I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14297

Opening up Military Innovation: Causal Effects of 'Bottom-up' Reforms to U.S. Defense Research

Sabrina T. Howell
Jason Rathje
John Van Reenen
Jun Wong

APRIL 2021

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14297

Opening up Military Innovation: Causal Effects of 'Bottom-up' Reforms to U.S. Defense Research

Sabrina T. Howell
NYU Stern and NBER

Jason Rathje
U.S. Air Force

John Van Reenen
LSE, MIT, NBER and IZA

Jun Wong
NYU Stern

APRIL 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Opening up Military Innovation: Causal Effects of 'Bottom-up' Reforms to U.S. Defense Research*

When investing in research and development (R&D), institutions must decide whether to take a top-down approach – soliciting a particular technology – or a bottom-up approach in which innovators suggest ideas. This paper examines a reform to the U.S. Air Force Small Business Innovation Research (SBIR) program that transitioned from "Conventional topics," which solicit specific technologies, to "Open topics," which invite firms to suggest any new technology that may be useful to the Air Force. The reform seeks to address challenges facing military R&D, in particular a less innovative defense industrial base. We show that the Open program attracts new entrants, defined as younger firms and those without previous defense SBIR awards. In a regression discontinuity design that offers the first causal evaluation of a defense R&D program, we show that winning an Open award increases future venture capital investment, non-SBIR defense contracting, and patenting. Conventional awards have no effect on these outcomes but do increase the chances of future defense SBIR contracts, fostering incumbency. The bottom-up approach appears to be a mechanism behind Open's success. For example, winning has a positive effect on innovation even in less specific Conventional topics. The results suggest that government (and perhaps private sector) innovation could benefit from more bottom-up, decentralized approaches that reduce barriers to entry, minimize lock-in advantages for incumbents, and attract a wider range of new entrants.

JEL Classification: O31, O32, O38, H56, H57
Keywords: innovation, defense, R&D, procurement

Corresponding author:
John Van Reenen
Department of Economics
London School of Economics
Houghton Street
London, WC1E 2AE
United Kingdom
E-mail: j.vanreenen@lse.ac.uk

* We thank Chris Benson, David O'Brien, Susan Celis, Charles Chimento, Gregory Coleman, Allen Franke, Steve Lauver, Charles Perla, David Shahady, Molly Walsh, Will Roper and many others at the U.S. Air Force who have helped make this research possible. We also thank Pierre Azoulay, Liat Belinson, Adam Jaffe, Saul Lach, Josh Lerner, Danielle Li, Ramana Nanda, Trang Nguyen, Jacquelyn Pless, Claudia Steinwender, Noam Yuchtman, Tom Wollman and seminar participants at the AEA, Census Bureau, LSE, MIT, Rotterdam, NYU and Utah for comments. Howell served as an unpaid Special Government Employee of the U.S. Department of Defense to perform this research, and would like to thank the Kauffman Foundation for financial support. Van Reenen would like to thank the ESRC for financial support.

1 Introduction

Concern regarding the slowdown in productivity growth has escalated in recent decades (Decker et al. 2016, Syverson 2017). This slowdown appears due in part to declining innovation, especially among high-growth new firms, making the design of innovation incentives particularly important (Arora et al. 2018, Bloom et al. 2019, Bloom et al., 2020, Lerner and Nanda 2020). One relatively overlooked but crucial dimension facing a government or private sector research funder is whether to take a centralized "top-down" approach, tightly specifying the desired innovation, or a more open "bottom-up" approach, giving much more latitude to firms to define their research proposals. The bottom-up approach may be useful if the research funder has much uncertainty over what opportunities exist. We compare these two strategies using a major reform to the Small Business Innovation Research (SBIR) program at the U.S. Air Force.

Defense R&D, and the SBIR program in particular, is a useful setting for several reasons. First, the U.S. Department of Defense (DoD) was historically an important financier and early customer for transformational technologies such as jet engines, cryptography, nuclear power, and the Internet (Mowery and Rosenberg 1991, Mazzucato and Semieniuk 2017, Gross and Sampat 2020). Frontier defense technologies have historically had dual-use components that can lead to large private sector spillover opportunities. Second, DoD is one of the largest single investors in R&D in the world and comprises about 60% of total federal government R&D (CRS 2018). Third, the SBIR program is among the world's largest and most influential government programs to spur innovation in small business, spending \$3.11 billion across 11 Federal agencies in 2018. Of this, the DoD accounted for \$1.32 billion, and the Air Force had the largest single SBIR program, at \$664 million. Finally, the defense setting enables us to study the government as a customer rather than a regulator and financier. While there is extensive literature on the latter two roles (e.g. Jaffe and Palmer 1997, Bloom et al. 2002, Denes et al. 2020), the former is quantitatively important in the U.S. and to an even larger degree in many other countries.

Top U.S. policymakers have suggested that the nation's defense R&D increasingly lags behind the private sector, with one reason being that innovation procurement is narrowly specified and siloed in a small group of defense specialist firms, leaving little room for more radical thinking (Cox et al. 2014, Griffin 2019). We document that the U.S. defense sector has indeed been growing *less* innovative compared to the rest of the U.S. economy since the early

1990s, a period which coincides with extensive M&A activity that consolidated the defense industrial base.[1] From DoD's perspective, it is problematic if the best technologies are no longer marketed to the military. From a broader social perspective, there may be significant productivity growth implications from DoD's attenuated role in funding frontier ideas.

To address these issues, the Air Force experimented with "Open" topics in its SBIR contracts starting in 2018. The goals of Open topics are to reach non-traditional firms with frontier dual-use technology and to source ideas that the Air Force may not yet know it needs. The SBIR program is a useful venue for experimentation because it is flexible, with more ability than other procurement methods to adjust contract types and lengths. The program also has faced criticisms of being dominated by incumbent contractors who repeatedly apply and win many contracts, apparently relying on SBIR for revenue and failing to produce technology that is useful for military operations (Edwards 2020). In an Open topic, a firm can propose *any* idea or technology that may be relevant to the Air Force. By contrast, the traditional Conventional SBIR topics are more narrowly specified. Like most mission-oriented R&D programs, they are top-down with particular research ideas generated within the Air Force and firms then invited to complete them.

The Open topic reforms aim to revive DoD's historical role as a large, early customer for risky new technologies from new firms. Sourcing innovative ideas via open solicitations is not unique to this reform, as other government agencies both in the U.S. and overseas have developed similar programs.[2] Companies are also increasingly using bottom-up approaches through customer-driven, outsourced, or open innovation, especially in R&D-intensive industries (Chesbrough 2003, de Villemeur and Versaevel 2019).[3] Whether a bottom-up approach to innovation can be successful is a longstanding economic question (Azoulay and Li 2020). In many cases, a research funder cannot spell out exactly what promising projects will look like, making a bottom-up approach attractive. At the same time, there are potential downsides, especially in the defense context. For example, companies oriented

[1] See Figures 1 and A.1-A.3. Carril and Duggan (2020) show that the substantial consolidation among major defense contractors in the mid-1990s reduced competition.

[2] These include the U.K.'s Defense and Security Accelerator Open Call for Innovation (UK Defence and Security Acceleator 2020), DoD's DARPA, and the U.S. Department of Energy's ARPA-E (Advanced Research Projects Agency-Energy 2020).

[3] For example, Unilever's Open Innovation platform, launched in 2010, invites the public to submit ideas for potential adoption by the company in broad product areas. Successful submitters may be offered a commercial contract for their solution, and today more than 60% of Unilever's research projects involve external collaboration. See https://www.unilever.com/about/innovation/open-innovation/ and https://www.warc.com/newsandopinion/news/open_innovation_boosts_unilever/30488

towards private sector commercialization may not deliver technologies that are useful to the Air Force.

In this paper, we assess how these reforms affected selection into applying for an SBIR award as well as the causal effect of winning an award. This is to our knowledge the first causal analysis of a defense R&D program. We use administrative data on applications and evaluations of Air Force SBIR proposals over the 2003-2019 period (and outcomes through January 2021). We focus on 2017-2019 to facilitate comparison of Open and Conventional topics, which were run simultaneously in 2018 and 2019. In the baseline 2017-19 sample, the data include 7,300 proposals across 3,200 firms. The larger sample of applications from 2003 includes 19,500 proposals from 6,500 firms.

We first show that the Open topics reached a dramatically different type of firm. Compared to firms applying to Conventional topics, Open topic applicants are about half as old, half as large, less likely to have previous Air Force SBIR awards, and more likely to be located in an entrepreneurial hub.[4] Their technologies are also more often software- rather than hardware-based. In sum, Open applicants appear much more representative of potentially high-growth startups in the U.S.

Next, we assess the effect of winning an SBIR contract on two main outcomes: Future venture capital (VC) funding and non-SBIR DoD contracts. These capture benefits to the wider private sector as well as to the DoD, and correspond to the program administrators' key success metrics. VC represents high-growth innovation potential and leads to spillovers (Gornall and Strebulaev 2015, Lerner and Nanda 2020). DoD contracts indicate that the technology may be useful in an operational DoD mission. Our design compares firms by evaluators' rank around the cutoff for an award. The rank that determines the award decision is constructed by the forced ordering of independent scores from three evaluators, and the cutoff is independent of the evaluation process, making manipulation of any firm around the cutoff extremely unlikely. We document a smooth density around the cutoff and continuity in baseline covariates.

We find that winning an Open topic competition increases the probability of subsequent VC investment by 5.4 percentage points, which is 68% of the mean among Open applicants. There is no significant effect of winning a Conventional competition. Second, we find that winning an Open award increases the chances of a subsequent non-SBIR DoD contract by 7.5 percentage points (51% of the mean). Again, there is no effect for Conventional awards.

[4]Entrepreneurial hubs are defined as the San Francisco Bay area, greater Boston, and New York City.

The Open program does not seem to have crowded out the causal effects of the Conventional program because we also find no effects of winning a Conventional award before the Open program was introduced.

We also consider patenting as an alternative measure of commercial innovation to VC. Winning an Open award increases the chances of a patent by about twice the mean. There is no significant effect in the Conventional program. Winning an Open award also increases patent originality, while Conventional does not and also has no effect on citations in the longer period, where there is enough time for citations to accrue. The final outcome is future Air Force SBIR awards. We find that winning an initial Conventional award increases the chances of winning a future SBIR award, while there is no effect in the Open program. This implies that Open competitions do not appear to suffer from the same lock-in and persistent dominance of recurring SBIR-winners that seem prevalent in the Conventional program.

In sum, the Open SBIR program has strong effects on measures of innovation (VC and patenting) and on converting these ideas into new non-SBIR contracts. By contrast, the strongest causal effect of winning an initial Conventional award is to help the firm win more SBIR awards in the future (which is not a particularly desirable outcome from the DoD perspective). Having shown that the Open reforms appear successful, we investigate whether these effects primarily reflect a different composition of applicants or whether openness – that is, the bottom-up approach – plays an independent role as well.

