Final Project Report - Team 3

# **Impact of U.S. National Science Foundation funding on patenting output: Sociodemographic snapshots**

Authors

Ernest Alema-Mensah; Morehouse School of Medicine; Georgia

Rachael Allbritten; U.S. National Science Foundation; Virginia

Abdel Lawani; Tennessee State University; Tennessee

Christine Swanson; U.S. National Science Foundation; Virginia

Linge Yang; University of Connecticut; Connecticut

Karina Vega-Villa; U.S. National Science Foundation; Virginia

# **Introduction**

Innovation is a critical driver of economic growth, and understanding the factors that contribute to disparities in innovation output is essential for shaping effective science and technology policies. This report describes the project our group (Team 3) completed for the Coleridge Initiative’s Advanced Data Analytics course entitled, “Exploring the Diverse Paths to Invention of U.S.-Trained Doctoral Scientists and Engineers.” Our team shared a keen interest in understanding differences in innovation output levels produced by different groups of people working in Science, Technology, Engineering, and Math (STEM). This curiosity stemmed from observing that some groups have higher innovation output than others, and we were eager to explore the factors contributing to this variability. Therefore, the goal of this project was to identify the correlations among these differences and their relationship to funding from the U.S. National Science Foundation (NSF). A better understanding of these relationships could help inform paths to fostering a more diverse, inclusive, and equitable STEM community.

Team 3 was composed of individuals from a variety of professional backgrounds, bringing a rich diversity of perspectives and expertise to our work. Among us, we had a graduate student just beginning her journey in the world of STEM, bringing fresh ideas and a strong enthusiasm for innovation. We also had an assistant professor in the early stages of his academic career, contributing to the academic society. Additionally, our team included a tenured professor who provided valuable insights and mentorship. Finally, we had staff members from NSF, who bring a deep understanding of the broader scientific landscape and policy implications. This blend of backgrounds allowed us to tackle complex problems from multiple angles and foster a collaborative environment that achieves our goals.

By linking patenting data to the 2015 Survey of Doctorate Recipients (SDR), our group provides novel insights into the role of NSF funding in promoting high-impact innovation and highlights areas where policy intervention may be necessary. Specifically, we investigated the relationships between patent productivity — measured by the number of U.S. patents and the mean patent citations — and funding from NSF across all respondents of the 2015 SDR. Further, we examine whether relationships between NSF funding and patent productivity vary based on demographic characteristics (e.g., gender, race/ethnicity, discipline, and job sector). This investigation aligns with our broader goal of understanding the variability in innovation output among different groups in STEM, as mentioned earlier. We believe that by identifying and addressing any disparities, we can foster a more sustainable U.S. trained STEM workforce. Our finding not only aims to enhance our understanding of innovation dynamics but also to promote policies and practices that support diverse and underrepresented groups in STEM.

# **Literature Review**

The relationship between government funding and innovation output has been previously explored, particularly in the context of NSF funding. Previous studies highlighted how NSF funding influences patenting activity. Huang et al (2006), a group associated with NSF’s National Nanotechnology Initiative, investigated the correlation of NSF funding in Nanotechnology with both the number of resulting patents and number of citations of those patents. For nanotech at least, they found that NSF-funded patents had higher "impact" or "influence" than other patents based on higher citation rates. Their work also found that the rate of NSF-funded patent citations was increasing over time (2001-2004). The research team interpreted this as an indicator that NSF-funded nanotechnology research was more fundamental, with more staying power in the long term, with implications that NSF investments in innovation are worthwhile.

Further examining the differential effects of government funding, Shelton and Leydesdorff (2011) demonstrated that the goals of innovation funding vary by sector. In higher education, government funding tends to encourage the production of publications, which are a key metric of academic success. Conversely, in the business sector, funding is more likely to drive patenting activity, reflecting the sector’s focus on monetizing discoveries. This sectoral divergence in the goals of funding highlights the need for tailored government policies that can balance the objectives of fostering broad knowledge dissemination through publications and promoting commercialization through patents.

