ONLINE APPENDIX

"The use and misuse of patent data: Issues for corporate finance and beyond"

by Josh Lerner and Amit Seru

This section provides additional information on methodology, data, and variable constructions, as well as empirical extensions.

Appendix A: The Patent Award Process

The formal patenting process begins with the filing of an application. (In the U.S., applicants can also file provisional applications since 1994. These allow applicants to ensure that they have priority over other applicants. These filings are not recorded in the major patent databases.)

A patent application to the U.S. Patent and Trademark Office (USPTO) or any other major patent office essentially consists of a series of claims and supporting documentation. Some of the claims in a patent application will be cast in very specific terms; others may be sweeping. A supervising primary examiner reviews each incoming patent application, and assigns it to one of the over 120,000 U.S. patent subclasses. This classification determines which examining group reviews the application. During the review process, the patent will also be assigned to secondary subclasses, as well as classified according to the International Patent Classification (IPC) scheme. Unlike the U.S. scheme, which has grown in an organic (and quite pell-mell) manner over the centuries, the IPC typology has been periodically restructured. It is a nested approach, where each lower level is a subclass of the previous one.

A crucial part of the patent application is the citations. These indicate prior patents, as well as unpatented work, that bear upon the invention. Thus, like citations in scientific journal articles, these may be an indication of influence or knowledge flows. But citations in patents are different from those of articles, as they also play a legal role, serving as "property markers" delineating the scope of the granted claims. They are also sometimes included to defend against others who may seek to invalidate a patent.

There are other differences as well. Citations in articles are largely left to the authors to select. Patent citations are also suggested by the inventors (more likely, their attorneys). But about 60% of the citations are added by the patent examiners (Alcacer and Gittelman, 2006; Thompson, 2006).

The distribution of the patents that get cited is highly skewed. About one-quarter of all patents garner no citations; while 0.01% garner 100 or more citations. Patents that are subsequently more cited are typically interpreted as having more impact or as being more important than less cited patents. Trajtenberg (1990) shows that citations are associated inventions with greater social welfare, but they also appear to be associated with private returns to the inventors: JHT (2005) show that citation-weighted awards are far more correlated with firm value than simple patent counts, and subsequent studies have suggested that more cited patents are also more likely to be litigated, licensed, or otherwise involved in major transactions. The identities of the cited and citing patent firm have also been used to deduce knowledge flows, a line of inquiry that has been validated in a clinical study (Jaffe, et al., 2000).

In addition, the distribution of citations is also important. Patents that cite other patents in a broader array of technology classes are often viewed as having more "originality." Patents that are themselves cited by a more technologically dispersed array of patents are viewed as having greater "generality." Both "originality" and "generality" have been interpreted as measures of the fundamental importance of the research being patented.

After a delay of months or years, a patent examiner in the assigned group then evaluates the proposed patent. To assess the novelty of the application, he searches previous patents issued in the original and related subclasses and several on-line databases. To be entitled to utility patent protection (the most common form of U.S. grant), an innovation must satisfy three criteria. Under 35 U.S.C. 101-103 and 112, it must be:

- A process, machine, manufacture, or composition of matter.
- New, useful, and nonobvious.

- Disclosed in sufficient detail that a skilled person could build and operate it.

Patentable subject matter has traditionally not included fundamental scientific discoveries. A frequently invoked rationale for this omission is that many scientists care little for monetary rewards, and would consequently have pursued the discoveries in any case. To grant patent awards for purely scientific discoveries would consequently be socially wasteful.

If the application appears to conform to the other standards for patentability, the patent examiner will then determine whether the claimed innovation conflicts with any in-process applications or recent patent awards. Unlike almost every other nation, the United States until 2012 granted patents to the party that was "first to invent" a new product or process rather than the one that was "first to file" for an award.

The process of granting an application averages about 34 months.[1] But this number is deceptively low. In many cases, applicants divide, refine, or "continue in part" their patent filings, steps which entail the designation of a new application date. (The USPTO web page provides the history of the patent filing sequence, but most other databases, including the NBER's, only indicate the latest application date.) Were the original filing date to be used for continued applications, the pendency would be much longer. The distribution of review times have a very long tail, with some applications (particularly in overtaxed examining groups such as those handling software, business method, and biotechnology awards) taking ten years or more.

U.S. patent applications and awards only provide protection in the United States. To receive European patent protection, U.S. firms must file an application at the European Patent Office (EPO) within one year of the U.S. application, and similarly in Japan, Australia, China, and so forth. Similarly, foreign firms that seek protection in their home markets must undertake a U.S. filing within a year. (The World Intellectual Property Office provides a way to undertake filings in multiple countries on an expedited basis.) Each of these offices will

[1] http://www.uspto.gov/dashboards/patents/main.dashxml (accessed December 21, 2016).

conduct a separate examination, and decide whether to issue the patent. While the criteria they use are similar, nations differ in the way in which they define the scope of the awards: for instance, the Japanese Patent Office has traditionally been reputed to grant relatively narrow awards. The combination of filing fees, translation costs, and legal service mean that pursuing global patent protection can be very expensive, totaling many hundreds of thousands or even millions of dollars for a single award. As a result, firms are strategic in terms of which patents they file where.

Patent grants since 1994 in the U.S. have been for twenty years from the date that the patent is filed. (Many other nations had converged to this standard earlier.) It allows the owner not only to use the protected product or process, but also to keep others from doing so. Patent awards, rather than applications, have been the typical focus of economists for two reasons. First, it is here that the economic value lies. Second, and more pragmatically, patent applications are not always well documented, while awards are. Patent applications prior to 2000 were kept confidential by the USPTO, while they were published after 18 months by the EPO and Japanese office. Beginning in that year, the USPTO began publishing applications after 18 months as well, as long as a filing had also been made in another office that published the application as well. While these applications are included on the USPTO's web site, apparently the office has in the past removed applications once the patent has either been issued or withdrawn. One important feature of the award process is that the assignment of a patent to an owner is not done until shortly before the patent issues. As is discussed in the paper, this practice makes the use of patents in research more difficult.

To enforce the patent, the primary forum for formally resolving disputes is the Federal courts. (There are also a variety of procedures for resolving questions about patent quality within the USPTO.) The Federal courts have exclusive jurisdiction over disputes involving the infringement of patents, as well as over appeals of USPTO decisions. (Other disputes—e.g., a disagreement between a firm and an employee over a royalty—are routinely referred back to the state courts.) If a firm believes that a patent is being infringed, it may sue the infringer for

damages and/or injunctive relief (a judgment ordering the defendant to cease infringing the patent). Conversely, the alleged infringing firm may preemptively sue the other firm for declaratory relief (a judgment that the plaintiff is not infringing any patent held by the defendant).

In either event, the initial litigation must be undertaken in a district court. Prior to 1982, appeals were heard in the court of appeals of the district in which the case was tried. These circuit courts varied considerably in their interpretation of patent law, and the resolution of these differences through appeals to the U.S. Supreme Court was a lengthy and uncertain process. Consequently, the Court of Appeals for the Federal Circuit (CAFC) was established as the appellate court for all patent-related Federal cases. CAFC decisions may still be appealed to the U.S. Supreme Court, but the latter seldom agrees to hear such appeals.

At any point in the litigation process, the adversarial parties may settle their dispute. This agreement may or may not be accompanied by compensation for retroactive relief and/or a patent license or cross-license agreement. If the settlement is reached before the filing of a suit, the existence of the dispute is unlikely to become public knowledge. Even after a suit has been filed, the nature of the settlements are rarely announced, and certainly not in any systematic manner.

References

Alcacer, Juan, and Michelle Gittelman, 2006, "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations," *Review of Economics and Statistics* 88, 774-779.

Hall, Bronwyn H., 2005. "Exploring the Patent Explosion," *Journal of Technology Transfer*, 30 (2_2), 35-48.

Jaffe Adam B., and Manuel Trajtenberg, 1996, “Flows of Knowledge from Universities and Federal Laboratories: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries,” *Proceedings of the National Academy of Sciences* 93, 12671–12677.

Thompson, Peter, 2006. "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-added Citations," *Review of Economics and Statistics* 88, 383-388.

Trajtenberg, Manuel, 1990, "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *RAND Journal of Economics* 21, 172-187.

Appendix B: Detailed Descriptions of the Fixed Effect and Quasi-Structural Adjustments

HJT (2001) propose three main approaches to help adjust for truncation problems in patents and citations.

B.1 Time Fixed Effect Adjustment:

The Time Fixed Effect adjustment relies on "re-scaling" the patent information for public firms with data about patent population during a certain period. In the adjustment for patents, the annual heterogeneity is removed by dividing the number of patents assigned for each firm in a year by the total number of successful patents applied in corresponding year:

$$Adj\ Patent_{ft} = \frac{n_{ft}}{N_t},$$

where $Adj\ Patent_{ft}$ is the adjusted number of patents for firm f applied in year t, n_{ft} is the total number of patents applied by firm f in year t, and N_t is the total number of patents applied in year t.

Similarly, when adjusting citations, the annual heterogeneous component is corrected by dividing the number of citations received by each firm with average number of citations received by patent cohorts in the same year:

$$Adj\ Citation_{ft} = \frac{\sum_i^{n_{ft}} Citation_i}{\sum_j^{N_t} Citation_j/N_t},$$

where $Adj\ Citation_{ft}$ is the adjusted number of citation received by firm f applied in year t, $Citation_i$is the number of citations received by ith patent from firm f, n_{ft} is the total number of patents applied for by firm f in year t. $Citation_j$ is the number of citations received by jth patent applied in year t, and N_t is the total number of patent applied in year t.

B.2 Time and Tech Class Fixed-Effect Adjustment:

In the Time and Tech Fixed Effect adjustment, the information about different patent classes in different years is also used in the adjustments. The number of patents in different class assigned for each firm in a year is adjusted with total number of successful patents applied in corresponding year and class:

$$Adj\ Patent_{ft} = \sum_{k}^{M} \frac{n_{fkt}}{N_{kt}}$$

where $Adj\ Patent_{ft}$ is the adjusted number of patents for firm f applied in year t, n_{fkt} is the total number of patents for firm f applied in year t in class k, and N_{kt} is the total number of patents applied in year t in class k. M is the total number of patent classes in the data. In our analysis, M is 6 based on the HJT classification.

The same is true for citation adjustment: we divide the number of citations by patents in different classes for each firm in a year by average number of citations for patent cohorts in each patent class in the same year.

$$Adj\ Citation_{ft} = \sum_{k}^{M} \frac{\sum_{i}^{n_{fkt}} Citation_i}{\sum_{j}^{N_{kt}} Citation_j / N_{kt}}$$

where $Adj\ Citation_{ft}$ is the adjusted number of citation received by firm f applied in year t, $Citation_i$is the number of citations received by ith patent from firm f, n_{fkt} is the total number of patent for firm f applied in year t in class k, $Citation_j$ is the number of citations received by jth patent applied in year t, and N_{kt} is the total number of patent applied in year t in class k. M is the total number of patent classes in the data. Again, M is 6 based on the HJT classification.

B.3 "Quasi-Structural" Adjustment:

In this method, HJT (2001) propose a quasi-structural model for citations. We follow them and implement the adjustment as:

$$\frac{C_{kst}}{P_{ks}} = a'_0 a'_s a'_t a'_k \exp[f_k(L)], \qquad (1)$$

or equivalently,

$$\log\left[\frac{C_{kst}}{P_{ks}}\right] = a_0 + a_s + a_t + a_k + f_k(L), \qquad (2)$$

where $f_k(L)$ is given as follows:

$$f_k(L) = \exp(-\beta_{1k}L)(1 - \exp(-\beta_{2k}L)) \qquad (3)$$

P_{ks} is the total patents observed in technological field k in year s, C_{kst} is the total number of citations to patents in year s and technology field k flowing from patents in year t, a'_0, a'_s, a'_t, a'_k are the constant and fixed effects for year s, year t and technology field k, and L is the lag t-s. We also normalized $a_{t=1} = a_{k=1} = a_{s=1} = 0$, and group the cited year effects for five-year interval in the same way as HJT (2001).