We probe this question with three strategies. First, we conduct heterogeneity tests to assess whether the Open program's ability to attract a different mix of applicants explains the results. We find that the treatment effect on VC is driven by new entrants, both in the sense of young firms and those without previous Air Force SBIR awards. This is consistent with Howell (2017)'s evaluation of U.S. Department of Energy (DoE) SBIR grants.[5] There are, however, no effects of these interactions on other outcomes or in the Conventional program for any outcome. Thus, while attracting new entrants clearly plays a role in the success of Open for VC, there appear to be other forces at work. Second, to more directly assess the role of selection, we use data from two other Air Force SBIR reforms. These had specific topics but, due to other features such as faster contracting and outreach to startup hubs, attracted firms similar to those in the Open program. We find that the Open program had significantly larger

[5] However, the DoE SBIR program seems to have much stronger average effects than the Conventional Air Force program. There are many possible reasons for the difference, but an obvious one is DoD's massive procurement capability and much larger SBIR program, which allow firms to focus solely on the defense market. We show that DoD's SBIR program has more repeat awardees than the DoE program.

effects than these other programs.

Our final exercise on mechanisms uses machine learning techniques on application abstract texts to characterize the degree of specificity for each topic (essentially through examining the similarity of text across applications within a given topic). We show that when a Conventional topic is less specific – and thus closer to the Open program's bottom-up approach – winning a Conventional award topic has a significantly larger positive effect on innovation, using measures of patent quality and quantity.

The Open reforms seem to work because they provide firms with an avenue to identify needs the Air Force did not previously know they had. The Open contracts may represent an entry point to much larger DoD contracts, which helps to explain their large effect, especially on VC. Startups with a successful Open Phase 1 can bring evidence to VCs that large defense customers are interested in their commercially-driven development efforts, which appears to improve their odds of raising funds. While high-growth startups appear poorly aligned with the Conventional SBIR program, which requires awardees to produce a particular, ex-ante specified technology, the Open program allows firms to bring something to the SBIR program that is their own existing idea oriented primarily to the civilian commercial market.

Our results suggest that a more bottom-up approach to innovation that encourages new entrants can have significant payoffs to firms, the military, and ultimately consumer welfare through enhanced innovation. We cannot be sure that a wholesale switching to Open would be best for the DoD or socially optimal, because there may be unobserved benefits from the Conventional program. However, we conduct a simple cost-benefit analysis to explore the magnitude of the net benefits to DoD from running a marginal additional competition, comparing Conventional to Open topics, and find that the net benefits of a simplified marginal Open competition, at about $1 million, are much larger than those of an equivalent Conventional competition, at $1 thousand.

The structure of this paper is as follows. First, we describe our contribution to the literature in Section 1.1. Next, Section 2 documents the decline in U.S. defense contractor innovation in recent decades and discusses the SBIR program and its recent reform. Section 3 contains the data and descriptive statistics. Section 4 explains the empirical design. We describe the results in Section 5, the mechanisms is Section 6 and conclude in Section 7. Online Appendices give more details on the slowdown in U.S. defense innovation (A), institutional details on the the reforms (B), data (C), further results (D), Phase 2 (E) and the cost-benefit analysis (F).

1.1 Contribution to the Literature

This paper contributes to several strands of literature. First, defense is unique because the ultimate buyer is a monopsonistic government agency providing a public good. This implies a narrow market, but one with potentially high risk tolerance and – particularly in the U.S. – essentially unlimited buying power in the event of success. This situation is quite different from other R&D-funding agencies seeking to promote innovation for private sector deployment in a particular location or sector. There are few attempts to unravel the causal impact of military spending or defense research on economic performance, and to our knowledge no causal evaluation of a defense R&D program. The literature has used military spending as an exogenous shock to demand.[6] Our focus is on the supply side where there is rich historical evidence, but it is mainly anecdotal (e.g. Braddon 1999, Senor and Singer 2009, and Mazzucato 2013). This paper is closely related to work on the crowd-in effects of defense R&D, including Lichtenberg (1984; 1988; 1995), Middleton et al. (2006), Moretti et al. (2020), and Belenzon and Cioaca (2021). Draca (2013) estimates the impact of U.S. defense spending on firm-level innovation and finds that increases in procurement contracts are associated with increases in patenting and R&D.

Second, we contribute to the literature on innovation policy, especially the role of direct government subsidies (Bloom et al. 2019). The literature focuses on two types of R&D policies. There are tax-based policies encouraging R&D analyzed in Hall (1993), Bloom et al. (2002), Moretti and Wilson (2014), Rao (2016), Chen et al. (2018), and Dechezlepretre et al. (2019). These studies generally find positive effects (see Akcigit and Stantcheva (2020) for a recent survey). There is also a body of empirical research on the effect of direct government subsidies on private firms (e.g. Goolsbee 1998, Rathje 2019, Fleming et al. 2019, Pless 2019, Rathje and Katila 2020). Especially related to this paper are evaluations of how R&D grants affect innovation, which includes Lach (2002), Jacob and Lefgren (2011), Bronzini and Iachini (2014), and Azoulay et al. (2019). Analysis of the SBIR program includes Lerner (1999), Wallsten 2000, Howell (2017), and Lanahan and Feldman (2018). Bhattacharya (2018) presents a structural model of R&D procurement in the context of the U.S. Navy SBIR program, suggesting that increasing the number of firm applicants (and forced sharing of ex-post patents) would be socially beneficial, but would not necessarily benefit the DoD.

Finally, our paper is related to studies of how to motivate innovation, such as Manso (2011),

[6] This work includes Ramey (2011), Nakamura and Steinsson (2014), Barro and Redlick (2011), Perotti (2014), and Akcigit et al. (2017).

Azoulay et al. (2011), Nanda, Younge and Fleming (2014), and Krieger et al. (2018). Azoulay and Li (2020) explore how research grant programs can best stimulate innovation. We offer the first effort to compare a bottom-up with a top-down strategy for funding frontier innovation and explore how these strategies interact with firm incumbency.

2 Institutions and Policy: Defense R&D and SBIR Reforms

This section first explains how military investment in innovation is relevant to the rest of the economy (Section 2.1). We then discuss challenges facing defense innovation (Section 2.2) and the SBIR context at the Air Force (Section 2.3).

2.1 The Relationship between Defense and Commercial Innovation

Military R&D has shaped technological advances since antiquity, and defense R&D can serve to both "push" and "pull" civilian innovation.[7] In the U.S., spillovers from defense R&D to commercial applications occur through two primary channels. First, DoD both conducts and funds basic R&D, and is an important source of basic, open-ended funding for university research. This "pushes" private sector innovation by creating new pools of general engineering or scientific human capital and knowledge, for example by funding MIT's Lincoln Lab, or basic research on crucial technologies such as GPS (Belenzon and Schankerman 2013, Babina et al. 2020). Second, the military procures new technologies, creating an early market that might otherwise not exist, and shaping the direction of private sector R&D through its vast spending power. DoD has been willing to fund extremely risky, capital-intensive new technologies that have a potential military application.

Since World War II, the U.S. military has invested in innovation primarily through procurement contracts. Much more so than other Western countries, the U.S. has procured defense technologies from an industrial base that also supplies commercial markets (Flamm 1988). In the 1950s and 1960s, large orders for early-stage technologies such as transistors and integrated circuits were crucial to reducing their prices while improving quality, such that they could ultimately be applied to commercial products (Mowery 2012). Dual-use technologies have many attractions. As a monopsonist in the defense market, it is difficult for

[7] For example, many historians (e.g. Polybius' *Histories)* credit Archimedes in inventing many new technologies in the defense of Syracuse against the Romans in 213–212 BC such as cranes (the "Archimedes' Claw" dragged ships out of the sea).

DoD to create competition among defense contractors. A dual-use technology can be exposed to the discipline of the private market, reducing cost inflation and leading to higher quality.

The theory of procurement, as applied to defense, highlights a hold-up problem in production with large fixed costs in technology innovation and development. As the only customer, once the firm invests, the government can potentially eliminate profits by refusing to pay a high price once the technology is available (Tirole 1986). Furthermore, innovation is a defining characteristic of defense procurement, so incentivizing it effectively is crucial. Other key factors in the government's regulatory problem for defense procurement beyond R&D and monopsony include uncertainty and economies of scale in production (Rogerson 1994). Together, these forces create a rationale for DoD to fund the development stage, in which an innovation is developed for use, tested, and scaled.

2.2 Declining Innovation in the Defense Sector

In recent decades, there have been increasing concerns that the "virtuous cycle" in which American defense R&D investment yields powerful commercial applications and enables unrivaled military supremacy is failing. There are at least four challenges. First, procurement regulations have become more complex and onerous, raising barriers to entry for new firms and contributing to the dominance of the prime contractors (Cox et al. 2014). Second, relevant frontier technologies do not seem to be marketed to DoD. Third, the national innovation ecosystem has shifted away from areas most relevant to defense (Sargent and Gallo 2018). Fourth, prime defense contractors have consolidated, often serve only the defense market, and are perceived as increasingly less innovative. Several senior defense officials have highlighted this last challenge publicly. For example, in 2019, an Under Secretary of Defense tasking memo noted that

> "The defense industrial base is showing signs of age. The swift emergence of information-based technologies as decisive enablers of advanced military capabilities are largely developed and produced outside of the technologically isolated defense industrial base" (Griffin 2019).

Despite these concerns, to our knowledge the evolution of defense contractors' innovation has not been previously documented.

Therefore, we document innovation trends in Appendix A and summarize the results here. We begin with the top eight contractors ("primes") of the past 20 years and document their historical acquisition of other defense contractors.[8] At the beginning of our sample in 1976, there are 225 contractors, which accounted for about 35% of total defense contract value. By 2019, there are just six, which continue to account for 35% of contracts (Figure A.1).[9] This dramatic consolidation was accompanied by a decline in innovation quality. Figure 1(a) shows patent activity for the primes described in Figure A.1, weighted by future citations.[10] Citations are normalized by the average number of citations for all patents in the same CPC3 technology class-by-year cohort so that a number above one indicates the patent is more impactful than the average patent in its class-year. The solid blue line includes all forward citations, and we see a secular decline across the unit threshold where defense patents changed from being relatively more innovative to relatively less innovative within their narrow technology areas.

This pattern is even starker when we include only citations from non-defense contractors, shown in the dashed line, which offers a proxy for knowledge spillovers to the broader economy, versus being insular to the defense industrial base. Here, we also exclude self-citations, where the company cites one of its own previous patents.[11] On this measure, defense contractor patents decline even more steeply, from having 15% more non-prime citations than the average patent in 1976 to 60% fewer in 2019. This relationship is not mechanical from consolidation because we do not count cites from a future acquirer as self-cites. Also, the prime and target share of patents in a class-year has declined over time, so there are not "fewer outside patents to cite" in a class-year (see Figure A.3 Panel A).

This falling relative innovation trend in defense has occurred despite a substantial increase in prime contractors' profits and assets, which is discussed in Appendix A, and shown in Figure A.3 Panels C and D). It is also robust to alternative ways of measuring innovation such as the number of patents and amount spent on R&D (Figure A.3 Panels A and D).

