Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13-15, 2016

Federal Funding and the Rate and Direction of Inventive Activity

Rafael Corredoira
Ohio State University - Fisher College of Business
Management and Human Resources
corredoira.2@osu.edu

Brent Goldfarb
University of Maryland
Department of Management and Organizations
bgoldfarb@rhsmith.umd.edu

Yuan Shi
University of Maryland
Department of Management and Organization
yuanshi@rhsmith.umd.edu

Abstract

Leveraging a new measure of patent citation trees (Corredoira & Banerjee, 2015), we demonstrate that research funded by the federal government is likely to spark more active technological trajectories. Our findings tie government funding to the generation of breakthrough inventions. The differences are only evident at the upper percentiles of the distribution of long term patent influence and stem primarily from research conducted at universities and academic medical centers that is sponsored by the DoD, HHS and NSF. Additional analysis indicates that federal programs invest in many technological areas that private corporations eschew. Moreover, we also demonstrate that an exogenous shock to the funding availability will impact the quality of the inventions funded afterwards. In these senses, the government affects both the rate and direction of inventive activity.

Jelcodes:O38,M20

Federal Funding and the Rate and Direction of Inventive Activity*

"Generally speaking, the scientific agencies of Government are not so concerned with immediate practical objectives as are the laboratories of industry nor, on the other hand, are they as free to explore any natural phenomena without regard to possible economic applications as are the educational and private research institutions." - Vannevar Bush, Science the Endless Frontier*"*, 1945.

1. Introduction

30% of US research and development (R&D) is funded by the federal government. However, since 2010, US government spending on research and development has remained flat and funding in basic research has actually declined (Sargeant, 2015). The basic logical argument for government funded R&D is articulated in Vannevar Bush's 1945 Report to the President. The welfare reasons to support Bush's policy recommendations were explored by economists in 1957 (Nelson) and Arrow (1962). Nelson suggested and Arrow formalized the idea that the government might play a role in sponsoring R&D because the private sector is likely to underinvest in R&D due to difficulties in appropriating returns to particular projects. The appropriability problem is exacerbated by the technological riskiness of innovative projects that reduce expected private value. In general, while most methods and assumptions show a strong positive private return to R&D spending, and some evidence of spillovers from the public to the private sector, systematic research based on disaggregated analysis of fine-grained, multi-sector data has been scarce (Hall, Mairesse, and Mohnen, 2009).

As a recent article in Science (Wang, Song & Barabási, 2013) reports, the long-term impact of research defies simple modeling and prediction efforts have progressed very little beyond direct citations. Our analysis leverages a new measure, influence, that tracks entire citation trees of patents over time (Corredoira & Banerjee 2015). This measure is well designed to detect differences in long-term value. That is, instead of using citations to a focal patent to proxy that patent's economic value, the measure takes into account all the subsequent patents that cite the patents citing the focal patent. It then counts the patents citing those citing patents and so on. In the empirical analysis in this study, influence and first generation citations are only weakly correlated. This indicates that influence is a superior measure for understanding how a particular invention affects technological progress over the long term. In this sense, this work complements the recent work that uses a text-based measure of novelty in patents that is also found to not be highly correlated with citations (Kaplan & Vakili, 2015). Our analysis asks whether the influence of patents stemming from research funded by the US federal government is different than that resulting from the private sector's R&D efforts.

In addition to the influence measure, and to address the issues raised by Thompson & Fox-Keane (2005), we designed a comprehensive matching scheme which matches a federally funded patent with a control patent as close as possible in terms of technology subclass and granting time. The matching scheme also allows for the possibility of no match as well as randomness in multiple matches. The empirical strategy bears immediate fruit as it allows us to examine whether patented technological inventions funded by federal government agencies are more influential than their private counterparts. We are also able to examine whether differences are due to funding mechanisms, the type of recipient or the funding agency. Based on analysis of a unique sample of federally funded US patents in the period of 2001-2004, our results indicate that government funded projects are more likely to lead to "home-run" inventions that make long-lasting and large contributions to technological trajectories. However, the median government sponsored and privately sponsored patents are different in neither impact nor influence.

Besides the main result that government funded research is more influential, but only at the top percentiles, we present 3 additional findings.

1. Internally conducted research is less influential than externally conducted research.

The federal government funds research through a variety of instruments. Grants are used primarily to fund research in non-profit labs throughout the country, in particular university and federal research labs. However, through such programs as the SBIR research grants, the federal government also supports research and development efforts at private firms. We find that compared to their matches with no federal support, patents developed by external researchers funded through federal grants and contracts are generally more influential at 80% quantile and above. However, this result does not extend to federally funded research conducted intramurally. Patents developed internally by federal employees are less influential than their matches at 80% quantile and above, though the difference in influence disappears at 98% quantile. We are unable to determine whether this result is due to a difference in the type of research, or the many constraints put on the commercialization of research that is patented by the federal government.

2. The Department of Defense, Health and Human Services and the National Science Foundation sponsor the most influential work. These differences stem from the increased influence of breakthrough patents originated from research conducted extramurally at universities and research hospitals. Corporations appear less likely to produce highly influential patents.

The vast majority of federal research funding is funneled through 5 agencies: the Department of Defense (DoD), Health and Human Services (HHS), the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA) and the Department of Energy (DOE). Our results reveal that in our sample, three of the five major federal agencies contribute the baseline difference in influence between federally funded patents and control patents. Specifically, the DoD and HHS appear to have funded some very influential innovations at the top (98% quantile), while NSF-backed patents perform very well over a wider spectrum (80%, 90%, and 95% quantile). In 2013, these three agencies made up 76% of all federal R&D spending (Sargeant, 2015).

Our results regarding extra-mural funding appear to stem from higher influence of projects sponsored by the Army and to a lesser extent the Navy in the DoD as well as HHS. The NSF does not conduct intramural research. That is, the influential inventions funded by the federal government may be more likely to originate from US universities and hospitals. Corporations, as for-profit organizations, are far less effective in seeding long term technology trajectories - even when funded by the federal government. The power of these sub-agency tests are not great, and hence additional verification and replication will be necessary to confirm these conclusions.

3. We provide some evidence that Federal government agencies select funded technological areas in a non-random way, and exogenous shift in selection may lead to change in the influence of funded patents.

Our findings on the selection patterns of extramural federal funding suggest strong path dependencies within fields, especially by non-profit organizations. We also find that commercial value, as well as the uncertainty in approval and realized value may drive the presence of funding. Moreover, we also find that an unexpected change in the R&D budget available at an agency would impact the quality of the inventions it sponsored later, which is consistent with the selection mechanism as well. We find these results suggestive, though not conclusive.

2. Data and Sample

To examine the possible effect of federal funding on applied research, we first need to observe a population of applied research projects and whether those are funded by the federal government. We began by collecting data on granted patents and patent applications from USPTO. We limited our sample to those patents whose granting year and application year fall between 2001-2004 because bulk data files for US patent applications became available electronically in 2001. By limiting the grant year to 2004 we were able to measure influence of the sampled patents in the following ten years, which allows us to compare at least medium-term influence of the sample patents.

We then identified federally funded patents from this time period. We used three approaches to identify federally funded patents in the sample. First, we utilized a data field that is rarely used by prior research, "Federal Research Statement". We identified all patents that report information in this data field, and then dropped all cases of false positives by manually examining all indicated patents. Second, we located all patents assigned to federal government agencies from the assignment information reported in the NBER 2006 patent data. We were aware that the assignee type information reported in the NBER data is noisy so we went through all assignee names in all the government types, generated a list of US federal government assignees and extracted all patents associated with them. Third, we also included additional patents reported in the NIH and DOE's public patent database.[1] The process generated 4,311 federally funded patents.

Table 1 is a summary of the patent distribution among major federal agencies. As reported in the table, six major agencies, namely Department of Defense (DoD), Department of Health and Human Services (HHS), Department of Energy (DOE), National Aeronautics and Space Administration (NASA) and National Science Foundation (NSF) are responsible for the funding of more than 90% of all sample patents. Meanwhile, it should be noted that the total number of federally funded patents is about 1.6% of all patents filed and granted during the same period. Moreover, the US Navy funded the most inventions among the three branches of the DoD; most HHS funding reported in patents comes directly from National Institutes of Health. Table 2 summarizes the distribution of federally funded patents among six HJT categories (Hall et al., 2001). While Category 4 of electrics has the most patents (26.83%), all six categories are relatively evenly represented in the sample. The representation across all major technological categories suggests that

[1] We also contacted other major federal government agencies and requested data of their funded patents by email but could not obtain any additional data.

it is unlikely that our findings only apply to very few technological categories where federal R&D funding is concentrated.