We apply their method to both NBER 2006 dataset and our data. We estimate the model based on citation since 1976 to make results comparable across both datasets. The output is presented in Table B1. 'Tech Field Effect', 'Citing Year Effects' and 'Cited Year Effects' correspond to a_k, a_t and a_s in the equation (2). 'Obsolescence by Technology Field' and 'Diffusion by Technology Field' correspond to β_{1k} and β_{2k} in equation (3), where K is the technology field.

Based on HJT (2001), we construct citation deflator in table B2. Column 1 is the total number of granted patents applied each year since 1976. Column 2 is the patent index calculated as total numbers of granted patents applied each year divided by the total granted patents numbers in 1976. Coefficient a_t in full model is used as the citing year coefficient in column 3 in Table B2. Column 4 is calculated by dividing column 3 with column 2, and is defined as pure propensity to cite.

In adjusting total citations made to a firm, we can divide the citations with information from two panels in Table B2:

$$Citing\ Year\ Citation_{ft} = \sum_{i=t}^{Y_{ft}} \frac{Citation_{fti}}{Citing\ Coef_i},$$

where $Citing\ Year\ Citation_{ft}$ is the adjusted number of citations received by firm f for successful patents that were applied in year t, $Citation_{fti}$is the number of citations received by all successful patents applied in t by firm f from citing year i, Y_{ft} is the latest year when citation is created for the successful patent applied in year t by firm f, and $Citing\ Coef_i$ is the citing coefficient calculated in Table B2. Similarly, we can adjust for propensity to cite using:

$$Pro\ Citation_{ft} = \sum_{i=t}^{Y_{ft}} \frac{Citation_{fti}}{Pro\ Coef_i},$$

where $Pro\ Citation_{ft}$ is the adjusted number of citations received by firm f for successful patents that were applied in year t, $Citation_{fti}$is the number of citations received by all successful patents applied in t by firm f from citing year i , Y_{ft} is the latest year when citation is created for the successful patent applied in year t by firm f, and $Pro\ Coef_i$ is the propensity to cite coefficient calculated in Table B2. For brevity, we produce graphs for citation bias calculated based on propensity to cite adjustment in Figure B1 to B3, rather than in the main paper.

Figure B1: Distribution of Firm Citation Bias (Unadjusted and Adjusted) over Time

This figure presents the distribution for citation bias aggregated at the firm-year level from patents granted to public firms from 1976 through 2012. To compute the unadjusted citation bias for each firm-year, we compare the number of citations to all the patents for each firm filed in each application year in our data (i.e., citations in patents granted by 2012 to applications filed by a firm in a given year and granted by 2006) and in the NBER 2006 dataset (i.e., citations in patents granted by 2006 to applications filed by a firm in a given year and granted by 2006). We sum citation bias by year across publicly traded firms. The adjustments use the time fixed effect methodology, the time and technology class fixed effect methodology, and the propensity to cite adjustment using the quasi-structural approach (citing year), with details discussed in Appendix B. The lines for the time fixed effect methodology and the time and technology class fixed effect methodology are almost superimposed due to the scale. Sources: NBER 2006 patent and our datasets.

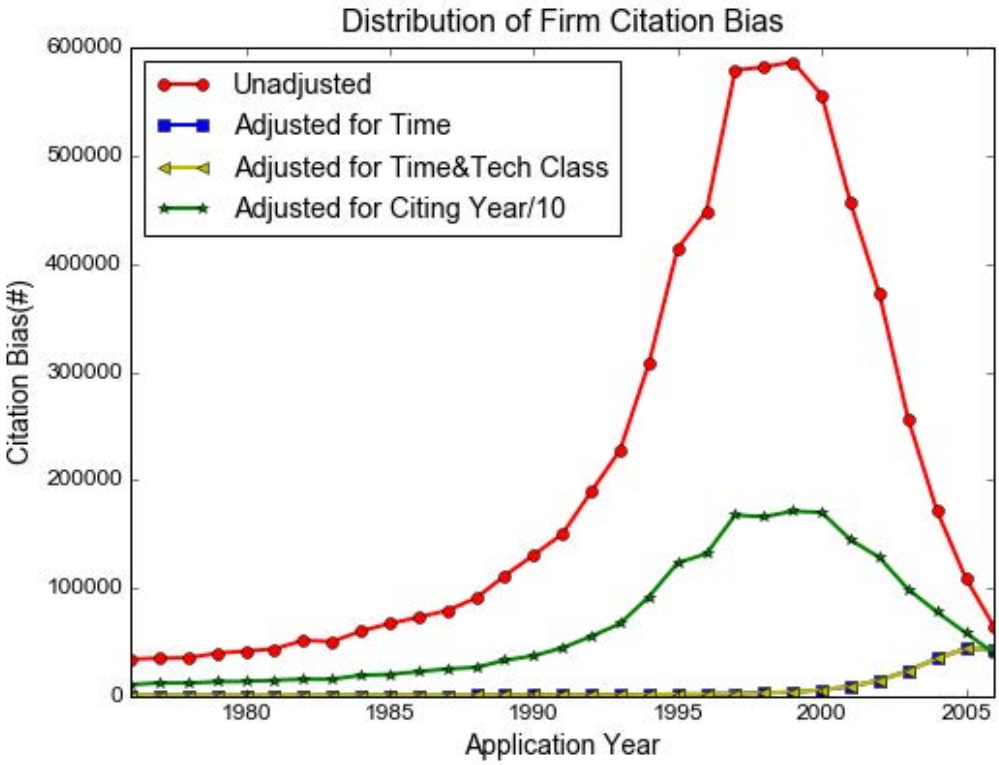


Figure B2: Firm Citation Bias (Adjusted with propensity to cite) across HJT Technology Classes

This figure presents the distribution for citation bias aggregated at the firm-year level from patents granted to public firms from 1976 through 2012 in different HJT technology classes. To compute the citation bias for each firm-year, we compare the number of citations to all the patents for each firm filed in each application year in our data (i.e., citations in patents granted by 2012 to applications filed by a firm in a given year and granted by 2006) and in the NBER 2006 dataset (i.e., citations in patents granted by 2006 to applications filed by a firm in a given year and granted by 2006 A firm is assigned to a particular technology class in a given year based on the modal primary patent class of patents produced by that firm in that year, based on the U.S. patent classification system. We use the propensity to cite adjustment using the quasi-structural approach, with details discussed in Appendix B. We sum citation bias in a technology class across publicly traded firms.

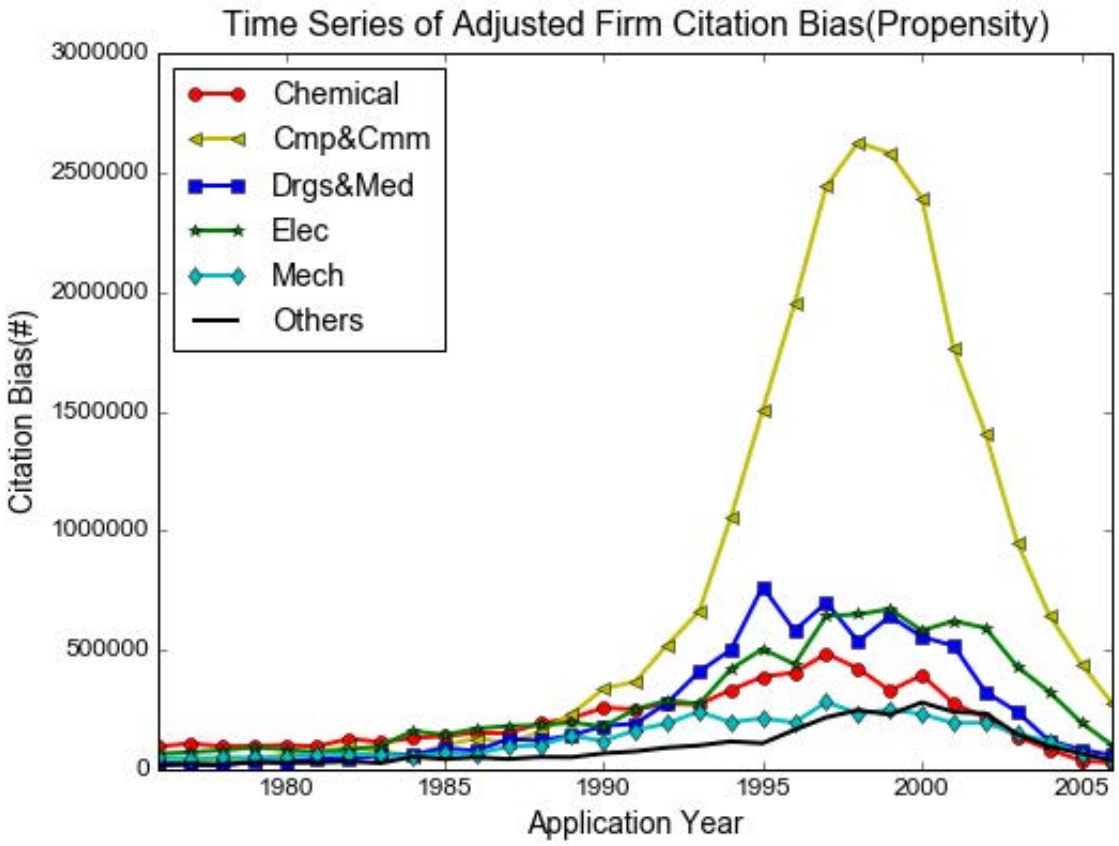


Figure B3: Firm Citation Bias (Adjusted with propensity to cite) across States

This figure presents the distribution for citation bias aggregated at the firm-year level from patents granted to public firms from 1976 through 2012 in different states. To compute the citation bias for each firm-year, we compare the number of citations to all the patents for each firm filed in each application year in our data (i.e., citations in patents granted by 2012 to applications filed by a firm in a given year and granted by 2006) and in the NBER 2006 dataset (i.e., citations in patents granted by 2006 to applications filed by a firm in a given year and granted by 2006). A firm is assigned to a particular state in a given year based on modal state of the assignee across patents granted to the firm at the time of the patent filing. We use the propensity to cite adjustment using the quasi-structural approach, with details discussed in Appendix B. We sum citation bias in each state across publicly traded firms.

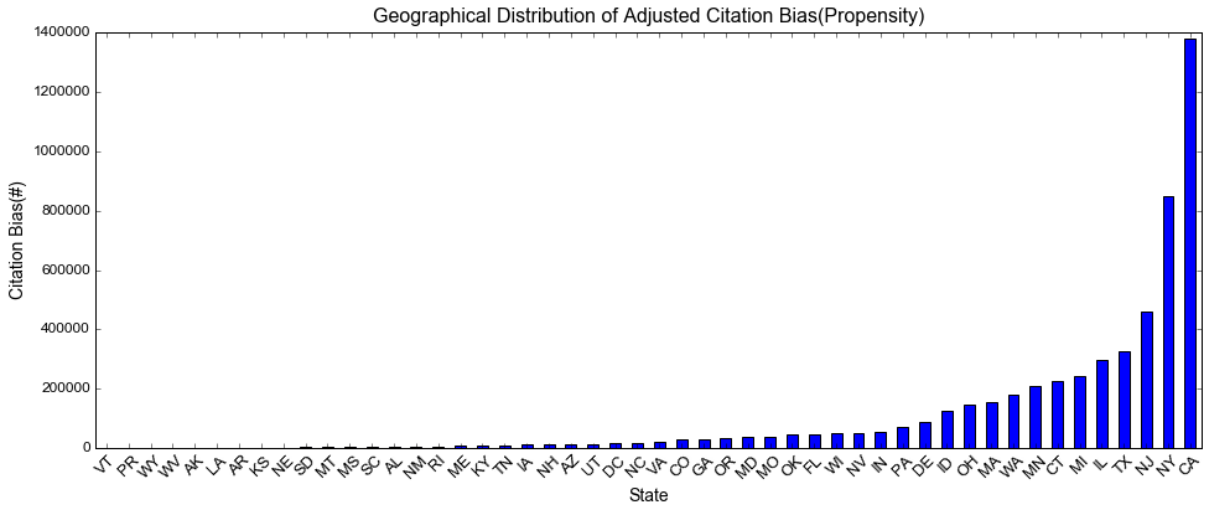


Table B1. "Quasi-Structural" Approach Model

The models are estimated based on "our data" and NBER 2006 data. They use citation since 1976 to make results comparable across both datasets. 'Tech Field Effect', 'Citing Year Effects' and 'Cited Year Effects' correspond to a_k, a_t and a_s in equation (2). 'Obsolescence by Technology Field' and 'Diffusion by Technology Field' correspond to β_{1k} and β_{2k} in equation (3).