In short, the evidence from this section suggests that senior policymaker concerns about

[8] The top eight contractors are Boeing, Raytheon, Lockheed Martin, Northrop Grumman, General Dynamics, United Technologies Corp, Harris, and L-3.

[9] The value of these contracts increased from around $70 billion (in 2019 dollars) for the 225 companies in the late 1970s to $115 billion awarded to just six companies in 2019. The remaining 65% of contracts did not become more dispersed and the total number of remaining contractors declined slightly over this period.

[10] Patent activity is only one proxy for innovativeness, but it is relevant to DoD-funded innovation. While a patent involves some disclosure, there are often trade secrets that prevent a competitor from copying the invention even once the patent is public. A patent can coexist with classified aspects of the research that do not appear in the patent itself.

[11] Self citation is calculated by matching the USPTO assignees of cited and citing patents. For example, if Boeing cited a McDonnell Douglas patent in 2000, it would not be counted as a self citation. If McDonnell Douglas and Boeing file for patents solely under "Boeing" as the assignee after the merger, then those citations will be counted as self citations. Note that we also exclude any citations from the firms in the figure (prime defense contractors and their acquisition targets).

falling U.S. defense innovation in recent decades are justified.

2.3 Air Force SBIR Program: Context, Process and Reform

2.3.1 SBIR Context and Challenges

Congress first authorized the SBIR program in 1982 to strengthen the U.S. high technology sector and support small firms. Congress requires SBIR to have two Phases: smaller Phase 1 awards fund proof-of-concept work, after which a firm may apply for a larger Phase 2 awards to support later stage demonstration.[12] SBIR applicant firms are typically small and high-tech. While this type of firm represents a small fraction of the economy, the subset of firms that succeed and grow become the primary drivers of future U.S. job creation, innovation, and exports (Haltiwanger et al. 2013). Amid a decline in large corporate R&D labs performing basic science (Arora et al. 2018), new firms have become an increasingly important source of innovation, especially VC-backed startups (Kortum and Lerner 2000, Foster et al. 2008, Decker et al. 2016, Howell et al. 2020). Startups possess frontier technology that is highly relevant for 21st century defense, including products based on AI, autonomous mobility, cyber-security, and biotechnology (Harrison 2016).

The SBIR program has different features and challenges than overall DoD procurement. Perhaps most importantly, firms must be small to participate so consolidation is not a primary concern. Instead, the main concern is lock-in and repeat contracts awarded to firms that are interested neither in commercializing innovation nor in seeking scale in the defense market. Such firms specializing in SBIR awards are sometimes derisively called "SBIR mills" (Edwards 2020). Figure 1(b) repeats the analysis from Section 2.2, but focuses on winners from the Conventional SBIR program. Since the 1990s, SBIR winners have experienced a decline in relative innovation like the prime contractors.[13]

A locked-in SBIR base appears to be more severe at DoD than at other agencies, in part reflecting the large size of DoD's SBIR program and the many similar types of R&D procurement contracts that DoD offers, which can be sustainably lucrative to a small

[12] The Small Business Technology Transfer (STTR) program is an add-on to the SBIR program and requires small business to collaborate with a research institution in the initial research phases. Our main findings do not differ across SBIR and STTR, so we refer to them jointly as "SBIR".

[13] Figure A.4 Panel A uses the HHI measure to show that the DoD SBIR program has become more concentrated over time, with more firms winning many awards in a single year. For both DoD overall and for SBIR, the diverse "markets" means the absolute concentration is very low and less relevant than in private sector markets.

research firm. To provide a benchmark, we consider the DoE, which does not procure SBIR technologies and is where Howell (2017) finds large positive effects on innovation of winning a Phase 1 grant. Each line in Figure A.4 Panel B shows the share of Phase 1 SBIR contract value awarded to firms that won no contracts in the previous three years from the agency. At the beginning of the sample, in the mid-1990s, the two lines are relatively close together, with about 35% (39%) of DoD (DoE) awards going to new firms. The series diverges subsequently, and during the 2010s only 20-25% of DoD SBIR Phase 1 awards went to new firms. This higher incidence of repeat contracts offers a parallel to the consolidation among prime contractors in the larger acquisitions program documented in Section 2.2.

2.3.2 SBIR Process at the Air Force (common to Conventional and Open)

What aspects of the Air Force SBIR process are common to the Conventional and reform programs like Open? First, the Air Force issues a public solicitation for applications. The solicitation describes one or more "topics," each of which represents a discrete competition. Once applications are received, the evaluation process has three steps. In the first step, ineligible applicants are disqualified. In the second step, multiple government evaluators with expertise in the topic area independently evaluate the application. Evaluators produce sub-scores on three criteria: Technology, Team, and Commercialization.[14] The commercialization sub-score reflects the potential to sell any derived product or service within and outside government. Firms' proposed cost is not a primary factor in the evaluation as long as the cost is below the maximum amount identified in the solicitation; that is, firms are not more likely to win if they submit a lower amount. This is different from an auction where firms compete on cost, which is used elsewhere in DoD procurement.

The three sub-scores are summed, and the winners are those whose overall scores are above a threshold determined by the amount of funding available. We will return to this point in the empirical design in Section 4, but this process implies that treatment (award) is exogenous to the running variable (score). While the overall score threshold is sometimes known to the evaluator in advance, because each evaluator produces an independent sub-score, no single

[14] The official description for the conventional program of these criteria are: "(1) Technical Merit – The soundness, technical merit, and innovation of the proposed approach and its incremental progress toward topic or subtopic solution. (2) Qualifications of the Principal Investigator (and Team) – The qualifications of the proposed principal/key investigators, supporting staff, and consultants. Qualifications include not only the ability to perform the research and development but also the ability to commercialize the results. (3) Potential for Commercial Application– The potential for commercial (Government or private sector) application and the benefits expected to accrue from this commercialization."

evaluator can manipulate a firm's position around the cutoff. In the final step, a contracting officer awards the contract and administers the award. This step does not disqualify applicants based on technical merit but does occasionally disqualify applicants for a business reason, such as a cost that is found to be ineligible, or if the proposal is found unrelated to R&D. After the awards are made, the winner identities are immediately public. The non-winner identities that we use in this study are never public, and the scores are never released beyond the evaluation team (i.e., no firms observe their own scores). After removing disqualified awardees, we obtain data for a sharp regression discontinuity design within each topic.

2.3.3 SBIR Reforms: Open vs. Conventional

The SBIR reforms have taken place within a new organization called Air Force Ventures. This is the business division of AFWERX, an office which seeks to foster innovation within the Air Force. The goals of Air Force Ventures are to leverage private capital to deploy new innovations for the military, to expand the industrial base interested in defense, and to grow the U.S. economy. That is, they hope to address the challenges identified in Section 2.2. We describe the context for Air Force Ventures, the contracting process, and the difference between Open and Conventional in more detail in Appendix B.

First conducted in May 2018, Open topics are bottom-up in that the solicitation contains no direction regarding the technology that the applicant may propose. This contrasts with Conventional topics, which are sourced primarily from the Air Force Research Laboratory (AFRL). One example is: "Develop Capability to Measure the Health of High Impedance Resistive Materials" (more examples are in Appendix B). With an explicit reference to seeking "unknown unknowns" in the solicitation, Open topics are designed to let the private sector do the work of identifying military applications for its technology. The firm's objective is to demonstrate the feasibility of developing a product or service with an Air Force partner interested in potentially procuring the firm's technology. The Phase 1 deliverable is a white paper or report describing the outcomes of research. The idea behind Open is that if the bottom-up approach is successful in this context it might be applied to the larger acquisition programs with the hope of garnering interest in the defense market among the large tech firms.

Each year, there are usually three solicitations, each of which has many Conventional topics but only one Open topic since 2018. All Open topics are the same; there are multiple topics because they are issued at different points in time (i.e., in different solicitations). The pool of competitors a given applicant faces in the Open topic depends on when it applies, as scoring and ranking are within-topic. This creates a different distributional structure in Open topics relative to Conventional, as there are many more applicants but also far more winners. The difference in topic structure should not bias the results towards favoring a stronger effect in Open because we estimate the effect of winning within each program, and the cutoff point for winning is lower in the score distribution for Open.

Open topic awards are smaller than Conventional ($50,000 vs. $150,000) and have shorter time frames (3 vs. 9 months). AFWERX's belief that offering many very small awards can be useful was in part informed by research finding strong positive effects on VC and patenting from small, early-stage Phase 1 awards (Howell 2017). This paper focuses on Phase 1, so we minimize the discussion of further awards (Phase 2 awards are described in Appendix B). The Phase 2 awards of $300,000 to $2 million are intended to last 12-24 months and fund a prototype, software, or dataset. In the later Open topics, AFWERX sought to encourage Phase 1 winners to access outside funding from private or government sources with a matching provision in Phase 2. Below, we evaluate the impact of match availability separately from openness.

In sum, we have documented that while U.S. defense R&D has been a historical powerhouse of innovation and military capability, since the mid 1990s, the innovation investment and knowledge spillovers of prime contractors have declined relative to the rest of the economy. At the same time, there has been consolidation and rapidly rising profits. In light of these trends, the Air Force reformed its SBIR program to make it more bottom-up and to encourage a wider diversity of entrants to their competitions. We turn now to describing data from these reforms and evaluating whether they were successful.

3 Data and Summary Statistics

3.1 Data Sources and Key Variables

This section summarizes our data sources, sample construction, and outcome variables. They are described in detail in Appendix C. Our starting point is a dataset of applications and awards to the Air Force SBIR program between 2003 and 2019. We observe complete evaluation data for all topics between 2017 and 2019, and further evaluation data for Conventional topics in select earlier years: 2003-2007, 2015, and part of 2016 (the remaining years' data were inadvertently destroyed in 2016). The focus of this paper is to compare Open and Conventional, so our main sample is restricted to the three years of 2017-2019, which makes the relevant economic environment and defense procurement factors similar across the sample. In 2017, all applicants are Conventional. In 2019, four-fifths of applicants are Open. We also assess the Conventional program using all available years. Figure A.6 shows the number of awards by program and year, and Table A.1 describes counts of topics, firms, and proposals for all programs. In the 2017-19 sample, there are 7,229 Phase 1 proposals from 3,170 unique firms.

There is no pre-classification of the applications into technologies or industries. To do so, we make use of the text in proposal abstracts. Rather than directly reading these and having to make a somewhat arbitrary expert-based classification, we employ a machine learning algorithm called k-means clustering (see Forgy (1965) or Bonhomme and Manresa (2015)) to classify each abstract based on its word "embedding." This is described in detail in Appendix C.3. The process essentially converts the text into vectors of numbers. Each application is represented by a vector whose elements reflect the words used in the application, so that we can then cluster applications into groups based on the similarity of the vectors (i.e. minimizing the total within-cluster variance using their vector representation). We present the five- and two-cluster model, as five is empirically the optimal number of clusters, but two clusters yield a clear dichotomy between software- and hardware-based technologies. The word clouds of keywords that form a cluster are in Figures A.8 and A.9, with the word's size scaled to reflect its prevalence in the cluster.