The issue of equitable participation in the innovation process is also a significant concern. Cook (2020) found persistent disparities in patenting activity among women and racially minoritized groups, despite their growing presence in STEM fields. This lack of increase in patent activity suggests that systemic barriers continue to hinder these groups from fully participating in the innovation economy. Both Cook (2020) and Bell (2019) argue that broader participation in innovation not only enhances individual well-being but also benefits the overall U.S. economy. These findings are supported by Shaw and Mariano (2021), who suggest that federal programs aimed at broadening participation could be instrumental in fostering the commercialization of inventions while addressing the systemic barriers that contribute to these disparities.

In addition to these specific studies, other literature has further elucidated the role of government funding in driving innovation across different contexts. Toole (2007) examined the pharmaceutical industry and found that public scientific research complements private research and development (R&D) investments, leading to increased innovation output. This is particularly important in industries where the risks and costs of initial research are high, underscoring the critical role of government funding in supporting foundational research that might not otherwise be pursued. David, Hall, and Toole (2000) provided a broader review of the interaction between public and private R&D, concluding that public R&D funding generally acts as a complement rather than a substitute for private investment. This complementarity enhances overall innovation activity, particularly in cases where private sector funding alone may not suffice. Moreover, Jaffe (1989) highlighted the significant contribution of academic research, often funded by government sources, to industrial innovation. By linking patent data to university research, Jaffe demonstrated that government-funded academic research is crucial for driving innovation in industries such as chemicals and electronics.

Czarnitzki and Hottenrott (2011) focused on small and medium-sized enterprises (SMEs) and found that government R&D subsidies were essential for overcoming financing constraints that limit these firms' ability to invest in innovation. This support was particularly vital for SMEs, which played a significant role in fostering innovation but often lacked the financial resources to sustain R&D activities independently. Finally, Feldman and Kelley (2006) emphasized the importance of government R&D policy in maximizing “knowledge spillovers”, or knowledge that benefits not only the immediate recipients of funding, but also other firms and industries or other projects and activities within the company receiving the funding. This is especially true if there are specific collaborations or linkages connecting organizations or projects. They argued that effective government policy should be guided by forecasting potential spillovers to enhance the overall innovation output.

# **Research Question**

On the NSF website homepage, NSF highlights three priorities for the agency: 1) Promote discovery in science and engineering; 2) Accelerate technology and innovation; and 3) Advance diversity in science and engineering. Given these NSF Priorities, our group wondered how these national aspirations were actually manifesting. These questions led to our research question: **Does NSF research funding relate to patent output across different populations of PhD-holders?**

Our project investigated the relationship between NSF funding and patenting output among PhD holders, with a focus on how these relationships vary across socio-demographic groups. Utilizing data from the 2015 Survey of Doctorate Recipients (SDR) and the U.S. Patent and Trade Office’s (USPTO) PatentsView, we found that NSF-funded patents generally receive more citations, indicating higher impact innovation. However, significant disparities exist in patenting rates across gender and racial/ethnic groups, reflecting broader inequities in the innovation economy. Our project supports the conclusion that the U.S. should continue work toward more inclusive innovation ecosystems.

# **Methods**

To better understand the implications of NSF funding on patent productivity, we combined the 2015 SDR with data from PatentsView. The SDR is conducted annually by the NSF’s National Center for Science and Engineering Statistics (NCSES) and is a representative survey of individuals who have earned a research doctorate at a U.S. institution. The survey collects detailed demographic information, educational history, funding sources, and postdoctoral plans from survey respondents. Though the SDR is a representative sample survey, it also contains pre-calculated population-representative weights, allowing us to extrapolate patterns to the overall population. PatentsView is maintained by the U.S. Patent and Trade Office (USPTO) and includes the names of individuals awarded patents, funding sources for patents disaggregated by U.S. government agencies, and the number of times each patent has been cited. PatentsView determines whether a patent was funded by a specific government agency based on whether that patent acknowledged funding from that agency. We used a linking method provided by the Coleridge Initiative team to combine the SDR with PatentsView data, thereby reducing the total pool of successful patent-holders to only those who had responded to the 2015 SDR. We refer to the linked SDR and PatentsView data as the “linked dataset”.