	Full Model for "our data"		**Full Model For NBER 2006**	
	Coef.	**SE.**	**Coef.**	**SE.**
Tech Field Effect (Base = others.)				
Chemical	0.403	0.048	0.381	0.085
Computer & Comm.	0.817	0.051	0.764	0.086
Drugs & Medical	0.607	0.042	0.512	0.078
Electrical & Electronics	0.496	0.044	0.518	0.076
Mechanicals	0.396	0.046	0.375	0.074
Citing Year Effects (Base = 1976)				
1977	0.109	0.124	0.110	0.191
1978	0.156	0.102	0.157	0.158
1979	0.187	0.089	0.189	0.138
1980	0.237	0.080	0.240	0.124
1981	0.241	0.073	0.247	0.113
1982	0.243	0.068	0.251	0.105
1983	0.229	0.064	0.237	0.099
1984	0.245	0.060	0.255	0.094
1985	0.257	0.058	0.268	0.090
1986	0.267	0.055	0.280	0.086
1987	0.281	0.053	0.296	0.082
1988	0.296	0.051	0.313	0.080
1989	0.300	0.049	0.318	0.078
1990	0.298	0.048	0.317	0.076
1991	0.287	0.047	0.307	0.074
1992	0.296	0.046	0.318	0.073
1993	0.309	0.045	0.333	0.073
1994	0.340	0.044	0.369	0.072
1995	0.442	0.044	0.455	0.072
1996	0.357	0.043	0.391	0.072
1997	0.410	0.043	0.452	0.072
1998	0.371	0.042	0.409	0.073
1999	0.386	0.042	0.422	0.073
2000	0.417	0.042	0.440	0.074
2001	0.433	0.042	0.447	0.075
2002	0.432	0.042	0.405	0.076
2003	0.413	0.043	0.278	0.078
2004	0.396	0.043	0.136	0.079
2005	0.370	0.043	0.036	0.081
2006	0.366	0.044	0.001	0.082
2007	0.351	0.044		
2008	0.264	0.045		
2009	0.208	0.045		
2010	0.143	0.046		
2011	0.069	0.047		
2012	0.009	0.047		

Table B1. "Quasi-Structural" Approach Model (contd.).

Cited Year Effects(base = 1976-1980)				
1981-85	1.254	0.021	1.225	0.037
1986-90	1.583	0.025	1.518	0.048
1991-95	2.015	0.031	1.806	0.062
1996-20	2.243	0.039	1.775	0.079
2000-05(06)	1.398	0.047	0.563	0.103
2006-12	0.672	0.059		
Beta1: Obsolescence by Technology Field(base=other)				
Chemical	0.013	0.004	0.011	0.008
Computer & Comm.	0.018	0.005	0.021	0.009
Drugs & Medical	-0.011	0.002	-0.002	0.006
Electrical & Electronics	0.010	0.003	0.009	0.007
Mechanicals	0.014	0.004	0.004	0.006
Beta2: Diffusion by Technology Field(base=other)				
Chemical	1.030	0.207	0.989	0.337
Computer & Comm.	0.900	0.180	1.425	0.568
Drugs & Medical	0.668	0.112	0.983	0.319
Electrical & Electronics	1.358	0.297	1.828	0.828
Mechanicals	1.265	0.272	1.850	0.838
Number of Observation	4218		2976	
Adjusted R-Squared	0.87		0.85	

Table B2. Potential "Deflators" for Citing Patent Totals.

Panel A. Our data 1976-2012				
Application Year	(1) Total Patents	(2) Index of Patent Total **Total patents/65827**	(3) Citing Year Coefficient	(4) Pure Propensity to Cite Effect **(3)/(2)**
1976	**65827**	1	1	1
1977	65998	1.003	0.109	0.109
1978	65615	0.997	0.156	0.156
1979	65731	0.999	0.187	0.187
1980	66499	1.010	0.237	0.234
1981	63930	0.971	0.241	0.248
1982	65040	0.988	0.243	0.246
1983	61586	0.936	0.229	0.245
1984	67093	1.019	0.245	0.241
1985	71477	1.086	0.257	0.237
1986	75118	1.141	0.267	0.234
1987	81518	1.238	0.281	0.227
1988	90191	1.370	0.296	0.216
1989	96167	1.461	0.300	0.206
1990	99400	1.510	0.298	0.198
1991	100290	1.524	0.287	0.189
1992	103945	1.579	0.296	0.187
1993	108353	1.646	0.309	0.188
1994	123361	1.874	0.340	0.182
1995	144697	2.198	0.442	0.201
1996	144885	2.201	0.357	0.162
1997	169693	2.578	0.410	0.159
1998	168890	2.566	0.371	0.145
1999	181878	2.763	0.386	0.140
2000	199583	3.032	0.417	0.137
2001	211406	3.212	0.433	0.135
2002	205214	3.117	0.432	0.138
2003	196149	2.980	0.413	0.139
2004	194363	2.953	0.396	0.134
2005	192391	2.923	0.370	0.127
2006	187945	2.855	0.366	0.128
2007	178969	2.719	0.351	0.129
2008	154136	2.342	0.264	0.113
2009	112333	1.706	0.208	0.122
2010	64512	0.980	0.143	0.146
2011	25160	0.382	0.069	0.179
2012	2926	0.044	0.009	0.197

Table B2. Potential "Deflators" for Citing Patent Totals (contd.).

Panel B: NBER data 1976-2006				
Application Year	(1) Total Patents	(2) Index of Patent Total **Total Patents/65813**	(3) Citing Year Coefficient	(4) Pure Propensity to Cite Effect **(3)/(2)**
1976	**65813**	1	1	1
1977	65998	1.003	0.110	0.110
1978	65610	0.997	0.157	0.157
1979	65728	0.999	0.189	0.189
1980	66505	1.011	0.240	0.238
1981	63936	0.971	0.247	0.254
1982	65038	0.988	0.251	0.254
1983	61585	0.936	0.237	0.254
1984	67096	1.019	0.255	0.250
1985	71484	1.086	0.268	0.247
1986	75122	1.141	0.280	0.245
1987	81520	1.239	0.296	0.239
1988	90218	1.371	0.313	0.228
1989	96188	1.462	0.318	0.217
1990	99412	1.511	0.317	0.210
1991	100298	1.524	0.307	0.201
1992	103949	1.579	0.318	0.201
1993	108342	1.646	0.333	0.202
1994	123325	1.874	0.369	0.197
1995	144523	2.196	0.455	0.207
1996	144792	2.200	0.391	0.178
1997	169360	2.573	0.452	0.176
1998	167826	2.550	0.409	0.160
1999	178560	2.713	0.422	0.155
2000	189530	2.880	0.440	0.153
2001	191722	2.913	0.447	0.154
2002	169599	2.577	0.405	0.157
2003	119770	1.820	0.278	0.153
2004	62396	0.948	0.136	0.143
2005	17177	0.261	0.036	0.139
2006	1039	0.016	0.001	0.094

Appendix C: Construction of Firm Characteristics

In this section, we provide details on the definitions and construction of firm characteristics used in the regressions. The year referenced below refers to fiscal year.

D.1 Valuation

Log_Size

$Size_t$ is defined as the log of market capitalization, which is calculated as the product of Common Shares Outstanding ($CSHO_t$) and Fiscal Annual Close Price ($PRCC\_F_t$) in year t, both from Compustat. Log_Size is defined as natural log of $Size_t$ in the regressions because of the skewness in distribution.

Log_M2B

Market to Book ratio ($M2B_t$) is calculated as Common Shares Outstanding ($CSHO_t$)*Fiscal Annual Close Price ($PRCC\_F_t$)/(Total Assets (AT_t) – Total Liabilities (LT_t)) in year t, where MKVALT, AT and LT come from Compustat. Log_M2B is defined as the natural log of $M2B_t$ in the regressions because of the skewness in distribution.

Log_Cash2Asset

Cash Flow to Total Asset ($Cash2Asset_t$) is calculated as Cash (CH_t)/Total Assets (AT_t) in year t, where CH and AT come from Compustat. Log_cash2asset is defined as the natural log of $Cash2Asset_t$ in the regressions because of the skewness in distribution.

D.2 Investment

Log_ RD2Sale

R&D Investment to Sales ($RD2Sale_t$) is calculated as Research and Development Expense (XRD_t)/ net sales ($SALES_t$) in year t, where XRD and SALE come from Compustat. Log_RD2Sale is defined as the natural log of $RD2Sale_t$ in the regressions because of the skewness in distribution.

Investment to Asset Ratio (IA)

Similar to Chen, et al. (2010), the IA is calculated as the annual change in Property, Plant and Equipment ($PPEGT_t$ -$PPEGT_{t-1}$) plus annual change in Inventories ($INVT_t$-$INVT_{t-1}$) divided by lagged Total Assets (AT_{t-1}) in year t, where PPEGT, INVT and AT come from Compustat.

D3.Earnings

Return on Assets (ROA)

Return on Assets (ROA) is defined as the ratio of Income before Extraordinary Items (IB_t) over Total Assets (AT_t) in year t, where IB and AT come from Compustat.

Return on Equity (ROE)

Return on Equity (ROE) is defined as the ratio of Net Income (NI_t) over Stockholders' Equity (SEQ_t) in year t, where both variables come from Compustat.

Sales Growth (SG)

Sales growth (SG) is the defined as the percent change in net sales ($SALE_t$–$SALE_{t-1}$)/$SALE_{t-1}$ in year t, where SALE is from Compustat.

D4. Financial Distress

Log_LEV

Market Leverage (LEV) is the ratio of Total Assets (AT_t) over the market value of equity ($CSHO_t$*$PRCC\_F_t$) in year t, where CSHO and PRCC_F come from Compustat. Log_LEV is defined as the natural log of LEV_t in the regressions because of the skewness in distribution.

Log_Spread

Bid-Ask Spread (Spread) is defined as annual mean value of ratio (Ask or High ($ASKHI_t$) - Bid or Low ($BIDLO_t$))/Ask or High($ASKHI_t$), where ASKHI and BIDLO come from CRSP. Log_Spread is defined as the natural log of Spread in the regressions because of the skewness in distribution.

D5. External Financing

Net Stock Issues (NSI)

Net Stock Issues (NSI) is defined as log of ratio of split-adjusted shares outstanding at fiscal year-ends t and t-1, calculated as log(Common Shares Outstanding ($CSHO_t$)*Cumulative Adjustment Factor by Ex-Date-Fiscal ($ADJEX\_F_t$)/Common Shares Outstanding ($CSHO_{t-1}$)*Cumulative Adjustment Factor by Ex-Date-Fiscal ($ADJEX\_F_{t-1}$)), where CSHO and ADJEX_F come from Compustat.

D6. Patent and Citation Information

Log(class patents)

Same Class Patent is defined as the total number of granted patents in the same class as the modal class of the firm's patents in year t. Log(class patents) is the natural log of same class patent.

Log(state patents)

Same State Patent is defined as the total number of granted patents in the same state as the modal state of the assignee on firm's patents in year t. Log(state patents) is the natural log of same state patent.