In addition to using this machine learning algorithm to assign each application to a technology cluster, we also use this approach to characterize the "specificity" of each topic based on the distribution of its applications. Open topics are by definition much less specific than Conventional topics, but there is considerable heterogeneity in the specificity of Conventional topics. To characterize this, we looked at the "distance" between every application and the average application (the centroid) for each topic. The average of these distances within the topic is an index of the non-specificity of the topic, analogous to variance in the scalar case. Intuitively, if the words describing the submissions are very different across applications within a topic, the non-specificity measure will take on a high value.

Table 1 Panel A describes competition characteristics at the proposal level. Conventional topics average 20 applicants and three winners (i.e., awardees). As they are open to any technology and all firms apply to the same request for proposals, Open topics have many more applicants and winners. After the end of our sample, in 2020 the number of Open applicants increased further to around 1600 applicants per topic. Table A.2 shows similar statistics for the whole Conventional sample. Summary statistics for Phase 2 are in Table A.3. The fact that Open topics' mean and median non-specificity measure is four times as large as Conventional Topics offers some validation of our algorithmically based index.

The two main outcomes of interest, in Table 1 Panel B, are subsequent VC investment and DoD non-SBIR contracts, which correspond to the two key metrics of success from AFWERX's perspective. The current Air Force leadership views commercial innovation as evidence of initial success, based on the idea that a strong U.S. industrial base (especially if its research has early-stage ties to DoD) will ultimately enable strong defense (Williams 2020). From an economic perspective, VC investment is a useful proxy for high-growth innovation potential. Although VC-backed startups make up only 0.11% of new firms, over 44% of public company R&D is performed by formerly VC-backed startups (Puri and Zarutskie 2012, Gornall and Strebulaev 2015). We obtain unique private financing deals from Pitchbook, CB Insights, SDC VentureXpert, and Crunchbase. The majority of deals come from Pitchbook, which we observe through January 10, 2021.

The second outcome is non-SBIR DoD contracts, representing defense procurement success in the sense that the research has led to a practical application for the military; in the DoD jargon this is often termed "transition to programs of record." An example of a successful Open applicant from the perspective of transitioning to Air Force operations is Aevum, which designs drone-launched rockets in a former textile mill. After winning a $50,000 Open Phase 1 award in July 2019, Aevum was awarded a $4.9 million Air Force launch contract in September 2019. An example in the Conventional program is Cornerstone Research, which won a $1.9 million SBIR Phase III contract from AFRL (the sponsor of its SBIR award) for further R&D in biotechnology. To construct this outcome, we use complete data from the Federal Procurement Data System through July 2020. We restrict to contracts worth at least $50,000 so that we do not capture very small add-on type awards or minor purchases, but the results are similar using all contracts.

We also consider two ancillary outcomes that are related to innovation and lock-in dynamics, respectively. First, we consider patents with application dates after the SBIR award date, and which were ultimately granted by the USPTO. A granted patent is a proxy for technical innovation with potential commercial applications. Patenting involves some amount of disclosure, but in practice secrecy orders are rarely issued for defense research inventions and all SBIR awardee technology abstracts are publicly available. Therefore, secrecy orders on patent applications are unlikely to affect our results.[15] Using the granted patents, we also build measures of patent quality, such as originality, forward citations, and generality. The patent outcomes go through December 2020. In Appendix C, we provide details on the patent data and patent-based variables.

The second ancillary outcome is subsequent Air Force SBIR awards, using data from the Small Business Administration (results are similar using all-DoD or any-agency SBIR awards). We examine whether winning one SBIR award causally increases the probability of winning a future one, to assess lock-in to the SBIR program.

3.2 Selection into the Programs

The Open reform aims to attract new types of firms to the defense market. This section assesses whether the reform was successful in this dimension. Panel A of Table 1 describes baseline company characteristics for the main analysis sample (2017-19) across the Conventional and Open programs. Open applicant firms are, on average, younger (9.8 vs. 18.2 years old) and smaller (26.9 vs. 60.8 employees) than Conventional applicants. They are more likely to be in the VC hubs of the San Francisco Bay area, greater Boston, and New York City (19.7% vs. 14.8%) and less likely to be in a county where there is an Air Force base (19.2% vs. 27.5%).[16] A lower fraction is female-owned (11.1% vs. 15.5%) and about the same fraction is minority-owned (12%).

Table 1 shows that across all applications from 2017-19, 24% of Open and 51.4% of Conventional applicants have technologies in the Hardware cluster, as opposed to the Training/Software cluster.[17] The clusters by program type and winner status is shown in Figure A.10. Not only do Training/Software proposals dominate Open topics relative to Conventional topics, but within Open they are a larger share of winning proposals. Moreover, new entrants in both programs are more likely to submit training/software proposals relative to incumbents in each topic (Figure A.11). This evidence is consistent with

[15] For statistics on these secrecy orders, see https://fas.org/sgp/othergov/invention/.

[16] To describe their geographic diversity, we map the location of applicants in Figure A.12, with larger bubbles indicating more firms and overlay the locations with VC activity. There is a greater concentration of Open topic applicants in Silicon Valley. Some of the otherwise improbable locations for both programs reflect defense spending hubs such as Washington DC and Ohio, where the AFRL is located. The same set of maps for awardees is in Figure A.13 and documents similar patterns.

[17] The patent data are consistent with the abstract modeling. The three most frequent patent classes for Open applicants are engines (F01D), data processing (G06F), and transmission of digital innovation (H04L), while the three most frequent classes for Conventional applicants (in the 2017-19 period) are semiconductors (H01L), optical systems (G02B), and materials analysis (G01N).

the Open program successfully targeting high-growth startups, since they are concentrated in the training/software sector. Closely aligned with the Open sample, 70% of unique companies that received VC investment between 2010 and 2020 had software rather than hardware products (based on the Pitchbook data platform).

We explore selection on our outcome variables in Panel B of Table 1. Open applicants are almost twice as likely to have previous VC financing (11.4% vs. 6%) and are less than half as likely to have previous DoD contracts or SBIR awards. Conventional applicants are more likely to have any patent before the award (47% vs. 25%). Finally, selection across all our observables is visually apparent in Figures A.14-A.17, which we use below to validate our empirical design, and in Figure 2, which we use below to compare Open to other reform programs.

In sum, firms applying to Open rather than Conventional topics are younger, smaller, and much more likely to have previous VC investment, be located in a VC hub, and have no experience with the DoD or SBIR markets. Thus, the Open program seems to have attracted a good deal of new entry into defense R&D procurement.

4 Empirical Design

The application and scoring processes for Open and Conventional topics are very similar (see Section 2.3.2). These institutional features allow us to use the same regression discontinuity design (RDD) for both programs. The RDD approximates the ideal experiment of randomly allocating awards among applicants. It is relevant in settings where treatment assignment is based on an applicant's location around a cutoff in a rating variable and is widely used for program evaluation. The intuition of the RDD is conceived either as a discontinuity at the cutoff (Hahn et al. 2001) or local randomization around the cutoff (Lee 2008). Regardless of this intuition, there are two implementations: sharp, in which the rating variable perfectly identifies treatment status, and fuzzy, in which there are crossover observations on one or both sides of the cutoff. Our setting permits a sharp RDD.

A valid sharp RDD has four conditions (Hahn et al. 2001, Lee and Lemieux 2010, Gelman and Imbens 2018). First, the rating variable must be established before treatment is assigned (i.e., treatment cannot cause the rating variable). This is the case in our setting, as evaluators score before the award decision is made. Also, as mentioned above, the cutoff (i.e., threshold for winning) is completely independent from the evaluation process and reflects budgets for the current SBIR cycle. Second, treatment assignment must be based solely on the combination of the rating variable and the cutoff. This is true for all the topics. To illustrate, Figure A.18 shows that the probability of treatment jumps from zero to one at a cutoff for four representative topics. As the scores and the cutoff vary across topics, we normalize scores into a rank around the cutoff, such that a rank of 1 is the lowest-scoring winner, and a rank of -1 is the highest-scoring loser.

The third condition for a valid RDD is that the cutoff must be independent of the rating variable. That is, the rating variable cannot be manipulated around the cutoff to ensure certain applicants receive treatment. The most important test for manipulation, common to all RDD settings, is to observe whether there is bunching around the cutoff. In Figure 3, we graph the density of the rating variable around the cutoff within each program. There is no bunching, consistent with no manipulation. The formal test also yields no evidence of manipulation, consistent with the figures (the p-value of manipulation test is over 0.6 in both groups). The second test is to assess the continuity of observable baseline covariates around the cutoff. Figures A.14-A.17 show 11 baseline covariates, including all the outcome variables, observed at the time of application. There are no discontinuities around the cutoff in any of the variables, consistent with an absence of manipulation.

We conduct a third test for manipulation, reflecting the concern in our setting that evaluators might manipulate sub-scores based on an ex-ante preference for which firms should win, potentially leading to scores that are not randomized around the cutoff. An intended benefit of three independent evaluators for three sub-scores is that this sort of manipulation is difficult. An individual evaluator cannot, in general, systematically sway applicants' award status. To confirm this, we examine sub-score variation within the topic. If the three sub-scores are usually correlated so that there is little variation in sub-scores around the cutoff, it might be easier for an evaluator to nudge applicants below or above the threshold. By contrast, if sub-scores exhibit substantial variation, such that often a winning firm has at least one sub-score that is lower than a loser sub-score, and vice versa, it will point to little scope for manipulation. Figure A.19 shows substantial variation in sub-scores around the cutoff. The red bars to the right side of zero show that many unsuccessful applicants (losers) have a sub-score that exceeds the lowest sub-score among winners. Similarly, the blue bars to the left side of zero show that many winners have sub-scores that are lower than the highest loser sub-score. Altogether, 81% of applicants have at least one sub-score that is a "crossover." This should make manipulation very unlikely. It is also worth noting that the evaluators are Air Force government officials (military officers and civilians), and manipulation would constitute a serious violation of acquisition rules.

The last condition for a valid RDD is to control for the rating variable in a well-specified functional form. Our primary model includes all ranks with linear controls for rank on either side of the cutoff. We use a triangular kernel to weight observations far from the cutoff less than those close to the cutoff, following DiNardo and Tobias (2001). Specifically, we use the formula $Kernel_{iT} = 1 - \frac{|Rank_{iT}|}{\max_j |Rank_{iT}| + (0.01)}$ for application i in topic T.[18] This kernel weighting approach weakens the parallel trends assumption for awardees and non-awardees.

We use two primary models in estimation. The first, presented in Equation 1, is estimated within either the Open or Conventional topic samples:

$$\begin{aligned} Y_{it} = \alpha + \alpha_T + \beta\left[1 \mid Rank_{iT} > 0\right] + \gamma_1\left[Rank_{iT} \mid Rank_{iT} > 0\right] \\ + \gamma_2\left[Rank_{iT} \mid Rank_{iT} < 0\right] + \delta PSBIR_{iT} + \varepsilon_{iT}. \end{aligned} \tag{1}$$

Here, the dependent variable Y_{it} is an indicator for the firm experiencing some event after the award decision, such as VC funding. To avoid truncation bias, we restrict the outcome variable to the 24 months after the award decision. However, we show that the results are very similar if we use ever-after outcomes or restrict further to 12 months. On the right-hand side of the equation, we control for whether the applicant has won a previous SBIR award ($PSBIR_{iT}$), though we show that the results are robust to using a wider array of controls, including pre-award outcome variables, or no controls at all. The coefficient of interest is β, on the indicator $[1 \mid Rank_{iT} > 0]$ for having a positive rank and thus receiving an SBIR award. Our primary models include only a firm's first proposal between 2017-19. We also report results using all proposals, which dramatically increases the sample but yields similar results. We show several further models in robustness tests, including a narrow bandwidth around the cutoff.