Table 1[1] Distribution of major federal agencies in federally funded patents

Agency Name	Patent Count	Percentage
DOD	1753	40.66%
NAVY	925	21.46%
ARMY	416	9.65%
USAF	308	7.14%
HHS	1068	24.77%
DOE	783	18.16%
NASA	283	6.56%
NSF	263	6.10%
USDA	87	2.02%
DOI	20	0.46%
EPA	18	0.42%
NSA	17	0.39%

[1] Some patents are developed by funding from multiple agencies.

Table 2[1] Distribution of HJT technology categories in federally funded patents

HJT Category	Patent	Percentage
1 Chemical	863	20.03%
2 Computer & Communication	542	12.58%
3 Drugs & Medicine	927	21.51%
4 Electric	1156	26.83%

5 Mechanic	515	11.95%
6 Others	306	7.1%
[1] 2 patents with missing category info.		

3. Method

3.1 Matching Techniques

There are two central ways in which federal funding might lead to different levels of influence. The first is that federally funded projects may be different along the ways described in the introduction. That is, federally funded projects might be more oriented towards uncovering fundamental principles and more oriented to long term ideas. This would then lead to a greater degree of measurable influence. It is also possible that given identical research outputs, federally funded inventions may receive greater influence due to the greater likelihood that a federally funded invention is assigned to a public sector or university assignee. Such actors may be more likely to emphasize diffusion of an invention over appropriation of returns. In this case, differences in influence would be caused not by differential motives of funders playing out in the choice of project types, but rather in differential motives leading to different institutional mechanisms that affect diffusion.

An ideal test would distinguish between these two mechanisms. For example, such a test might compare the influence of two identical inventions, one federally funded invention and the second funded by the private sector. Once identifying this post-invention institutional effect, we could then compare closely related patents to determine if the underlying potential for influence differs across projects of differing funding types. As such counterfactuals are not observable in reality, our strategy is to identify similar inventions. We match each government funded (sample) patent with a privately funded (control) patent that is introduced around the same period. These control patents represent a reference group of inventions that are similar in nature but do not receive federal support. This approach cannot precisely distinguish between a pre-invention quality or selection effect and a post-invention institutional or treatment effect.[2] While previous studies have utilized publications of simultaneous discoveries and offer more precision in equalizing the quality of match and control (Bikard, 2012), in theory it is impossible to identify two identical patents in the patent system due to requirement of novelty.

In order to identify the matched control groups, we primarily rely on the information of the primary subclass and granting date of the patent. Matching on the subclass, the most fine-grained patent classification, minimizes the possibility of spurious effects of unobserved differences within the aggregated 3-digit class and ensures that the matched patent embodies technology similar to that of the sample patent (Thompson & Fox-Kean, 2005). We first defined the universe of patents for matching as all US patents filed and granted in 2001-2004 in accordance with the sampling window of federally funded patents. We then located all the patents in the same primary subclass and the same grant year and keep those on the closest granting date.

In a perfect universe, the ideal outcome would be that we find a unique match for each sample patent. However, two problems emerge in the actual matching process. First, multiple patents may be matched to a sample patent. While this seems to be a good problem to have, and one solution is to randomly pick one match from all the candidates, potentially valuable information is lost when we drop all the equally good alternatives. Second, inventive activities are very sparse in many subclasses. Specifically in our case, we find no patent that meets the matching criteria for approximately 30% of the sample patents. It is worth noting that there are two scenarios in which no match is found for a sample patent. One is that all patents in the subclass are funded by the federal government. In this case a match simply does not exist. Another more complicated scenario is that there are not enough candidates of matches for all sample patents in the subclass. For instance, patents X and Y are potential matches for federally funded patents A, B, C and D, so patent X and Y may be matched to only two of A, B, C and D. Consequently, there will always be two sample patents left with no matches and the choice of these two sample patents is arbitrary.

Traditional remedies of matching pose a dilemma in this setting. While matching on more observable characteristics of the patent would alleviate the first problem of multiple matches, it would worsen the second problem of no admissible match because the universe of patents is further restricted by the additional criteria imposed. Meanwhile, relaxing certain criteria for the matching

[2] We can and do compare both government sponsored inventions assigned to private firms as well as government sponsored inventions assigned to public and research institutions with privately funded inventions to private firms. We have yet to compare between the two former groups.

process would alleviate the second problem of no match while aggravating the situation of multiple matches.

We adopted two approaches to tackle these problems. First, we took advantage of the hierarchical structure of classification scheme of USPTO to tackle the problem of no match.[3] Under the classification scheme, subclasses within a given 3-digit class are nested within one another, creating a multilevel structure where lower levels are aggregated to higher levels. For a patent without a "perfect" match in the same subclass, we searched the hierarchical structure of the subclass to locate its parent subclass (the one which can be identified as the immediate next aggregated level) and then repeat the search within the parent subclass. If no admissible result is found, the search is repeated within a subclass on the more aggregated level, until the universe of patents is reached. Once "imperfect" matches are found in the closest subclass available, we randomly pick one on the closest granting date.

For example, in Class 514 "Drug, Bio-affecting and Body Treating Compositions", Subclass 514/7.3 "Type I Diabetes" is a level 4 subclass nested within Subclass 514/6.9 "Diabetes" (level 3), which is further nested within Subclass 514/6.8 "Blood sugar affecting" (level 2). In a similar fashion, level 2 Subclass 514/6.8 can be aggregated to level 1 Subclass 514/1.1 "Peptide (e.g., protein, etc.) containing DOAI", which is a branch of the level 0 Subclass 514/1 "Designated organic active ingredient containing (DOAI)", the most aggregated subclass under Class 514. If no match is found in 514/7.3, a search will be conducted in the most immediate parent subclass, in this case, 514/6.9. If no match turns up in 514/6.9, the search will then be further expanded to the parent subclass of 514/6.9, i.e. 514/6.8. The process will be repeated to broader subclasses until one or more matches are identified.

Second, we repeated the matching process to generate multiple matching samples with some level of randomness, which may alleviate the concern that our estimation based on the matched sample may be driven by the "luck" stemming from arbitrary choice from multiple matches. The multiple matched control groups generated by the repeated matching process allow us to perform a meta-analysis of the effects of interest, which will be explained in Section 3.3. Randomness is built in the process in two ways. First, every time multiple matches are identified for the same sample patent, one match will be randomly selected. Second, every time a control patent is found to be a match for multiple sample patents, only one sample-control pair will be randomly created and the

---

[3] We use subclass information in master classification file (MCF) released in 2014 to identify matches, while the subclass hierarchy file is based on classification scheme of 2006 due to data constraint.

rest of the sample patents will be rematched to the universe of patents again. Note that the whole landscape of matching is reshuffled in each round of matching due to the simultaneous presence of the two aforementioned problems. Take the example of four sample patents (A, B, C, D) and two matches (X, Y) in the same subclass we introduced earlier, in one round of matching we may randomly choose two sample-control pairs A-X and B-Y, in which X and Y are perfect matches. Then patent C and D have to go through the process of "imperfect matching" in the parent subclass and more than two matches are likely to be found for C and D, which further increases the possible combinations of sample-match pairs.

The matching process was repeated thirty times independently, and accordingly thirty matched control groups of patents for the sample patents were generated. Using as many as thirty control groups would allow us to exploit the multiple matches available and increase overall sample size, which lowers the threat of Type I error. An immediate concern is whether the actual universe of patents would allow us to gain significant variance in matches and thus increase substantial statistical power over one-time matching. To verify the effectiveness of our matching scheme, we randomly picked three groups of matches, or 10% of all groups generated and compared the matching results. It is found that 1020, or 23.66% of the sample patents have more than one match, of which 229 patents are matched to a different patent in each round. Additionally, sample patents with extremely high influence score are critical to our findings in quantile regressions. Hence we restricted our comparison to top 5% of the sample patents, and found that 22.32% of these patents have two or three matches in three control groups. Taken together, the matching scheme does expand the sample size and increase the variance.

Another major concern with our differentiated matching method (as in perfect match in the same subclass vs. imperfect match in the closest subclass available) is that whatever the results we found may be the artifact of differences in the matching method. To address this concern, we constructed a dummy representing whether a sample patent is matched within its own subclass and added it as a moderator of the main independent variable in the baseline analysis, which will be elaborated in Section 4.1. We found no significant moderating effect. This suggests that our matching methods do not seem to confound the main effect of interest.

3.2 Measure of Influence

We followed Corredoira and Banerjee and measured the patent influence by an iterated sum capturing the size and the interconnections of the citation tree that followed the patent. To directly compare the influence of a single sample patent and its match, we constructed a forward citation tree which is composed of all the direct and indirect citing patents in the 10-year window. We then calculated the *influence* for the focal patent in this network as:

$$\text{Influence} = \left(\sum_{k=1}^{\infty} \alpha^k A^{\mathrm{T}k}\right) e \qquad \text{(Eq. 1)}$$

Where α is the attenuation factor, k is the citation generation, A^{T} is the transpose of the adjacency matrix defined by patent citations, and e is a vector capturing the significance of each patent for the technology.