Log(class cites)

Same Class Citation is defined as the total number of citations to granted patents in the same class as the modal class of the firm's patents in year t. Log(class cites) is the natural log of same class citation.

Log(state cites)

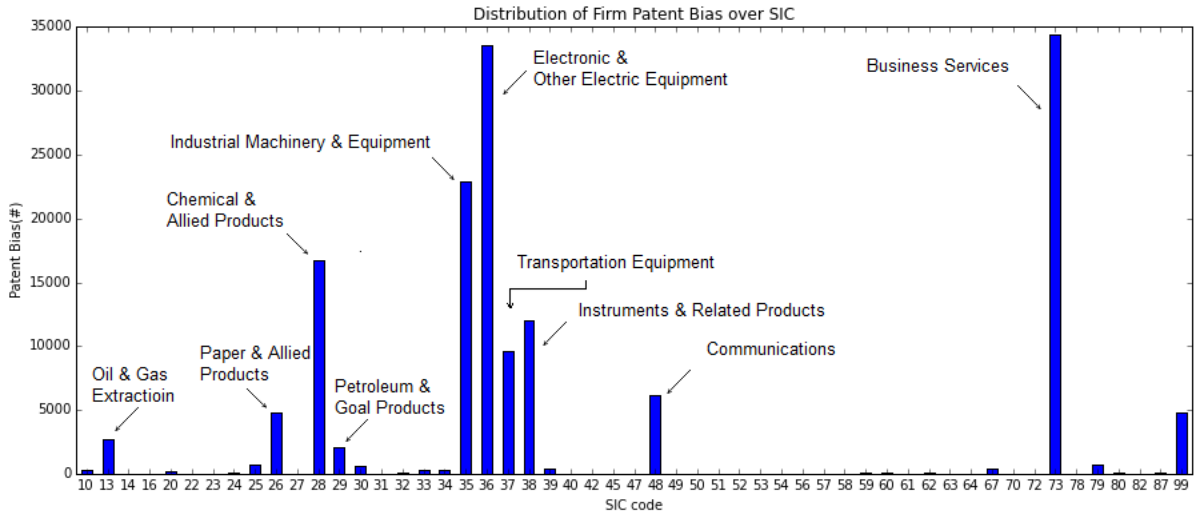
Same State Patent is defined as the total number of citations to granted patents in the same state as the modal state of the assignee on firm's patents in year t. Log(state cites) is the natural log of same state citation.

Appendix D: Firm Patent and Citation Bias across Industries (SIC)

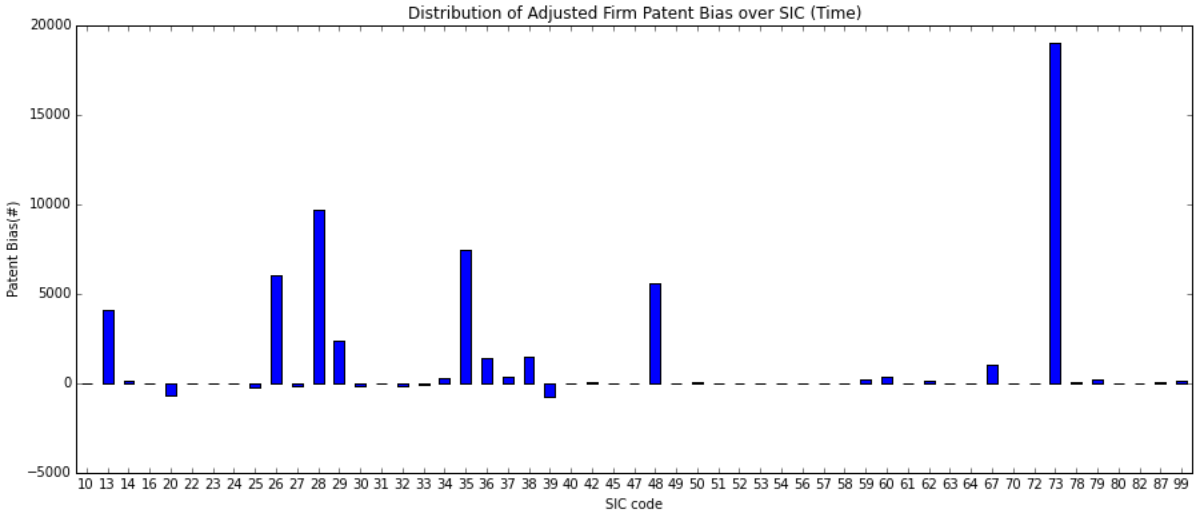
D.1 Firm Patent Bias across Industries (SIC)

Figure D1 shows the patent bias at the firm level across different industries, where industries are defined using two-digit SIC codes. Firms are assigned to industries, as discussed in Section 3. Similar to Figure 13, patent bias is found to be present in the Electronic and Industrial Equipment, Chemical, and Business Service (which includes advertising, computer programming, data processing, and other computer related services) industries. Adjustments in Panel B and C alleviate, but do not erase, the bias within certain industries.

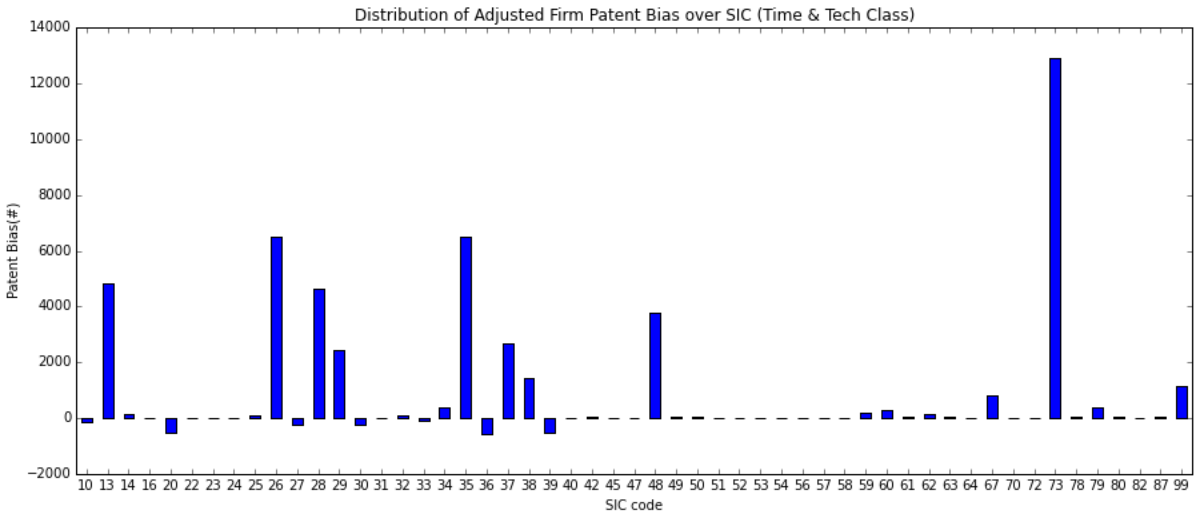
Figure D1. Firm Patent Bias (Unadjusted and Adjusted) across Industries (SIC)



(a)



(b)

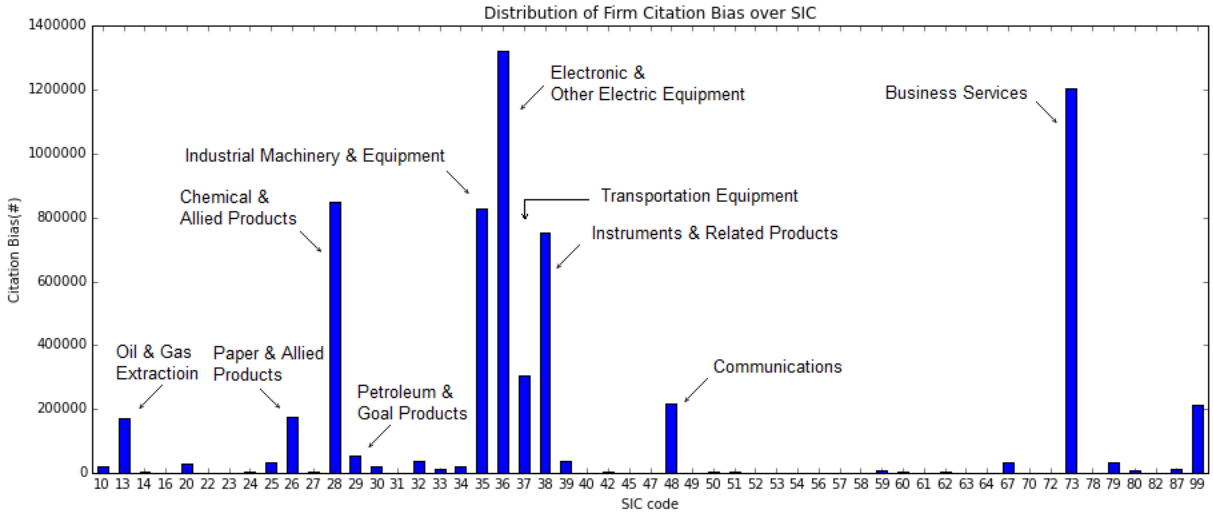


(c)

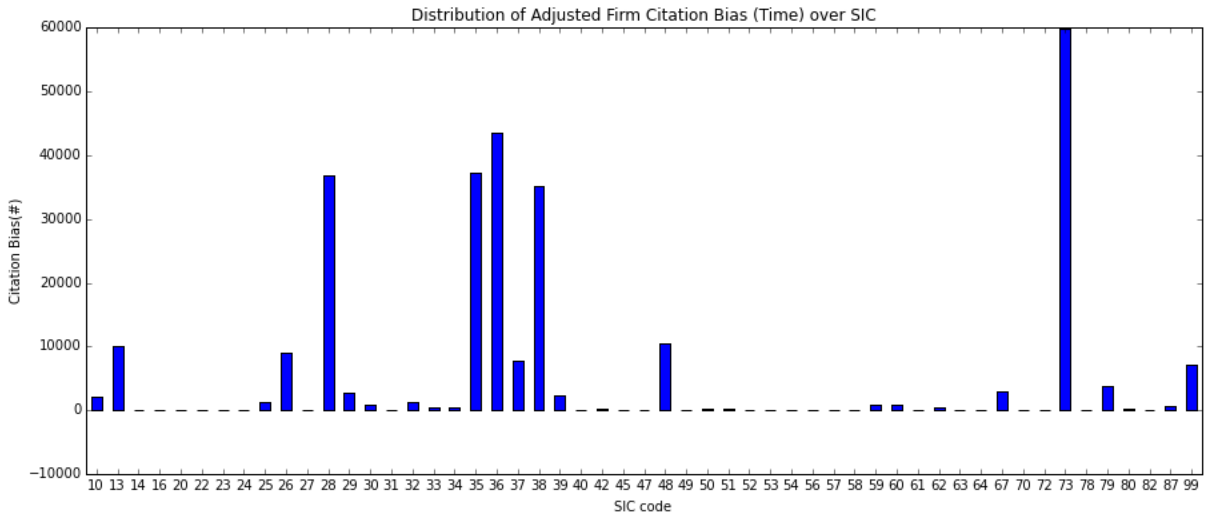
D.2 Firm Citation Bias across Industries (SIC)

Figure D2 shows the citation bias at the firm level across different industries. Again, citations bias is present in some industries, and remains even after adjustments.