We used the linked dataset to estimate the proportion of PhD-holders who have patented in the U.S., then performed the rest of our analyses on this subset of doctoral recipients. We performed our analyses using the number of patents and average citation count of patents for each patent-holder with a PhD and compared the differences between patents that cited NSF funding and those that did not. We also looked at how many times each patent was cited by other patents. Following Huang et al. (2006), we interpreted high levels of patent citation as indicative of “impact innovation”.

We described PhD patenting patterns for the overall population and for specific demographics, subdividing patent-holders by gender, race/ethnicity, disciplinary field, and job sector (i.e., academia, industry, government). We then examined differences in patent productivity and impact innovation for those patent-holders who were funded by NSF and those who were not funded by the agency. Though we intended to also conduct an intersectional analysis by combining demographic variables (e.g. gender and race/ethnicity, race/ethnicity and job sector, etc.), this severely restricted sample sizes therefore limiting any conclusions that could be made from the data.

We visualized the results of our analysis in an interactive Dashboard. We first showed the overall percentage of PhD patent-holders in the population and within each demographic category with donut charts. Then, we examined the differences in number of patents and average patent citation count for members of different demographic groups by using ridge plots of the distributions for each variable. For each ridge plot, the x-axis shows numbers of U.S. patents or average patent citation count, and the y-axis shows the proportion of survey respondents that had that number of patents or average citations. We used the same visualizations to compare patent productivity for holders of patents funded and not funded by NSF.

# **Key Findings**

Based on the 2015 linked dataset, only a small proportion (17.4%) of PhD-holders also go on to hold a U.S. patent. Of those patent-holders, the vast majority are men (85.8%). U.S. patent-holders in 2015 were also mostly non-hispanic whites (67.8%) or Asians (26.3%). Three fields dominated U.S. patenting by PhD-holders in 2015: engineering (38.5%), physical sciences (27.6%), and life sciences (21.5%), and 66.2% of patent-holders held a job in industry while only 5.76% had a government job. Patent productivity as measured by both the number of U.S. patents and mean patent citation count had a bimodal distribution (Figure 1). Patent-holders had a median of four patents and as many as 18 patents with a mean citation count between one and 41 citations per patent.

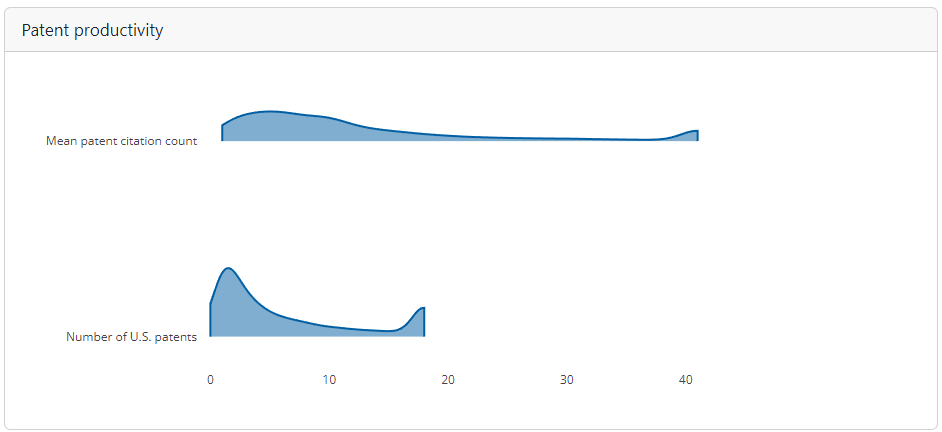
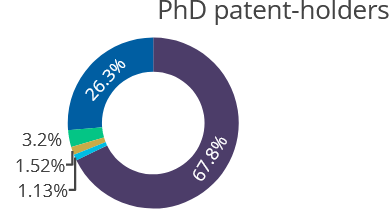


