBM did not seek any formal intellectual property protection for CICS. Rather, CICS was designed to only work with IBM devices, such as the IBM 360 mainframe and a small number of terminals. While the software was not priced, it was estimated by the IBM team that CICS led to over $60 billion in new hardware revenue for IBM.

Interestingly, IBM made application software like CICS open to its IBM customers (perhaps reflecting the computer giant's emphasis on hardware). Users made major contributions to the development of CICS, in some cases (e.g., oil giant Amoco) sharing the code with IBM to distribute to others (akin to a modern-day open-source project) and in others, customizing the code for their own purposes.

While IBM continues to offer CICS to this day, its economic importance has faded. The company failed to update CICS to reflect the radically different computing environment, in large part because the fear of cannibalization of existing sales (a problem that emerged in other product lines at IBM as well, such as relational databases). This failure created an opportunity that many software start-ups took advantage of, most notably BEA Systems, which was founded in 1995 to offer transaction-processing software for the finance and banking industries. Unlike CICS, BEA

[6] https://www.ibm.com/ibm/history/ibm100/us/en/icons/cics/.
[7] *Ibid.*

(which was ultimately acquired by Oracle) offered decentralized, web-based systems running on open standards.

Sources and References Not Cited in the Paper

"BEA Systems," *Wikipedia,* https://en.wikipedia.org/wiki/BEA_Systems.

"CICS," *Wikipedia,* https://en.wikipedia.org/wiki/CICS.

IBM Corporation. 2021. "CICS: Securing Online Transactions," https://www.ibm.com/ibm/history/ibm100/us/en/icons/cics/

Twenty-First Century Financial Innovations

Apple Pay

Apple Pay allows users to make purchases and bank transactions, with much greater security than traditional credit card transactions. In particular, when using Apple Pay, the merchant only receives a single use anonymized digital token from the purchaser. It runs on most Apple products (e.g., iPhone, Apple Watch, iPad, and Mac), but not on those devices running Android, Windows, or other operating systems. (On the other hand, it can work with virtually any merchant device that accepts contactless payments.)

Apple began developing the application in the early 2010s. In preparation for the effort, Apple acquired startups and hired executives related to payments. This was not the first such phone-based digital wallet. An earlier example was Google Wallet introduced in 2011. Google's offering was primarily a peer-to-peer payment system but had an optional (physical) debit card was also a mobile payment platform. While the program was announced in 2010, it did not launch until 2013 and folded soon thereafter.

Apple formally partnered with American Express, MasterCard, and Visa in early 2013. Each of these parties is said have delegated up to 750 engineers in designing the technological solution. Apple then approached several big banks in mid-2013. The service was announced by Apple in September 2014. Apple Pay was distinguished from its predecessors on the basis of ease of use and the extent of merchant coverage. It rapidly expanded its scope from U.S.-only to global.

Apple began building its patent portfolio relating to electronic payments in the early 2010s, well before Apple Pay launched. These included many filings relating to securely conveying payment information using local networks such as Bluetooth, systems for using the phone's location data to tailor coupon and rewards offers, and limits on the size of transactions by children or other accountholders.

Sources and References Not Cited in the Paper

"Apple Pay," *Wikipedia*, https://en.wikipedia.org/wiki/Apple_Pay.

"Google Pay Send," *Wikipedia*, https://en.wikipedia.org/wiki/Google_Pay_Send.

Jeffries, Adrianne. 2014. "Apple Pay Allows You to Pay at the Counter with your iPhone 6," *The Verge*, September 9, https://www.theverge.com/2014/9/9/6084211/apple-pay-iphone-6-nfc-mobile-payment.

"Softcard," *Wikipedia*, https://en.wikipedia.org/wiki/Softcard.

Blockchain

Proposals for a blockchain-like structures date back as far as at least as 1982. But many of the key features of the modern blockchain were not proposed until the famous white paper by the pseudonymous Satoshi Nakamoto in 2008. Nakamoto conceptualized a decentralized structure (blocks did not need to be authenticated by a trusted party and transactions were archived in a public digital ledger) where new blocks were added to the chain at a set rate. The design was implemented the following year by Nakamoto as an early version of the cryptocurrency bitcoin.

Nakamoto employed an open-source license for the bitcoin code, which meant that other users could access the code and use it for the foundation for their own projects. In particular, he or she chose the MIT open-source license, which unlike more restrictive licenses, gave the user flexibility to use the code either in other open source or proprietary projects as they saw fit. As a result, bitcoin's code served as the basis for other crypto projects, such as Litecoin and Dogecoin. Other projects, such as Ethereum—proposed in 2013 a way to build decentralized blockchain applications other than those relating to currency—employed licenses that imposed more restrictions on future developers. These restrictions may have served to assure potential new contributors that their code would not be "privatized" by a single corporation, which might limit the diffusion (and value) of the new currency (Lerner and Tirole, 2005).

As blockchain applications spread, commercial companies began paying more attention to this arena. Mastercard, under the prodding of a young software engineer, Steven Davis, began researching in the early 2010s how crypto currencies might disrupt their business and ways that the firm might respond. Davis and his peers, with the encouragement of senior management, undertook a series of patent filings that covered applications outside the core areas already covered by the code of existing cryptocurrency projects (which could not be patented, as they were already publicly disclosed). Many of Mastercard's awards related to secure cryptocurrency payment processing, conversions between crypto and fiat currencies, and the integration of public and private blockchains. Mastercard today ranks among the top ten blockchain patent holders in the world.

Sources and References Not Cited in the Paper

"Blockchain," *Wikipedia*, https://en.wikipedia.org/wiki/Blockchain.

Catalini, Christian, and Joshua S. Gans. 2019. "Some Simple Economics of the Blockchain." Working Paper 22952, National Bureau of Economic Research.

Hyman, Vicki. 2020. "How Mastercard's Blockchain Whiz Has Turned Risk into Opportunity." October 29, https://www.mastercard.com/news/perspectives/2020/how-mastercard-s-blockchain-whiz-has-turned-risk-into-opportunity/.

Lerner, Josh, and Jean Tirole. 2005. "The Scope of Open Source Licensing." *Journal of Law, Economics and Organization* 21: 20-56.

Lewis, Anthony, 2018. *The Basics of Bitcoins and Blockchains: An Introduction to Cryptocurrencies and the Technology that Powers Them*. New York: Mango.

Appendix C: Financial Database Validation Analyses

This appendix describes a variety of exercises we completed to validate the quality of the data and our methodologies.

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 classification for marginal patents——those patents for which a judgment call is difficult.

These results raised the concern that the initial classification 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) were 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 borderline patents were 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 Method to Identify Financial Patents

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.

As an illustration of the difficulties of using status as a "fintech" firm to identify financial patents a subsidiary of the payments firm Square, Weebly, held several patents. But Weebly was a website builder, rather than a financial company, and thus the bulk of their awards were associated with web site design and manipulation. Thus, it would be incorrect to assume that patents held by Square and its subsidiaries were 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.

The other rejected alternative approaches also had other challenges. Another problem with identifying financial patents solely by classification code is that the U.S. changed from the U.S. Patent Classification (USPC) to Combined Patent Classification scheme in January 2013, during our period under study. 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.

Issues with Proper Assignee Names

After downloading the 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 *Manual of Panel Examination Practice*, 8th Edition, 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 Matching

The Global Corporate Patent Dataset (GCPD) (Bena et al., 2017) allowed us to match 12,351 patents to a Compustat GVKEY, which could 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 the first assignee names with 12 million firm names in the Capital IQ database.[8]

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 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 a total of 19,406 patents, or 80% of the sample (nearly 88% of the patents not awarded to individuals).

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 SIFIs 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 were some errors in our dataset when it comes to matching patents with parent companies, they errors affected only a small percentage of the data and should not have affected the analysis materially.

[8] 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.

Appendix D: Corporate Venture Capital Database Construction

We looked at another way in which incumbent firms invested in new technologies, using data on corporate venture capital transactions. In corporate venturing programs, corporations typically designate a group of professionals to make investments in young firms. The team usually purchase minority stakes in entrepreneurial firms undertaken alongside other venture capitalists, with the hope that these expenditures will lead to more informed decisions about acquisitions, internal investments, or licensing arrangements (Ma, 2020).

We totaled the number and dollar volume of closed corporate venture investments in the United States, regardless of the nation of the investor, as reported by Capital IQ. We focused on the period between January 2000 and December 2019. We restricted the analysis to investments in firms classified in a primary industry class of Financials, Online Bill Payment Services, Internet Merchant Services, or Financial Services. We did not require that the companies in the corporate venture fund portfolios have (or ultimately be granted) financial patents, as many went bankrupt or were acquired before any patents issued.

Capital IQ's classification scheme allowed us to identify corporate venture investors. In particular, we included investments that Capital IQ declared as being by groups that Capital IQ classified as "corporate investments arms" and "financial institution investment arms." We then did extensive reviews using a wide variety of sources[9] on the investment groups that had undertaken two or more investments in finance portfolio companies, to eliminate investors that we did not consider to be true corporate venture investors that were nonetheless in these categories.

In particular, we eliminated investments by:

- Traditional private equity and venture capital funds without a corporate sponsor,
- Publicly traded entities that operated largely as traditional investment funds (for example, Softbank),
- Family offices,
- Government- or non-profit affiliated bodies (e.g., International Finance Corporation, European Bank for Reconstruction and Development),
- Subsidiaries of financial institutions that primarily invested funds for third parties, rather than internally (for instance, Norwest Capital, Goldman Sachs Principal Investment Arm), and
- Corporate groups investing internal capital but with explicitly stated financial (as opposed to strategic) objectives (e.g., GE Capital).

Some smaller investment and merchant banks doing primarily financial investments (whether proprietary or for third third-party clients) doubtless slipped through these screens, potentially overstating the investment amounts. Groups that occasionally made strategic investments off their balance sheet without a formal program may have been undercounted.

[9]Sources used include lists of CVCs compiled by Global Corporate Venturing, CB Insights, and Crunchbase. We also manually checked Capital IQ database entries, web sites, media reports, and filings with the U.S. Securities and Exchange Commission

Capital IQ, like most venture capital databases, did not provide a break-down of the amount of financing provided by each investor in each round, so we divided the total financing amount in each round by the number of investors, assuming each investor provided an equal amount of capital. We eliminated the largest 2% of investments, which appeared to be co-investments in buyouts that were accidentally included in the database. The industry assignments for the investors were based on the Capital IQ industry classifications and the authors' own research.

The computation of the share of total corporate venture capital investment was based on the data compiled above, the share of U.S. venture capital investment that was corporate venture capital computed by Akcigit et al. (2020) for the period between 2000 and 2016, and the estimates of total venture capital invested in the U.S. in those years by the National Venture Capital Association (https://nvca.org/research/nvca-yearbook/, which are based on PitchBook and Refinitiv VentureXpert data). For 2017 through 2019, we use National Venture Capital Association estimates of U.S. corporate venture capital activity.

The tabulation of this alternative manner of pursuing innovation was consistent with that of patenting in several significant respects:

- The level of activity increased over time.
- There was modest share of activity associated with banks, which fell over time as a share of all such investments, while the IT/other and (to a lesser extent) payments categories grew.
- The share of total corporate venturing activity in the financial sector was roughly similar to the shares in patenting seen in Figure 1. For instance, the share of total corporate venture activity devoted to financial services between 2000 and 2016 was 1.5%.

References Not Cited in the Paper

Ma, Song. 2020. "The Life Cycle of Corporate Venture Capital." *Review of Financial Studies* 33 (1): 358–394.

Appendix E: Supplemental Analyses of Patent Quality

Claim Length and Revision

We undertook additional analyses in Section 4 to examine the reliability of financial patents as an indicator of innovation. As noted in the text, we examined the quality of review in the 21st century by assessing the subset of finance patents whose original applications were published by the USPTO. We compared the crucial independent claims in the applications and awards and determined the extent to which the number and length of these claims were modified during the review process, following the methodology of Marco, Sarnoff, and deGrazia (2019).

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). These 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 prior to issue (https://www.uspto.gov/web/offices/pac/mpep/s1122.html#d0e120159). We did not include patents initially published outside their U.S., as these may have been modified by another patent office before USPTO review.

We determined the count and the 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 used the applications analyzed by Marco, Sarnoff, and deGrazia (2019) and archived at https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-claims-research-dataset.

Panels A and B of Figure A-10 present a comparison of 2.6 million non-finance patents and almost 16 thousand finance ones. Finance patents were 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). 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. Both of these results were consistent with more intensive reviews of finance patents since the mid-2000s. Table A-19 presents a more detailed tabulation and statistical comparison and finds consistent results.

Assignee Type

We looked as well at who was filing the finance patents. We examined the identity of the assignees of all utility patents applied for between 2000 and 2018 and awarded by February 2019. We used the classification of assignees provided by the USPTO and assumed that all unassigned patents were awarded to individuals.

Table 3 shows that 8.6% of finance patents since 2000 were assigned to individuals, similar to non-finance patents (7.8%). This share differs sharply from the 25% share in the pre-*State Street*

sample of finance patents collected by Lerner (2002), as reported in Table A-20.[10] Since many of the most problematic patents in the earlier era were those of individual inventors, this result was again consistent with the suggestion that patent awards filed in recent decades provide a valuable window into changing trends in financial innovation more broadly.

Earnings Calls

We also undertook another set of analyses to understand the relative importance of patents and trade secrecy for financial firms, and how that changed over time. This section provides additional details about these approaches.

The first followed the methodology in Hassan et al. (2019) and Bloom et al. (2021). To undertake the analysis, we looked at earning calls (ECs) of all publicly traded financial firms. We compiled all GVKEYs of what we considered to be finance firms that were publicly traded at any time between 2000 and 2018 and whose earning calls were included in the Refinitiv (formerly Thomson Reuters) database of earnings call transcripts.

To do so, we identified firms assigned in CapitalIQ to the GICS codes associated with Banks, Other Finance, and Payments firms, as defined in Section 3.2 of the paper. We did not examine the ECs of non-finance firms that may have pursued financial innovation, such as IT firms. This decision was made because (a) most references in conference calls to intellectual property were general in nature, and (b) we anticipated that in most cases, the bulk of the intellectual property owned by the non-finance firms would not be finance related (even if they were substantial financial innovators).

For these finance firms, we counted the number of earnings call transcripts in each quarter along several dimensions:

1. The cumulative count of earnings calls (ECs) involving these firms, and their average length in words.
2. The count of ECs mentioning the following keywords, as well as the number of mentions:
    a. "patent*"
    b. "trade secre*" or "proprietary knowledge*" or "commercial confidential*" or "business confidential*" or "confidential business information" or "industry confidential*"

We also compiled the count of ECs where there were references to secrets but not trade secrecy (("secret*" or "secrec*") and NOT ("trade secre*")"). When we audited these cases, however, we found that almost none of them dealt with trade secrets. Rather, they were less relevant comments, such as "[we did] little to no marketing, so were a bit of a well-kept secret," "not assuming anything major from Victoria's Secret contracts going forward," and "it was not a secret sauce; it's blocking and tackling."