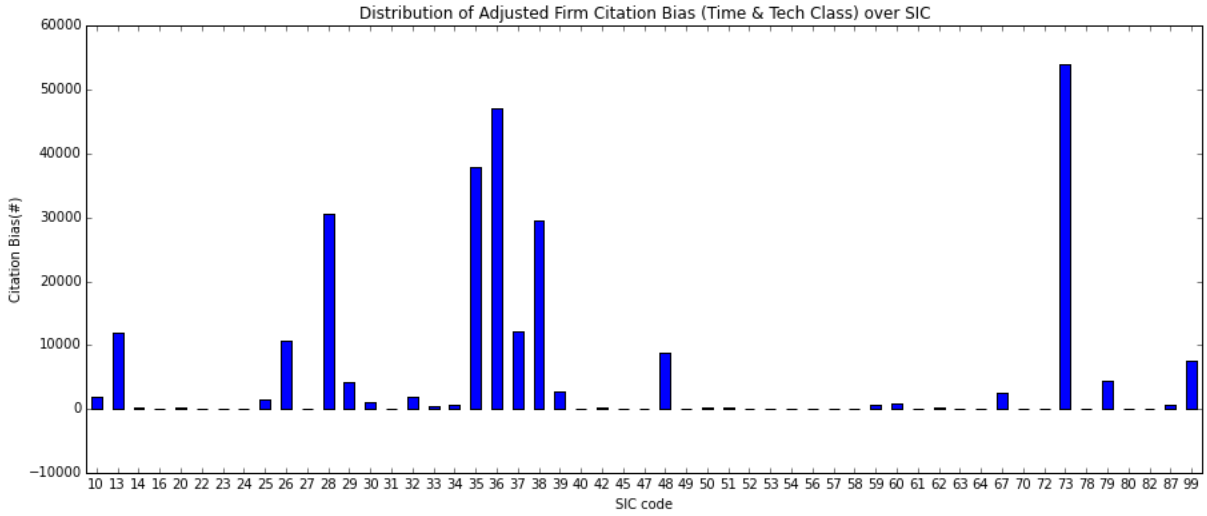
Figure D2. Firm Citation Bias (Unadjusted and Adjusted) across Industries (SIC)



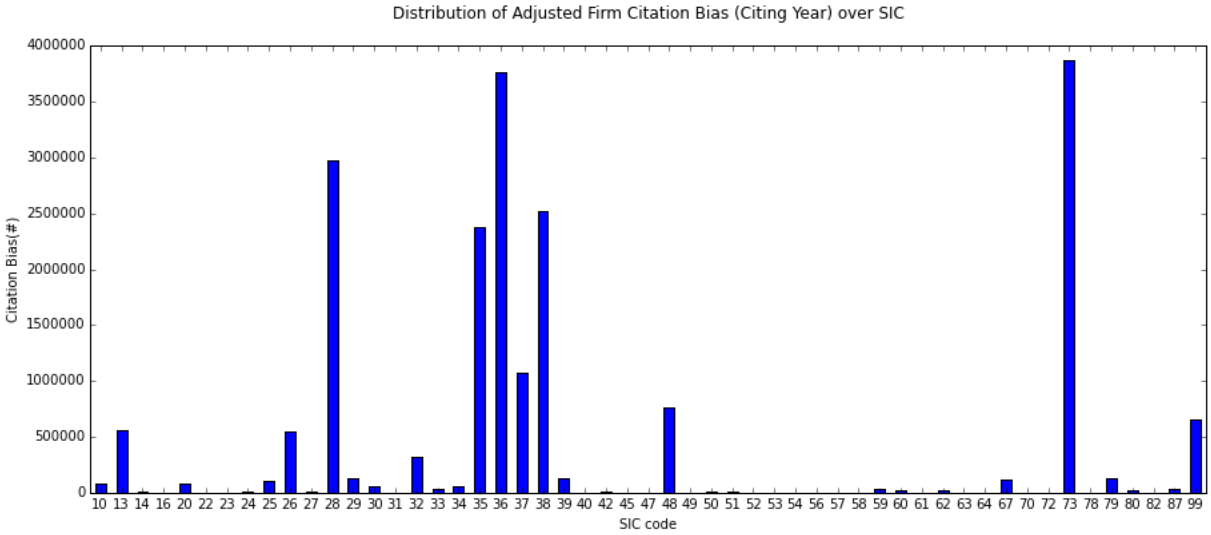
(a)


Distribution of Adjusted Firm Citation Bias (Time) over SIC
Citation Bias(#)
SIC code

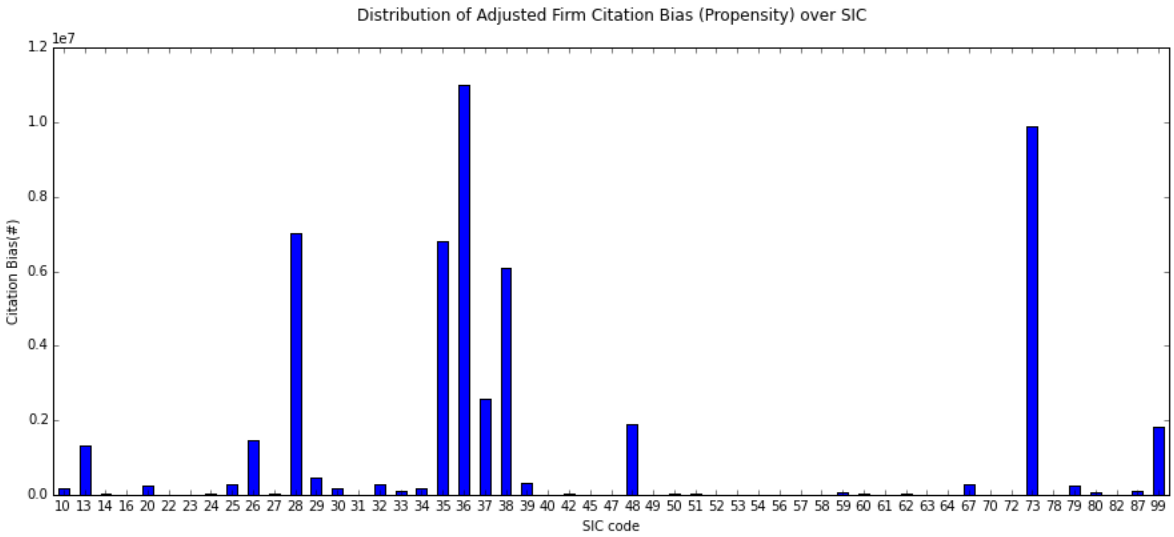
(b)


Distribution of Adjusted Firm Citation Bias (Time & Tech Class) over SIC
Citation Bias(#)
SIC code

(c)


Distribution of Adjusted Firm Citation Bias (Citing Year) over SIC
Citation Bias(#)
SIC code

(d)

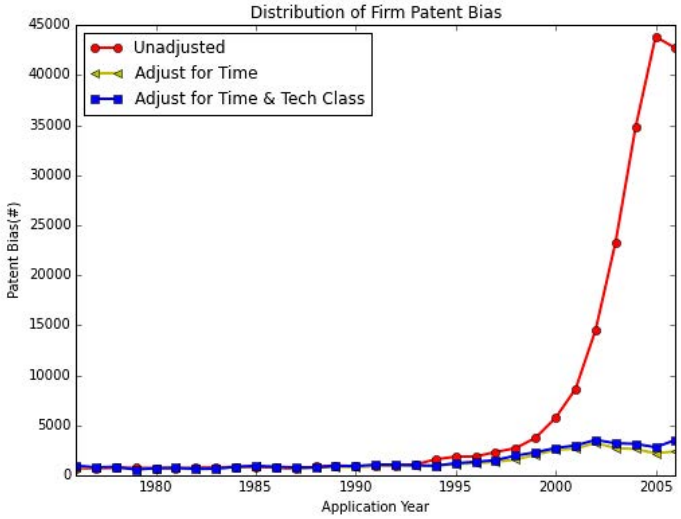

Distribution of Adjusted Firm Citation Bias (Propensity) over SIC
1e7
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Citation Bias(#)
10 13 14 16 20 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 42 45 47 48 49 50 51 52 53 54 56 57 58 59 60 61 62 63 64 67 70 72 73 78 79 80 82 87 99
SIC code

(e)

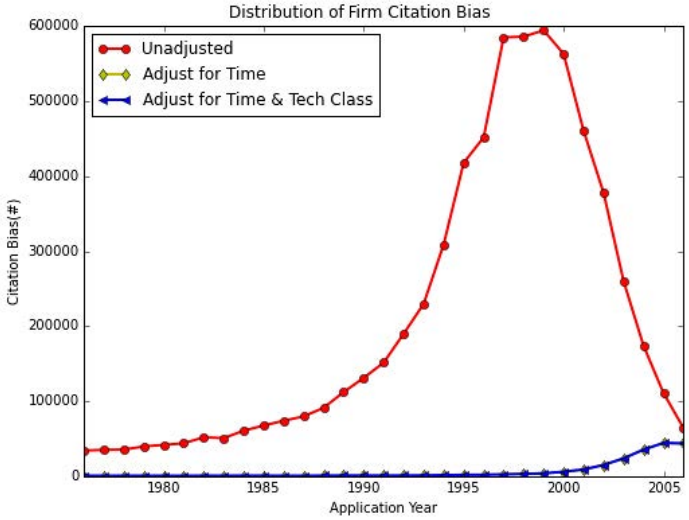
Appendix E: Other Adjustments

The figures below show patent and citation bias (both unadjusted and adjusted) computed using the methodology based on Jaffe and Rassenfosse (2016), where we consider only patents by publicly traded firms as the "group of patents." The figures are discussed in Section 3 in the paper.

Figure E1. Distribution of Firm Patent Bias and Citation Bias

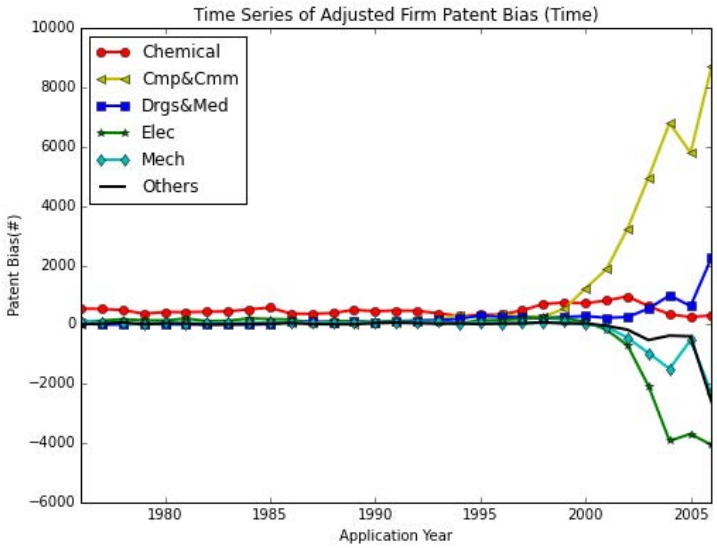


(a)

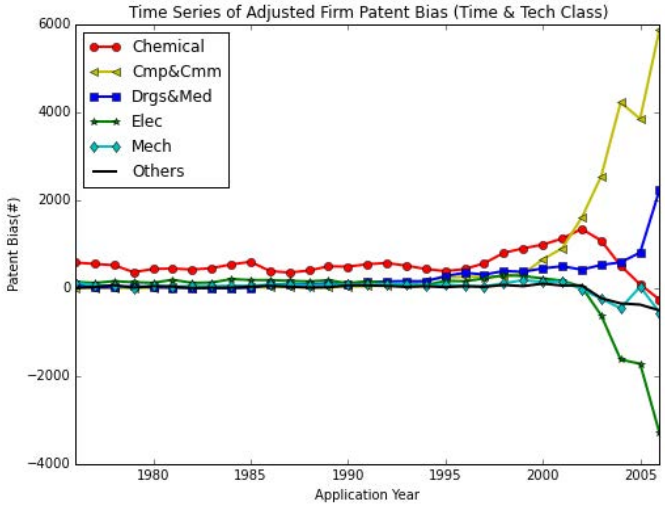


(b)