Figure 1. U.S. patent productivity in 2015 as measured by the number of U.S. patents (lower figure) and mean patent citation count (upper figure). The bimodal distribution indicates, for example, that the highest number of people have lower numbers of patents, but a moderately high number of people have 15-18 patents. Fewer people have total patents between the two extremes.

Only 7.1% of respondents to the 2015 SDR held NSF-funded patents. The majority of respondents with NSF-funded patents were men (85.9%). Interestingly, a lower proportion of non-Hispanic white patent holders had NSF-funded patents than non-Hispanic Asians did. Specifically, while non-Hispanic White patent holders made up 67.8% of all patent holders and 62.7% of NSF-funded patent holders, non-Hispanic Asians made up 26.3% of all patent holders, but 30.5% of NSF-funded patent holders (Figure 2).



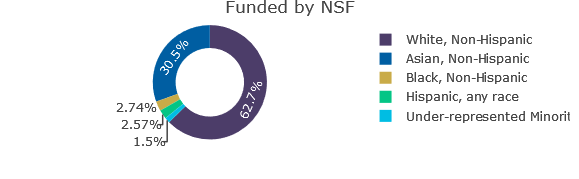


Figure 2. Proportion of Ph.D. patent-holders by race/ethnicity and proportion of NSF-funded patents by race/ethnicity of inventor.

Additionally, a higher proportion of patents in physical sciences had NSF funding than the proportion of total patents in that discipline category, with 31.5% with NSF-funding vs. 27.6% of total patents being in physical sciences. The same was also true of engineering (51% NSF-funded vs. 38.5% of the total). The converse was true of life sciences, with 10.1% being NSF-funded despite having 21.5% of the total share of patents. In addition, academics were much more likely to have an NSF-funded patent, representing 64.3% of patent-holders funded by NSF, even though they made up only 28.1% of the all patent-holders in the linked dataset. Patents acknowledging NSF funding were correlated with a slightly lower patent citation count (median with NSF funding: 8.3, median without funding: 9; Figure 3), which is counter to the findings of Huang et al. (2006). However, there were a larger number of NSF-funded patents (median: 6; Figure 3) compared to patents not funded by NSF (median: 4; Figure 3).

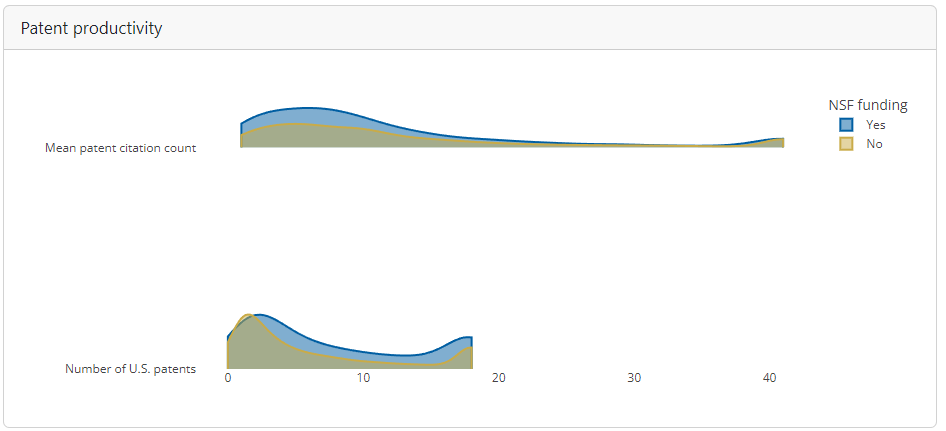


Figure 3. Patent productivity as defined by the number of U.S. patents and mean patent citation count for patents acknowledging NSF funding (blue) or not (gold). In this case, higher numbers of NSF-funded patents had low citation counts.