Additionally, for each sample and control patent, we also counted the number of forward citations in the 10-year window as the conventional measure of *impact*. Corredoira and Banerjee (2015) offer a comprehensive discussion on the patent influence measure and why it is different from direct impact. The construction of independent variables will be detailed in the next section as we walk through the findings.

3.3 Econometric Models

We used Quantile Regression with bootstrapped standard errors to estimate the effect of federal funding on patent influence and impact. Quantile regression models are chosen for two reasons. First, our dependent variables are highly skewed, and outliers were found to overwhelm any meaningful average effects yielded by traditional OLS models. Quantile regressions are known for their superior capability to embrace outliers because they analyze the relative distribution rather than the absolute mean of variables. Second, empirically we are interested in those outliers at the top, namely those extremely influential inventions that may be "homeruns" and 'breakthroughs", and fundamentally different from other inventions. Regressions based on average effects will hide the effects of outliers.

For each sample and control patent, we control for the number of claims and fixed effects of granting year and the primary 3-digit USPTO class. Note that the control patent is in the same or closest subclass available. Therefore, it is assumed that the technologies embodied in the sample and control patent are similar. Moreover, while it is possible that various unobserved characteristics of the patent may be related to patent influence, the estimates would be unbiased as long as these characteristics are distributed randomly across sample and control groups, which we hope to approximate by repeated matching.

As described in 3.1, randomness is built in our matching scheme. After we matched the sample patents to the universe thirty times, we paired the control patent identified in each round of matching with the sample patent and constructed thirty samples (each with a size of 4311 * 2 = 8622) for regression analyses.[4] Then we ran quantile regression on each of the samples and obtained thirty separate estimates on the model. Essentially, this is equivalent to running the identical study on the same population thirty times, with the inter-sample variance originating from the randomness in matching. Meta-analysis is an ideal estimation tool to combine the effects from individual samples and increase statistical power (Cohn & Becker, 2003). Following the common practice, for each variable of interest, we conducted fixed-effects meta-analysis and combined thirty individual coefficient estimates (effect size) by weighted average. Fixed-effects models are ideal in our setting because the exact same procedures are replicated for each individual analysis and our main concern is the sampling error from the same population. The weight applied is the inverse of the variance, which ensures greater contribution from estimates with more statistical power to the combined estimate. Accordingly, the variance for the combined estimate is the inverse of the sum of the weight.

4. Findings

4.1 Baseline analysis: are federally funded inventions more influential?

Using the data and methods we described in the previous two sections, we first compare the count of direct citations (*impact)*, and the logged "*influence*" of the focal patent for the federally funded patents and their multiple matches.[5] The correlation between *Impact* and *Influence* is 0.13 for the sample patents, suggesting that they are statistically distinctive despite the apparent relatedness. Due to the high skewness of the data, we used the logged term of the original alpha centrality value plus one as *influence*. Note that the distribution of influence score of sample patents is largely representative of the population. We compared 99%, 95%, 90%, 75%, 50% percentiles of the influence of the sample patents with that of the whole population of patents in the same period and found very similar values. The only noticeable difference is that sample patents have fewer zero values at the bottom.

[4] Because the number of unique matches for each sample patent is different, pooling all matches together would cause the data structure to be unbalanced.

[5] The mixture of direct counts of citations and the log of the influence measure is a bit strange, but our results will be unaffected due to our use of quantile regressions.

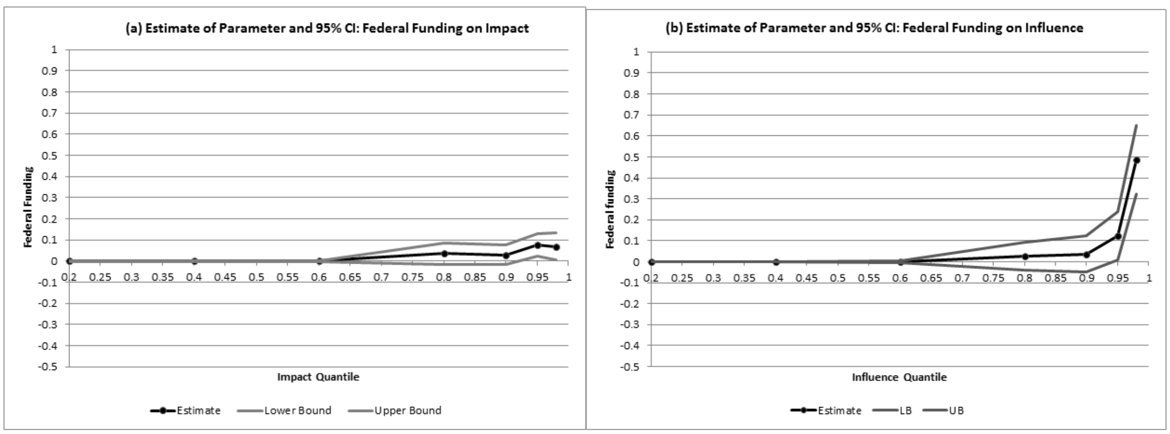
Given that the technological niche and time of introduction are matched as closely as possible between the sample group of federally funded patents and the control group of matches, we expect to see no difference in the impact/influence between two groups if federal funding has a null effect on the invention. In other words, if the inventions funded by the government are more influential than those funded by other sources in the same period and technological niche at conventional significance level, then we may infer that federal funding is unique in supporting certain types of R&D projects and may not be interchangeable with private funding or other types of public funding.

We first compare the impact score and influence score between federally funded patents and their matches to set the baseline for our further analyses. To identify the effect of federal funding, we create an independent variable, *Federal Funding*, which equals one if the focal patent is a federally funded patent and zero if it is a match. With this variable, we attempt to test whether on average inventions funded by the federal government outperform or underperform the imperfectly constructed counterfactuals.

Combined estimates of the coefficient and 95% confidence interval from thirty quantile regressions at 20%, 40%, 60%, 80%, 90%, 95% and 98% quantile for each outcome variable are plotted in Figure 1. Figure 1 (a) suggests that federally funded patents have more direct impact than their matches at the very top (95% and 98% quantile). Figure 1 (b) demonstrates that top federally funded patents are *much* more influential than their privately funded matches. In both cases, most federally funded patents, in particular, those below the 95% quantile are neither more impactful nor influential than their matches. Turning to the economic significance of the findings, the magnitude of the effect of federal funding on impact is found to be small but significant at 95% quantile in Figure 1 (a). Other conditions being equal, the 95% percentile of the impact of federally funded patents is 0.075 higher than that of the control patents. Nevertheless, as shown in Figure 1 (b), the magnitude of the effect on the patent influence is much larger. Taking into account that influence measure is logged, the 95% percentile of the original influence score for federally funded patents is 13.29% higher than that of the control patents. At the 98% quantile, federally funded patents are estimated to be 62.30% more influential, which is striking because no significant difference is found at 90% or lower quantiles or at the mean (in unreported OLS regressions).

As mentioned in Section 3.3, several alternative determinants of patent influence are accounted for in the regression model. First, we are able to capture the inter-class differences by employing the fixed effects of primary 3-digit USPTO class of the patent.[6] Second, it appears that year effects are also present based on the fixed-effects estimates of the patent granting years. Third, we also find that, not surprisingly, patents with more claims are more influential. To summarize the main findings, we detect a baseline difference between the group of federally funded patents and the matched control group. Specifically, the federally funded patents are more influential than their matches, but only at the very top. Moreover, the difference is much larger when the whole forward citation network of the focal patents is taken into consideration.

Figure 1



4.2 Analysis of mechanisms of federal funding: are inventions funded internally more (or less) influential than those funded externally?

Given the evidence of a baseline premium in high influence enjoyed by federally funded patents, our next question is whether this difference is driven by distinctive funding mechanisms, namely intramural and extramural funding. Research projects may be sponsored by the federal government through these different mechanisms. Intramural research is performed by R&D personnel of the federal agency and funded internally by the agency through its own budget. Extramural research is performed by researchers affiliated with external institutions and funded through contracts, grants and cooperative agreements (National Institute of Allergy and Infectious Diseases,2015; National Science Foundation, 1995).