[10]Another way to assess the importance of individual patentees is to look at the difference in the share of awards that were made to individuals between finance and all other patents. While this gap was less than 1% in patents filed in 2000 and later, it was 10% in the earlier period.

There were also a number of generic references to "intellectual property" in earnings calls that did not reference either patents or trade secrets explicitly (or the synonyms for trade secrets delineated above). In the majority of cases that we audited, these references were by firms that had been issued patents; in many cases, the firms appeared to be referring to these patents. But given the ambiguities, we did not count these cases as either ones that referenced patents or trade secrets.

We finally normalized the count of references in the finance calls to the patent- and trade secret-related keywords. To do so, we divided the count by the number of ECs by finance firms in that quarter and their average length in words, then multiplied by 1000.

In the nearly 26 thousand transcripts, 446 mentioned patents at least once, while the phrases associated with trade secrets appeared in only 23. Nor did mentions of trade secrecy become more frequent over time. The ratio of patent to trade secret-related mentions went from 17.5 in the pre-*Bilski* period (2002-09) to 21.4 thereafter (2010-19). The quarterly time series, normalized by the number of calls analyzed and their length, is depicted in Figure 3.

Intellectual Property Litigation

We also focused on federal litigation involving patents and trade secrets. The decision to focus on federal (and not state) court litigation reflected data availability. While services such as Lex Machina and Bloomberg Law have compiled federal filings for many years, the coverage of state court filings is at a much earlier stage. (For instance, Lex Machina did not begin coverage of state cases until the introduction of Delaware state cases in 2018 and Houston and Los Angeles area state cases in 2020.[11])

This limited coverage posed concerns. Traditionally, patent cases have been heard in federal cases. (Some contractual disputes involving patents were, and still are, heard in state courts, but all questions revolving around patent validity must be resolved in Federal courts.) Trade secret cases, on the other hand, are heard in both state and federal courts. Prior to 2016, most misappropriation and trade secrets lawsuits could be filed in federal court only through a diversity provision (i.e., where a plaintiff and defendant were citizens of different states and the amount in dispute exceeded seventy-five thousand dollars) or if the plaintiff asserted a federal claim in addition to the state law trade secret claim. This limitation was relaxed in 2016. Signed into law on May 11, 2016, the Defend Trade Secrets Act (DTSA) allowed firms to litigate trade secret cases more generally in the federal courts, by extending the Economic Espionage Act of 1996 to criminalize trade secret misappropriations.

Practitioner accounts suggest that firms turned rapidly to the federal courts after the passage of the DTSA to adjudicate additional trade secret cases. The advantages of litigating trade secrets in the federal courts were summarized in one legal blog as follows:

> Federal courts are accustomed to handling sophisticated civil litigation. They are experienced in dealing with complex discovery issues, including protective orders,

[11] Lex Machina, "Lex Machina Launches State Law Modules, Extending Its Groundbreaking Legal Analytics to State Courts in California and Texas," February 4, 2020, https://lexmachina.com/media/press/lex-machina-launches-state-law-modules-in-california-and-texas/.

> and issues regarding expert witness testimony. Alongside this, federal courts readily grant meritorious motions for summary judgment. Further, as Congress noted when it enacted the DTSA, trade secret theft today is often not confined to a single state and trade secret cases often require swift action by courts across state lines to preserve evidence. Federal courts can be better equipped to provide such relief.[12]

We identified federal trade secret litigation in two ways. First, we used the database of DTSA-related cases compiled by Professor Chris Seaman from Lex Machina and Bloomberg Law, who also downloaded the original complaints in these lawsuits. The database construction was described in Levine and Seaman (2018). We supplemented this list with a search of all non-DTSA related trade secret cases in the federal courts, which we identified using Lex Machina. For each supplemental case, we also obtained the original complaint. We reviewed all the complaints, whether DTSA-related or not, for evidence whether (a) the case involved a true innovation, and not a dispute over client lists/contacts or sales materials (which may also be covered by trade secret protection),[13]and (b) one of the parties was a financial institution (defined as above), or, if not, whether the dispute was over some financial innovation.[14] We downloaded from Lex Machina an indication of whether a patent claim was also asserted at some point in the litigation.

We wished to compare the volume of trade secret cases to patent ones. To do so, we used the Patent Litigation Dataset compiled by the USPTO Office of the Chief Economist and the University of San Diego Law School, which contains links between 81,350 unique district court cases filed during the period from 1963 to 2016 and the associated patent numbers (Schwartz, Sichelman, and Miller, 2019). We downloaded all litigation associated with the patents in our sample. We also downloaded from Lex Machina an indication as to whether a trade secret claim was also asserted at some point in the litigation.

Because we wished to focus on the period when the DTSA was active and patent litigation data available, we focused on lawsuits filed in the period from May 12 and December 31, 2016. We looked separately at the litigation involving finance patents and all other patent litigation, and trade secret cases about a financial innovation or another innovation. We found that the ratio of pure patent cases to trade secret ones for financial innovations was between 10.4 and 19.9 to 1. A similar pattern holds in non-finance cases: in fact, the ratios were almost twice as high.

---

[12] Holland & Knight, "The Impact of the New Federal Trade Secrets Act on Trade Secret Litigation: Holland and Knight Trade Secrets Blog," July 30, 2018, https://www.hklaw.com/en/insights/publications/2018/07/the-impact-of-the-new-federal-trade-secrets-act-on.

[13] More specifically, we identified cases that were unambiguously non-innovative in nature (where the theft/misappropriation was exclusively of customer contacts and marketing materials, which we refer to as definition 1) and ones that were likely non-innovative in nature (where the theft/misappropriation may have also included "software" or "samples," but no distinct claims are made that these materials contained information on novel products or processes, which we refer to as definition 2). In a small number of cases, we could not obtain information on the topic in dispute.

[14] In about 10% of the cases, the original complaint was not available in Lex Machina or did not provide the information to assess item (a) in the list above. These were typically cases that were transferred to or from another district. In most cases, we are able to find the information in other case filings or in the docket of the companion case. In the case of two financial disputes, we are unable to assess whether they were innovative or not.

	Finance	*Other*
DTSA cases	51	296
+Other Federal TS cases	57	399
=Total Federal TS cases	108	695
-Non-innovative TS cases (definition 1)	96	548
=Innovative TS cases (definition 1)	12	147
-Hybrid cases (TS + patent)	0	13
=Pure innovative TS cases (definition 1)	12	134
Total TS cases	108	695
-Non-innovative TS cases (definition 2)	101	619
=Innovative TS cases (definition 2)	7	76
-Hybrid cases (TS + patent)	0	11
=Pure innovative TS cases (definition 2)	7	64
Patent cases	125	2692
-Hybrid cases (TS + patent)	0	30
=Pure patent cases	125	2662

Information Technology Spending

The He et al. (2022) analysis employs the Harte Hanks Market Intelligence Computer Intelligence Technology database, which provides detailed information on specific spending categories. The paper computes for U.S. commercial banks their annual expenditures for two categories in the database: Software and Communications.

The authors compared for us on the bank-year level between 2010 and 2017 the ratio of patent applications (provided by us) to revenue (the later taken by the authors from the Federal Reserve's Call Report database and matched to the Harte Hanks data) and that of IT expenditures in these two categories to revenue. They did so by regressing the patent ratio on the IT spending ratio using four specifications: with no fixed effects, with year fixed effects, with bank fixed effects, and with year and bank fixed effects.

Due to the confidentiality constraints around the Harte Hanks database, we were restricted in the results that we could report. The key coefficient and standard error on the IT spending variable were 0.090 (0.027) (with no fixed effects), 0.110 (0.032) (with year fixed effects), 0.038 (0.011) (with bank fixed effects), and 0.034 (0.018) (with year and bank fixed effects). The R-squared of the regressions ranged from 0.049 (with no fixed effects) and 0.924 (with year and bank fixed effects).

References Not Cited in the Paper

Bloom, Nicholas, Tarek A. Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun. 2021. "The Diffusion of Disruptive Technologies." National Bureau of Economic Research Working Paper Series No. 28999, https://www.nber.org/papers/w28999.

Levine, David S., and Christopher B. Seaman. 2018. "The DTSA at One: An Empirical Study of the First Year of Litigation Under the Defend Trade Secrets Act." *Wake Forest Law Review* 53, 106*ff.*

Schwartz, David L., Ted M. Sichelman, and Richard Miller. 2019. "USPTO Patent Number and Case Code File Dataset Documentation." USPTO Economic Working Paper No. 2019-05, https://ssrn.com/abstract=3507607.

Appendix F: Construction of Data for Economic Activity and Patenting Comparison

U.S. gross output

We took annual gross revenue from https://apps.bea.gov/iTable/index_industry_gdpIndy.cfm.[15] To facilitate the comparison to the patent data by industry, we made two simplifying consolidations of the BEA industries. In particular, we aggregated (a) the three insurance-related BEA industries (all within NAICS codes 5341 and 5242), and (b) the BEA industry "Non-depository credit intermediation and related activities" (NAICS codes 5222 and 5223), which largely consists of payments companies, consumer finance firms, and non-bank banks, with three categories that consist largely of lessors: the consumer-facing auto finance firms (NAICS code 5321) and two commercial ones ("Commercial and industrial machinery and equipment rental and leasing" (5324) and "Lessors of nonfinancial intangible assets" (533)). We also renamed some of the adjusted BEA industries to make their nature clearer.

These changes are summarized in the table that follows.

Adjusted BEA Industry	*NAICS Codes*	*BEA Industry(ies)*
Non-bank credit and payments	5222-23, 5321, 5324, 533	Non-depository credit intermediation and related activities; Automotive equipment rental and leasing; Commercial and industrial machinery and equipment rental and leasing; Lessors of nonfinancial intangible assets
Banks	521, 5221	Monetary authorities and depository credit intermediation
Investments and funds	5329	Other financial investment activities
Securities intermediation	5231-32	Securities and commodity contracts intermediation and brokerage
Insurance	5241-42	Direct life insurance carriers; Insurance carriers, except direct life insurance; Insurance agencies, brokerages, and related activities
Passive funds and trusts	525	Funds, trusts, and other financial vehicles
Accounting	5412	Accounting, tax preparation, bookkeeping, and payroll services

U.S. value added

The data for value added use the same 405-industry scheme as above and are available only on a quinquennial basis. Thus, unlike the other series reported here, the value-added series presents activity in one particular year, not over the entire period.

[15] To find the relevant data, we select "Access Underlying Detail Tables" in the "Additional information" section. These tables are at the very bottom of the Gross Output section and have "Detail Level" appended to the end of the table title.

The benchmark tables for 2007 and 2012 are updated by the BEA to ensure that they are conceptually consistent with each other. 2007 and 2012 data were found at https://www.bea.gov/industry/input-output-accounts-data under "Use Tables," in a sheet labelled as Use_SUT_Framework_2007_2012_DET.xls.

Historical benchmark tables, including 2002, use a slightly different variant of the industry scheme that has not been updated. To compute value added for 2002, we added three main subcomponents: compensation, taxes less subsidies, and gross operating surplus (the three "commodities" with codes V00100, V00200, and V00300). 2002 data are in a file labelled REV_NAICSUseDetail 4-24-08.txt file (which is included in the download under "2002 Standard Make and Use Tables at the detailed level" folder) at https://www.bea.gov/industry/historical-benchmark-input-output-tables.

2017 data were taken from "Use Tables" for 2017 at https://www.bea.gov/industry/input-output-accounts-data. 2017 data uses the 71-industry scheme (2017 405-industry data were not scheduled to be released until late 2023.) In cases where some of the 405 industries were aggregated in the 2002 and 2017 data, we assigned value added to the individual BEA industries proportionate to the relative activity in the closest year with 405-industry level data (2007 or 2012).

Patenting by using industry

We assign the patents in the sample to industries based on the classification of patent types described in the paper. The following table shows the mapping we use between the BEA industries and the patent types, repeating the relevant NAICS codes for reference:

Adjusted BEA Industry	*NAICS Codes*	*Patent Type*
Non-bank credit and payments	5222-23, 5321, 5324, 533	Real estate; payments
Banks	521, 5221	Commercial banking; retail banking
Investments and funds	5329	Wealth management; currency; cryptocurrency; active funds
Securities intermediation	5231-32	Investment banking/exchanges
Insurance	5241-42	Insurance
Passive funds and trusts	525	Passive funds
Accounting	5412	Accounting

This mapping is inexact by necessity. In particular:

- Two patent type categories are cross-cutting, and do not lend themselves to assignment to a single category: communications and security. In these cases, we assigned the patents to other industries, using the same proportions as the industries that were jointly assigned to (a) communications and/or security on the one hand and (b) another industry or industries on the other. Because the composition of the industries changed over time, we did this calculation separately for patents applied for in the 2000-04, 2005-09, 2010-2014, and 2015-18 periods.

- The relatively few finance patents classified as real estate largely focused on securitization, so seemed best classified with non-bank credit.
- A number of patents classified under commercial and retail banking applied to credit analysis or repayment schemes in general, and thus could also be included under non-bank credit. This may have led to an undercount of non-bank credit patents.
- The very few currency-related patents related to portfolio management, corporate hedging, and liability management applications, and thus could be classified in multiple categories.

Appendix G: CSA Database Construction and Supplemental Regional Analysis

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." 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) for venture capital financings (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, the data were 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 the count of households in the county 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.

Regional Analysis

Figure A-11 provides another view of the overall patterns, focusing on activity across U.S. Census regions over time. We constructed the analysis sample at the application year – U.S. Census region level, for a total of 171 observations (19 years x 9 census regions). We estimated the following specification to examine the pattern of financial patenting in U.S. Census regions over time:

$$Patent\ Count_{rt} = \beta_0 + \beta_1 (Region_r \times Time\ Period_t) + \mu_r + \gamma_t + \epsilon_{rt} \quad (30)$$

The dependent variable was the number of finance patents applied for in census region r in year t. As before, we divided the application years into four periods. The key independent variables were application period indicators $Time\ Period_t$ interacted with the U.S. Census region dummies $Region_r$, using the Middle Atlantic region and the 2000-04 period as the baseline. We also included census region fixed effects μ_r and year fixed effects γ_t as controls in our regression.

The figure presents the coefficients of the above regression for two specific regions: the Pacific and South Atlantic (which includes Charlotte) regions. Financial patenting in these two regions increased sharply over time relative to the Middle Atlantic region, suggesting that the locations of financial patenting gradually shifted from the east coast to the west and south. These results were consistent with the rise of patenting in the San Jose-San Francisco and the Charlotte-Concord CSAs and the decline in the importance of New York reported in Table 8. Table A-21 further presents the detailed share of patenting by region for the nine U.S. Census regions between 2000 and 2018.