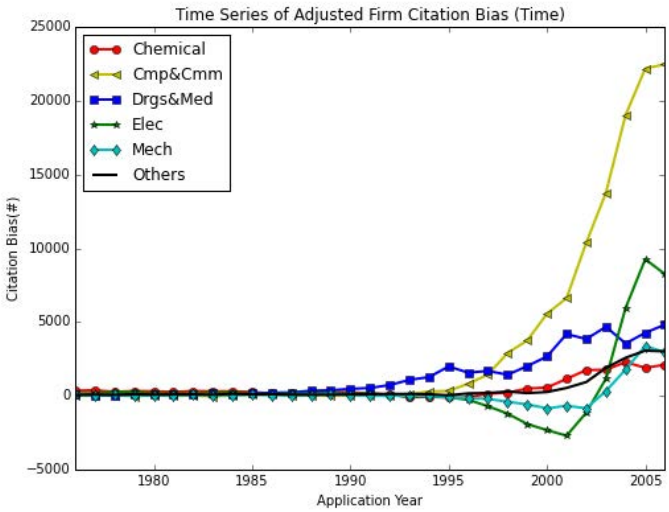
Figure E2. Distribution of Firm Patent Bias and Citation Bias across Time



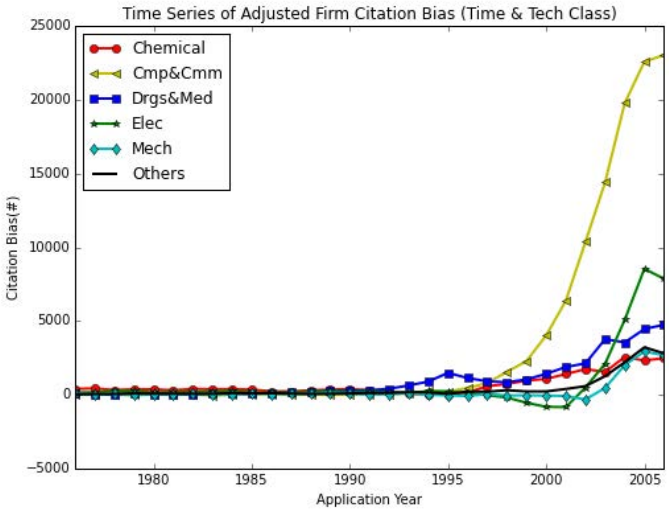
(a)



(b)

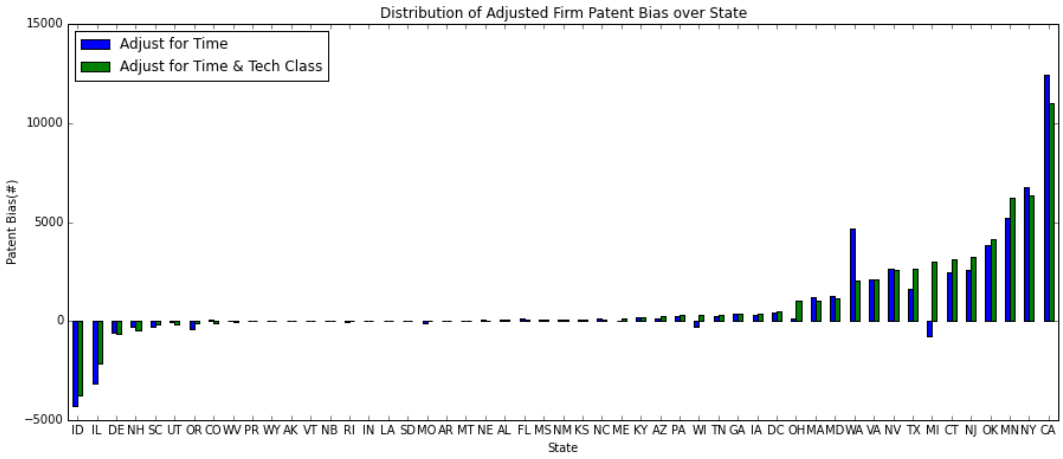

Time Series of Adjusted Firm Citation Bias (Time)
Chemical
Cmp&Cmm
Drgs&Med
Elec
Mech
Others
Citation Bias(#)
Application Year
25000
20000
15000
10000
5000
0
−5000
1980
1985
1990
1995
2000
2005

(c)

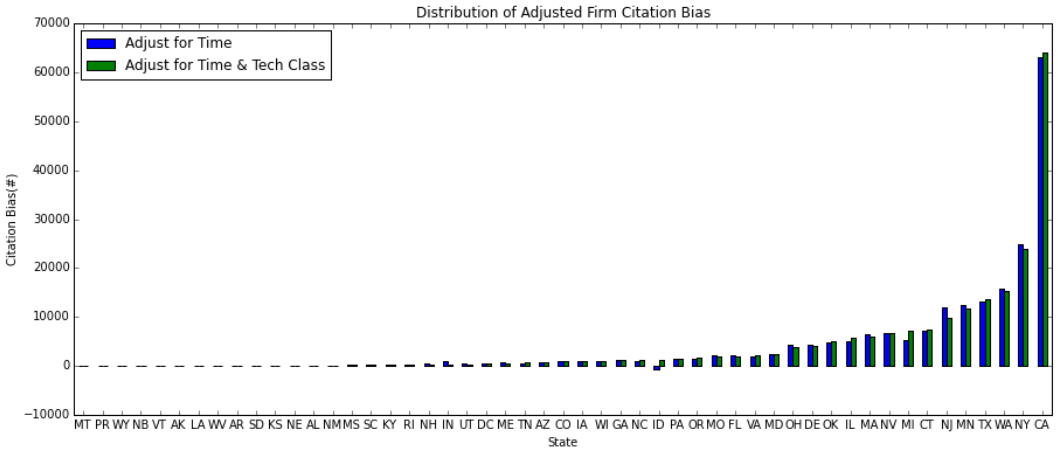

Time Series of Adjusted Firm Citation Bias (Time & Tech Class)
Chemical
Cmp&Cmm
Drgs&Med
Elec
Mech
Others
Citation Bias(#)
Application Year
25000
20000
15000
10000
5000
0
−5000
1980
1985
1990
1995
2000
2005

(d)

Figure E3. Distribution of Firm Patent Bias and Citation Bias across Regions

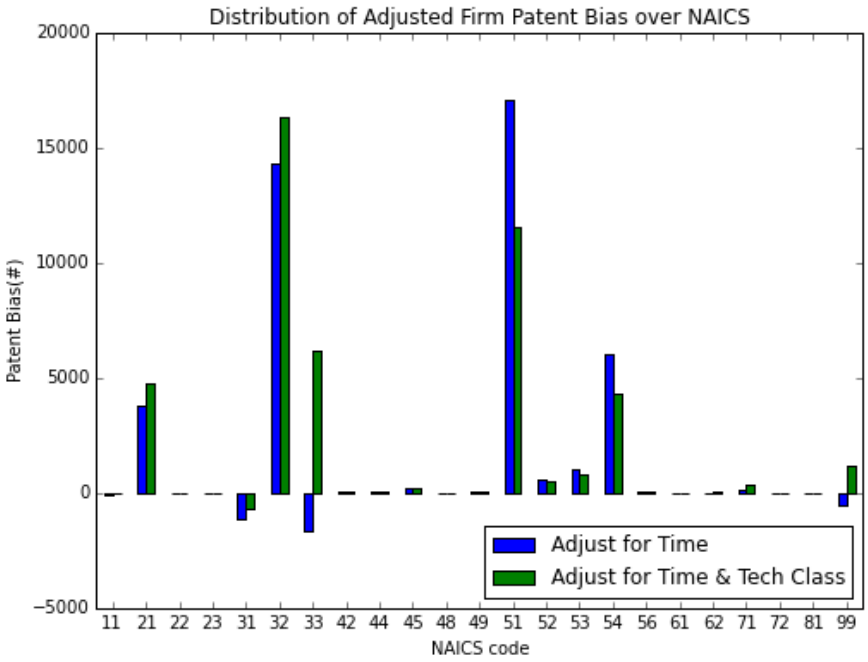


(a)

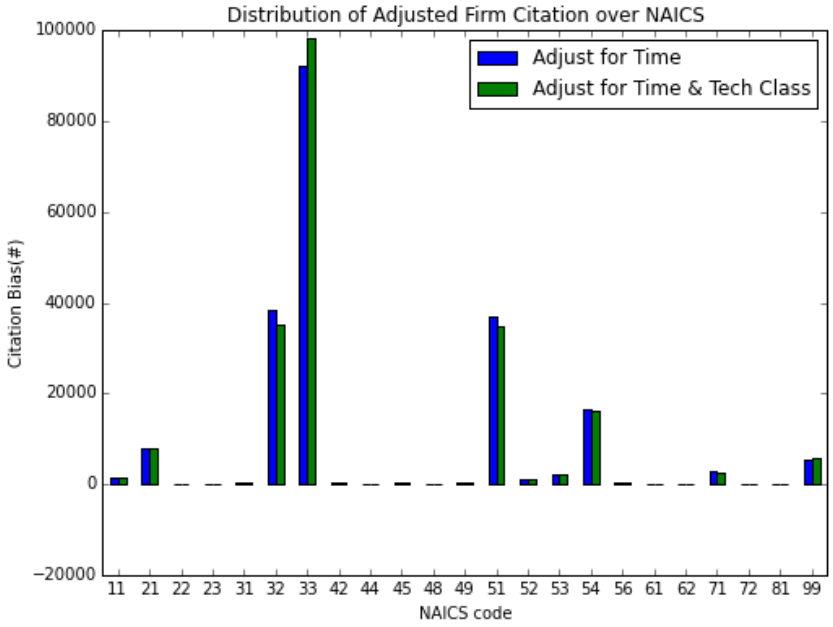


(b)

Figure E4. Distribution of Firm Patent Bias and Citation Bias across Industries (NAICS)



(a)



(b)

Table E1: Patent Bias and Firm Characteristics

This table presents OLS regressions relating time fixed effect method adjusted patent bias at the firm level with different firm characteristics. The dependent variable is the time fixed effect adjusted patent bias of a given firm in that year for years 1976-2006 (columns 1-3) and for subsamples, 1976-1996 (columns 4-6) and 1997-2006 (columns 7-9). The dependent variable is computed as the difference in log of one plus number of successful patents filed by a firm in a given year as of 2012 ("our data") and log of one plus number of successful patents filed by that firm in the same year as of the end of sample in the NBER 2006 dataset. Logs are taken to account for skewness in patenting activity. Control variables and their construction are described in Appendix C. Details about time fixed effect adjustment can be found in Appendix B, noting that here we follow Jaffe and Rassenfosse (2016) and consider only patents by publicly traded firms. Robust t-tests are reported in the parenthesis. Sources: NBER 2006 patent and our datasets.