The pattern of decreased patent citation count for those patents citing NSF funding is consistent across all demographics, except for patent-holders in the government (Figure 4). For this job sector, NSF-funded patents had a median of 8.7 average patent citations compared to a median of 7 average patent citations for patents not funded by NSF.



Figure 4. Patent productivity as defined by mean patent citation count for patent-holders who were funded by NSF (blue) or not (gold) across different job sectors.

Increases in the number of patents with NSF funding were seen for all genders, races/ethnicities, and job sectors. This increase was particularly large for non-hispanic black patent-holders, whose median number of patents increased from two for those who did not cite NSF funding to 6.5 for those who did (Figure 5).



Figure 5. Patent productivity as defined by the number of U.S. patents for patent-holders funded by NSF (blue) or not (gold) across different races/ethnicities.

If NSF determines that encouraging patents is a priority as a measure of STEM innovation, it may need to consider outreach to demonstrate the value of patenting to NSF-supported STEM fields outside Engineering and Physical Sciences. NSF may also want to fund initiatives to better understand the gender gap in STEM research and patenting, then implement evidence-based changes that incentivize groups underrepresented in patenting. Fostering an inclusive environment that values diverse perspectives will lead to transformative research and decision making processes, thereby driving global progress, as well as addressing the challenges of the 21st century.

# **Caveats**

We present our results and conclusions with certain caveats that are important to note. First, given that we only had a single year of data (2015), we were unable to evaluate the data longitudinally to examine patterns in relationships between NSF funding and patent productivity over time. Expanding the dataset across multiple years would allow us to better understand whether relationships between NSF funding and patenting productivity are dynamic and how they have evolved along a timeseries. The current dataset’s binary indication of NSF funding also restricted the scope of the analysis, as we were unable to explore how differences in funding level, for instance, may be related to patent outcomes. Additionally, we did not conduct any significance testing on the observed differences, which limits the conclusions we can draw from these analyses.

Moreover, we do not have data on the geographical location of inventors so we are unable to assess any geospatial factors that could also affect patent productivity. For instance, some states have more advanced industries, a higher STEM workforce concentration, and/or higher average educational achievement, all of which could affect patenting output. Likewise, not all states are equally funded by NSF, which has led to the development of initiatives such as EPSCoR to increase funding rates in underfunded states. Understanding the geospatial variation in patenting output, especially as it relates to NSF funding is crucial to having a full view of the relationships between the two.

# Perhaps most importantly, our ability to conduct intersectional analysis was limited by the small sample size of some sociodemographic intersections. Understanding the nuances in patent productivity for people at the intersection of different genders, races/ethnicities, disciplines, and job sectors is critical to targeting any policy measures that might be developed to spur innovation across diverse groups. While we maintain that some of the most interesting correlations might come out of such an analysis, doctoral patenters with NSF-funded patents represent such a small percentage of the overall patenting pool to obtain meaningful conclusions related to intersectional identities.

# **Possible Extensions**

There are several interesting extensions to the work from this project. Perhaps foremost, it would be interesting to disaggregate our category of patents without NSF funding into patents that cite funding of any kind and those that do not cite funding at all. It is possible that patents without any type of funding had the lowest citation rates. If so, this would lower the average for the entire group. It is also possible that patents with U.S. government funding from any agency would result in higher citation rates. These important caveats mean that this project cannot really conclude that NSF-funding is strongly related to patenting output and innovation impact.

In future directions, adding data for the dollar amount of NSF awards would be an interesting way to create a metric of patent citation per dollar. This would help illuminate the impact: does more research money lead to a higher number of patents and/or citations?