Supplemental Analyses of Switchers

Table A-16 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 2000-04; and
- Continuing innovators, who filed an (ultimately successful) financial patent in 2000-04 and in 2015-18.

For the third category, we also broke out firms that shifted their modal CSA for patenting between these two periods. Location-switching continuers are relatively few in number (28 firms), but very significant when patents are tabulated: these firms represent 32% of the awards by continuing innovators, and 22% of the awards across all three categories. (Note we did not include firms that did not patent in 2000-04 and 2015-18, but just in intermediate years.)

Panel B looks at the 28 location-switching continuers in more depth. Nine of the firms (representing 2778 patents in total) moved their modal location from New York-Newark; no other CSA was close to this volume of losses. Meanwhile, the destination of these firms was much more diversely spread. These results suggested the importance of location-switching continuers in the location analyses.

Table A-22 looks at which continuing financial innovators were switchers in a probit analysis. We use all 129 continuing innovators as observations. We estimated:

$$\mathbf{Pr}\ (\textit{Firm is Switcher}_i = 1) = \mathbf{\Phi}(\beta_0 + \beta_1\ (\textit{Modal 2000-04 Location}_i) + \beta_2\ (\textit{2000 Finance VC in Modal 2000-04 Location}_i) + \mu_1 + C_i'\mathbf{B} + \epsilon_i) \quad (31)$$

Pr (·) denoted probability and **Φ** was the cumulative distribution function of the standard normal distribution. *Firm is Switcher*$_i$ was an indicator for whether a firm shifted its modal location for innovation from 2000-04 to 2015-18. *Modal 2000-04 Location*$_i$ were dummy variables indicting whether the firm's modal patent applied for between 2000 and 2004 was in the New York or the San Jose/San Francisco CSAs. *2000 Finance VC in Modal 2000-04 Location*$_i$ was the dollar volume of venture financing of finance firms in 2000 in the modal location for the firm's patenting in 2000-04. We also included firm industry dummies and a vector of firm controls C_i, such as whether the firm was publicly traded or venture backed. The results suggested that banks and payments firms were consistently more likely to switch than IT and other firms. Firms with the modal early patenting location in the greater New York area, as well as those that were privately held, were more likely to switch.

Appendix H: Returns Analysis

Before launching the analysis, we undertook a variety of preparatory steps. We started with the population of firms that were awarded at least one successful financial patent in the sample. We restricted the analysis to those firms that were publicly traded in the U.S. and had at least one year of non-missing or non-zero R&D expenditures reported in Compustat. This gave us a total of 278 firms, each of whom was awarded at least one non-financial patent as well during this period. While modest in number, these firms were substantial innovators: collectively, these firms were awarded 1,203,145 patents during the period (32% of the total awards in the sample period), among which 5,369 are financial patents (22% of all finance patents in the sample).

Table A-17 looks at the 278 publicly traded firms included in the returns analysis in Panel A. It shows that the industry mixture leans towards software and telecommunications. Similarly, the firms are very research intensive, as Panel B demonstrates. The cumulative ratio of R&D to sales in the sample was 4.8%, which can be compared to the ratio for U.S. firms of 4.4% in 2019.[16] (At the beginning of the sample period, in 2000, the corresponding ratio was 3.8%.[17])

Table A-18 compares the patents included in the returns analysis sample with the other patents used in the analysis (i.e., all the other patents examined in analyses such as Table 1). Many of the patterns are inconsistent and economically modest (even if statistically significant, reflecting the large sample sizes). For instance, we see in Panel A that the return analysis sample has slightly more citations; on a class and year adjusted basis, the citation score is modestly lower for these patents. One of the consistent patterns is that the return analysis patents are more likely to stem from the New York and Bay areas. Panels B and C show that patents in the life sciences are underrepresented in the return analysis sample (presumably due to the absence of any financial patents at these firms), but patents in information technology are overrepresented.

The 2019 version of the Kogan et al. extended data was then used for the matching between patents awarded to public firms with the corresponding Kogan values. For each firm in every year from the application year of their first financial patent through 2018, we computed the total number of ultimately successful patent awards filed in that year, the total adjusted citations of those patents (adjusted by mean citation level for patents filed during the same application year by all firms), and the ratio of the mean Kogan value of these patents and R&D expenditure. We made these calculations separately for financial and non-financial patents. R&D expenditures were typically available only on a firm-year level, so we apportioned them based on the proportion of successful financial and non-financial patent applications for each firm in every year.

We then calculated the R&D stock, patent award stock, and citation-weighted patent stock as follows:

[16] National Center for Science and Engineering Statistics and Census Bureau, *Business Enterprise Research and Development Survey*, 2019, https://ncses.nsf.gov/pubs/nsf22303.

[17] National Science Foundation, Division of Science Resource Statistics, Survey of Industrial Research and Development. 2000, https://wayback.archive-it.org/5902/20150627201523/http://www.nsf.gov/statistics/infbrief/nsf03306/.

$$K_t = (1-\delta)K_{t-1} + C_t \quad (32)$$

with δ being the depreciation rate, set at 15% (following Hall, Jaffe, and Trajtenberg, 2005), C_t being the innovation measure at time t (i.e. R&D, patent applications, and citation-weighted filings), and K_t being the corresponding stock measure at time t.[18] The market value and book value of equity were calculated using raw data from Compustat.[19] In the end, our panel data used in the following analyses consisted of 2,808 observations at firm-year level from 246 firms.

Table A-23 gives the summary statistics of the key variables used in the analysis. From the table, it is clear that measures of innovation were highly positively skewed, as the means were much larger than the medians with large standard deviations. This pattern was consistent with the observations in Hall, Jaffe, and Trajtenberg (2005). The mean market-to-book value was 5.29, which is higher than the 1.73 in Hall, Jaffe, and Trajtenberg (2005). One possible reason was that Hall, Jaffe, and Trajtenberg (2005) uses firms in the manufacturing industry only, while our data sample covers all firms with at least one financial patent. The difference may also reflect the overall appreciation in the valuation of technology firms in recent years. In addition, the distribution of mean Kogan/R&D ratio was more skewed for financial patents, with a higher standard deviation, mean, and median. The mean Kogan/R&D ratio for financial patents is 2.27 and the median 0.60. Non-financial patents had a mean Kogan/R&D ratio of 0.10 on average and 0.02 at the median.

We then emulated Hall, Jaffe, and Trajtenberg (2005) and explored how Tobin's q is affected by the stocks of R&D, patents, and citation-weighted awards for financial and non-financial firms. Following Table 3 of that paper, Table A-24 gives the non-linear regression results, but now examining financial and non-financial patents separately. This shows the market value of firms as a function of assets and the stock of R&D, patents, and citations:

$$\begin{aligned} logQ_{it} = {} & logq_t \\ & + log\left(1 + \gamma_1^{Fin}\frac{R\&D_{it}^{Fin}}{A_{it}} + \gamma_2^{Fin}\frac{PAT_{it}^{Fin}}{R\&D_{it}^{Fin}} + \gamma_3^{Fin}\frac{CITES_{it}^{Fin}}{PAT_{it}^{Fin}} + \gamma_1^{Nonfin}\frac{R\&D_{it}^{Nonfin}}{A_{it}}\right. \\ & \left. + \gamma_2^{Nonfin}\frac{PAT_{it}^{Nonfin}}{R\&D_{it}^{Nonfin}} + \gamma_3^{Nonfin}\frac{CITES_{it}^{Nonfin}}{PAT_{it}^{Nonfin}}\right) + \varepsilon_{it} \end{aligned} \quad (33)$$

Emulating column (2) in Table 3 of the Hall paper, we included dummy variables indicating whether financial or non-financial R&D expenditures in that year were zero and application year fixed effects. We report the results using firms with at least one, five, and ten financial patents applied between 2000 and 2018.

[18] We do not use observations prior to 2000 in calculating the stock measures.

[19] The market value was obtained using price times common shares outstanding at the end of fiscal year. The book value of equity was calculated as the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credits, minus the book value of preferred stock. If some of these variables were missing, book equity was calculated as the book value of assets minus total liabilities. This method of calculating the market (Q) and book values (q) of equity follows Fama and French (1992) and was slightly different from that used in Hall, Jaffe, and Trajtenberg (2005) because not all variables listed could be obtained from Compustat. We dropped observations with negative book value. This is because a negative book value and hence a negative market-to-book ratio gives a missing logarithmic result, which is omitted in the regression analysis described in estimating equation (33).

We then computed the elasticity of firm value to citation intensity. We calculated this in two ways. We first looked at the direct effect of citation intensity on firm value:

$$\frac{\partial logQ}{\partial(CITES^{Fin}/PAT^{Fin})} = \frac{\widehat{\gamma_3{}^{Fin}}}{1+\hat{X}}, \quad \frac{\partial logQ}{\partial(CITES^{Nonfin}/PAT^{Nonfin})} = \frac{\widehat{\gamma_3{}^{Nonfin}}}{1+\hat{X}} \quad (34)$$

with

$$\hat{X} = \widehat{\gamma_1{}^{Fin}}\frac{R\&D^{Fin}}{A} + \widehat{\gamma_2{}^{Fin}}\frac{PAT^{Fin}}{R\&D^{Fin}} + \widehat{\gamma_3{}^{Fin}}\frac{CITES^{Fin}}{PAT^{Fin}} + \widehat{\gamma_1{}^{Nonfin}}\frac{R\&D^{Nonfin}}{A} + \widehat{\gamma_2{}^{Nonfin}}\frac{PAT^{Nonfin}}{R\&D^{Nonfin}} + \widehat{\gamma_3{}^{Nonfin}}\frac{CITES^{Nonfin}}{PAT^{Nonfin}} \quad (35)$$

Second, as described in the text, we looked at the change in market value with respect to change in R&D through the impact of R&D on mean citation intensity.

We examined the robustness of the results in Table 12 in several ways. The first of these was to address concerns about the measures of patent and citation stock by only evaluating yearly observations through 2013, in order to ensure that all patents have had sufficient time to garner citations. We also evaluated the semi-elasticities at the median, rather than the mean. We found the changes made little difference to the results. Third, we reran the regressions, now using a common R&D measure in the regressions (rather than the imputed financial and non-financial amounts) but allowing the patent and citation measures to differ for financial and non-financial analyses. We again got similar results. Tables A-25 and A-26 summarize two such robustness analyses.

We also examined the robustness of the analysis depicted in Figure 7, as depicted in Figure A-12. As one sensitivity check, we changed the calculation of financial and non-financial fractions when computing the benefit of social and private returns. When the benefit is scaled by the ratio of the number of finance patents granted to the total patents granted in Panel A (rather than weighted patents), the ρ measure was numerically slightly higher than our baseline measures, but the trend was similar. In Panel B, the firm set included all firms with R&D information available. The magnitudes of social return for both financial and non-financial social returns were still similar to our baseline measure, which only considered financial innovators. The non-financial social return was still higher than financial social return before the GFC, and the trends afterward were similar to the baseline measure. Panel C shows the private return of financial vs. non-financial innovation when the firm set included all firms with R&D information available (not only financial innovators). Financial innovation still had a much higher private return than non-financial innovation, even when including firms that never innovated in finance.

As another sensitivity check, we also calculated the financial social return using a different source for the cost measure: the gross fixed investment of R&D input of the financial sector from the St. Louis Federal Reserve Bank. The macro measure could overestimate the cost of financial innovation, as the Fed's financial sector has a more comprehensive scope, including spending by government, research institutes, and non-profit organizations. The macro measure could also underestimate the benefit of financial innovation, as Fed's financial sector R&D measure does not

include innovations by IT firms, whom (as the paper demonstrates) are important contributors to financial innovation.

These alternative analyses are depicted in Figure A-13. Panel A shows what happens when we used the Fed R&D input measure in equation (27), but otherwise left the expression unchanged. Panel B undertakes a similar calculation, but we now excluded financial patents awarded to IT firms from these calculations (reflecting the fact that the Fed's financial sector R&D measure did not include IT firms). Particularly in the later analysis, the financial social return was much lower. This suggests that the calculations using Fed's macro data may have underestimated the benefit of financial innovation.

References Not Cited in the Paper

Fama, Eugene F., and Kenneth R. French, 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance* 47 (2): 427-465.

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.

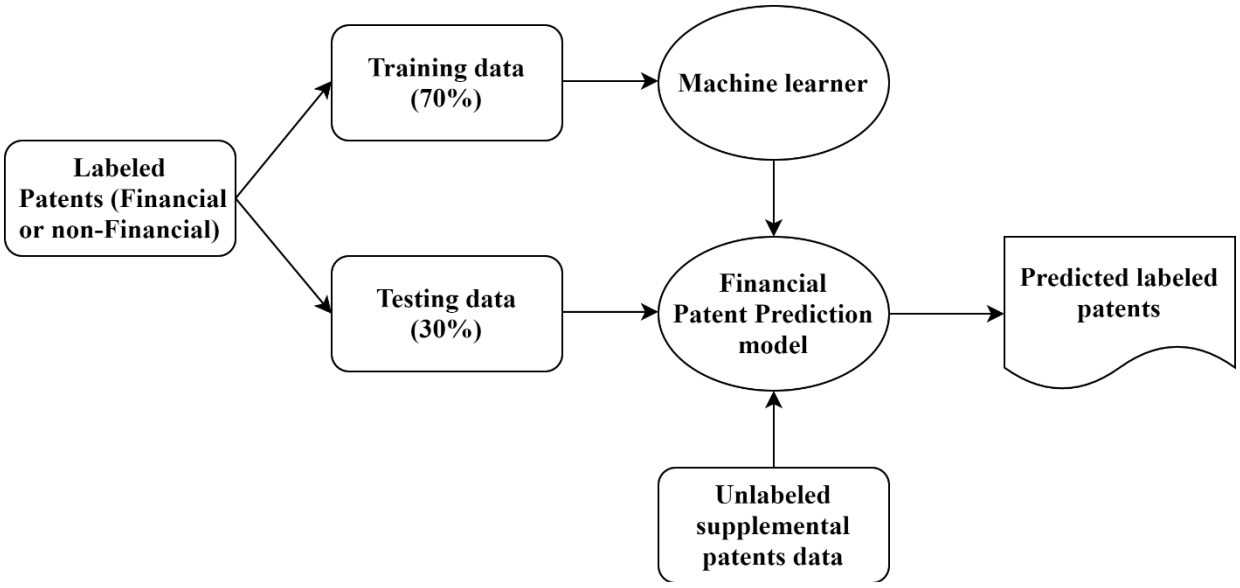


Figure A-2. Financial patents machine learning model architecture. The figure presents the structure of our final machine-learning model. Compared to the text-only model, the text-inventor model slightly decreases sensitivity from 91.3 to 89.9 percent (a drop of 1.4 percentage points), but significantly 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.