For there to be a difference between the two, one of three mechanisms might be at play: First, the projects done intramurally and extramurally might be different in kind due to different

6 Further controlling for subclass difference is not possible due to constraints on the degree of freedom.

selection processes those projects are subject to. It may be the case that secrecy, instead of patent is preferred for sensitive projects within the government. Second, the quality of researchers in the two settings might differ. Finally, the post-invention institutional settings might lead to different diffusion and citation patterns.[7]

We leverage aspects of the institutional process of patenting to identify intramural and extramural funding. Remember that federally funded patents were identified through both the government mandated disclosure on patent applications as well as the patent assignment information of federal government agencies, based on which we are able to empirically observe inventions supported by both intramural and extramural federal funding. The Bayh-Dole Act and Executive Order 12591 requires that "in applying for a patent, the organization (who is the contractor or grantee) must add a government interest statement that discloses the government's rights to the invention" (page 4, United States General Accounting Office, 1999). Meanwhile, it seems to be the case that intramural research not funded through external grants and contracts are not subject to the mandated disclosure, since 82.88% of all the patents assigned to the federal government do not report sponsorship information in the federal research statement. Accordingly, we distinguish extramural funding from intramural funding based on whether the patent application discloses funding information in the federal research statement. *Extramural Federal Funding* is defined as one if the relevant funding support information is reported in federal research statement of the patent replication and zero otherwise. *Intramural Federal Funding* is defined as one if no information on federal support is disclosed in the patent application and the patent is assigned to the federal government, and zero otherwise. These two dummies would replace the dummy *Federal Funding* which lumps both funding mechanisms together in the estimation.

Among all 4311 federally funded patents in the sample, 3023 patents disclose funding support information in the application files, and 1554 patents are assigned to the federal government, which largely correspond to extramural and intramural funding mechanisms. For 266 patents that are assigned to the federal government and yet discloses federal support information, we classify them as recipient of extramural funding because we assume these patents are developed by intramural researchers (assignment to the federal government) with support from other federal

[7] As noted in Goldfarb (2008), creating more influential knowledge is not necessarily the goal of all or even most federally funded research. Most federally funded research is channeled through mission oriented agencies, and many of those missions require very concrete and applied projects and solutions. Moreover, to the extent that such goals work against the production of influential research, and given that it is easier to align these goals inside the agency than across organizational boundaries (Aghion & Tirole, 1994), we expect extra-mural projects to be more influential. This prediction is reinforced if we expect more talented researchers to work at universities.

agencies (disclosure of federal government support as a grantee/contractor). The classification may not be entirely accurate, but the noise in the measurement would bias the estimation downward and thus make the estimation more conservative.

Figure 2-a and Figure 2-b report the combined estimates of the effect of extramural funding on patent impact and influence. Starting at 80% quantile, patents supported by extramural federal funding outperform other patents in the similar technological space and the same period both in direct impact and overall influence, and the performance advantage is greater as we focus on quantiles closer to the top. On average, extramurally funded patents are 15.68% more influential at 80% quantile and almost twice as influential at 98% quantile, as suggested in Figure 2-b. Figure 3-a and Figure 3-b report the results on intramural funding. As it turns out, inventions funded internally underperform their matches from 80% quantile, though the difference in influence becomes insignificant at 98% quantile. The 95% percentile of the influence for intramurally funded patents is only 69.07% of the that of the matches.

Taking Figures 1-3 together, we may infer that the effects of federal funding at the 80% and 90% quantiles in Figure 1 are insignificant because the positive effect of extramural funding and the negative effect of intramural funding counteract each other. Meanwhile, the positive effect at 95% and 98% quantile in Figure 1 is mainly driven by the superior performance of extramurally funded patents. In summary, we find that alternative mechanisms of federal funding yield significant and divergent implications for performance. Compared to their matches with no federal support, patents developed by external researchers funded through federal grants and contracts are generally more influential at 80% quantile and above. In contrast, patents developed internally by federal employees are less influential than the matches at 80% quantile and above, though the difference in influence disappears at 98% quantile.

Figure 2

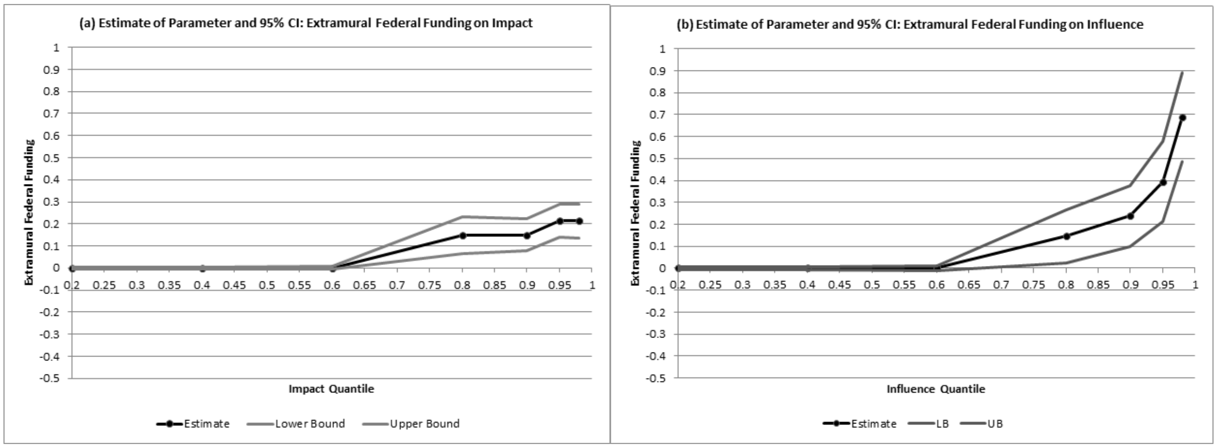
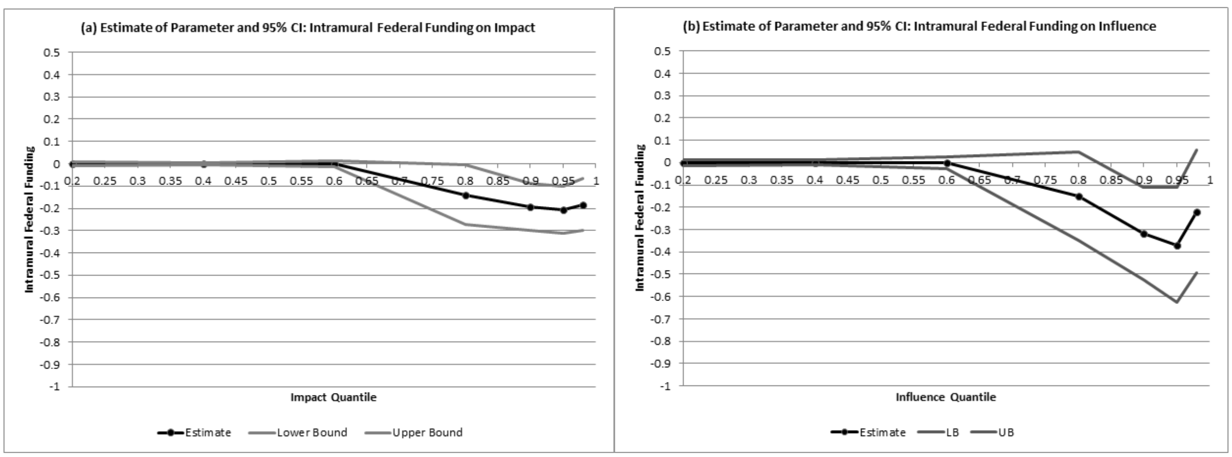


Figure 3



4.3 Analysis of funding agencies: do some federal agencies fund more influential inventions?

The baseline analysis performed in Section 4.1 also lumps together the potential effects of different funding agencies, which we seek to separately measure in this section. While one may argue that these agencies specialize in very different technological fields and thus agency differences are an artifact of underlying technologies, our empirical approach matches patents in the same or similar technological fields together and thus allows us to make reasonable comparisons between different agencies. We will focus on the "Big Five" federal agencies because each agency accounts for more than 5% of all patents in the sample and together they fund about 90% of the sample patents. Given the dominant presence of Department of Defense (*DOD*) funding in the sample (1753 or or 40.66% patents funded), we are able to further unpack the differences among its three major sub-agencies, US Army (*ARMY*), US Navy (*NAVY*), and US Air Force (*USAF*). An agency dummy, labeled by the abbreviation of the agency, is defined as 1 if agency information is either identified from the federal research statement or the assignee name. We will further disentangle effects of different funding mechanisms by individual agencies in Section 4.4. Additionally, we create a dummy, *Other Agencies*, to represent funding support from other agencies or records where agency information is not identifiable. For the analysis on DoD sub-agencies, we also create a dummy, *DOD/Other*, to capture the effect of funding from other DoD sub-agencies (e.g., DARPA).

Table 3 reports our findings on the relationship between funding agency and patent influence.[8] First, we replace the dummy *Federal Funding* with six agency dummies and find that only three agencies fund more highly influential patents compared to the control groups. Specifically, DoD funded patents beat the matches in influence by 34.72% at 98% quantile, where HHS funded patents beat the matches by 147%. While the influence of patents supported by the NSF outperforms that of the matches by a margin as large as 177.32% at 80%, 90% and 95% quantile. The difference disappears at 98% quantile, indicating that NSF has yet to outperform the private sector in terms of homeruns DoD and HHS have.