Panel A: Adjusted Patent Bias (Time)

	Adjusted Patent Bias (Time) (1976-2006)			Adjusted Patent Bias (Time) (1976-1996)			Adjusted Patent Bias (Time) (1997-2006)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log_Size	0.0128***	0.00644	0.0137	-0.0145**	-0.0178***	0.0127*	0.0353***	0.0277***	0.0739***
	(2.61)	(1.28)	(1.28)	(-2.46)	(-2.95)	(1.68)	(4.85)	(3.72)	(3.44)
Log_M2B	0.0345**	0.0413**	0.0314*	0.0995***	0.0968***	0.0271**	-0.00952	-0.00298	-0.00677
	(2.00)	(2.38)	(1.77)	(3.98)	(3.85)	(2.04)	(-0.41)	(-0.13)	(-0.26)
Log_RD2Sale	0.0126*	0.0283***	0.0288**	-0.0105	-0.0127	-0.00101	0.0270***	0.0565***	0.0591***
	(1.73)	(3.73)	(2.18)	(-1.14)	(-1.33)	(-0.12)	(2.62)	(5.23)	(2.58)
Log_Cash2Asset	0.0111	0.0105	0.00149	0.0275***	0.0298***	-0.00283	0.00334	-0.000461	0.00216
	(1.62)	(1.53)	(0.21)	(3.58)	(3.90)	(-0.70)	(0.32)	(-0.04)	(0.17)
Log_LEV	0.0229	0.0222	0.0207	0.0713***	0.0662***	0.0321**	-0.00131	-0.00651	0.0804**
	(1.33)	(1.28)	(0.98)	(3.04)	(2.81)	(2.07)	(-0.06)	(-0.27)	(2.40)
IA	-0.0441	-0.0364	-0.00800	0.0482	0.0761	0.0207	-0.0558	-0.0555	-0.0452
	(-0.52)	(-0.43)	(-0.10)	(0.50)	(0.78)	(0.47)	(-0.43)	(-0.43)	(-0.35)
ROA	-0.00622	0.0631	0.0705	-0.430***	-0.425***	0.109	0.151	0.274***	0.0831
	(-0.08)	(0.82)	(0.85)	(-3.53)	(-3.48)	(1.61)	(1.51)	(2.72)	(0.71)
ROE	0.0537	0.0600	0.0950**	0.148**	0.145**	-0.00806	-0.00141	-0.000455	0.0719
	(1.31)	(1.47)	(2.51)	(2.33)	(2.28)	(-0.29)	(-0.03)	(-0.01)	(1.35)
SG	-0.0194	-0.0232	-0.0750***	-0.00852	-0.0133	-0.0226	-0.0272	-0.0319	-0.0774**
	(-0.72)	(-0.86)	(-3.03)	(-0.23)	(-0.35)	(-1.32)	(-0.75)	(-0.89)	(-2.19)
NSI	0.0901	0.0508	-0.152*	-0.0294	-0.0672	-0.0901	0.122	0.0732	-0.0848
	(0.91)	(0.51)	(-1.67)	(-0.23)	(-0.53)	(-1.60)	(0.88)	(0.53)	(-0.62)
Log_Spread	-0.0352	-0.0268	-0.0839***	-0.0258	-0.0190	-0.00890	-0.0689*	-0.0531	-0.142***
	(-1.58)	(-1.20)	(-3.71)	(-1.06)	(-0.79)	(-0.78)	(-1.94)	(-1.48)	(-3.39)
Log(state patents)	-0.0108**	-0.0103**	-0.00518	-0.00753	-0.00849	0.0246***	-0.00995	-0.00954	0.00269
	(-2.44)	(-2.31)	(-0.60)	(-1.31)	(-1.48)	(4.47)	(-1.60)	(-1.53)	(0.16)
Log(class patents)	0.116***	0.120***	0.0534***	0.0808**	0.0868**	0.0231*	-0.0625	-0.0823	-0.0165
	(4.72)	(4.94)	(2.92)	(2.14)	(2.33)	(1.86)	(-0.69)	(-0.92)	(-0.49)
Observation	14503	14503	14503	5908	5908	5908	8595	8595	8595
R^2	0.218	0.227	0.516	0.056	0.076	0.877	0.250	0.265	0.548
Firm Fixed Effect			Yes			Yes			Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class Fixed Effect	Yes	Yes		Yes	Yes		Yes	Yes	
NAICS Fixed Effect	Yes			Yes			Yes		

Table E1: Patent Bias and Firm Characteristics (contd.)

This table presents OLS regressions relating time and tech class fixed effect method adjusted patent bias at the firm level with different firm characteristics. The dependent variable is the time and tech class fixed effect adjusted patent bias of a given firm in that year for years 1976-2006 (columns 1-3) and for subsamples, 1976-1996 (columns 4-6) and 1997-2006 (columns 7-9). The dependent variable is computed as the difference in log of one plus number of successful patents filed by a firm in a given year as of 2012 ("our data") and log of one plus number of successful patents filed by that firm in the same year as of the end of sample in the NBER 2006 dataset. Logs are taken to account for skewness in patenting activity. Control variables and their construction are described in Appendix C. Details about time and tech class fixed effect adjustment can be found in Appendix B, noting that we follow Jaffe and Rassenfosse (2016) and consider only patents by publicly traded firms. Robust t-tests are reported in the parenthesis. Sources: NBER 2006 patent and our datasets.

Panel B: Adjusted Patent Bias (Time and Tech Class)

	Adjusted Patent Bias (Time & Tech Class) (1976-2006)			Adjusted Patent Bias (Time & Tech Class) (1997-2006)			Adjusted Patent Bias (Time & Tech Class) (1997-2006)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log_Size	0.0185***	0.0125**	0.0161	-0.0139**	-0.0171***	0.00849	0.0435***	0.0367***	0.0700***
	(3.75)	(2.46)	(1.51)	(-2.34)	(-2.82)	(1.11)	(5.96)	(4.90)	(3.25)
Log_M2B	0.0370**	0.0445**	0.0322*	0.104***	0.101***	0.0303**	-0.00638	0.000975	-0.00299
	(2.14)	(2.56)	(1.82)	(4.13)	(3.99)	(2.25)	(-0.27)	(0.04)	(-0.11)
Log_RD2Sale	0.0134*	0.0287***	0.0310**	-0.00699	-0.00970	0.00432	0.0266**	0.0553***	0.0593***
	(1.83)	(3.77)	(2.34)	(-0.75)	(-1.01)	(0.49)	(2.57)	(5.09)	(2.59)
Log_Cash2Asset	0.0141**	0.0138**	0.00470	0.0270***	0.0295***	-0.00340	0.00768	0.00424	0.00676
	(2.06)	(2.01)	(0.65)	(3.50)	(3.82)	(-0.84)	(0.73)	(0.41)	(0.53)
Log_LEV	0.0339**	0.0342**	0.0212	0.0800***	0.0747***	0.0337**	0.0133	0.00990	0.0753**
	(1.96)	(1.97)	(1.00)	(3.38)	(3.15)	(2.15)	(0.56)	(0.41)	(2.25)
IA	-0.0475	-0.0407	-0.0120	0.0367	0.0685	0.0177	-0.0655	-0.0705	-0.0517
	(-0.55)	(-0.47)	(-0.15)	(0.37)	(0.70)	(0.40)	(-0.50)	(-0.54)	(-0.40)
ROA	0.00612	0.0735	0.0747	-0.408***	-0.407***	0.111	0.151	0.271***	0.0887
	(0.08)	(0.95)	(0.90)	(-3.32)	(-3.31)	(1.62)	(1.50)	(2.69)	(0.75)
ROE	0.0500	0.0573	0.0851**	0.154**	0.151**	-0.00980	-0.00536	-0.00391	0.0683
	(1.21)	(1.39)	(2.25)	(2.41)	(2.36)	(-0.34)	(-0.10)	(-0.07)	(1.28)
SG	-0.0173	-0.0213	-0.0655***	-0.0106	-0.0165	-0.0224	-0.0233	-0.0283	-0.0643*
	(-0.64)	(-0.79)	(-2.65)	(-0.28)	(-0.44)	(-1.30)	(-0.64)	(-0.78)	(-1.82)
NSI	0.101	0.0637	-0.152*	-0.0335	-0.0722	-0.0812	0.146	0.100	-0.0887
	(1.01)	(0.64)	(-1.68)	(-0.26)	(-0.57)	(-1.43)	(1.05)	(0.72)	(-0.65)
Log_Spread	-0.0237	-0.0153	-0.0774***	-0.0228	-0.0163	-0.0121	-0.0541	-0.0376	-0.139***
	(-1.06)	(-0.68)	(-3.42)	(-0.93)	(-0.67)	(-1.05)	(-1.52)	(-1.04)	(-3.33)
Log(state patents)	-0.0129***	-0.0127***	-0.0102	-0.00754	-0.00859	0.0247***	-0.0127**	-0.0127**	-0.00803
	(-2.90)	(-2.83)	(-1.18)	(-1.31)	(-1.49)	(4.43)	(-2.02)	(-2.03)	(-0.49)
Log(class patents)	0.0637***	0.0683***	0.0138	0.0574	0.0632*	0.0160	-0.0525	-0.0708	-0.0579*
	(2.59)	(2.79)	(0.75)	(1.51)	(1.68)	(1.27)	(-0.58)	(-0.78)	(-1.71)
Observation	14503	14503	14503	5908	5908	5908	8595	8595	8595
R^2	0.221	0.230	0.522	0.054	0.075	0.876	0.249	0.263	0.549
Firm Fixed Effect			Yes			Yes			Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class Fixed Effect	Yes	Yes		Yes	Yes		Yes	Yes	
NAICS Fixed Effect	Yes			Yes			Yes		

Table E2: Citation Bias and Firm Characteristics

This table presents OLS regressions relating time fixed effect adjusted citation bias at the firm level with different firm characteristics. The dependent variable is the time fixed effect adjusted citation bias of a given firm in that year for years 1976-2006 (columns 1-3) and for subsamples, 1976-1990 (columns 4-6) and 1991-2006 (columns 7-9 The dependent variable is computed as the difference in log of one plus number of citations to all patents of a firm applied for in a given year and granted by 2006 in our data and log of one plus number of citations to the same set of successful patents of that firm in the same application year in the NBER 2006 dataset. Restricting the successful patents from our data to only those that are granted by 2006 allows for comparison with successful patents in the NBER 2006 data. Logs are taken to account for skewness in citation activity. Control variables and their construction are described in Appendix C. Details about time fixed effect adjustment can be found in Appendix B, noting that we follow Jaffe and Rassenfosse (2016) and consider only patents by publicly traded firms. Robust t-tests are reported in the parenthesis. Sources: NBER 2006 patent and our datasets.

Panel A: Adjusted Citation Bias (Time)

	Adjusted Citation Bias (Time) (1976-2006)			Adjusted Citation Bias (Time) (1976-1990)			Adjusted Citation Bias (Time) (1991-2006)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log_Size	0.0788***	0.0761***	0.00847	-0.0139*	-0.0107	0.0230*	0.110***	0.107***	0.0242
	(13.62)	(12.80)	(0.65)	(-1.75)	(-1.30)	(1.93)	(15.55)	(14.73)	(1.35)
Log_M2B	0.0606***	0.0667***	0.0604***	0.133***	0.129***	0.00135	0.0136	0.0218	0.0239
	(2.98)	(3.26)	(2.80)	(3.34)	(3.21)	(0.06)	(0.58)	(0.93)	(0.91)
Log_RD2Sale	0.0290***	0.0386***	0.0607***	-0.00297	-0.00937	-0.00687	0.0393***	0.0527***	0.0743***
	(3.39)	(4.31)	(3.77)	(-0.22)	(-0.67)	(-0.54)	(3.90)	(4.99)	(3.55)
Log_Cash2Asset	0.0192**	0.0193**	0.0222**	0.0294***	0.0347***	-0.00297	0.0162*	0.0145	0.0242**
	(2.39)	(2.40)	(2.53)	(2.72)	(3.20)	(-0.52)	(1.65)	(1.48)	(2.08)
Log_LEV	0.0456**	0.0408**	0.0615**	0.0680*	0.0568	-0.0115	0.0300	0.0258	0.0696**
	(2.25)	(2.00)	(2.39)	(1.89)	(1.56)	(-0.49)	(1.27)	(1.09)	(2.19)
IA	0.0279	0.0644	0.0548	-0.0170	0.0342	0.0893	0.0852	0.126	0.134
	(0.28)	(0.64)	(0.57)	(-0.12)	(0.25)	(1.49)	(0.70)	(1.03)	(1.12)
ROA	-0.0253	0.000778	-0.0630	-0.532***	-0.566***	-0.0543	0.0319	0.0705	0.0226
	(-0.28)	(0.01)	(-0.63)	(-2.74)	(-2.91)	(-0.47)	(0.31)	(0.68)	(0.19)
ROE	0.0449	0.0500	0.120***	0.239**	0.217**	0.0201	-0.0144	-0.00830	0.0921*
	(0.93)	(1.04)	(2.62)	(2.48)	(2.26)	(0.48)	(-0.26)	(-0.15)	(1.70)
SG	0.0360	0.0244	-0.0163	-0.0482	-0.0563	-0.0626**	0.0402	0.0255	-0.0141
	(1.13)	(0.77)	(-0.54)	(-0.79)	(-0.93)	(-2.35)	(1.12)	(0.71)	(-0.40)
NSI	-0.332***	-0.365***	-0.317***	0.0569	0.00434	-0.0654	-0.378***	-0.414***	-0.302**
	(-2.83)	(-3.12)	(-2.87)	(0.29)	(0.02)	(-0.77)	(-2.80)	(-3.07)	(-2.29)
Log_Spread	-0.0503*	-0.0471*	-0.0982***	-0.0183	-0.0102	0.00614	-0.0635*	-0.0560*	-0.150***
	(-1.91)	(-1.78)	(-3.57)	(-0.56)	(-0.31)	(0.42)	(-1.93)	(-1.69)	(-3.99)
Log(state cites)	0.00243	0.00118	-0.0156	0.0135	0.0123	0.0125	0.00399	0.00185	-0.0290**
	(0.47)	(0.23)	(-1.49)	(1.60)	(1.44)	(1.53)	(0.66)	(0.30)	(-2.09)
Log(class cites)	0.348***	0.351***	0.229***	0.0711	0.107	0.00479	0.507***	0.508***	0.276***
	(12.83)	(13.00)	(10.73)	(0.67)	(1.01)	(0.28)	(11.49)	(11.56)	(9.65)
Observation	14503	14503	14503	3090	3090	3090	11413	11413	11413
R^2	0.348	0.355	0.570	0.086	0.109	0.890	0.349	0.357	0.569
Firm Fixed Effect			Yes			Yes			Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class Fixed Effect	Yes	Yes		Yes	Yes		Yes	Yes	
NAICS Fixed Effect	Yes			Yes			Yes		

Table E2: Citation Bias and Firm Characteristics (contd.)