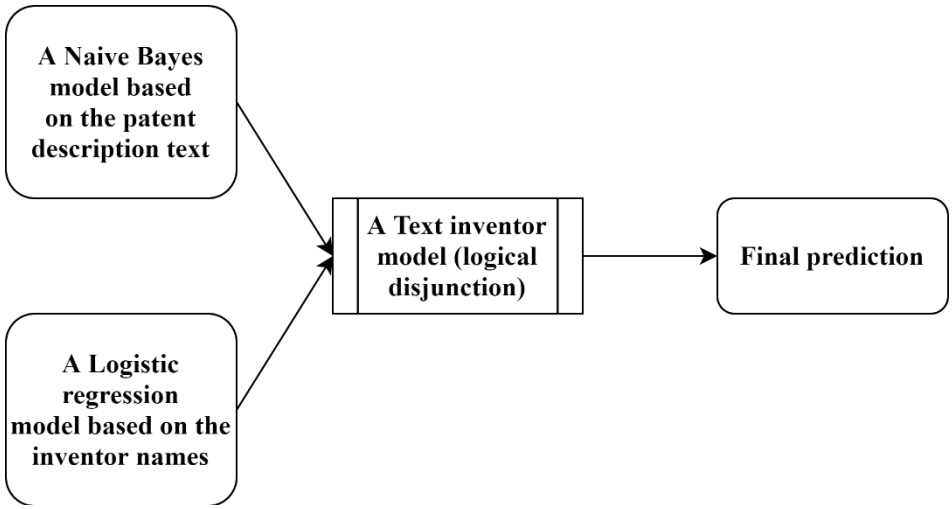


Figure A-3. Fuzzy name matching between assignee names and Capital IQ names. This figure presents how we use a Levenshtein distance-based fuzzy name matching techniques to match the unmatched assignee names with 12 million firm names in the Capital IQ database. The Capital IQ database was divided into three 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 were merged.

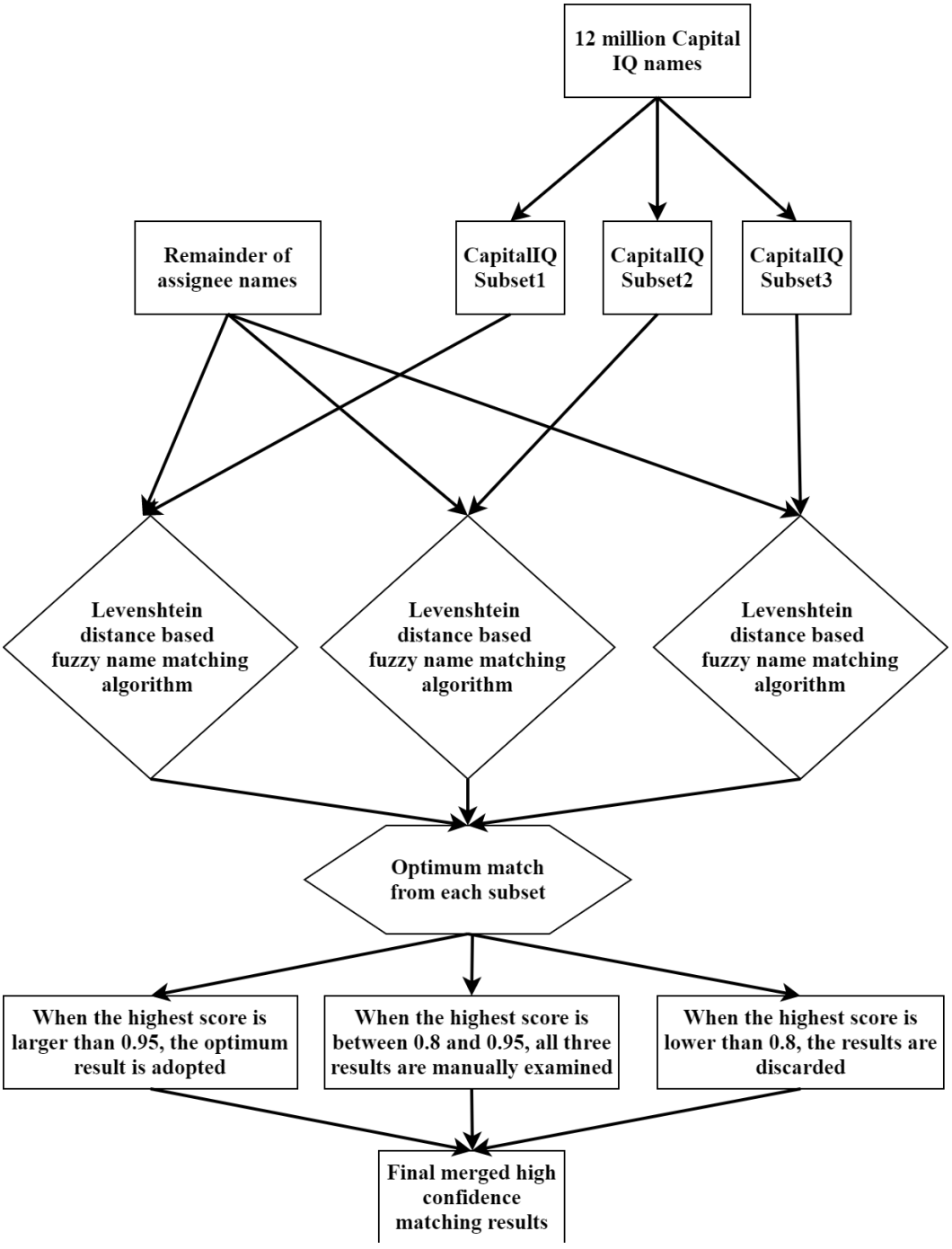


Figure A-4. An overview of the financial dataset construction procedure. The first step in our process was to obtain additional patent-level data on financial patents from Derwent. We obtained from Patentsview the patent assignee type and a host of other information. Then the assignee's Capital IQ ID 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.

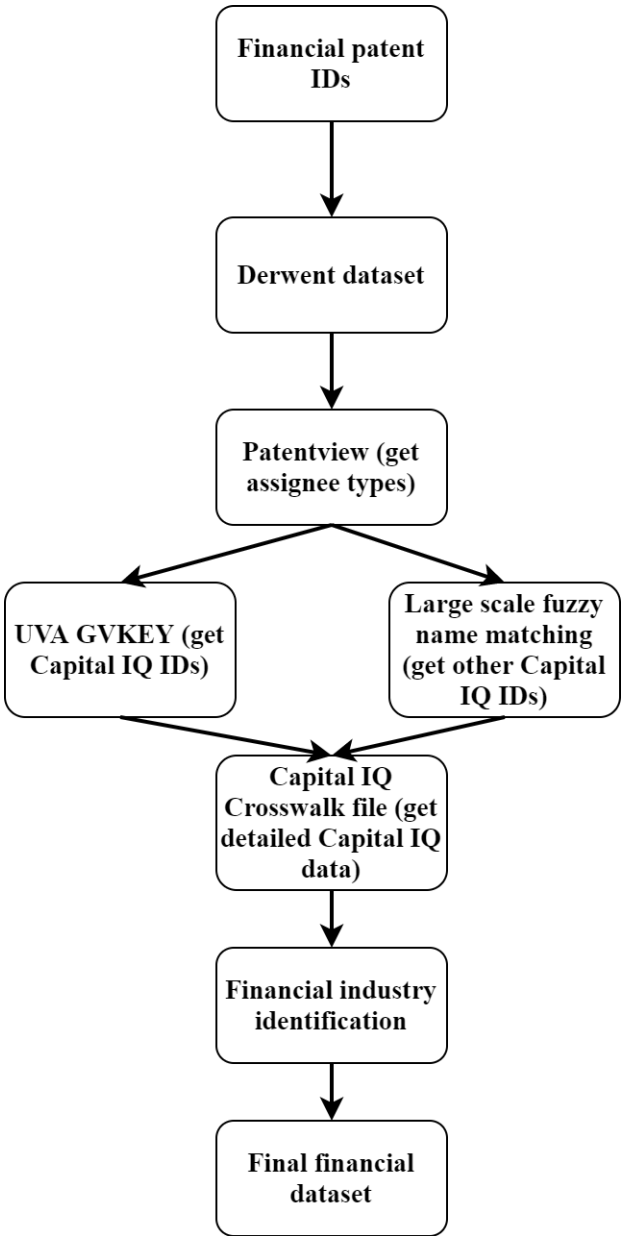
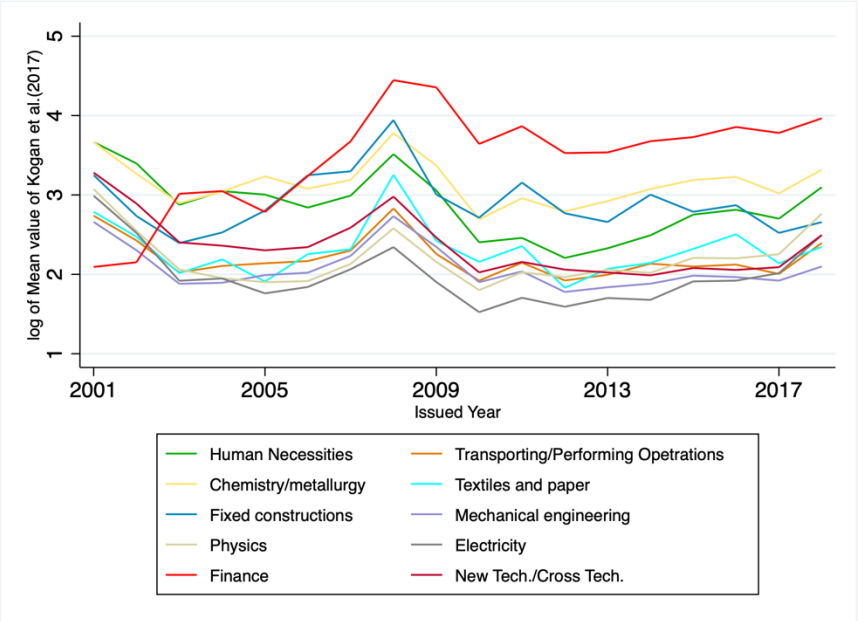
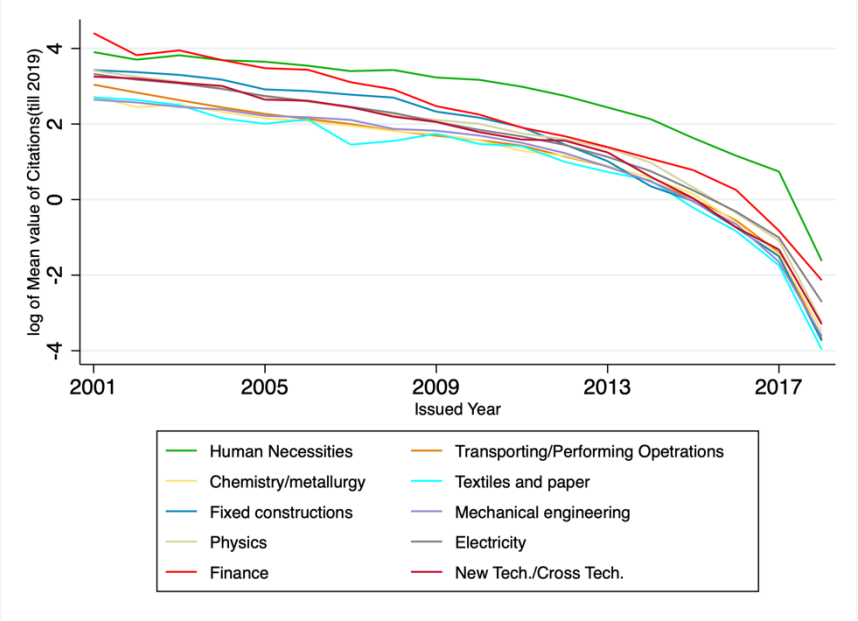


Figure A-5. Trends in Kogan et al. (2017) value and patent citations by cooperative patent classification (CPC) category and award year. We use all patents applied for between 2000 and 2018 and awarded by February 2019. There are nine main categories under the CPC scheme. We separate all of our finance patents and classify them into a new category. Panel A depicts the log of the mean Kogan et al. (2017) value by CPC category over time, and Panel B depicts the log of the mean patent citations (through October 2019) by CPC category over time. Panel C depicts the log of the top 5th percentile of Kogan et al. (2017) value by CPC category over time, and Panel D depicts the log of the top 5th percentile of patent citations (through October 2019) by CPC category over time.

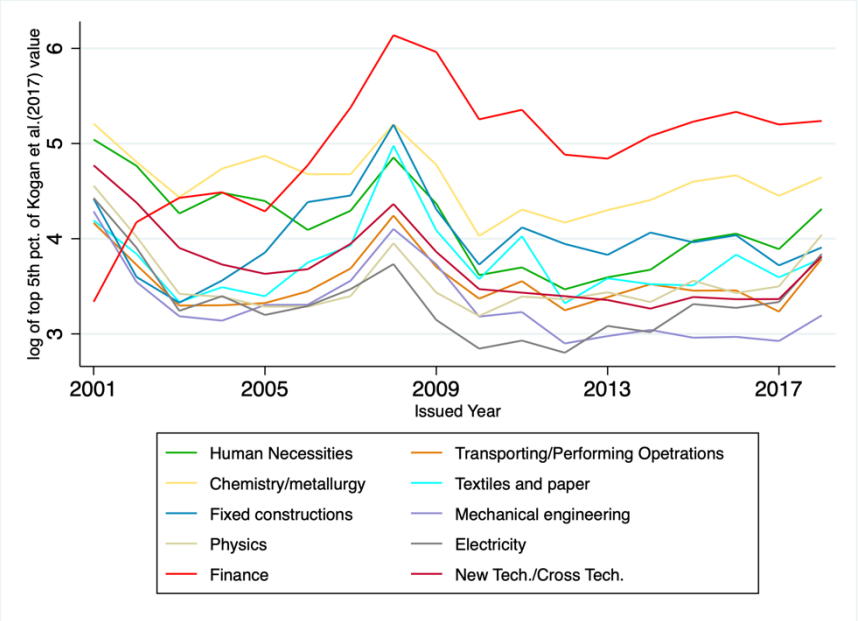
Panel A: Mean of Kogan et al. (2017) value over time, by patent's CPC category.



Panel B: Mean of patent citations over time, by patent's CPC category.



Panel C: Top 5th percentile of Kogan et al. (2017) value over time, by patent's CPC category.



Panel D: Top 5th percentile of patent citations over time, by patent's CPC category.