# Of particular importance for addressing patent gaps based on gender and/or race and ethnicity, we may be missing measures that could lead to illuminating relevant correlations. For example, we did not include information about the patenter’s stage, or “academic age,” usually defined by the number of years since an SDR respondent earned a Ph.D. Are academic inventors less likely to patent in early stages of research compared to industry inventors? The ability to address gaps might also be enhanced if we had data on the benefits of patenting for individuals. For example, is there an increased likelihood of earlier promotion/tenure, a raise, or other personal accomplishments? In addition, while other research has established that men patent at higher rates than women (e.g. Cook 2020), it is also true that women are underrepresented in the fields of Engineering and Physical Sciences. Given that a large proportion of NSF-funded patents fall into these fields, normalizing the data for the representation in these fields could yield a different and interesting result.

# **References**

Bell, A., et al. (2019) Who becomes an inventor in America? The importance of exposure to innovation. *The Quarterly Journal of Economics*. 647-713.<https://academic.oup.com/qje/article/134/2/647/5218522>

Cook, L. D.(2020) Policies to broaden participation in the innovation process. *Brookings*.<https://www.brookings.edu/wp-content/uploads/2020/08/Cook_PP_LO_8.13.pdf>

Czarnitzki, D. and Hottenrott, H. (2011) R&D investment and financing constrains of small and medium-sized firms. *Small Business Economics* (36) 65-83 <https://www.jstor.org/stable/41486077>

David P. A., Hall, B. H., Toole, A. A. (2000) Is public R&D a complement or substitute for private R&D? A review of the economic evidence. *Research Policy.* (29) 497-529. <https://doi.org/10.1016/S0048-7333(99)00087-6>

Feldman, M. P. and Kelley, M. R. (2006) The *ex ante* assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Research Policy.* (35) 1509-21. <https://doi.org/10.1016/j.respol.2006.09.019>

Hock, A., Carll, E., & Rajouria, A. (2024, May 28). *Women in engineering and STEM: A review of the 2023 literature*. Society of Women Engineers. <https://swe.org/magazine/2023-lit-review>

Huang, Z, et al. (2006) Connecting NSF Funding to patent innovation in nanotechnology (2001 - 2004). *Journal of Nanoparticle Research.* (8) 859-79<https://link.springer.com/article/10.1007/s11051-006-9147-9>

Jaffe, A. B. (1989) Real effects of academic research. American Economic Review. (77) 957-70. <https://www.researchgate.net/publication/4729106_The_Real_Effects_of_Academic_Research>

Liao Y. (2016) The relationship between STEM educational attainment and utility patent conferrals: A state-level analysis [Master’s thesis, Georgetown University].<https://repository.library.georgetown.edu/bitstream/handle/10822/1040777/Liao_georgetown_0076M_13190.pdf?sequence=1>

Nugent A., et al. (2022) Government funding of university-industry collaboration: exploring the impact of targeted funding on university patent activity. *Scientometrics*. (127) 29-73<https://link.springer.com/article/10.1007/s11192-021-04153-0>

Shaw E. and Mariano, H. (2021) Tackling the gender and racial patenting gap to drive innovation. *Institute for Women’s Policy Research.* 22. <https://iwpr.org/wp-content/uploads/2021/07/Key-Findings_Tackling-the-Gender-and-Racial-Patenting-Gap.pdf>

Shelton, R. D. and Leydesdorff L. (2012) Publish or patent: Bibliometric evidence for empirical trade-offs in national funding strategies. J*ournal of the American Society for Information Science and Technology.* (63) 498-511<https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/asi.21677>

Toole, A. A. (2007) The impact of public basic research on industrial innovation: Evidence from the pharmaceutical industry. *Research Policy* (41) 1-12 <https://www.sciencedirect.com/science/article/pii/S004873331100117X>