Second, we further replace the dummy for DoD with four dummies representing DoD sub-agencies and discover that influential patents may be developed and supported by sub-agencies outside the three major branches of DoD or we simply lack statistical power to make an inference.[9] The analysis cannot identify a difference between any of the DoD subagencies, even though when aggregated there is a measurable difference at the 98th percentile. To sum up, our results reveal that three of the five major federal agencies contribute to most of the baseline difference in influence between federally funded patents and control patents. Specifically, DoD and HHS appear to have funded some very influential inventions at the top (98% quantile), while NSF-backed patents perform very well over a wider spectrum (80%, 90%, and 95% quantile). Meanwhile, we do not find any evidence that sample patents funded by DOE, NASA and other federal agencies perform differently from their matches in similar technological space, nor do we find evidence that the effect emanates from any particular sub-agency within the DoD.

Table 3. Quantile regression of funding agency on patent influence[1 2 3 4]

Quantile	80%	90%	95%	98%

[8]Due to space constraint, we will not report plots of estimates of all quantiles and results on patent impact in this and all following sections. They are, as usual, available from the authors upon request.

[9] We also attempt to obtain information on sub-agencies of HHS (mainly various research institutes under NIH), but the data is not as reliable because there is no standard format of reporting in the statement.

DOD	0.098	(0.101)	0.113	(0.116)	0.097	(0.131)	**0.298***	**(0.141)**
DOD/ARMY[5]	0.116	(0.216)	0.218	(0.269)	0.067	(0.352)	0.418	(0.338)
DOD/NAVY[5]	0.002	(0.139)	-0.172	(0.161)	-0.023	(0.19)	0.077	(0.199)
DOD/USAF[5]	0.338	(0.25)	0.474	(0.34)	0.475	(0.335)	0.596	(0.396)
DOD /Other[5]	**1.003***	**(0.465)**	0.846+	(0.433)	0.4	(0.684)	1.144	(1.071)
HHS	0.071	(0.157)	0.257	(0.214)	**0.631***	**(0.271)**	**0.903*****	**(0.241)**
DOE	0.011	(0.149)	0.119	(0.2)	0.384	(0.248)	0.311	(0.249)
NASA	-0.114	(0.265)	-0.454	(0.324)	0.12	(0.454)	0.63	(0.417)
NSF	**0.76***	**(0.307)**	**0.989***	**(0.468)**	**1.02***	**(0.508)**	0.717	(0.465)
Other Agencies	0.002	(0.191)	-0.075	(0.23)	-0.056	(0.241)	-0.014	(0.294)

[1] Weighted average estimates from 30 separate regressions based on a group of 4311 federally funded patents and 30 alternative matched control groups. All estimates below 80% quantile are insignificant.

[2] Bootstrapped error reported in parentheses.

[3] + p<0.10, * p < 0.05, ** p < 0.01, *** p < 0.001.

[4] Number of claims and dummies for 3-digit class and granting year included.

[5] Sub-agency effects of DOD branches estimated by replacing the dummy for DOD with sub-agency dummies in the model.

4.4 Analysis of funding agencies by funding mechanisms: do effects of funding mechanisms vary by agency?

The question that follows naturally from the findings on funding mechanisms in 4.2 and funding agencies in 4.3 is whether the effectiveness of extramural and intramural funding vary across different agencies. This section will explore the answers to this question by testing if the significant effects of DoD and HHS reported in 4.2 are contingent upon the funding mechanism.[10] In light of the previous findings, we are curious to know i. whether the null findings on three major branches of DoD are due to the opposite effects of extramural funding and intramural funding, and ii. whether the positive effects of HHS is primarily driven by research conducted internally or externally.

[10] No patent in the sample is assigned to NSF. This is not surprising considering that NSF does not operate research facilities of its own and thus most of the funding should be extramural.

We split each agency dummy into two dummies representing extramural and intramural funding of the focal agency. Take US Army as an example, "*DOD/ARMY: Extramural*" is coded as one if the federal research statement of the patent application acknowledges the support from US Army and zero otherwise. "*DOD/ARMY: Intramural*" is coded as one if the federal research statement does not report any information related to US Army, and the patent is assigned to US Army. Other variables in the regression model are kept the same. Table 4 reports the findings on the effect of DoD branches and HHS split by funding mechanisms. We find strong evidence that extramural funding of US Army and HHS is effective in sponsoring highly influential inventions. In contrast, inventions developed in-house by either agencies do not have more influence, and in fact they may be *less* influential. The results also proffer a suggestion that external research funded by US Navy might be more influential than its in-house counterpart. Further replications with larger samples or more precise measurement will be required to draw any firm conclusions.

In sum, our results suggest that extramural funding is associated with inventions that have higher influence in the case of DoD (Army, and to some extent, Navy) and HHS. We do not find evidence that intramural funding has the same effect. In contrast, we discover moderate evidence that intramural funding might be less effective in supporting highly influential patents at least in US Army and US Navy. This is by and large consistent with the aggregate findings in Section 4.2.

Table 4. Quantile regression of funding agency on patent influence by funding mechanism[1,2,3,4,5]

Quantile	80%		90%		95%		98%	
DOD/ARMY: Extramural	0.346	(0.308)	0.882+	(0.489)	**1.533****	**(0.589)**	0.793	(0.544)
DOD/ARMY: Intramural	-0.145	(0.276)	-0.297	(0.325)	-0.532+	(0.321)	-0.682	(0.581)
DOD/NAVY: Extramural	0.293	(0.23)	0.449	(0.321)	0.52+	(0.31)	0.372	(0.419)
DOD/NAVY: Intramural	-0.305	(0.192)	-0.451+	(0.231)	-0.508+	(0.264)	-0.538+	(0.309)
DOD/USAF: Extramural	0.43	(0.325)	0.665	(0.445)	0.538	(0.485)	0.956	(0.755)

DOD/USAF: Intramural	0.146	(0.383)	0.182	(0.528)	-0.119	(0.649)	0.461	(0.621)
HHS: Extramural	0.118	(0.165)	0.277	(0.224)	0.556+	(0.329)	**1.023*** **	**(0.26)**
HHS: Intramural	-0.073	(0.377)	-0.036	(0.748)	0.836	(0.513)	-0.265	(0.456)

[1] Weighted average estimates from 30 separate regressions based on a group of 4311 federally funded patents and 30 alternative matched control groups. All estimates below 80% quantile are insignificant.
[2] Bootstrapped errors reported in parentheses.
[3] + $p<0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[4] Number of claims and dummies for 3-digit class and granting year included.
[5] Agency-mechanism effects of DOD branches and HHS estimated separately by replacing one dummy for a (sub-)agency with dummies for agency-mechanism in each model.

4.5 Analysis of funding recipients: do some types of funding recipients produce more influential inventions?

Our focus in all the previous analyses has been on the federal agencies, which mainly act as the source of the funding. In this section, we shift our attention to the receiving side of federal funding, and aim to identify the types of external organizations who are more likely to create the the most influential inventions with the support of the federal government. We attempt to capture the effect specific to the recipient type by interacting the dummy for federal funding with four dummies representing major types of patent assignees, which we use as proxies for the corresponding types of recipient.[11] The baseline comparison group mainly consists of patents assigned to the federal government, most of which are classified as intramurally funded.[12] Table 5 reports the findings from the preliminary analysis on recipient types.[13]

As a whole, the findings demonstrate that highly influential inventions are more likely to occur in some types of funding recipients but not others, and it depends on the where you draw the line for high influence. US universities, as the most visible beneficiary of Bayh-Dole Act, perform better in producing inventions of great influence at 95% quantile and higher than other types of

[11] The assumption is that outside funding recipients retain the property rights of patents developed with federal support. While this assumption may not be completely accurate, it may increases the chance of Type II error and thus work against any significant findings.
[12] Occasionally, federally funded patents may also be assigned to other type of entities, such as individuals, but the number of cases in both sample and control groups is too few to warrant a separate comparison. We include dummies for all types of assignees in the model as a remedy.
[13] Bootstrapping with more than 450 independent variables in the model is very time-consuming. Due to time constraint, we calculate robust errors instead of bootstrapped errors in Table 5 using “qreg2” in Stata. The error is robust to heteroskedasticity and misspecification. We will test the sensitivity of the results with bootstrapping models in future.

funding recipients. More remarkably, the 98% quantile of the influence of federally sponsored inventions at US hospitals is estimated to be more than twelve times of the same quantile of the control group. Nevertheless, this impressive lead at the top seems to be at the expense of funding many inferior inventions in the lower quantiles. Moreover, US research institutes constantly underperform other funding recipients except US corporations. While the gap narrows at higher quantile, inventions from US research institutes fail to beat the baseline group at 98% quantile.