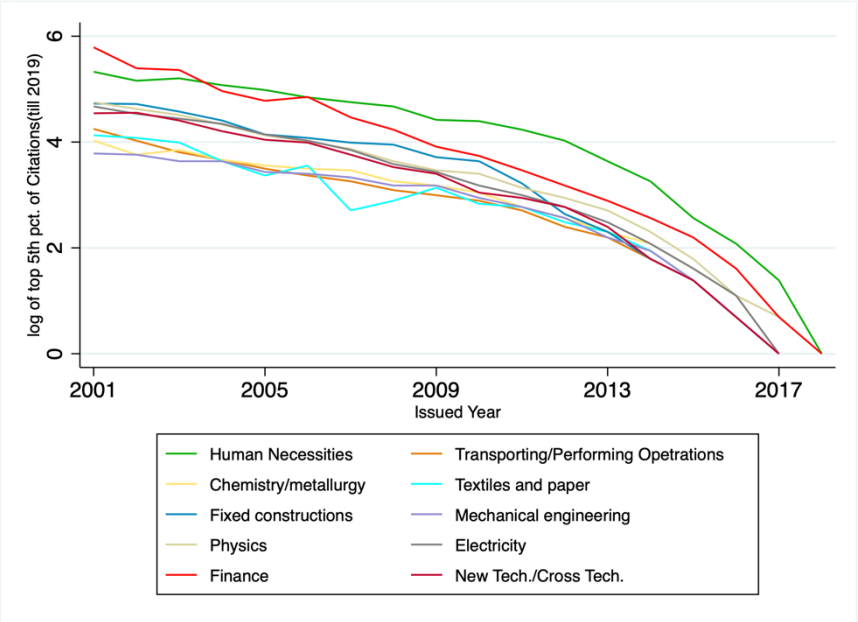
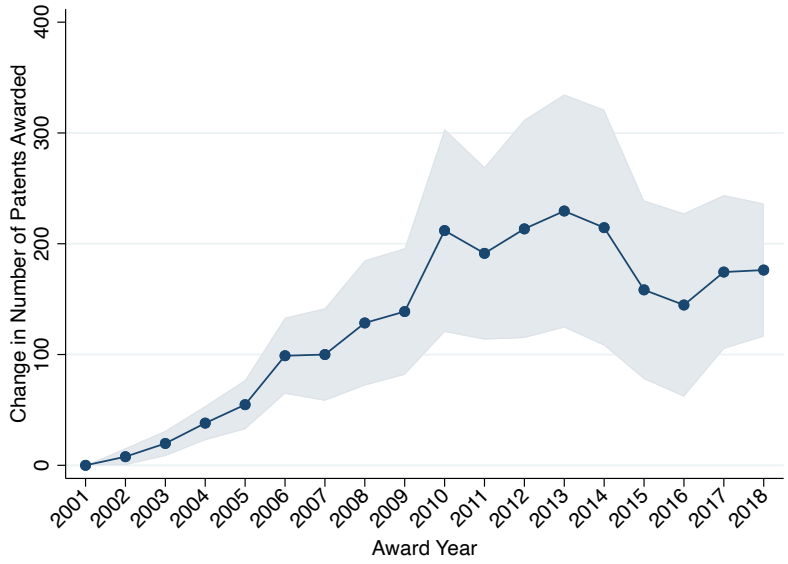
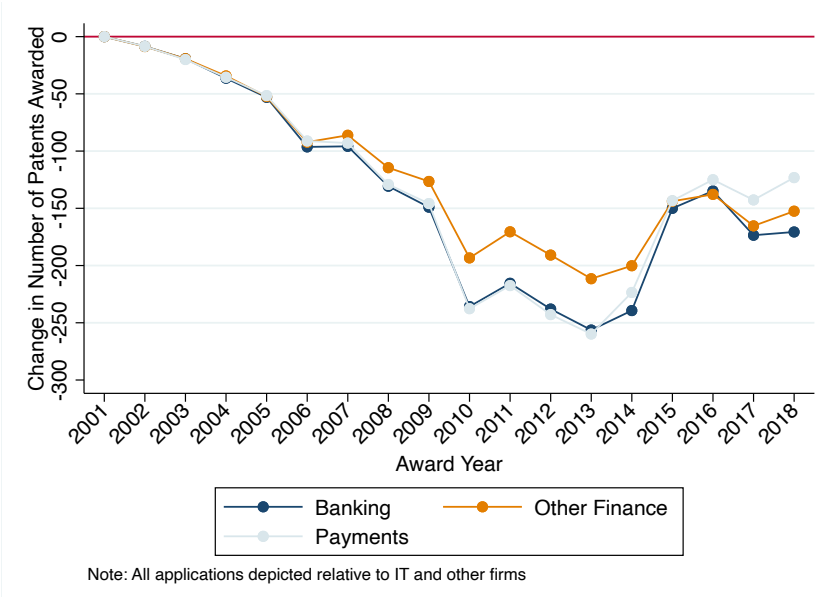


Figure A-6. Decomposition of financial patenting. The charts depict the results of a regression analysis, where the dependent variable is the number of financial patents awarded in each year-assignee firm industry-patent type-inventor location cell. The charts depict the annual fixed effects with 95% confidence limits (Panel A), the interactions between year and assignee industry (Panel B, relative to "IT and Other Industries"), and inventor location (Panel C, relative to "Non-U.S. Inventors").

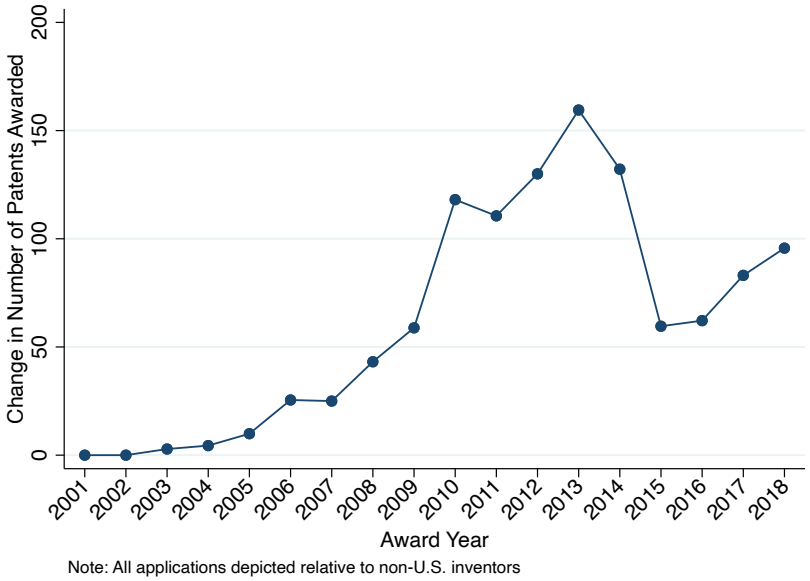
Panel A: Financial patenting by award year.



Panel B: Financial patenting by assignee industry.



Panel C: Financial patenting by geography.



Note: All applications depicted relative to non-U.S. inventors

Figure A-7. Trends in patent citations to academic articles in finance patents. The figure presents the number of academic citations per finance patent over time, to publications in business, economics, and finance, information technology, and other fields, by application year, normalized by the number of academic citations in non-finance patents. Each series is set to 100 for applications in the year 2000.

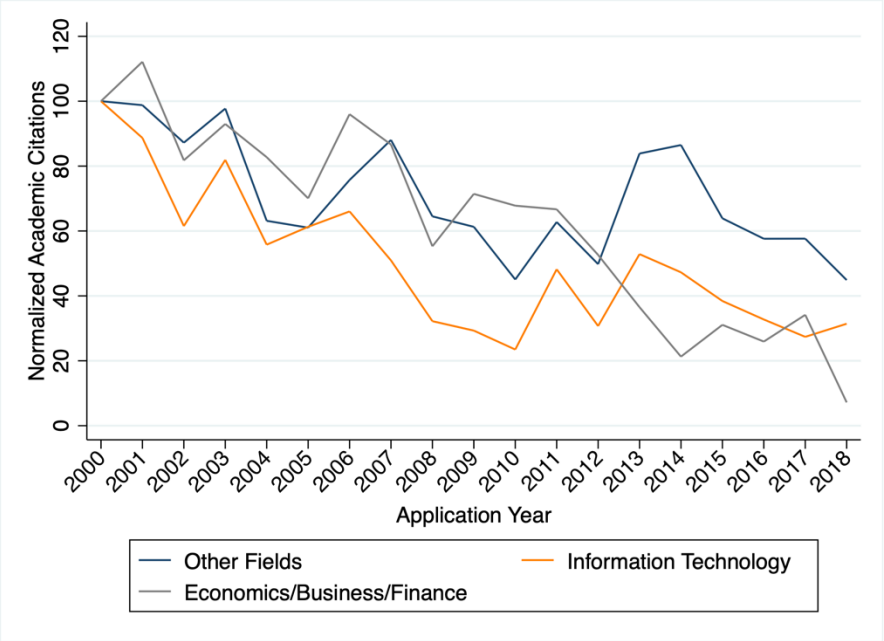
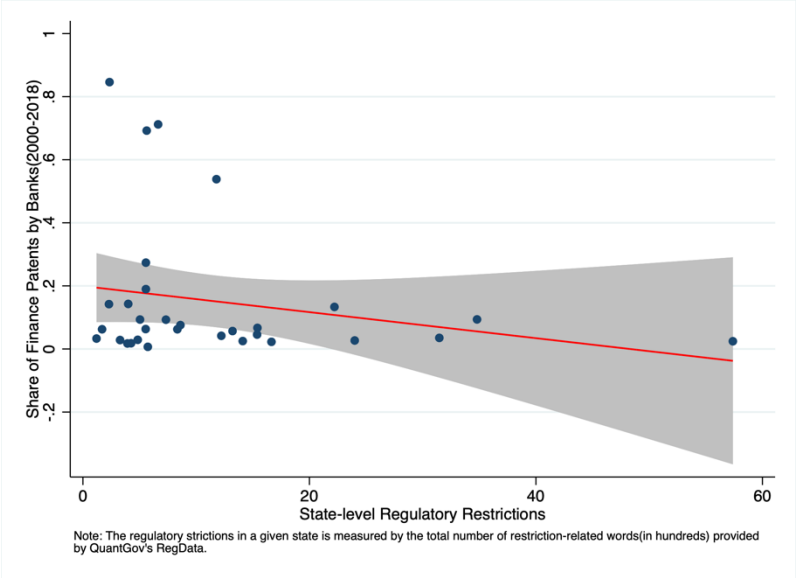


Figure A-8. Banks' finance patenting share and a state's regulatory restrictions. The x-axis reports the state-level regulatory restrictions. measured using the total number of restriction-related words (in hundreds) provided by QuantGov's RegData; the y-axis, a given state's total share of finance patent applications (consumer-oriented finance patents only in Panel C) from banks between 2000-2018 (Panels A and C) and 2008-2018 (Panel B), calculated as the total number of finance patents by banks divided by the total number of finance patents by all kinds of firms.

Panel A: Share of finance patents by banks between 2000 and 2018.



Panel B: Share of finance patents by banks between 2008 and 2018.

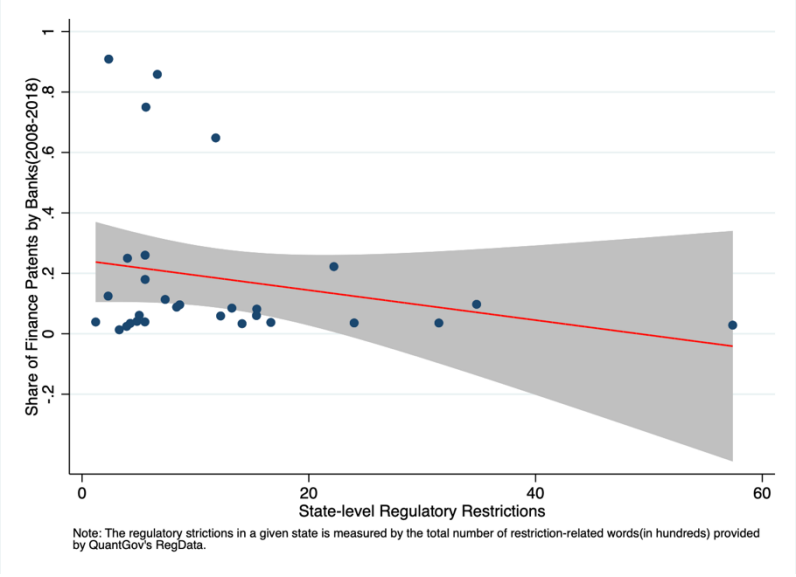


Figure A-8 (continued).

Panel C: Share of finance patents (consumer-only) by banks between 2000 and 2018.

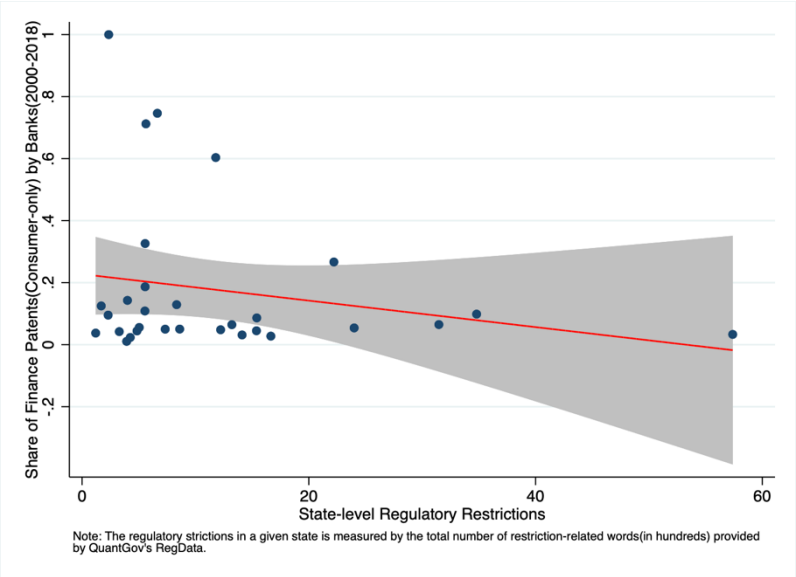
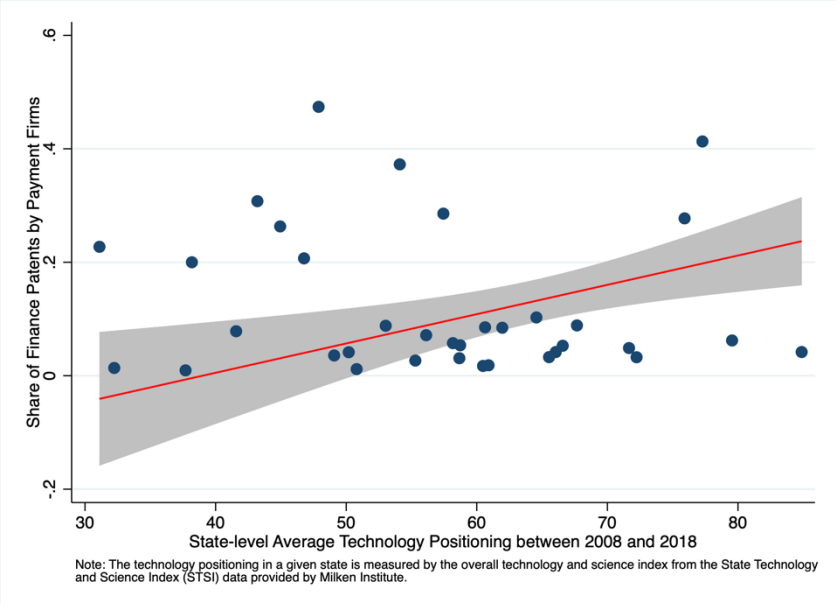


Figure A-9. Payments firms' finance patenting share and a state's technological positioning (measured by two measures used in Table 10). The x-axis reports a given state's average technological positioning between 2008 and 2018 (calculated using average Overall Technology (Panel A) and R&D Input (Panel B) indices between 2008 and 2018 from the STSI data provided by Milken Institute); the y-axis, a given state's total share of finance patent applications from payments firms between 2008 and 2018 (calculated as the total number of finance patents by payments firms divided by the total number of finance patents by all kinds of firms).