The second model, presented in Equation 2, allows us to compare program effects to one another by fulling interacting the right-hand side variables with the vector $\mathbf{Program'_T}$:

$$\begin{aligned} Y_{it} = \alpha + \alpha_T + \beta\left[1 \mid Rank_{iT} > 0\right] \cdot \mathbf{Program'_T} + \gamma_1\left[Rank_{iT} \mid Rank_{iT} > 0\right] \cdot \mathbf{Program'_T} \\ + \gamma_2\left[Rank_{iT} \mid Rank_{iT} < 0\right] \cdot \mathbf{Program'_T} + \delta PSBIR_{iT} \cdot \mathbf{Program'_T} + \varepsilon_{iT}. \end{aligned} \tag{2}$$

[18] We add .01 so that the observations with the maximum absolute rank do not end up with a weight of zero (which would cause them to drop out of the regression).

In the main tables, $\mathbf{Program'_T}$ takes one of two values for Open and Conventional. In supplementary analysis, we add two other program types: NSIN and Pitch Day (Section 6.2 describes these in detail). The Conventional program is always the omitted base group. Otherwise, all variables are defined as above. The fixed effects for the topic (α_T) control for the independent effect of program type and the date of award. When we estimate Equation 2, we cluster standard errors by topic. When we estimate Equation 1, we cluster by firm because Open has too few topics to cluster on this variable. All results from that model are similar if we cluster by topic for Conventional.

Our main analysis comparing Open and Conventional topics assesses the Phase 1 award. Phase 2 awards for Open are too recent to observe outcomes. We evaluate the Conventional Phase 2 in an extension. In the Phase 1 analysis, we do not consider the award amount because it is co-linear with winning. In the Phase 2 analysis, we do consider the award amount separately from winning because there is substantial variation in Phase 2 award amounts.

5 Results

This section describes the causal impact of winning Open and Conventional competitions on innovation and contracting outcomes, and describes key robustness tests. The effects of winning a Phase 2 award are in Appendix E.

5.1 Venture Capital (VC) Investment

We first examine the effect of winning an award on receiving any VC after the award decision.[19] Figure 4 shows the topic-effect adjusted mean by rank around the cutoff, using a firm's first application in the 2017-19 period. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. Panel A shows that subsequent VC investment rises just to the right of the cutoff for Open. By contrast, we see no relationship for Conventional topics in Panel B.[20]

Table 2 examines these results in regression format using Equations 1 and 2. Here and subsequently, the mean of the dependent variable is reported at the bottom of the table.

[19]In unreported models, we also consider the level and log amount of VC funding with qualitatively similar results.

[20]The probability of VC before and after the award decision increases in rank for the Open program (Figure A.16 Panel B and Figure 4 Panel A), pointing to the need to control for rank in the analysis.

Column (1) of Panel A shows that winning an Open award increases the probability of VC by 5.4 percentage points, which is 68% of the mean (7.9%) among Open applicants. This compares to a near-zero effect for Conventional, whether using the 2017-19 sample (column (2)) or the whole 2003-2019 sample (column (3)). Column (4) pools the data and estimates Equation 2. The interaction between winning and Open demonstrates that there is a significantly larger effect on VC funding of winning an Open relative to a Conventional award. Column (5) repeats the specification of the previous column but on the longer time period of column (3) from 2003 onwards, and again shows a significantly larger effect of Open vs. Conventional on VC. The final column uses all applications instead of only a firm's first proposal, increasing the sample size to 21,432 compared to 8,769 applications in the previous column.[21] The interaction is again positive and significant, with a very similar coefficient to the previous column.[22]

As explained in Section 2.3.3, an additional reform in the Open topics was to offer matching in Phase 2. Phase 2 applicants could request additional funds to match private or government money that they secured during the Phase 1 period. Several features of the program's implementation facilitate evaluation, for example that matching was not available for the earlier Open topics. We discuss these factors and evaluate the role of matching in Appendix D.2. The main finding is that while matching does increase the probability of VC, winning an Open competition significantly increases VC even without the possibility of matching. Hence, we conclude that something over and above matching in the structure of Open made it more successful than Conventional.

In sum, Panel A of Table 2 shows that winning the Open program has a strong positive causal effect on subsequent VC investment, while winning Conventional does not. This positive effect of an Open topic SBIR contract on VC serves as an early-stage proxy for success and indicates private sector interest in the firm. Why would small procurement contracts have such a large effect? One reason is that the Phase 1 SBIRs may be an entry point to much larger contracts in the future. A goal of the Open Phase 1 program is to find a large customer in the Air Force. Startups with a successful Open Phase 1 process can bring evidence of large potential defense customers to VCs, which appears to improve their odds of raising funds.

[21] While only 17% of Open applications are from firms that previously applied to Open, the comparable statistic over the same period is 52% for Conventional, consistent with greater lock-in and recurring SBIR-winner presence in the Conventional program. From a policy perspective, it is relevant to consider the effect using all applications because the program is not permitted to exclude firms that have applied before.

[22] The full set of interactions mean that the specification in column (4) is identical to those of the first two columns. Thus, the linear coefficient (-0.005) is equal to that of column (2), and the coefficient on the interaction is the difference between the coefficients in the first two columns (i.e. 0.059 = 0.054 - (-0.005)).

The expectation of a Phase 2 award, which averages about $830,000 (see Table A.3 Panel A), may also help to explain the large Phase 1 treatment effect on VC. VCs may be responsive to Phase 1 because they expect it to be associated with substantially more non-dilutive cash. If Phase 2 is important for VC but only through a dynamic channel – via its implications for the Phase 1 treatment effect – this would help explain why the Phase 1 award is so impactful. This channel might be present only for Open if the type of firms that select into Conventional are unlikely to seek VC funding.

5.2 Non-SBIR Defense Contracts

One goal of the SBIR reforms is to enable more firms to "transition" out of the SBIR program to operational programs of record (i.e. beyond only R&D awards). We therefore consider the effect of winning on an indicator for subsequent non-SBIR DoD contracts. The visual results are in Figure 5 where we observe that there clearly is a level shift upwards to the right of the cutoff, implying that winning an Open award increases the probability of a future non-SBIR DoD contracts. There is no such jump for Conventional topics.

In Panel B of Table 2 we replicate the specifications of Panel A, except we use non-SBIR DoD contracts as an outcome. In column (1) we see an effect of winning in an Open topic of 7.5 percentage points, or 51% of the mean of 14.8%. By contrast, in columns (2) and (3), we see no significant effect of winning a Conventional award. The next three columns show that the effect of winning an Open award is always larger than winning a Conventional award, as the interaction coefficient is positive. However, this difference is significant only in the largest sample in the final column.[23] Although weaker than the VC results in Panel A, Panel B suggests that winning an Open topic has a positive causal effect on a firm subsequently creating something of practical value for the DoD.

5.3 Patents

We next turn to patenting for an alternative type of innovation. The visual analysis in Figure 6 shows a clear jump to the right of the cut-off in the probability of a granted patent applied for after winning an Open award. By contrast, there is no jump among Conventional applicants.

[23] In unreported analysis, we find that the effects in Open are not different across R&D and non-R&D contracts. They are, however, driven by two contract types: definitive contracts and blanket purchase agreements, and there is no effect for delivery and purchase orders.

Panel A of Table 3 confirms the effect of winning an Open award on subsequent patenting, with an effect of 5.1 percentage points (column (1)), which is almost twice the mean of 2.7%. Column (2) shows that there is also a positive treatment effect on patents in the Conventional program, although this is statistically insignificant. The coefficient itself is larger than in the previous column (0.077) but is much smaller proportionally at 52% of the mean.[24] Although there is no significant difference in the 2017-19 period (column (4)), the last two columns indicate that winning an Open award has a significantly larger effect than winning a Conventional award.

Columns (3) and (6) of Table 3 contain negative coefficients on winning a Conventional award over the longer 2003-2019 time period. One possible reason – especially when we do not restrict the sample to a firm's first application (column (6)) – is that if the firm is oriented primarily toward getting the next SBIR award, there is no reason to invest in patents, which measure intent to commercialize an invention. Consistent with this, we show below that receiving an initial Conventional SBIR locks a firm into SBIR; this patenting result suggests it may turn them away from private markets.

We also examine several other patent-related variables in Appendix D.1. There is a strong positive effect of winning in Open on originality (i.e. patenting in new technology classes as in Trajtenberg et al. (1997), see Appendix C), but no effect on any of the measures – including citations – in Conventional. Below, we find that the positive Conventional coefficient on patents in the recent period does not survive basic robustness tests. These results, together with the discouraging visual analysis (Figure 6 Panel B), lead us to conclude that while there may be some effect of winning in Conventional, it is much less compelling than the evidence of a strong effect in the Open program. That said, it is intuitive that patents are a more promising outcome for Conventional than VC, since the program attracts more mature firms that are more likely to be in the hardware space.

5.4 Future SBIR contracts

Our final outcome variable is the probability of subsequent Air Force SBIR contracts. This is akin to looking at the impact of the lagged dependent variable. The RDD helps us to overcome the usual difficulty of separating state dependence (the causal impact of the lagged dependent variable) from unobserved heterogeneity (maybe the best firms can keep winning

[24]The positive coefficient is surprising given the visual analysis. Figure 6 uses a narrower bandwidth than Table 3, and we show in Section 5.5 that the coefficient falls to 0.015 in a narrow bandwidth model.

contracts). Interestingly, the effect on this outcome is rather different than the previous three. Figure 7 shows a strong jump above the threshold for winning a Conventional award, but no change for Open.[25] This suggests there is a strong dynamic towards incumbent advantage for Conventional topics. Table 3 Panel B shows that there is no effect of winning an Open award on obtaining future SBIR contracts in column (1). By contrast, in columns (2) and (3) we see a significant positive effect of winning Conventional awards on getting another SBIR award in the future. In the 2017-19 period in column (2), the effect of 17 percentage points is 55% of the mean, while over the full time period in column (3) the effect is more than double the mean. The interaction with Open in the last three columns is consistently negative and significantly so in columns (5) and (6). Whether through reputation, dedicated staff, or some other channel, the traditional SBIR contract gives birth to recurring SBIR-winners. By contrast, Open topics seem to have thus far avoided this lock-in effect.

5.5 Robustness Tests

This subsection describes the results of key robustness tests, most of which are reported in Table 4. First, in Panel A we add a vector of control variables, including the pre-award outcome variables.[26] Here and in subsequent panels, for each of the four outcome variables the first column shows the effect in the Open program and the second column shows the effect in the Conventional program. Consistent with a valid RDD, the coefficients are very similar to the main results and remain significant. In Panel B we omit all controls and continue to find similar results, though the Open effect on contracts becomes insignificant.