Now we turn to US corporations, the only type of for-profit recipients of federal funding examined here. Although we cannot determine whether federal funding offers corporations the proper incentive to develop more fundamental inventions, the results do suggest that corporate R&D projects that receive federal support do not produce inventions as influential as other recipients. This is evidenced in the significant and negative results on the influence of corporate patents that receive federal funding at all quantiles examined. More importantly, the gap in influence widens as analysis shifts to higher quantiles. To conclude, the influential inventions funded by the federal government appear more likely to originate from US universities and hospitals. However, we do not find similar evidence for research institutes. Corporations, on the other hand, appear to be far less effective in transforming federal funding into influential technologies. The superior performance of universities may have something to do with its unique position in creating and disseminating basic knowledge as opposed to the other types of assignees. For example, Fleming and Sorenson (2004) find that university-owned patents are more likely to cite science literature.

Table 5. Quantile regression of funding recipient type on patent influence[1 2 3 4 5]

Quantile	80%		90%		95%		98%	
Federal Funding	**0.581*** **	**(0.054)**	**0.771*** **	**(0.067)**	**0.917*** **	**(0.047)**	**1.073*** **	**(0.068)**
Federal Funding × US Corporation Assignee	**-0.589*** **	**(0.056)**	**-0.794*** **	**(0.072)**	**-0.812*** **	**(0.074)**	**-1.141*** **	**(0.086)**
Federal Funding × US University Assignee	-0.132+	(0.08)	0.037	(0.097)	**0.448*** **	**(0.09)**	**0.703*** **	**(0.102)**
Federal Funding × US Hospital Assignee	**-0.347*** **	**(0.104)**	**-1.725*** **	**(0.12)**	**-1.078*** **	**(0.182)**	**1.437*** **	**(0.139)**
Federal Funding × US Institute Assignee	**-1.431*** **	**(0.107)**	**-0.749*** **	**(0.17)**	**-0.553*** **	**(0.151)**	-0.186	(0.171)

[1] Weighted average estimates from 30 separate quantile regressions based on a group of 4311 federally funded patents and 30 alternative matched control groups.
[2] Robust error reported in parentheses.
[3] + $p<0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.
[4] Number of claims and dummies for 3-digit class, granting year and assignee type included.
[5] All moderating effects estimated in one model.

4.6 Auxiliary analysis of funded subclass: what kind of technological niches are selected by extramural federal funding?

So far all our findings are conditioned on the specific technological subclasses selected by federal funding, and in consequence most subclasses with no federal presence are filtered out by the matching algorithm. Hence, we have not been able to address the question of whether the set of subclasses, or "niches" in the technological space supported by the federal government is substantially different. This is an important question as it directly relates to how we can explain the main findings from the selection side.

We will focus on the presence of extramural federal funding in patent applications in a given technological niche (among all niches in the patent universe defined in Section 2) and a given application year (from 2001 through 2004), and examine how various characteristics of the subclass may influence the chance of federal funding presence in the class in the subsequent year. First, we use the number of patents in a given year (*Patent Count*) as well as the average growth rate in the number of patents ("Avg. Growth in Patent Count (t-3, t-1)") to capture subclass demographics. We further decompose the body of patents in a given subclass-application year by calculating the proportion categorized to each assignee type (US Corporation, US Federal Government, US State Government, US University, and US Institute). We also create a dummy that equals one if no patent is assigned to a certain type of assignee to capture lack of activities ("*No US Corp. Patent*", for example) and zero otherwise. Second, we create several measures to characterize the past impact of the subclass, such as the mean (*Avg. Impact*), normalized standard deviation (*Heterogeneity in Impact*) and average growth rate over the past 3 years ("*Avg. Growth in Impact (t-3, t-1)*") of the 10-year citation count. We also include the average of the diameter of the 10-year forward citation trees to represent average number of generations of technology the focal subclass spawns (*Avg. Generations*). Third, we account for uncertainty in the granting process by including the mean ("*Avg. Time Gap btw. Application and Granting*") and normalized standard deviation (*Heterogeneity in Time Gap*) of the time lapse between application and granting. Last but not least, we measure the commercial value of a subclass by the proportion of the patents that expired later ("*% Expired*"), as well as the proportion of patents whose maintenance fees were paid at the 12th year ("*% Renewed 12th Year: Small Entity*", "*% Renewed 12th Year: Large Entity*").

We use a fixed-effects logit regression to predict the presence of funding in a subclass to account for intrinsic differences among various subclasses. Regression results are displayed in Table 6. Overall, our tentative analysis portrays interesting patterns of selection by the federal agencies. First, extramural funding is likely to be present in subclasses where past patenting activities are present, though this is a decreasing marginal relationship. Second, lack of past inventions from non-profit organizations, including the federal government per se, appears correlated with future federal funding. Meanwhile, federal funding does not seem to be sensitive to the presence of corporation-backed inventions. Third, there is also evidence that federal funding may be more likely to flow into areas of high uncertainty in granting (as measured by time lapse from application to granting), but low uncertainty in realized impact (as measured by the heterogeneity in impact). Fourth, it appears that federal funding is also most likely to be present in areas of either very high or very low commercial value. In conclusion, based on the fixed-effects logit model, we have some primary findings on the selection patterns of extramural federal funding. It appears that the past activities, especially by non-profit organizations, commercial value, as well as the uncertainty in approval and realized value may drive the selection of extramural federal funding.

Table 6. Fixed-effects logit regression of subclass selection of extramural federal funding

DV: Extramural Funding = 1	Coefficient	Std. Error
Patent Count	**0.032*****	**(0.007)**
Patent Count Squared	**-0.000*****	**(0.000)**
Avg. Growth in Patent Count (t-3, t-1)	0.111	(0.082)
% US Corp. Patents	-0.158	(0.217)
No US Corp. Patent = 1	0.248	(0.181)
% US Govt. Patents	**-1.616****	**(0.653)**

No US Govt. Patent = 1	0.042	(0.175)
% State Govt. Patents	-9.657	(7.015)
No State Govt. Patents = 1	-0.072	(0.854)
% US Univ. Patents	**-1.189*****	**(0.416)**
No US Univ. Patent = 1	**0.586*****	**(0.134)**
% US Inst. Patents	**-2.035****	**(0.825)**
No US Inst. Patent = 1	-0.240	(0.205)
% Small Entity Status Patents	-0.044	(0.205)
Avg. Impact	0.012+	(0.007)
Heterogeneity in Impact	**-0.310*****	**(0.089)**
Avg. Growth in Impact (t-3, t-1)	0.009	(0.042)
No Impact in t-3 to t-1	-0.163	(0.238)
Avg. Generations	0.094+	(0.052)
Avg. Time Gap btw. Application and Granting	**0.522*****	**(0.085)**
Heterogeneity in Time Gap	**0.598****	**(0.252)**
% Expired	**0.486*****	**(0.164)**
% Renewed 12th Year: Small Entity	**0.944****	**(0.401)**
% Renewed at 12th Year: Large Entity	**0.972*****	**(0.189)**
N	3752	
[1] + p<0.10, * p<0.05, ** p<0.01, *** p<0.001		

4.7 Auxiliary analysis of exogenous shock to funding availability and patent influence

While the analysis performed in 4.6 offers some insight into the selection patterns of federal funding, it still leaves readers who seek rigorous causal implication unsatisfied. Given the data constraint and complexity of the funding and knowledge production process, the issue of identifying causal mechanisms of federal funding on innovation is a crucial, yet thorny one. Specifically, in order for causal inference to be valid, the unobserved characteristics of treated and control patents are assumed to be uncorrelated with the funding-related variables in our earlier analyses. To alleviate concerns regarding the unobservables and offer more convincing evidence of causality, we need to identify conditions where there is exogenous shift in the funding decisions while there is no reason to expect a shift in the quality of general patent pool.

We attempt to exploit the surge in R&D budget of Department of Defense in the post-911 period and conduct an approximate dif-in-dif estimation to compare post-911 DoD funded patents with both the pre-911 patents (pre-treatment group) and the matched sample (control group). After the September 11th terrorist attack in 2001, the R&D budget of DoD soared to over $64 billion in Fiscal Year 2002, a 14.65% increase that surpassed the budget growth of the five previous years combined (American Association for the Advancement of Science, 2015). The budget increase in fiscal year 2002 is followed by an additional, bigger increase of 16.66%, boosting DoD's R&D budget to $75 billion in fiscal year 2003, a record for last several decades.