Panel A: Overall technology index. (updated from 0527 draft)



Panel B: R&D input index. (updated from 0527 draft)

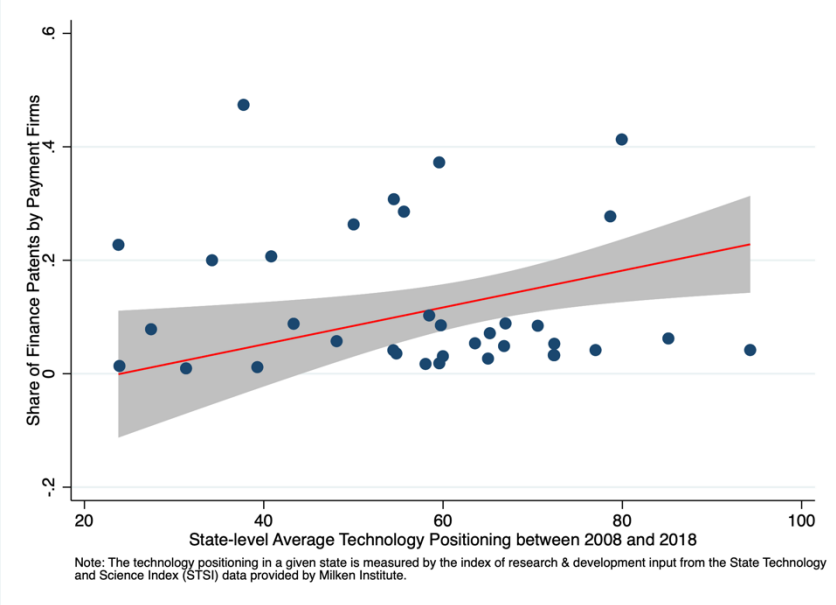
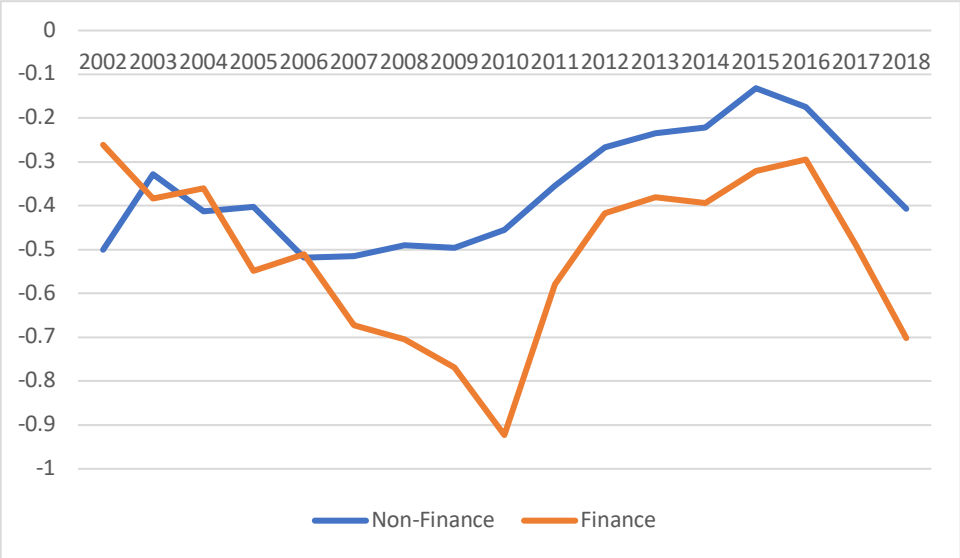


Figure A-10. The extent of patent revision between application publication and award, over time. Panel A reports the change in the number of independent claims at the time of application publication and award, for finance and non-finance patents. Panel B reports the change in the length of the shortest independent claim between these two points, for finance and non-finance patents. The mean values are presented by year of award.

Panel A. Change in independent claim count.



Panel B. Change in independent claim length.

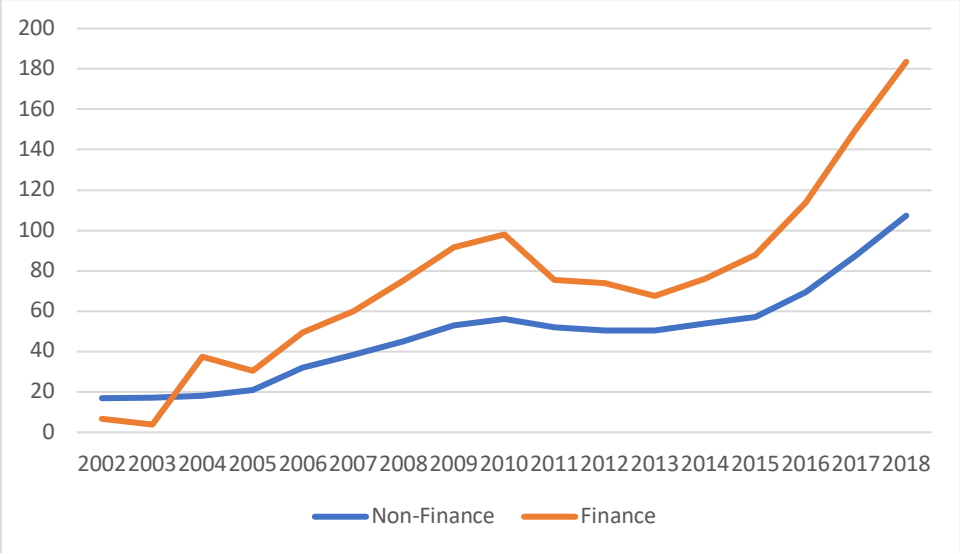


Figure A-11. Financial patenting in U.S. Census regions over time. The chart depicts the results of an OLS regression analysis of financial patenting across U.S. Census regions over time. Using observations at the application year-census region level, the dependent variable is the number of financial patents in a given cell. The chart presents coefficients on the interactions of the application time period fixed effects with fixed effects for two specific census regions: Pacific and South Atlantic regions. The Middle Atlantic region and the 2000-04 period are the baselines**.** Robust standard errors (90% level) are denoted with shadowed areas.

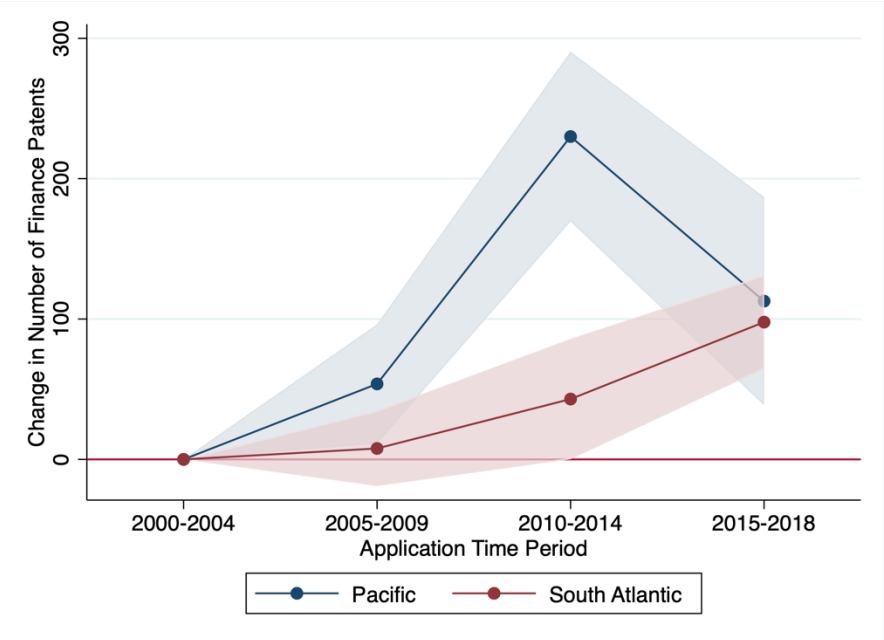
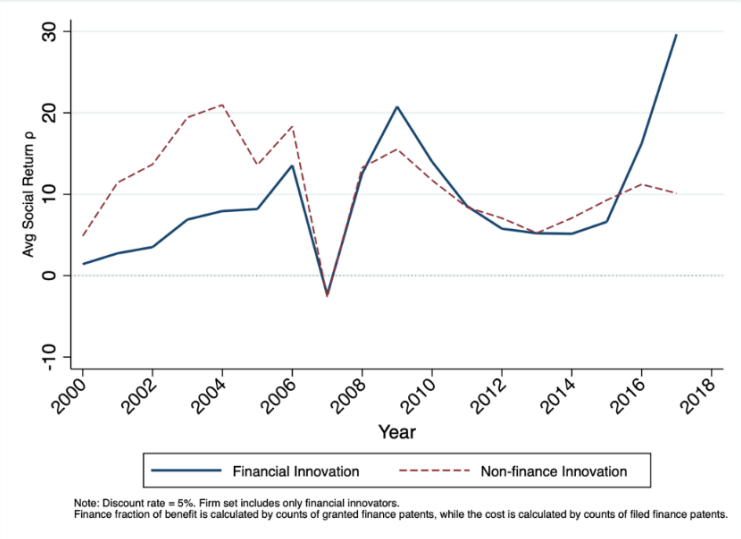
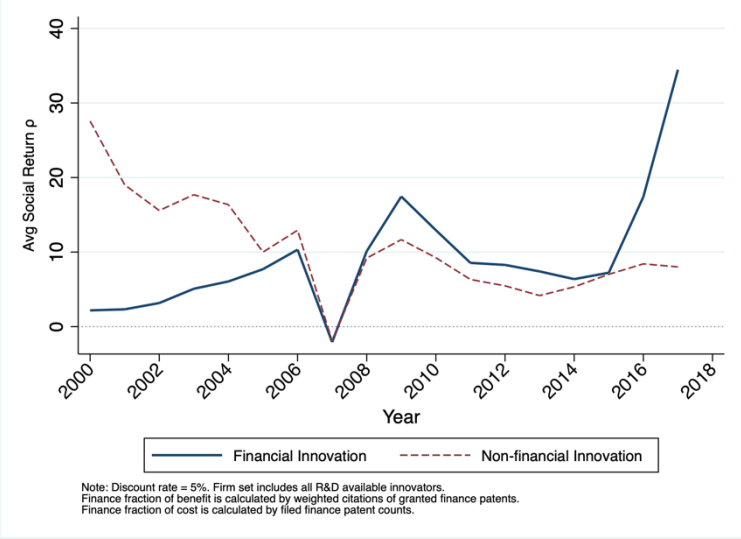


Figure A-12: Sensitivity checks of social and private returns of financial vs. non-financial innovation. Panel A shows the social return of financial vs. non-financial innovation when the benefit is scaled by the ratio of the (unweighted) number of finance patents granted to the total patents granted. Panel B shows the social return of financial vs. non-financial innovation when the firm set includes all firms with R&D information available. Panel C looks similarly at the private returns.

Panel A: Social return, scaling by patent count.



Panel B: Social returns, including all R&D-performing firms. (updated from 0406 draft.)



Panel C: Private returns, including all R&D-performing firms. (updated from 0406 draft)

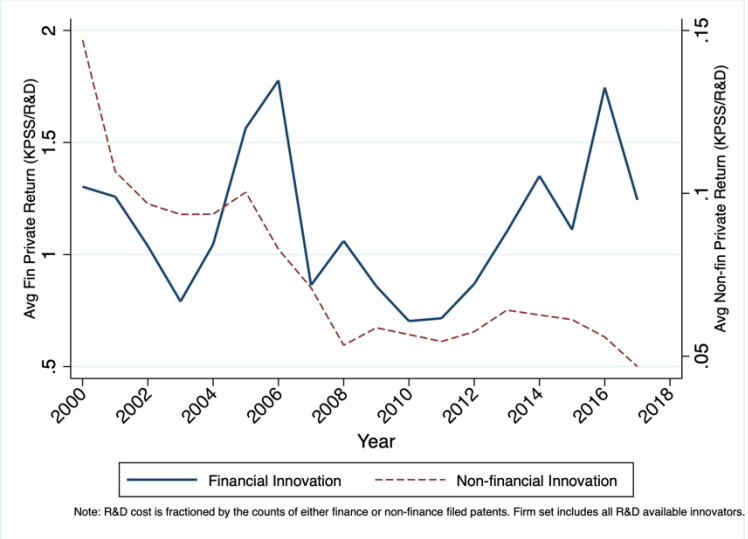
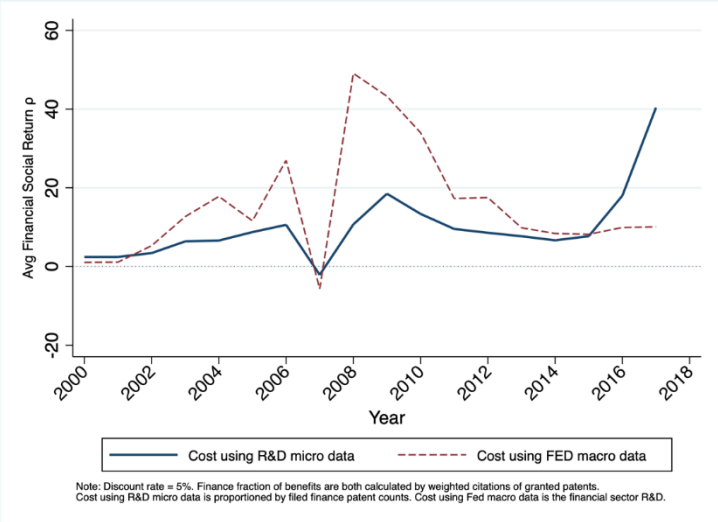


Figure A-13: Comparison of financial social returns using R&D firm-level data and using St. Louis Federal Reserve Bank macro R&D. Panel A used the Fed R&D input measure in equation (27), but otherwise left the expression unchanged. Panel B undertook a similar calculation, but excluded financial patents awarded to IT firms from these calculations.