This table presents OLS regressions relating time and tech class fixed effect adjusted citation bias at the firm level with different firm characteristics. The dependent variable is the time and tech class fixed effect adjusted citation bias of a given firm in that year for years 1976-2006 (columns 1-3) and for subsamples, 1976-1990 (columns 4-6) and 1991-2006 (columns 7-9). The dependent variable is computed as the difference in log of one plus number of citations to all patents of a firm applied for in a given year and granted by 2006 in our data and log of one plus number of citations to the same set of successful patents of that firm in the same application year in the NBER 2006 dataset. Restricting the successful patents from our data to only those that are granted by 2006 allows for comparison with successful patents in the NBER 2006 data. Logs are taken to account for skewness in citations activity. Control variables and their construction are described in Appendix C. Details about time and tech class fixed effect adjustment can be found in Appendix B, noting that we follow Jaffe and Rassenfosse (2016) and consider only patents by publicly traded firms. Robust t-tests are reported in the parenthesis. Sources: NBER 2006 patent and our datasets.

Panel B: Adjusted Citation Bias (Time & Tech Class)

	Adjusted Citation Bias (Time & Tech Class) (1976-2006)			Adjusted Citation Bias (Time & Tech Class) (1976-1990)			Adjusted Citation Bias (Time & Tech Class) (1991-2006)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log_Size	0.0792***	0.0755***	0.00965	-0.0144*	-0.0113	0.0221*	0.110***	0.106***	0.0274
	(13.53)	(12.53)	(0.73)	(-1.76)	(-1.33)	(1.83)	(15.44)	(14.46)	(1.51)
Log_M2B	0.0619***	0.0686***	0.0630***	0.146***	0.144***	0.00944	0.0128	0.0216	0.0280
	(3.00)	(3.31)	(2.89)	(3.54)	(3.46)	(0.45)	(0.54)	(0.91)	(1.06)
Log_RD2Sale	0.0317***	0.0413***	0.0592***	-0.00242	-0.00928	-0.0116	0.0430***	0.0566***	0.0772***
	(3.65)	(4.56)	(3.64)	(-0.17)	(-0.64)	(-0.90)	(4.21)	(5.30)	(3.64)
Log_Cash2Asset	0.0176**	0.0169**	0.0178**	0.0291***	0.0343***	-0.00440	0.0132	0.0107	0.0189
	(2.16)	(2.07)	(2.01)	(2.60)	(3.05)	(-0.77)	(1.33)	(1.08)	(1.61)
Log_LEV	0.0530***	0.0480**	0.0718***	0.0829**	0.0721*	-0.00199	0.0363	0.0318	0.0827***
	(2.58)	(2.32)	(2.76)	(2.22)	(1.91)	(-0.08)	(1.52)	(1.33)	(2.58)
IA	0.0203	0.0526	0.0433	-0.0140	0.0357	0.101*	0.0702	0.107	0.113
	(0.20)	(0.51)	(0.45)	(-0.10)	(0.25)	(1.68)	(0.57)	(0.86)	(0.93)
ROA	-0.0184	0.0114	-0.0595	-0.532***	-0.568***	-0.0722	0.0371	0.0806	0.0287
	(-0.20)	(0.12)	(-0.59)	(-2.65)	(-2.83)	(-0.62)	(0.36)	(0.77)	(0.24)
ROE	0.0328	0.0384	0.111**	0.255**	0.233**	0.0319	-0.0283	-0.0222	0.0844
	(0.67)	(0.79)	(2.39)	(2.56)	(2.35)	(0.76)	(-0.51)	(-0.40)	(1.54)
SG	0.0245	0.0132	-0.0299	-0.0461	-0.0551	-0.0670**	0.0274	0.0131	-0.0309
	(0.76)	(0.41)	(-0.98)	(-0.73)	(-0.88)	(-2.49)	(0.75)	(0.36)	(-0.87)
NSI	-0.306***	-0.337***	-0.289***	0.0465	-0.00778	-0.0827	-0.343**	-0.377***	-0.267**
	(-2.58)	(-2.85)	(-2.59)	(0.23)	(-0.04)	(-0.96)	(-2.51)	(-2.76)	(-2.00)
Log_Spread	-0.0379	-0.0360	-0.0949***	-0.0158	-0.00708	0.00802	-0.0508	-0.0450	-0.147***
	(-1.42)	(-1.34)	(-3.42)	(-0.47)	(-0.21)	(0.55)	(-1.53)	(-1.34)	(-3.87)
Log(state cites)	0.00171	0.000209	-0.0152	0.0134	0.0120	0.0133	0.00293	0.000522	-0.0273*
	(0.32)	(0.04)	(-1.43)	(1.53)	(1.37)	(1.62)	(0.48)	(0.08)	(-1.95)
Log(class cites)	0.282***	0.285***	0.193***	0.0389	0.0765	0.0329*	0.410***	0.411***	0.226***
	(10.28)	(10.41)	(8.94)	(0.35)	(0.70)	(1.92)	(9.19)	(9.23)	(7.82)
Observation	14503	14503	14503	3090	3090	3090	11413	11413	11413
R^2	0.340	0.346	0.566	0.079	0.103	0.894	0.341	0.349	0.563
Firm Fixed Effect			Yes			Yes			Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Class Fixed Effect	Yes	Yes		Yes	Yes		Yes	Yes	
NAICS Fixed Effect	Yes			Yes			Yes		

Appendix F: Patent Lag Distribution

This table represents the statistical information for the distribution of lag (in years) between application year and grant year, and citing and cited year for granted patents during 1976 to 2012 in different HJT technology categories. The table shows there are longer lags in the "Computer & Communications" and "Drugs & Medical" classes compared with other four classes. Sources: Our dataset.

Panel A: Application-Grant Lag Distribution for Granted Patent Cohorts from 1976 to 2012

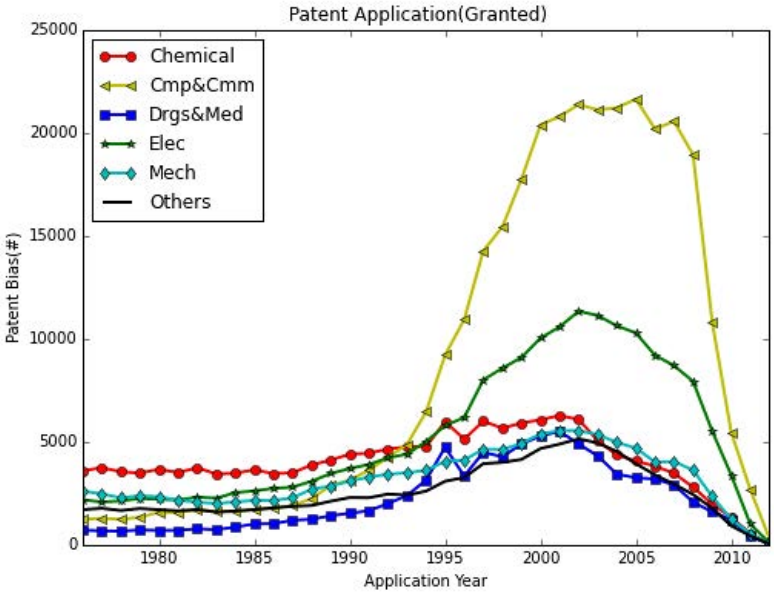
Statistics	Application-Grant-Lag in years					
	Chemical	Computer& Communication	Drugs& Medical	Electrical& Electronic	Mechanical	Others
Min	0	0	0	0	0	0
25%	1	2	2	2	1	1
50%	2	3	2	2	2	2
75%	3	4	4	3	3	3
Max	59	67	28	58	48	30
Mean	2.36	3.27	2.87	2.36	2.20	2.17
Std	1.26	1.76	1.69	1.30	1.23	1.26

Panel B: Cited-Citing Lag Distribution for Filed Patent Cohorts from 1976 to 2012

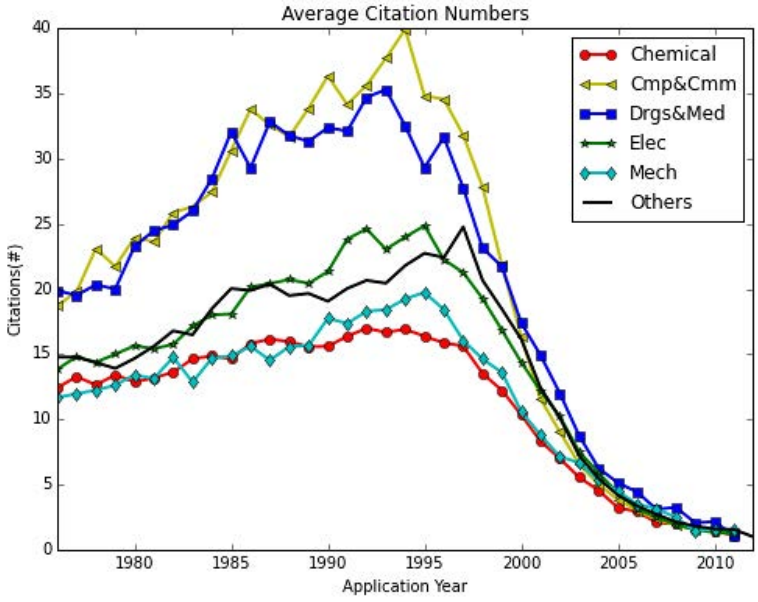
Statistics	Cited-Citing-Lag(For Cited Patents) in years					
	Chemical	Computer& Communication	Drugs& Medical	Electrical& Electronic	Mechanical	Others
Min	0	0	0	0	0	0
25%	5	4	5	4	4	5
50%	8	6	8	7	8	8
75%	13	10	13	11	13	14
Max	36	36	36	36	36	36
Mean	9.62	7.45	9.51	8.43	9.37	9.85
Std	6.546	5.171	5.959	6.105	6.660	6.778

Appendix G: Distribution of Granted Patents and Average Citations per Patent in Different HJT Technology Classes

This figure presents the distribution for successful patent applications and the average citation received by a patent in each HJT technology class. Firms are assigned to a particular technology class in a given year based on modal primary class of patents produced by the firm in that year based on the U.S. patent classification system. We then aggregate the granted patents and citations per patent in a technology class across publically traded firms by year. Sources: NBER 2006 patent and our datasets.



(a)



(b)