Our premise is that the unexpected surplus in available R&D funding a ther DOD caused an exogenous change in funding decisions. The risky and fundamental projects with potential for high influence received financial support as the concerns over budget constraint were alleviated, whereas without the increase in funding they would not have been funded. Empirically, we analyze the matched sample of patents intramurally funded by DoD, or those patents funded and developed in-house within DoD. By restricting the analysis to intramural sample, we are able to rule out the alternative explanations of how the shock simultaneously impact other institutions where extramural patents are developed. We use the year of application to identify the pre- and post-shock period. While the year of application does not map to the year of funding for the originating R&D project of the patent application perfectly and thus may be noisy for our purpose, it does mark the pre-treatment period clearly because patent applications in 2001 by definition are most likely to be funded prior to the 911 shock. In the post-treatment period (2002-2004), some patent applications may still be the outcome of funding from pre-treatment period (2001 and earlier), which may bias the estimates downward. Nevertheless, it is reasonable to expect that more recent years would see higher proportion of the inventions funded in the post-911 period.

We interact the variable for DoD's intramural funding with each application year in the post-911 period in the sample to allow for flexible parametric assumptions on the timing of the funding and application and report our results in Table 7. As the results demonstrate, both direct impact and long-term influence of post-911 patents developed by DoD are significantly superior to their matches and to the pre-911 counterparts, which is consistent with our argument that exogenous change in funding constraints causes the shift in the long-term value of the projects selected.

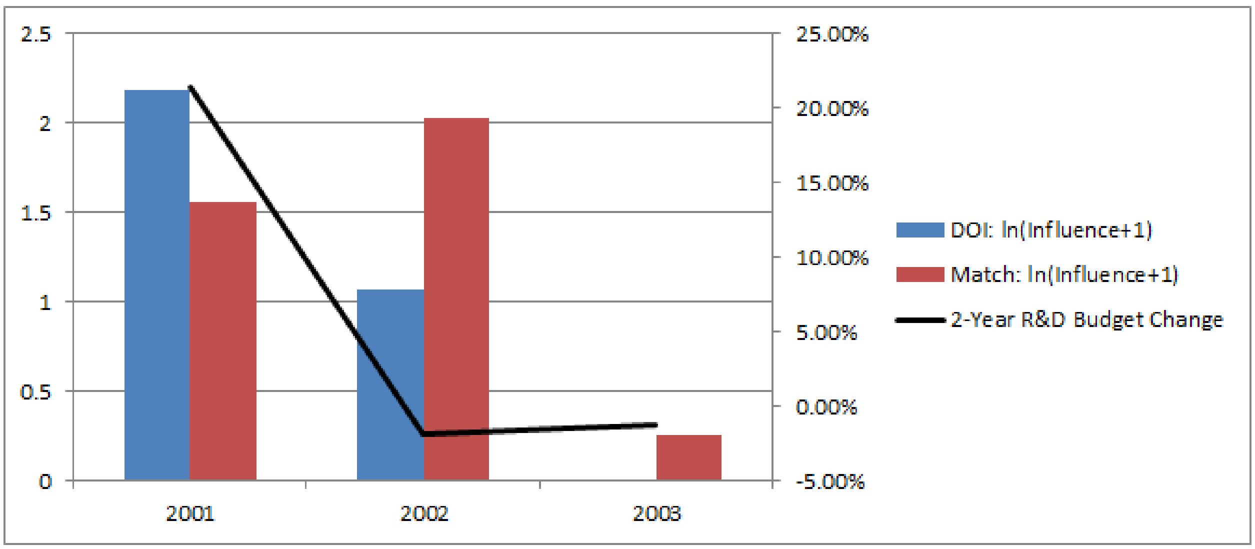
Table 7. Difference-in-Difference Analysis of the Impact and Influence of Post-911 DoD's Intramural Inventions[1 2 3]

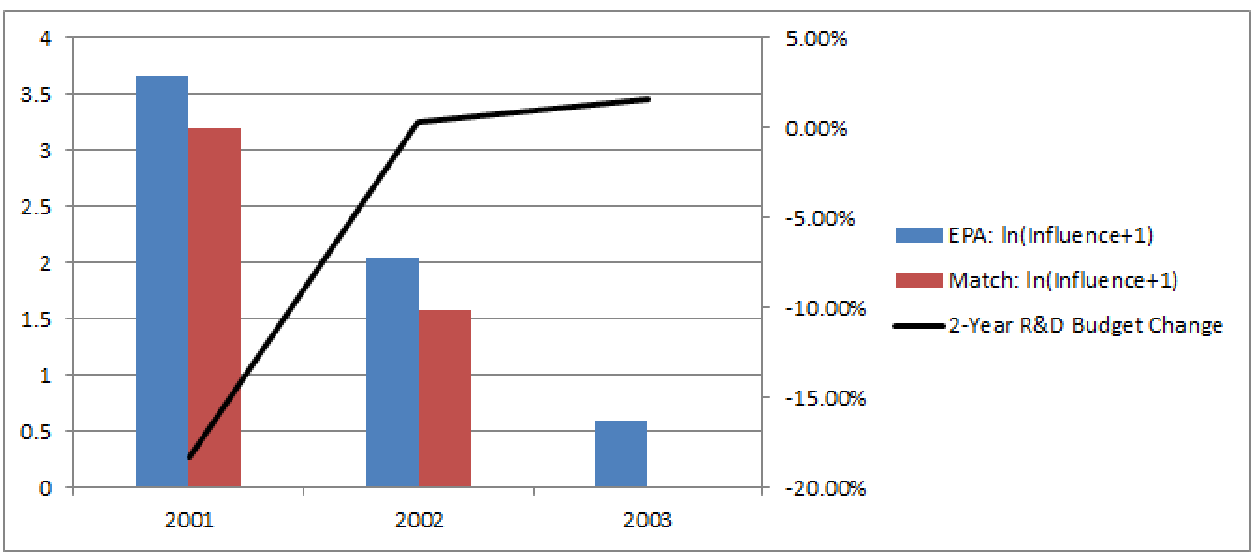
DV	ln(Impact+1)		ln(Influence+1)	
	Coefficient	Std. Error	Coefficient	Std. Error
DOD Intramural	-0.1255243	(0.011)***	-0.2095558	(0.023)***
DOD Intramural * App. Year = 2002	0.0835929	(0.017)***	0.1381722	(0.034)***
DOD Intramural * App. Year = 2003	0.2155291	(0.021)***	0.386998	(0.040)***
DOD Intramural * App. Year = 2004	0.2709044	(0.022)***	0.4763767	(0.042)***
Grant Year Dummies	Yes		Yes	
Application Year Dummies	Yes		Yes	

[1] Weighted average estimates from 30 separate OLS regressions based on a sample of 987 patents sponsored by DoD intramural funding and 30 alternative matched control groups.
[2] Robust error reported in parentheses.
[3] + $p<0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

While we cannot further examine the implications of other major R&D budget shifts in major agencies (e.g., budget sequestration in NASA and NIH) for innovation due to data limitations, we are able to qualitatively compare and contrast the trend of average patent influence at two smaller agencies, Department of Interior (DOI) and Environmental Protection Agency (EPA), which are comparable in size, similar in research mission, and subject to the same appropriation subcommittee under the U.S. House of Representatives. More importantly, the trend of R&D budgets in 1999-2002 were always the opposite at the two agencies. Due to the scarcity of patent records we are able to obtain for these two agencies (20 from DoI and 18 from EPA) and thus the power of statistical tests will be very limited, we will simply plot the average influence of patents applied in year 2001-2003 and the percentage budget change in the previous two years for descriptive purpose. Interestingly, the rough pattern we can observe from Figure 4 below agrees with the findings of Table 7. Budget change seems to precede the change in the relative performance of the patents applied later, suggesting that what we could detect at DoD around the 911 shock may be quite generalizable to broader contexts.

Figure 4





5. Discussion & Conclusions

Our results indicate that Federally funded inventions are more influential in the sense that they are associated with larger citation trees. The result is driven by the greater likelihood for federally sponsored research to be associated with patents that have seminal influence in citation trees. Controlling for technological class and patenting year, the most productive technological trajectories are likely to build on government funded inventions. However, the median government funded and privately funded patents are no different in terms of long term citations.

The results are consistent with the notion that *"Generally speaking, the scientific agencies of Government are not so concerned with immediate practical objectives as are the laboratories of industry nor, on the other hand, are they as free to explore any natural phenomena without regard to possible economic applications as are the educational and private research institutions."* Research results are hard to predict, and the long term implications of results or even entire fields of research are exceptionally difficult to foresee. If science policy is fulfilling its role, we should expect to see this particularly in the long term. Our results suggest that for all its imperfections, US science policy remains successful in supporting the long term productivity of inventive activity. In preliminary findings, we also find that the government is likely to invest in higher risk technological areas that corporations do not find attractive. Our results are consistent with the notion that the federally sponsored research leads to an increased rate of inventive activity and also the support of areas of research that would otherwise be orphaned or neglected by the private sector. In this sense, federally funded research affects both the rate and direction of inventive activity.