We next exploit the intuition of randomization around the cutoff and restrict the sample to the ranks immediately on either side of the cutoff, in which case no control for rank is needed. Specifically, in Panel C we use two ranks (to keep the sample size reasonably large) above and below the threshold ($\pm$1,$\pm$2). The results remain robust, except that the effect on SBIR of Conventional disappears; this reflects it being more tenuous in the 2017-19 period than in the full sample, where this narrow bandwidth model yields a robust result. In Panels D and E we document similar effects to the main model when including all proposals in the 2017-19 period rather than only a firm's first, and when using "ever after" outcomes, which essentially gives the Conventional topics an extra year.

[25] For Conventional, we include the whole period to highlight the large lock-in effect.

[26] The controls are indicators for previous patents, previous VC, previous non-SBIR DoD contract, previous Air Force SBIR award, location in a VC hub city, location in a county with an Air Force base, software rather than hardware-based technology, and a continuous variable for firm age.

The zero effects for Conventional in the recent period might reflect the Open program crowding out good Conventional projects. While the total pot of funding has expanded significantly since 2017, there has also been a replacement of Conventional with Open (Figure A.6). Therefore, the reform has been a reallocation rather than a pure addition. In practice, however, crowding out does not seem to play a role because we find no effect of Conventional when we restrict to earlier periods when Open did not exist. Panel F uses the 2003-17 period and shows no significant effects except for subsequent SBIR as above. In unreported analysis we find similar results for 2015-17.

Further unreported exercises find that the effect of Open does not differ significantly by year, and that the results are similar to the main model controlling for rank quadratically, omitting the kernel weighting, and using alternative vectors of controls for baseline characteristics. Finally, we assess whether the differential effect in the Open program reflects a larger number of winners by interacting winning with the topic's number of applicants being above median (or above the 75th or 90th percentile). In unreported results, we find the interactions to be small and insignificant. In other words, the effects are not larger in large Conventional topics.

We close this section by addressing a challenge facing the Open program. Both the applicant and the evaluators may struggle to know whether a new technology will be useful to the government. This is not a problem in the Conventional targeted competitions that have an articulated need. To probe this idea, we assess whether the Commercialization sub-score predicts innovation outcomes. If so, this would suggest that evaluators can predict whether the technology will be useful. In Table A.4 we show how the total score and three sub-scores predict the main outcomes. Reassuringly from an identification perspective given our coarse running variable, the total score that is used to force rank applicants does not predict any of our outcomes (Panel A). However, the Commercialization score predicts both VC and DoD contracts (Panel B). In contrast, the Team score negatively predicts VC (Panel C). These results suggest that while overall scores are not especially informative, the Commercialization score does predict whether the technology will be useful.

6 Mechanisms: New Entrants and Bottom-Up Innovation

The upshot from the RDD analysis on our four outcomes is clear. Winning an Open award increases the chances of future VC funding, patenting and supplying the DoD with future technologies. By contrast, the only consistent effect of winning a Conventional SBIR contract is increasing the firm's chances of winning *another* SBIR contract in the future. We are now interested in whether these patterns reflect a different composition of firms with larger treatment effects or are the result of bottom-up R&D across all firm types. In this section, we use three strategies to probe this question. The first two focus on the role of selection. The last focuses specifically on the role of openness.

6.1 Heterogeneity of the treatment effects

The summary statistics in Table 1 and Figure 2 revealed that the Open program did attract, as was the intention, a different applicant composition relative to the Conventional program. If the causal effects of Open reflect its ability to attract new entrants – in the sense of young firms and firms that are new to the SBIR market – we expect larger effects for these groups. We also might expect larger effects among software-based firms, since this characteristic is associated with startups and is more prevalent among Open applicants.

Focusing on VC with the specification from Table 2 column (1), the first two columns of Table 5 interact the treatment effect with whether a firm is an incumbent. In column (1) we measure incumbency as an indicator for having won a previous Air Force SBIR award. The positive effect of Open is driven by firms without previous Air Force SBIR awards; the effect is 6.3 percentage points for this group. The total effect of winning for incumbent firms is near zero (0.009=0.063-0.072). The second incumbency measure is an indicator for the firm being more than five years old (i.e. above the median age). In column (2) we see the effect is 8.3 percentage points for younger firms, and again a nearly-zero negative impact for older firms.

Beyond these significant results, we find no other meaningful heterogeneity effects, either using these covariates for the other outcomes or using alternative covariates. For example, in column (3) of Table 5, we show that the impact of winning is not significantly different for applicants in hardware compared to software technologies. The last three columns repeat these specifications for Conventional topics and find no significant differences.

These results suggests that the effort to attract entrants to apply for Open is well-placed: the positive effect on VC is stronger for these firms. Although our results are robust to selection as they are identified from the discontinuity and entrant status is continuous around that discontinuity (Figures A.14 and A.17), attracting a wide range of new entrants is likely to be beneficial to overall program success because they experience larger treatment effects.[27] At the same time, the absence of interaction effects for other outcomes and in Conventional suggests that another mechanism besides selection may be at play.

6.2 Other Reforms: Pitch Day and NSIN

We next compare openness and selection using two other reform SBIR programs: Air Force Pitch Day and National Security Innovation Network (NSIN) topics. These programs attracted firms with characteristics similar to the Open applicants, but the topics were more specific. They were instituted at roughly the same time as the Open program.

Pitch Days were held in VC hubs such as Boston, New York City, and Austin starting in 2019. They sought to bring mission programs with large procurement needs directly in contact with promising startups, with senior officers from the mission programs serving as pitch judges. The evaluation process has the same structure as all the other topics, but a key difference is that the evaluators are physically present at the pitch and make their decisions in real time. Winners are immediately notified and are expected to sign a contract at the event. An example of a topic is "Battlefield Air Operations Family of Systems Technologies." NSIN topics, which start in 2018, come from a central DoD office, rather than one of the services. These topics share with Open topics a focus on dual-use viability; in particular, identifying commercial technologies that can provide immediate solutions in the field for the Air Force (e.g. "Machine Learning for Defense Applications"). Summary statistics for these other reform programs are in Table A.2.

We compare selection into all four programs in Figure 2. It reports coefficients in the form of relative risk ratios from a multinomial logistic model that estimates whether ex-ante firm characteristics predict applying to a particular program. They indicate that firms applying to the three reform programs have markedly different characteristics relative to firms applying to the Conventional topics (the baseline group, denoted in the horizontal red line). For example, applicants in Open, NSIN, and Pitch day are all much more likely to have an IT/Software-based technology and much less likely to have previous patents or above-median employment. In almost all cases, the coefficients for the three reform programs are clustered tightly together

[27] A related point on selection concerns those firms that apply to both Open and Conventional topics. In this small group, we find no effects of winning conditional on applying to both programs with and without firm fixed effects. Firms that apply to both are more likely to be SBIR specialists who apply to as many topics as possible. The larger effect of Open stems from firms who have only ever applied for Open.

with a large difference from Conventional. These common differences on key variables relative to the Conventional program make the other reform topics useful for exploring the role of selection in any causal effects.

We compare the differential effects of the reform programs in Table 6 using Equation 2, with VC as the dependent variable. As explained above, we run fully interacted models to understand whether the program effects are significantly different from one another. A caveat is that the samples for the two additional reform topics are small. However, we observe a significant effect in Open topics on VC even with a narrow bandwidth, which has a similar sample size as the other reform topics. We first include NSIN proposals in column (1). The interaction between Award and NSIN indicates that relative to Conventional topics, there is no significant effect of winning an NSIN topic. We next add Pitch Day topics in column (2), and similarly find no significant interaction. As expected given the complete interactions, the coefficients are the same when all programs are included in column (3). We consider Pitch Day and NSIN topics alone in columns (4) and (5) respectively, where both coefficients are negative and insignificant. We similarly find no effects of these two reform programs on the other main outcomes. These results, like those in Section 6.1, suggest that selection alone does not appear to explain the strong effects of the Open reform.

6.3 Topic Specificity Within the Conventional Program

Some Conventional topics are more specific than others in identifying the technology DoD wishes to procure. If a bottom-up, technology-neutral approach drives the large effects of winning the Open program, there should be larger effects of winning Conventional when the topic is more technology-neutral, encouraging a broader range of ideas. As noted in Section 3.1, it is difficult to manually code the specificity of topics and doing so would entail a high degree of subjectivity. Therefore, we assess how specific a topic is by evaluating the abstract text of the applications. If the content across proposals is more diverse, then the topic is less specific (more open). To create a non-specificity measure, we use the abstract embeddings that we developed in order to assign abstracts to technology clusters. The process is explained in detail in Appendix C.3. Since the composition of applicants in non-specific and specific Conventional topics is very similar (Figure A.20), any differential effect of winning in specific topics cannot reflect selection along the dimensions that we can observe.

In Table 7, we assess whether the effect of winning in a Conventional topic is modulated by the specificity of the topic. We use continuous demeaned non-specificity, though the results are similar using indicators, such as a dummy for the topic having above-median non-specificity. For outcome variables, we focus on patenting, because this is the only outcome for which Conventional was found to have some evidence of a possible effect (Section 5.3). All columns include topic fixed effects, which absorb the measure of specificity.

We consider three outcomes. The first is any granted patents as in Panel A of Table 3 above. Columns (1) and (2) of Table 7 show that for both first and all proposals, there is a strong positive coefficient on the interaction between non-specificity and winning an award, indicating that more open topics yield a relatively larger positive effect of winning on patenting even amongst Conventional topics. In particular, column (1) implies that a one standard deviation increase in non-specificity (0.58, from Table A.2) implies a 20.2% (= (0.58*0.055)/0.158)) increase relative to the mean. Since non-specificity is demeaned, the coefficient on Award represents the effect at mean specificity. We find a similar relationship for measures of patent quality – above-median forward citations and originality – in columns (3)-(6). See Appendix C.2 and D.1 for details on how these variables are constructed and average effects.

We next test whether heterogeneity in openness is more important than the differential effect of the Open topics. Recall that consistent with attracting diverse ideas, the Open topics have a non-specificity index that is on average about four times that of the Conventional topics (Table 1). If the bottom-up approach is more important to the success of Open than some other characteristic, we expect the interaction with non-specificity to be more important than the interaction with Open. In Panel B, we include the Open applications and adjust the main models from Table 3 Panel A columns (5)-(6) by adding an interaction between award and non-specificity. We continue to find that the effects of the interaction between winning and non-specificity load positively, with statistical significance in three of the four models. Meanwhile, the interaction between winning and Open becomes insignificant.