Panel A: Using the Fed R&D input measure in equation (27).



Panel B: Using the Fed R&D input measure in equation (27) and excluding IT firms' innovation from the calculation.

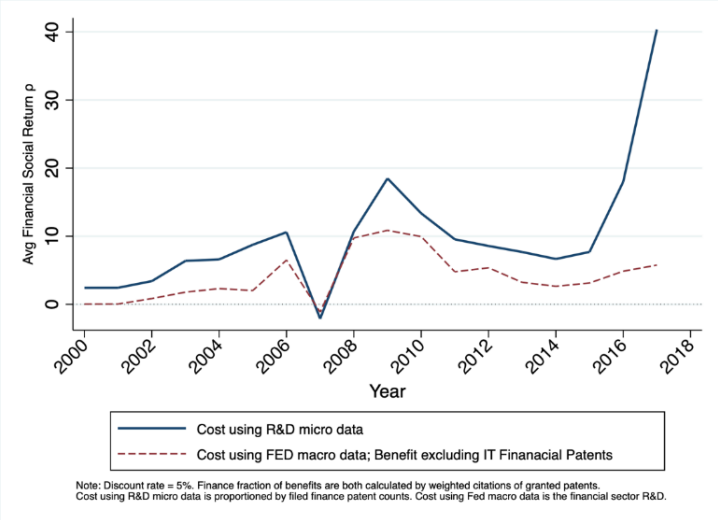


Table A-1. List of keywords.

Accounting	**Consumer Banking**	**Communications**	**Cryptocurrencies**	**Currency**	**Funds**	**Investment Banking**
Accounting	Bridge Finance	Broadcast	Altcoin	Currency Conversion	ETF	Asset Analysis
Accounts Payable	Commercial Loan	Broadcasts	Bitcoin	Exchange Rate	Exchange Traded Fund	Asset Characterization
Accounts Receivable	Covenant	Communication	Blockchain	Foreign Exchange	Hedge Fund	Bid Ask
Audit	Debtor Finance	Communications	Cryptocurrency	Forex	Mutual Fund	Bond
Auditor	Debtor in Possession	Message	Distributed Ledger	Swap	Private Equity	Call Option
Bookkeeper	Default		Initial Coin Offering		Venture Capital	Chinese Wall
Budget	Event	News Feed	Token			Derivative
Budgeting	Indicator Lending Rate	News Feeds				Dummy Order
Cash Flow	Interest Coverage					Gilt
Controller	Letter Of Credit					Hair Cut
FIFO	Line of Credit					Hidden Liquidity

Financial Controls	Material Adverse Change					Initial Public Offering
First in First Out	Sweep Account					Liquidity Pool
Forecasting	Term Loan					Liquidity Provider
Free Cash Flows	Zero Balance Account					Margin
GAAP						Moving Average
Generally Accepted Accounting Principles						Option
Gross Margin						Order Book
Information System						Price Level
Interest Coverage						Put Option
Inventory						Short Selling
Last In First Out						Trading Protocol
LIFO						Valuation
Net Present Value						

Net Working Capital						
Payable						
Payback						
Payroll Taxes						
Quick Ratio						
Working Capital						

Table A-1 (continued).

Insurance	Passive Funds	Payments	Real Estate	Retail Banking	Security	Wealth Management
Actuarial	Index Fund	Authorized	Appraisal	ATM	Authentic	Active Management
Auto Insurance	Passive Fund	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	Credit Score	Cryptographic	Custodian
Disability		Keypad	Dual Agency	Direct Deposit	Decipher	Financial Industry Regulatory Authority
Driving Behavior		Kiosk	Easement	Direct Payroll Deposit	Decrypt	FINRA
Driving Environment		Merchant	Eminent Domain	Interbank Fee	Decryption	Front-End Load

Earned Premium		NFC	Escrow	Money Market	Detection	Individual Retirement Account
Home Insurance		Payment	Eviction	NOW Account	Encrypt	Prospecti
Homeowners Insurance		Point of Sale	Foreclosure	Online Banking	Encryption	Prospectus
Indemnity		POS	Home Equity	Overdraft	Fraud	Target Date Fund
Insurance Risk			Home Warranty	Passbook	Fraudulent	Tax Avoidance
Life Insurance			Jumbo Loan	Savings	Identifier	Tax Benefit
Life Settlement			Loan To Value	Student Loan	Identity	Tax Cost
Long-Term Care			Mortgage	Time Deposit	Public Key	Tax Deduction
Malpractice			Non-Conforming Loan	Withdrawal Fee	Secure Key	Wrap Fee
Reinsurance			Prepayment		Security	
Structured Settlement			Real Estate Investment Trust		Spoofing	
Term Insurance			Realtor		Symmetric Key	
Umbrella Liability			Refinancing		Theft	

Vehicle Damage			REIT		Token	
			Tax Lien		Verify	
			Title Search			
			Zoning			

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. The impact of finance patents and other academic-oriented patents, by assignee type. The table presents the citation weights, the Kogan et al. (2017) weights, and the Kelly et al. (2021) weights for finance patents and all other patents applied for between 2000 and 2018 and awarded by February 2019, restricting the control group to all patents in the 53 four-digit CPC patent classes in which universities most frequently filed patents. The table also presents the results of t-tests and nonparametric k-sample tests of the equality of medians. The table also presents the differences in the percentile ranks of the means and medians of the finance and non-finance patents using the distribution of all patents in the sample.

	Citation weights		Kogan et al. weights		Kelly et al. weights		# of patents
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	
Finance Patents	1.25	0.28	53.61	17.50	0.86	0.99	24,255
Other Patents	1.00	0.19	11.76	3.78	0.81	0.90	1,823,388
p Value, equality test	0.000	0.000	0.000	0.000	0.000	0.000	
Difference percentile	+4	+5	+21	+35	+5	+13	
	Citation weights		Kogan et al. weights		Kelly et al. weights		# of patents
	Mean	*Median*	*Mean*	*Median*	*Mean*	*Median*	
Finance Patents	1.25	0.28	53.61	17.50	0.86	0.99	24,255
Other Patents	1.00	0.19	11.76	3.78	0.81	0.90	1,823,388
p Value, equality test	0.000	0.000	0.000	0.000	0.000	0.000	
Difference percentile	+4	+5	+21	+35	+5	+13	

Table A-5. The assignee types of financial and other academic-oriented non-financial patents. The sample consists of finance and non-finance patents applied for between 2000 and 2018 and awarded by February 2019, restricting the control group to all patents in the 53 four-digit CPC patent classes in which universities most frequently filed patents. We compare the distribution of assignees of finance and non-finance patents in t-tests. * 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	*Other patents*
Assignee Type:		
U.S. corporation	74.96%	47.07%***
Foreign corporation	16.05%	46.09%***
Individual	8.65%	3.38%***
U.S. government	0.08%	0.34%***
Foreign government	0.01%	0.11%***
U.S. university	0.19%	2.04%***
Foreign university	0.06%	0.96%***

Table A-6. The assignees of financial patents. The table presents the share of applications with assignees below various employment size thresholds in the application year, as a share of all corporate applications with employment data in that period.

Employment threshold	*2000-04 patent applications*	*2015-18 patent applications*
<250	2.4%	1.8%
<500	5.8%	2.1%
<1000	7.8%	3.2%

Table A-7. Decomposition of financial patenting. The table presents results of a regression analysis of finance patenting, where the dependent variable is the number of financial patents awarded in each year-assignee industry-patent type-inventor location cell. The table reports the results of F-tests of the joint significance of the 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-8. Keywords associated with finance patents designated as consumer oriented.

401k or 401(k)
Annuity or annuities
ATM or teller machine
Auto[mobile] insurance or car insurance
Auto[mobile] loan
College savings
Credit card
Credit report
Credit score
Customer
Debit card
Defined benefit
Defined contribution
e-Commerce
Financial adviser
Financial literacy
Health insurance
Home equity
Homeowner's insurance
Identity theft
Individual
Life insurance
Lottery payment
Medical loan or medical debt
Mobile phone
Mutual fund
Payday loan
Pension
Prepaid card
Policy holder or policyholder
Renter's insurance
Retail
Retirement account
Reverse mortgage
Savings account
Social security
Student loan or student debt
Unemployment insurance

Table A-9. Software patents. The sample consists of all finance patents applied for between 2000 and 2018 and awarded by February 2019. The table presents OLS regression analyses. The dependent variable is a dummy variable denoting if the patent is a software one. The key independent variables are the application year, dummies for whether the patent was assigned to a bank or an information technology, payments, or other non-finance firm, and the interaction between the application year and the assignee type. We also include unreported controls for firm characteristics (see text for details). Robust standard errors in brackets; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Software patent?		
	(1)	*(2)*	*(3)*
Application year	0.010***	0.010***	0.006***
	[0.0005]	[0.001]	[0.001]
Assignee is bank		-0.061***	-15.674***
		[0.012]	[4.678]
Assignee in IT, payments, or other		-0.036***	-8.888***
		[0.006]	[2.748]
Bank * Application year			0.008***
			[0.002]
IT/payment/other * Application year			0.004***
			[0.001]
Bank = IT/payment/other		0.063	
Bank*year = IT/payment/other*year			0.101
Observations	24,123	17,552	17,552
R-squared	0.019	0.061	0.062
Assignee characteristic controls	No	Yes	Yes

Table A-10. Most frequently cited academic journals in finance patents. The table present the journals most frequently cited in finance patents applied for between 2000 and 2018 and awarded by February 2019. 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*.

Journal Name	*Number of Citations*
Communications of the ACM	1166
Journal of Finance	701
Journal of Animal Science	499
Financial Analysts Journal	381
IEEE Computer	367
Journal of Portfolio Management	288
Social Science Research Network	281
ABA Banking Journal	277
Computers & Security	246
Lecture Notes in Computer Science	242
IBM Systems Journal	238
IEEE Spectrum	216
Management Science	213
ACM Computing Surveys	206
Journal of Financial Economics	197

Table A-11. 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, economics, and finance journals, all business, economics, and finance 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 reported for finance patents, all patents, and all patents in the 53 four-digit CPC patent classes in which universities most frequently filed 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	2.33	3.69***	6.19***
Total Citations to High-Impact Factor Journals	0.27	1.53***	2.82***
Total Citations to Business/Economics/Finance Journals	0.54	0.02***	0.02***
Total Citations to High-Impact Bus/Econ/Fin 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)	8.75	9.62***	9.25***
Number of Observations	24,255	3,781,439	1,823,420

Table A-12. OLS regression analyses 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 dummy (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 brackets; * 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	-4.46***	0.71***	0.07***	-0.35***
	[0.34]	[0.01]	[0.001]	[0.12]
Panel B				
Financial patent	-1.33	0.47***	0.04***	-0.42
	[1.04]	[0.02]	[0.003]	[0.42]
U.S. corporation	4.78***	0.04***	0.0001	-0.95***
	[0.12]	[0.002]	[0.0003]	[0.06]
Foreign corporation	2.94***	0.02***	-0.0001	-1.60***
	[0.18]	[0.004]	[0.0005]	[0.07]
U.S. university	26.50***	0.04***	0.0001	-1.15***
	[0.24]	[0.005]	[0.0006]	[0.08]
Financial * U.S. corporation	-2.88***	0.28***	0.037***	0.03
	[1.10]	[0.02]	[0.003]	[0.44]
Financial * Foreign corporation	-1.09	0.05	-0.01***	-0.49
	[2.00]	[0.05]	[0.005]	[0.69]
Financial * U.S. university	-20.37***	0.37***	0.03***	0.98
	[5.63]	[0.13]	[0.01]	[1.43]
Panel C				
Financial patent	-4.24***	0.80***	0.08***	-0.04
	[0.40]	[0.01]	[0.001]	[0.13]
Venture-backed firm	7.19***	0.02***	-0.0003	0.53***
	[0.16]	[0.004]	[0.0004]	[0.05]
Financial * Venture-backed	-6.10***	-0.02	0.05***	-4.57***
	[1.48]	[0.03]	[0.004]	[0.48]

Table A-13. Academic citations. The sample consists of finance patents applied for between 2000 and 2018 and awarded by February 2019. The table reports the mean citation weight, the Kogan et al. (2017) weight, and the Kelly et al. (2021) weight for patents that do and do not cite any academic output, cite publications with an above-median impact factor, and cite publications of various types (all business, economics, and finance journals, all business, economics, and finance journals with an above-median impact factor, and "Top 3" finance journals).

	Mean, weighted citations		Mean, Kogan et al. value		Mean Kelly et al. value	
	Yes	*No*	*Yes*	*No*	*Yes*	*No*
Academic Citation(s)?	1.49	1.11***	59.9	50.2***	0.89	0.83***
Citation(s) to High-Impact Factor Journals?	1.88	1.16***	69.1	51.7***	0.87	0.86
Citation(s) to Business/Economics/Finance Journals?	1.38	1.22***	85.3	48.0***	0.90	0.85***
Citation(s) to High-Impact Bus/Econ/Fin Journals?	1.52	1.24**	96.6	52.3***	0.91	0.85**
Citation(s) to Top 3 Finance Journals?	1.30	1.25	184.2	52.1***	0.93	0.86***

Table A-14. 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 Charlotte-Concord CSA. 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 define mid-sized firms as those where the firm's revenue in the application year was more than $100 million but less than $10 billion, and small and large firms similarly. We then 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.

Table A-14 (continued).

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

Table A-14 (continued).

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

Table A-14 (continued).