Our results also indicate several areas of concern. We are not concerned that the mean and median federally funded patent is as ordinary as its commercial counterpart. Technological advance is a small tail phenomenon in which influential inventions are rare. However, it is disconcerting that seminality appears confined to a few agencies and is only present with government funding of academic research. Further consideration of when patent influence is a reasonable way to measure the long term effects of federally funded research conducted by for-profit firms as well as federal labs is needed to draw sharp policy recommendations. Moreover, it does not appear to be an apparent trade-off between immediate impact and long-term influence in the tail, indicating that innovations with high level of rigor and relevance might be archived simultaneously under the right circumstances.

The results are also intriguing because they do not map perfectly with the emphasized type of research (applied vs. basic) of the agencies. On the one hand, the National Science Foundation - the sole agency whose mission is to support basic research, is more likely than some of its peers to produce useful research results that are critical building blocks that support generations of patents. In contrast, the Department of Energy and National Aeronautics and Space Administration whose emphasis is more applied, appear less likely to generate path-breaking technologies. Of course, due to their sheer size, the bulk of influential results come from research funded by the Department of Defense and the National Institutes of Health. A more in-depth program level analysis will be necessary to understand the sources of these differences. We must also consider that the influence of the DOD might be underestimated if the most advanced technologies under development are classified.

We should be careful in drawing broad conclusions from our research. For example, we are constrained by the limited timeframe of our analysis. We examine only patents from 2001 through 2004, and this in turn limits the time window with which we can measure influence. Although this still allows us to track patent trees across several generations, patterns in some slower moving fields may be obscured. It should also be noted that some of our tests in the quantile regressions rely on a small population of home run inventions. While the quantile regression methodology is not sensitive to the precise magnitudes of outliers, perhaps small shifts in the influence of a few inventions may affect our results in many of the more fine grained sub-agency analyses. Replication using other samples of government funded patents will be necessary to draw stronger conclusions. In addition, our ability to sample earlier patents was limited due to difficulty in observing federal sponsorship prior to 2001. However, in future analysis we can abstract away from funding sources and examine university versus corporate assignees directly. This will allow a more direct reexamination of Henderson, Jaffe and Trajtenberg (1998).

Although we are able to associate greater influence with university-assigned patents funded by particular government agencies, we cannot identify the mechanism by which these differing citation patterns occur. In some preliminary findings, we note that the qualitative differences in the types of classes populated by federally funded patents suggests that federally funded research is different in kind. However, there are likely institutional differences in the propensity to patent as well as the strategic nature of technological choices post patenting. For example, David and Hall (2000) point out that positive spillovers (or more negative crowding out) of private sector R&D efforts by publicly sponsored R&D may operate across many mechanisms. Federally funded research may generate spillovers through a selection effect: by increasing the marginal productivity of privately funded research due not the direct contribution of a unique type of knowledge only likely to be generated through federal support direct contribution of knowledge, or through a treatment effect associated with greater ease in the use of federally funded research results as compared to privately funded research results. We leave it to future researchers to attempt to further disentangle these mechanisms.

In addition, due to data constraints, we cannot offer further insight on the implications of project characteristics. While the federal research statements in some patent applications further report information on the specific grants or contracts sponsored by the agency, such information is noisy and proves to be difficult to be matched to the data on federal grants and contracts reliably despite our various efforts. Moreover, even if the information on the amount of project funding is available for all federally funded patents, it is challenging to draw implications on how much money is allocated to the specific invention as compared to other inventions and scientific papers from the same project. As a result, with our current data it is difficult to shed light on issues such as what kind of federally funded projects are more likely to spawn influential inventions or the efficiency of federal and non-federal R&D dollars in spurring future technological development. Future researchers who can overcome these data challenges may be able to address these important questions. It may also be possible to pursue alternative strategies based on inventor name and geography.

By definition the influence of a patent application yet to be granted is not possible. For this reason, our analyses are conditioned on the success of a given patent application. Nevertheless, by focusing on granted patents, we may overlook the pre-selection process of USPTO. The working assumption for our main analyses is that patent applications are examined based on merit only, and should not be subject to different treatment based on the source of funding support. While thoughtful readers may argue that patent examiners may interpret the information of federal government sponsorship as a certification or signal of quality and act differently towards those federally funded applications, it cannot explain the difference in all indirect citations to the federally funded patents, which is well incorporated in the influence measure. future research may explore the possibility of learning about the ex ante assumption held by the patent examiners and estimating how much of the difference in granted patents can be attributed to the application process. Nevertheless, such a bias would likely work against our findings, as it would lead to the lowering of at least the average quality of federally funded inventions.

References

Aghion, P., & Tirole, J. (1994). The management of innovation. *The Quarterly Journal of Economics*, 1185-1209.

Alcacer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics* 88(4) 774-779.

Bikard, M. A. (2012). Is academic science trapped inside the ivory tower? Universities and the diffusion of new knowledge. *Academy of Management Proceedings*, 2012(1), 1-1.

Budish, E., Roin, B. N., & Williams, H. (2015). Do firms underinvest in long-term research? Evidence from cancer clinical trials. *American Economic Review*, *105*(7), 2044-2085.

Cohen, L., & Noll,, R. (1991). *The Technology Pork Barrel.* Washington, D.C.: The Brookings Institution.

Cohn, L. D., & Becker, B. J. (2003). How meta-analysis increases statistical power. *Psychological Methods*, *8*(3), 243.

Corredoira, R. A., & Banerjee, P. M. (2015). Measuring patent's influence on technological evolution: A study of knowledge spanning and subsequent inventive activity. *Research Policy*, 44(2), 508-521.

David, P. A., & Hall, B. H. (2000). Heart of darkness: modeling public–private funding interactions inside the R&D black box. *Research Policy*, *29*(9), 1165-1183.

Fleming, L., & Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, 25(8- 9), 909-928.

Goldfarb, B., (2008). The effect of government contracting on academic research: Does the source of funding affect scientific output?, *Research Policy*, 37(1), Pages 41-58, ISSN 0048-7333, http://dx.doi.org/10.1016/j.respol.2007.07.011.

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2001). *The NBER patent citation data file: Lessons, insights and methodological tools*. NBER working paper 8498.

Hall, B., Mairess, J and Mohnen, P. (2009), Measuring the Returns to R&D. NBER working paper 15622.

Hegde, D., & Mowery, D. C. (2008). Politics and funding in the US public biomedical R&D system. *Science*, *322*(5909), 1797-1798.

Henderson, R., Jaffe, A. B., & Trajtenberg, M. (1998). Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. *Review of Economics and statistics*, *80*(1), 119-127.

Howell, Sabrina. (2014). "Financing constraints as barriers to innovation: evidence from R&D grants to energy startups" Working paper, Harvard University.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3), 577-598.

Lichtenberg, F. R. (1987). The effect of government funding on private industrial research and development: a re-assessment. *The Journal of Industrial Economics*, 36(1),97-104.

Mowery, D. C., Sampat, B. N., & Ziedonis, A. A. (2002). Learning to patent: Institutional experience, learning, and the characteristics of US university patents after the Bayh-Dole Act, 1981-1992. *Management Science*, 48(1), 73-89.

Mowery, D. C., & Ziedonis, A. A. (2002). Academic patent quality and quantity before and after the Bayh-Dole act in the United States. *Research Policy*, 31(3), 399-418.

Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. *Handbook of the Economics of Innovation*, *2*, 1083-1127.

National Institute of Allergy and Infectious Diseases. (2015). Extramural and Intramural Research Questions and Answers. http://www.niaid.nih.gov/researchfunding/qa/pages/extraintra.aspx. Accessed on Aug. 30, 2015.

National Science Foundation. (1995). Federal Funds Glossary. http://www.nsf.gov/statistics/fedfunds/glossary/def.htm. Accessed on Aug. 30, 2015.

Sampat, B. N. (2006). Patenting and US academic research in the 20th century: The world before and after Bayh-Dole. *Research Policy*, *35*(6), 772-789.

Sampat, B. (2010). When do applicants search for prior art? *Journal of Law and Economics* 53(2), 399-416.

Sargent, John F. (2015). Federal Research and Development Funding: FY2015. Congressional Research Service Report, R43580, 7-5700.

Thompson, P., & Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, 95(1), 450-460.

United States General Accounting Office. (1999). Reporting Requirements for Federally Sponsored Inventions Need Revision, GAO/RCED-99-242, 1-35.

Wang, D., Song, C., & Barabási, A. L. (2013). Quantifying long-term scientific impact. *Science*, *342*(6154), 127-132.