Non-specificity is only one characteristic that distinguishes topics from one another, so we checked whether the results in Table 7 were robust to including other observable characteristics of competitions. In unreported analysis, we find that the differential effect of winning in a non-specific topic is robust to including interactions between winning and other topic characteristics such as topic competitiveness (winners per applicant) and topic size (number of applicants).[28]

[28]For example, we re-ran column (1) of Panel A in Table 7, but included an interaction of winning an award with the topic-specific fraction of applicants who win awards (which Table 1 shows is much higher for Open than Conventional). The coefficient on our key interaction fell only slightly from 0.055 to 0.053 (with a standard error of 0.025) and the new "competition" variable was insignificant (coefficient of 0.049 with a

To summarize, the effect of winning a Conventional topic on subsequent patenting is significantly larger for non-specific – i.e. more open – topics. While these patterns do not exist for other outcomes (though the coefficients go in the same qualitative direction), they suggest that when Conventional topics attract diverse ideas, their effects on patenting are indistinguishable from the effects in Open topics.[29] That is, the Conventional program appears to be more impactful when it takes a more bottom-up approach, consistent with openness being important independently from selection or other characteristics of the Open reform program.

6.4 Implications for Competitive Procurement

Overall, this section has provided evidence consistent with the hypothesis that the "bottom-up" nature of Open topics is important to the success of Open. While selection of firms with more startup-like characteristics is also relevant, it appears to be the case that there is something about openness itself that affects outcomes. Practically, the qualitative findings of the paper point to benefits from the move towards the Open program and away from Conventional. In Appendix F, we conduct a simple analysis to explore the magnitude of the net benefits from replacing Conventional topics with Open ones, where benefits are based on amounts of VC raised and non-SBIR DoD contracts signed, and costs include the award amount and administrative costs. Under a wide range of specifications and assumptions we find that the net benefits of a marginal Open program are much larger than those of a Conventional competition; in our baseline analysis the net benefit of an Open award is $1,104,000, while the net benefit of a Conventional award is $1,000. This dramatic difference is driven by the much higher treatment effect in Open and somewhat by its lower costs.

The success of Open relates to the broad question of how to procure defense innovation. The main mechanisms are (i) government contracting with private organizations (ii) design competitions; and (iii) in-house R&D at government laboratories. In recent decades, the U.S. has emphasized the first channel, but before World War II and in certain parts of the defense establishment associated with transformational inventions (such as DARPA), the second channel of competitions has been important (Mansfield 1971, Jacobsen 2015). Competitions can enable more government flexibility and encourage contractor risk-taking (Lichtenberg

standard error of 0.033).

[29] We do not include the citations outcome because there is not enough time after the Open awards for citations to accrue.

1984). Historically, as Lichtenberg (1995) explains, most of the expected revenue comes not from the prize but from follow-on, noncompetitive contracts. The contractor incurs losses on initial development for the competition as a means to "buy-in" to future contracts.

The SBIR program is one of the largest single competitive innovation procurement programs aimed at small firms in the world. In a traditional competition, the government identifies a need for a certain product, and firms must privately invest in initial R&D to compete for a prize. A downside of this approach is that the technology must be ex-ante specified, while in direct contracting it can be more ambiguous and evolve over time. The Open program mitigates the downside of ex-ante specification by allowing firms to present their own ideas but evaluating them according to the same metrics (technology quality, team, and potential for defense or civilian application). This potentially offers a new template for other competitive R&D procurement efforts.

7 Conclusion

U.S. defense R&D is often held up as an example to the world of how to stimulate aggregate innovation through mission-driven research. This paper shows that the luster has faded somewhat in recent decades, with the prime defense contractors becoming less innovative than the rest of the U.S. economy on a number of dimensions. One response of the U.S. Air Force to this concern is to introduce the Open SBIR program, which seeks to stimulate bottom-up innovation from firms who would not normally apply to the traditional, more tightly specified topics in DoD's Conventional program. Though the SBIR program differs from mainline procurement, it faces parallel problems of stale innovation and lock-in of repeat contractors. Skeptics of the innovation benefits of military R&D have noted that while there is a surfeit of anecdotes, there is a dearth of rigorous evaluations of U.S. defense R&D programs. This paper helps to address the lacunae by causally evaluating the Air Force SBIR program, with a focus on the Open reform.

We show that the Open program successfully attracted a cohort of high-tech startups to the defense market. Our two primary outcomes are proxies for innovation benefits for civilian use (VC investment and patents) and a benefit for the military (winning future DoD contracts). Using a regression discontinuity design, we find that winning an Open topic award has positive effects on these measures of military and non-military success, whereas winning a Conventional topic award does not. By contrast, winning a Conventional award increases the chances of a subsequent SBIR contract, creating a lock-in effect for incumbents. Of course, the government may have objectives beyond the outcomes we examine, and the implications of any effects on those outcomes are beyond the scope of this paper. It may be that some mixture of Open and Conventional-style topics is optimal for DoD, but our research suggests that the Open reform was well warranted.

Why has the Open reform been successful? One reason is that it attracts new entrants. But we believe the bottom-up nature of the program design also plays an important role for two reasons. First, the effect of winning on patent quality and quantity increases in Conventional topics that are less specific. Second, a comparison with other U.S. Air Force SBIR reforms which were less bottom-up, but also attracted firms with startup-like characteristics, finds that the other reforms were less successful than the Open program.

While our context is specific, it is important because the U.S. DoD funds more R&D than any other single entity in the world. Beyond defense, our findings relate to efforts at bottom-up or open innovation, ranging from the large pharmaceutical companies outsourcing innovation to biotech startups (Schuhmacher et al. 2013), to LEGO Ideas, which has led to 30 LEGO model kits based on externally submitted ideas.[30] An important avenue for future work is whether causal evaluations of bottom-up non-military R&D programs reveal similar patterns.

[30] See https://hbr.org/2020/01/turn-your-customers-into-your-community.

References

Advanced Research Projects Agency-Energy, "OPEN," 2020.

Akcigit, Ufuk and Stefanie Stantcheva, "Taxation and Innovation: What Do We Know?," Technical Report, National Bureau of Economic Research 2020.

_ , John Grigsby, and Tom Nicholas, "The rise of american ingenuity: Innovation and inventors of the golden age," Technical Report, National Bureau of Economic Research 2017.

Arora, Ashish, Sharon Belenzon, and Andrea Patacconi, "The decline of science in corporate R&D," *Strategic Management Journal*, 2018, *39* (1), 3–32.

Azoulay, Pierre and Danielle Li, "Scientific Grant Funding," Technical Report, National Bureau of Economic Research 2020.

_ , Joshua S Graff Zivin, and Gustavo Manso, "Incentives and creativity: evidence from the academic life sciences," *The RAND Journal of Economics*, 2011, *42* (3), 527–554.

_ , Joshua Zivin, Danielle Li, and Bhaven Sampat, "Public R&D Investment and Private Sector Patenting: Evidence from NIH Funding Rules," *Review of Economic Studies*, 2019.

Babina, Tania, Alex Xi He, Sabrina Howell, Elisabeth Perlman, and Joseph Staudt, "Does Funding Source Matter for University R&D? The Effect of Government vs. Industry Grants," *Working Paper*, 2020.

Barro, Robert and Charles J. Redlick, "Macroeconomic Effects from Government Purchases and Taxes," *Quarterly Journal of Economics*, 2011.

Belenzon, Sharon and Larisa Cioaca, "Guaranteed Markets and Corporate Scientific Research," Technical Report, Duke University 2021.

_ and Mark Schankerman, "Spreading the word: Geography, policy, and knowledge spillovers," *Review of Economics and Statistics*, 2013, *95* (3), 884–903.

Bhattacharya, Vivek, "An empirical model of R&D procurement contests: An analysis of the DOD SBIR program," Technical Report, Mimeo, Northwestern University 2018.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb, "Are ideas getting harder to find?," *American Economic Review*, 2020, *110* (4), 1104–44.

_ , John Van Reenen, and Heidi Williams, "A toolkit of policies to promote innovation," *Journal of Economic Perspectives*, 2019, *33* (3), 163–84.

Bloom, Nick, Rachel Griffith, and John Van Reenen, "Do R&D tax credits work? Evidence from a panel of countries 1979-1997," *Journal of Public Economics*, 2002.

Bonhomme, Stephane and Elena Manresa, "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 2015, *83* (3), 1147–1184.

Braddon, "Commercial Applications of Military R&D: US and EU Programs Compared," *University of West England mimeo*, 1999.

Bronzini, Raffaello and Eleonora Iachini, "Are Incentives for R&D Effective? Evidence from a Regression Discontinuity Approach," *American Economic Journal: Economic Policy*, 2014.

Carril, Rodrigo and Mark Duggan, "The impact of industry consolidation on government procurement: Evidence from Department of Defense contracting," *Journal of Public Economics*, 2020, *184*, 104141.

Chen, Zhao, Zhikuo Liu, Juan Carlos Suárez Serrato, and Daniel Yi Xu, "Notching R&D investment with corporate income tax cuts in China," Technical Report, National Bureau of Economic Research 2018.

Chesbrough, Henry William, *Open innovation: The new imperative for creating and profiting from technology*, Harvard Business Press, 2003.

Cox, Amy G, Nancy Y Moore, and Clifford A Grammich, "Identifying and eliminating barriers faced by nontraditional Department of Defense suppliers," Technical Report, RAND National Defense Research Institute 2014.

CRS, "Government Expenditures on Defense related R&D by the US and other OECD Countries," Technical Report, Congressional Research Service, https://fas.org/sgp/crs/natsec/R45441.pdf 2018.

de Villemeur, Etienne Billette and Bruno Versaevel, "One lab, two firms, many possibilities: On R&D outsourcing in the biopharmaceutical industry," *Journal of health economics*, 2019, *65*, 260–283.

Dechezlepretre, Antoine, Elias Einio, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen, "Do fiscal incentives increase innovation?," *CEP Discussino Paper 1413*, 2019.

Decker, Ryan A, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "Declining business dynamism: What we know and the way forward," *American Economic Review*, 2016, *106* (5), 203–07.

Denes, Matthew, Sabrina Howell, Filippo Mezzanotti, Xinxin Wang, and Ting Xu, "Investor Tax Credits and Entrepreneurship: Evidence from US States," Technical Report, NBER Working Paper 27751 2020.

DiNardo, John and Justin L Tobias, "Nonparametric density and regression estimation," *Journal of Economic Perspectives*, 2001, *15* (4), 11–28.

Draca, Mirko, "Reagan's innovation dividend? Technological impacts of the 1980s US defense build-up," *CAGE Working Paper Series No. 168*, 2013.

Edwards, Colin, "SSTI analysis reveals SBIR mills take outsized portion of the programs awards," Technical Report, SSTI 2020.

Flamm, Kenneth, *Creating the computer: government, industry, and high technology*, Brookings Institution Press, 1988.

Fleming, Lee, Hillary Greene, G Li, Matt Marx, and Dennis Yao, "Government-funded research increasingly fuels innovation," *Science*, 2019, *364* (6446), 1139–1141.

Forgy, E, "Cluster Analysis of Multivariate Data: Efficiency vs. Interpretability of Classifications," *Biometrics*, 1965, *21* (1), 768–780.

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: selection on productivity or profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

Gelman, Andrew and Guido Imbens, "Why high-order polynomials should not be used in regression discontinuity designs," *Journal of Business & Economic Statistics*, 2018, pp. 1–10.