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	4.5%	25.4%	38.1%	58.2%	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-15. The impact of technological positioning on financial patenting. The table is similar to Table 10, but with the key independent variables being interactions between (a) the other four STSI technology indexes in a given state in year t and (b) assignee industry. All regressions include fixed effects for time, state, patent type, and assignee industry. Only selected interactions are reported. Clustered standard errors (at the state-year level) are in brackets; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Patent count			
	(1)	*(2)*	*(3)*	*(4)*
Technology Concentration x Payments Firms	0.038***			
	[0.015]			
Technology Concentration x IT/Other Firms	0.179***			
	[0.035]			
Entrepreneurial Capacity x Payments Firms		0.047***		
		[0.017]		
Entrepreneurial Capacity x IT/Other Firms		0.240***		
		[0.041]		
Technology Workforce x Payments Firms			0.035**	
			[0.014]	
Technology Workforce x IT/Other Firms			0.179***	
			[0.034]	
Human Capital Investment x Payments Firms				0.032***
				[0.009]
Human Capital Investment x IT/Other Firms				0.109***
				[0.024]
Observations	6,600	6,600	6,600	6,600
R-squared	0.395	0.402	0.390	0.362
Time FEs	Yes	Yes	Yes	Yes
State FEs	Yes	Yes	Yes	Yes
Patent type FEs	Yes	Yes	Yes	Yes
Assignee industry FEs	Yes	Yes	Yes	Yes
Data sample period	2008-18	2008-18	2008-18	2008-18
Test Equality of Coefficients (F Statistic Reported)				
Interaction with Payments vs. Bank	6.90***	7.12***	6.16**	12.23***
Interaction with IT/Other vs. Bank	26.52***	33.94***	27.09***	20.87***

Table A-16. 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 largest (patent-weighted) departure and destination CSAs are 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	129	11026
Of these, firms that shifted modal CSA	28	3640

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	2778
Denver-Aurora, CO	1	297
San Jose-San Francisco-Oakland, CA	3	188
Three most frequently arrived 2015-18 CSAs:		
Charlotte-Concord, NC-SC	1	652
Rochester-Austin, MN	1	589
St. Louis-St. Charles-Farmington, MO-IL	1	418

Table A-17. Returns analysis sample. The table presents the distribution of most frequently represented industries (using four-digit Standard Industrial Classification codes) (Panel A) and R&D expenditures and the ratio of R&D to sales for the 278 firms in the return analysis sample (Panel B). Each firm-year is an observation; the R&D/sales ratio is weighted by firm revenue.

Panel A: Most frequently represented industries.

Industry Code	*Industry Name*	*Share*
7372	Prepackaged software	11.8%
7370	Computer programming, data services, etc.	11.4%
3674	Semiconductors and related devices	11.2%
3663	Radio and TV broadcasting and communications equipment	5.3%
7373	Computer integrated system design	4.9%
4813	Telephone communications	4.0%
3577	Computer peripheral equipment, etc.	3.3%
4812	Radiotelephone communications	2.4%
3711	Motor vehicles and passenger car bodies	2.4%
3752	Computer storage devices	2.3%

Panel B: Distribution of R&D spending and R&D/sales ratio.

	Mean	*1%*	*5%*	*10%*	*25%*	*50%*	*75%*	*90%*	*95%*	*99%*
R&D ($MM)	1009	0	1	9	43	208	963	3200	5151	9275
R&D/sales	4.8%	0.0%	0.1%	0.2%	1.1%	3.6%	5.8%	13.1%	15.5%	21.5%

Table A-18. Comparison of patents included in the returns analysis with other awards in the sample. Panel A compares the features of the 1.2 million patents included in the analysis with the other 2.6 million awards in the sample. Panel B summarizes the most underrepresented primary four-digit CPC patent classes in the analysis (as a difference in share of all patents); Panel C the most overrepresented classes.

Panel A. Comparison of returns analysis sample with other patents in sample.

	Return analysis sample	*Other patents*
Number of total citations	7.4	7.1***
Normalized citations	0.97	1.01***
Assignee in Bay Area	12.0%	6.1%***
Assignee in New York area	3.8%	2.7%***
Application date	Oct. 23, 2008	Dec. 29, 2008***
Award date	Jan. 29, 2012	Jan. 20, 2012***

Panel B. Most underrepresented patent subclasses in return analysis sample.

CPC subclass	*Title*	*Difference*
A61K	Preparations for medical, dental, or toilet purposes	-2.6%
C07D	Heterocyclic compounds	-1.8%
A61B	Diagnosis; surgery; identification	-1.8%
G01N	Investigating or analyzing materials by determining their chemical or physical properties	-1.2%
C07K	Peptides	-1.1%
A61F	Filters implantable into blood vessels	-1.0%
C12N	Microorganisms or enzymes	-0.9%
A61M	Devices for introducing media into, or onto, the body	-0.9%
A63B	Apparatus for physical training, gymnastics, swimming, climbing, or fencing	-0.8%
B65D	Containers for storage or transport of articles or materials	-0.7%

Panel C. Most overrepresented patent subclasses in return analysis sample.

CPC subclass	*Title*	*Difference*
G06F	Electric digital data processing	12.2%
H04L	Transmission of digital information, e.g. telegraphic communication	6.3%
H04W	Wireless communication networks	3.5%
H01L	Semiconductor devices; electric solid state devices not otherwise provided for	3.1%
H04N	Pictorial communication, e.g., television	2.8%
G11B	Information storage based on relative movement between record carrier and transducer	1.2%
G11C	Static stores	1.1%
G06T	Image data processing or generation, in general	0.9%
H04B	Transmission	0.9%
H04M	Telephonic communication	0.8%

Table A-19. The extent of patent revision between application publication and award. The table reports the number of independent claims at the time of the application publication and award, the length of the shortest independent claim at these two points, and the change in these measures for finance and non-finance patents. The sample consists of all patents applied for between 2000 and 2014 and issued by February 2019 with an original review by the USPTO. 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 A-20. Comparison of the finance patent samples in Lerner (2002) and this paper. Information is derived from Patentsview, as well as the methodologies described in the paper.

	Lerner (2002) sample	*This sample*
Number of patents:	445	24,255
Patent age:		
First Application Year	1968	2000
Last Application Year	1999	2018
First Award Year	1971	2001
Last Award Year	2000	2019
Median Application Year	1995	2009
Median Award Year	1998	2013
First inventor foreign:	13.9%	21.0%
First inventor U.S. location:		
East North Central	10.4%	11.1%
East South Central	0.3%	0.5%
Middle Atlantic	27.4%	13.2%
Mountain	4.2%	5.6%
New England	10.4%	5.8%
Pacific	23.0%	21.8%
South Atlantic	15.7%	12.1%
West North Central	2.6%	3.5%
West South Central	6.0%	5.5%
Assignee type:		
U.S. corporation	51.2%	75.0%
Foreign corporation	11.9%	16.0%
Individual	35.6%	8.7%
U.S. government	0.4%	0.1%
Foreign government	0.0%	0.0%
U.S. university	0.9%	0.2%
Foreign university	0.0%	0.1%
Assignee corporate type:		
Banking	18.5%	6.6%
Payments	3.9%	9.8%
Other finance	29.2%	14.2%
IT	33.8%	37.4%
Other	14.6%	32.1%
Mean impact:		
Citation weight	1.97	1.25
Kogan et al. weight	63.41	53.61
Kelly et al. weight	2.64	0.86

Top 3 assignees:		
	Merrill Lynch	Bank of America
	Citigroup	Trading Technologies International
	Hitachi	Visa

Note: The assignment of patentee type differs slightly from Lerner (2002), as this classification is now based on USPTO reporting in Patentsview. The 2002 paper classified patents based on the author's own research. In particular, a small number of patents that were assigned to holding companies associated with a single inventor were classified in that paper as being individual patents, but by the USPTO (and Patentsview) as corporate ones.

Table A-21. Financial patenting by U.S. region over time. The table presents the share of financial 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 computes shares 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%

Table A-22. Probit regression analysis of the determinants of the movement of financial patentees. The sample consists of 129 firms that filed financial patents in 2000-04 and 2015-18. The dependent variable is a dummy indicating if the firm shifted its modal CSA for patent applications filed in these two periods. The independent variables include dummies for firm industry (payments is the omitted category), whether the firm is venture-backed or publicly traded (both as of the time of the first patent filing in the 2000-04 period), and whether its modal patenting location in 2000-04 were the New York or San Francisco CSAs, as well as the volume of finance venture capital investments in 2000 (in billions of U.S. dollars) in the modal CSA. The observations are weighted by the number of patents filed by the firm in 2000-04. Robust standard errors are in brackets; * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	Did the firm switch CSAs?		
Is firm a bank?	0.71***	-0.40***	0.30**
	[0.15]	[0.13]	[0.14]
Is firm other financial service?	-0.13	-1.34**	-0.64***
	[0.13]	[0.15]	[0.16]
Is firm IT or other?	1.02**	-1.38***	-0.69***
	[0.09]	[0.10]	[0.11]
Is firm venture-backed?	-0.46	-0.23	0.27
	[0.49]	[0.55]	[0.47]
Is firm publicly traded?	-0.22***	-0.57***	-0.49***
	[0.07]	[0.10]	[0.11]
Is modal patent in 2000-04 in NY CSA?		2.13***	1.83***
		[0.10]	[0.10]
Is modal patent in 2000-04 in SJ/SF CSA?		0.20**	-1.80***
		[0.10]	[0.22]
2000 Finance VC investments in modal CSA			2.02***
			[0.22]
Number of unweighted observations	129	129	129
Weighted observations	2176	2176	2176
p-Value, χ^2-test	0.000	0.000	0.000
Pseudo R^2	0.145	0.399	0.417

Table A-23: Summary statistics for private return analysis. The table consists of 2,808 observations at firm-year level of 246 firms covering the period from the year of their first financial patent application through 2018. Market values and book values of equity, R&D expenditures, and Kogan et al. (2017) values are in millions of U.S. dollars. R&D/A is the ratio of R&D stock and book value of equity. Patents/R&D is the ratio of patent stock and R&D stock. Citations/ patent is the ratio of citation stock and patent award stock. Mean Kogan/R&D is the ratio of mean Kogan value of successful patent applications in that year and the R&D expenditure of a firm in a particular year. The dummy variables indicating whether financial or non-financial R&D expenditure in that year for a particular firm is 0 are also reported. The mean, median, minimum, maximum and standard deviation in the data are reported for financial and non-financial patents separately.

	Mean	Median	Minimum	Maximum	Standard Deviation
Market value ($M)	36,404.22	8,598.24	2.97	1,073,390.50	78,477.64
Book value ($M)	13,922.05	2,775.00	0.79	241,948.00	29,210.92
Market-to-book value	5.29	2.58	0.18	2,027.99	41.72
Fin R&D stock ($M)	32.06	5.00	0.03	1,045.76	83.96
Non-fin R&D stock ($M)	4,577.11	813.77	0.23	78,639.42	8,792.22
Fin patent award stock	9.38	1.62	0.05	240.04	24.24
Non-fin patent award stock	1,677.26	299.91	0.32	39,774.17	3,578.09
Fin citation stock	15.84	1.73	0	625.03	47.79
Non-fin citation stock	1,649.97	327.37	0	34,640.07	3,361.69
Fin R&D/Assets	0.03	0	0	5.59	0.18
Non-fin R&D/Assets	1.11	0.37	0	717.81	13.99
Fin patents/R&D	1.00	0.40	0	113.36	4.51
Non-fin patents/R&D	0.87	0.32	0	107.54	4.01
Fin citation/patent	1.75	0.94	0	64.88	3.33
Non-fin citation/patent	1.46	1.01	0	36.98	2.07
Fin mean Kogan/R&D	2.27	0.60	0	82.20	6.17
Non-fin mean Kogan/R&D	0.10	0.02	0	18.43	0.57
D(Fin R&D = 0)	0.63	1.00	0	1	0.48
D(Non-fin R&D = 0)	0.09	0	0	1	0.29

Table A-24: The market value as a function of financial and non-financial R&D, patents, and citations, 2000 – 2018. The table presents the results of the estimation of a nonlinear model with the dependent variable log Tobin's q. The table presents the results from estimating equation (33) in Appendix H relating the market value of firms and innovation stocks from 2000 to 2018 using nonlinear least squares. In columns (1), (2), and (3), we report the results for firms with at least one, five, and ten financial patents applied for from 2000 to 2018. For financial and non-financial patents, we include the following independent variables: R&D stock (million USD) over book value of equity (million USD); patent award stock over R&D stock (million USD); adjusted (for the mean citations in that application year) citation stock over patent award stock; application year fixed effects; and dummy variables indicating whether financial and non-financial R&D expenditures in that year are zero for a particular firm. The number of observations in column (1), (2), and (3) are 2,808 from 246 firms, 1,440 from 107 firms, and 1,069 from 71 firms respectively. Heteroskedastic robust standard errors are reported in brackets. * denotes statistical significance at the 10% level; ** at the 5% level; and *** at the 1% level.

	(1)	(2)	(3)
$R\&D^{Fin}/A$	3.792***	7.806***	6.315***
	[0.683]	[1.537]	[1.539]
$R\&D^{Nonfin}/A$	0.438***	0.671***	0.736***
	[0.0623]	[0.130]	[0.150]
$PAT^{Fin}/R\&D^{Fin}$	-0.0205***	-0.0231**	-0.0242**
	[0.00513]	[0.0100]	[0.00957]
$PAT^{Nonfin}/R\&D^{Nonfin}$	0.0128*	0.00866	0.0136
	[0.00659]	[0.0121]	[0.0119]
$CITES^{Fin}/PAT^{Fin}$	0.0237***	0.0453***	0.0267***
	[0.00848]	[0.0149]	[0.00826]
$CITES^{Nonfin}/PAT^{Nonfin}$	0.110***	0.488***	0.428***
	[0.0227]	[0.0830]	[0.0896]
Observations	2,808	1,440	1,069
R-squared	0.248	0.322	0.317
Year FEs	Yes	Yes	Yes
D(Fin R&D = 0)	Yes	Yes	Yes
D(Non-fin R&D = 0)	Yes	Yes	Yes
Minimum number of finance patents	1	5	10

Table A-25: Measuring the private returns to financial and non-financial innovation. The table presents two measures of the (private) return on social innovations, the elasticity of firm value to citations and the semi-elasticity of R&D through citations. The analysis is identical to that in Table 12, but in each case, only observations through 2013 are used (rather than 2018), to address concerns about the truncation of citation and patent counts.

	Financial Patents			Non-financial Patents		
	(1)	(2)	(3)	(1)	(2)	(3)
$\frac{\partial logQ}{\partial(CITES/PAT)}$	0.036	0.042	-0.011	0.135	0.236	0.240
$\frac{\partial logQ}{\partial(CITES/PAT)}\frac{\partial(\frac{CITES}{PAT})}{\partial(R\&D)}$	0.275	0.183	-0.101	10.934	9.664	12.414
Minimum number of finance patents	1	5	10	1	5	10

Table A-26: Measuring the private returns to financial and non-financial innovation. The table presents two measures of the (private) return on social innovations, the elasticity of firm value to citations and the semi-elasticity of R&D through citations. The analysis is identical to that in Table 12, but in each case, the semi-elasticity is evaluated at the median rather than the mean.

	Financial Patents			Non-financial Patents		
	(1)	(2)	(3)	(1)	(2)	(3)
$\frac{\partial logQ}{\partial(CITES/PAT)}$	0.061	0.096	0.061	0.176	0.309	0.284
$\frac{\partial logQ}{\partial(CITES/PAT)}\frac{\partial(\frac{CITES}{PAT})}{\partial(R\&D)}$	0.523	0.330	0.191	11.696	21.618	15.846
Minimum number of finance patents	1	5	10	1	5	10