Goolsbee, Austan, "Does Government R&D Policy Mainly Benefit Scientists and Engineers?," *American Economic Review*, 1998.

Gornall, Will and Ilya A Strebulaev, "The economic impact of venture capital: Evidence from public companies," 2015.

— and — , "Squaring venture capital valuations with reality," *Journal of Financial Economics*, 2020, *135* (1), 120–143.

Griffin, Michael D., "Terms of Reference - Defense Science Board Task Force on 21st Century Industrial Base for National Defense," 2019.

Gross, Daniel P and Bhaven N Sampat, "Inventing the endless frontier: The effects of the World War II research effort on post-war innovation," Technical Report, National Bureau of Economic Research 2020.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw, "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, 2001, *69* (1), 201–209.

Hall, Bronwyn, "R&D Tax Policy during the 1980s: Success or Failure?," *Tax Policy and the Economy*, 1993.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda, "Who creates jobs? Small versus large versus young," *Review of Economics and Statistics*, 2013, *95* (2), 347–361.

Harrison, Adam Jay, "DOD 2.0: high tech is eating the Pentagon," *Proceedings Magazine (US Naval Institute), February*, 2016.

Howell, Sabrina T, "Financing innovation: Evidence from R&D grants," *American Economic Review*, 2017, *107* (4), 1136–64.

—, Josh Lerner, Ramana Nanda, and Richard R Townsend, "Financial Distancing: How Venture Capital Follows the Economy Down and Curtails Innovation," Technical Report, National Bureau of Economic Research 2020.

Jacob, Brian and Lars Lefgren, "The impact of research grant funding on scientific productivity," *Journal of Public Economics*, 2011.

Jacobsen, Annie, *The Pentagon's Brain: An Uncensored History of DARPA, America's Top-secret Military Research Agency*, Little, Brown, 2015.

Jaffe, Adam B and Karen Palmer, "Environmental regulation and innovation: a panel data study," *Review of economics and statistics*, 1997, *79* (4), 610–619.

— **and Manuel Trajtenberg**, *Patents, citations, and innovations: A window on the knowledge economy*, MIT press, 2002.

Kortum, Samuel and Josh Lerner, "Assessing the contribution of venture capital to innovation," *RAND journal of Economics*, 2000, pp. 674–692.

Krieger, Joshua L, Danielle Li, and Dimitris Papanikolaou, "Missing novelty in drug development," Technical Report, National Bureau of Economic Research 2018.

Lach, Saul, "Do R&D subsidies stimulate or displace private R&D? Evidence from Israel," *Journal of Industrial Economics*, 2002.

Lanahan, Lauren and Maryann P Feldman, "Approximating exogenous variation in R&D: Evidence from the Kentucky and North Carolina SBIR State Match programs," *Review of Economics and Statistics*, 2018, *100* (4), 740–752.

Lee, David S, "Randomized experiments from non–random selection in US House elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

— **and Thomas Lemieux**, "Regression discontinuity designs in economics," *Journal of economic literature*, 2010, *48* (2), 281–355.

Lerner, Josh, "The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program," *The Journal of Business*, 1999, *72* (3), 285–318.

— , *Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed–and what to do about it*, Princeton University Press, 2012.

— **and Ramana Nanda**, "Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn," *Journal of Economic Perspectives*, 2020, *34* (3), 237–61.

Lichtenberg, Frank, "The Relationship between Federal Contract R&D and Company R&D," *American Economic Review*, 1984.

— , "Assessing the Impact of Federal Industrial R&D Expenditure on Private R&D Activity in the U.S.," in Philip Gummettm and Judith Reppy, eds., *The Relation Between Defense and Civil Technologies*, 1988.

— , "The Economics of Defense R&D," in K Hartkey and Sandler T, eds., *Handbook of Defense Economics, Volume 1*, 1995.

Mansfield, Edwin, *Technological change*, Norton, 1971.

Manso, Gustavo, "Motivating innovation," *The Journal of Finance*, 2011, *66* (5), 1823–1860.

Mazzucato, Mariana, *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*, Anthem Press, 2013.

— **and Gregor Semieniuk**, "Public financing of innovation: new questions," *Oxford Review of Economic Policy*, 2017, *33* (1), 24–48.

Middleton, Andrew, Steven Bowns, Keith Hartley, and James Reid, "The effect of defence R&D on military equipment quality," *Defence and Peace Economics*, 2006, *17* (02), 117–139.

Moretti, Enrico and Daniel Wilson, "State Incentives for Innovation, Star Scientists and Jobs: Evidence from Biotech," *Journal of Urban Economics*, 2014.

— **, Claudia Steinwender, and John Van Reenen**, "The Intellectual Spoils of War? Defense R&D, Productivity and International Spillovers," *NBER Working Paper No. 26483*, 2020.

Mowery, David C, "Defense-related R&D as a model for Grand Challenges technology policies," *Research Policy*, 2012, *41* (10), 1703–1715.

— **and Nathan Rosenberg**, *Technology and the pursuit of economic growth*, Cambridge University Press, 1991.

Nakamura, Emi and Jon Steinsson, "Fiscal stimulus in a monetary union: Evidence from US regions," *American Economic Review*, 2014.

Nanda, Ramana, Ken Younge, and Lee Fleming, "Innovation and entrepreneurship in renewable energy," in "The changing frontier: Rethinking science and innovation policy," University of Chicago Press, 2014, pp. 199–232.

Pennington, Jeffrey, Richard Socher, and Christopher Manning, "GloVe: Global vectors for word representation," *Proceedings of the 2014 conference on empirical methods in natural language processing*, 2014, pp. 1532–1543.

Perotti, Roberto, "Defense Government Spending is Contractionary, Civilian Government Spending is Expansionary," *NBER Working Paper No. 20179*, 2014.

Pless, Jacqueline, "Are 'Complementary Policies' Substitutes? Evidence from R&D Subsidies in the UK," Technical Report, MIT Working Paper 2019.

Puri, Manju and Rebecca Zarutskie, "On the life cycle dynamics of venture-capital-and non-venture-capital-financed firms," *The Journal of Finance*, 2012, *67* (6), 2247–2293.

Ramey, Valerie, "Identifying Government Spending Shocks: It's All in the Timing," *Quarterly Journal of Economics*, 2011.

Rao, Nirupama, "Do tax credits stimualte R&D spending? Revisiting the effect of the R&D tax credit in its first decade," *Journal of Public Economics*, 2016.

Rathje, Jason, "Essays on Public Funding for Private Innovation." PhD dissertation, Stanford University 2019.

— **and Riita Katila**, "Enabling Technologies and the Role of Private Firms," *Strategy Science*, 2020, (Forthcoming).

Rogerson, William P, "Economic incentives and the defense procurement process," *Journal of Economic Perspectives*, 1994, *8* (4), 65–90.

Sargent, John F. and Marcy E. Gallo, "The Global Research and Development Landscape and Implications for the Department of Defense," Technical Report, Congressional Research Service 2018.

Schuhmacher, Alexander, Paul-Georg Germann, Henning Trill, and Oliver Gassmann, "Models for open innovation in the pharmaceutical industry," *Drug discovery today*, 2013, *18* (23-24), 1133–1137.

Senor, Dan and Saul Singer, *Start-up Nation*, Council on Foreign Relations, 2009.

Syverson, Chad, "Challenges to mismeasurement explanations for the US productivity slowdown," *Journal of Economic Perspectives*, 2017, *31* (2), 165–86.

Tenenbaum, Joshua B, Vin De Silva, and John C Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, 2000, *290* (5500), 2319–2323.

Tirole, Jean, "Procurement and renegotiation," *Journal of Political Economy*, 1986, *94* (2), 235–259.

Trajtenberg, Manuel, Rebecca Henderson, and Adam Jaffe, "University versus corporate patents: A window on the basicness of invention," *Economics of Innovation and new technology*, 1997, *5* (1), 19–50.

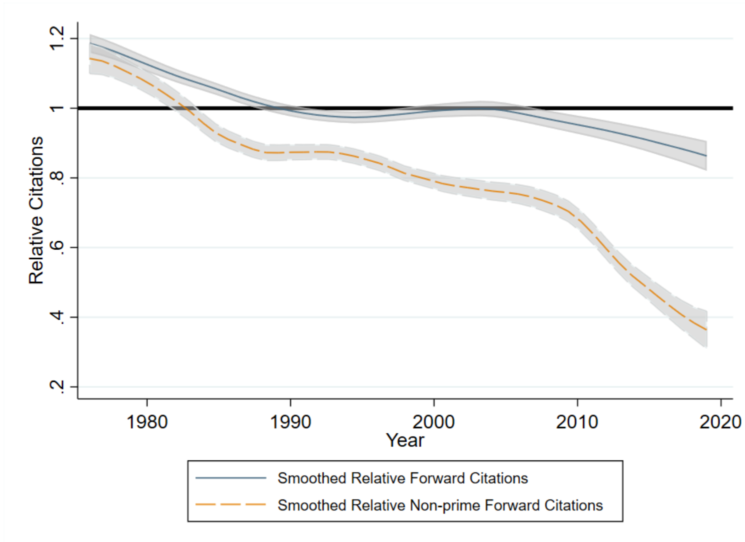
UK Defence and Security Acceleator, "Open Call for Innovation," 2020.

Wallsten, Scott, "The effects of government-industry R&D programs on private R&D: the case of the Small Business Innovation Research program," *RAND Journal of Economics*, 2000.

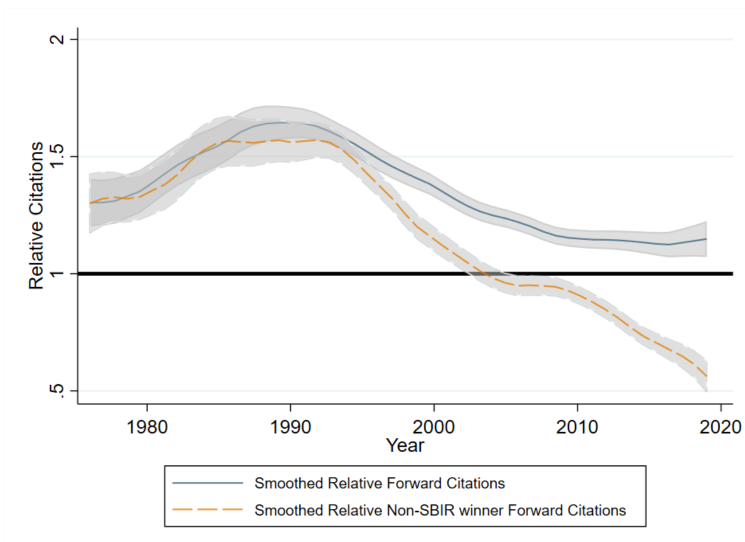
Williams, Lauren, "Air Force acquisition chief wants to reshape defense industrial base," Technical Report, FCW, https://fcw.com/articles/2020/06/09/williams-roper-rethink-dib.aspx 2020.

Figure 1: Declining Relative Innovation Among U.S. Prime Defense Contractors

(a) Prime Patent Citations

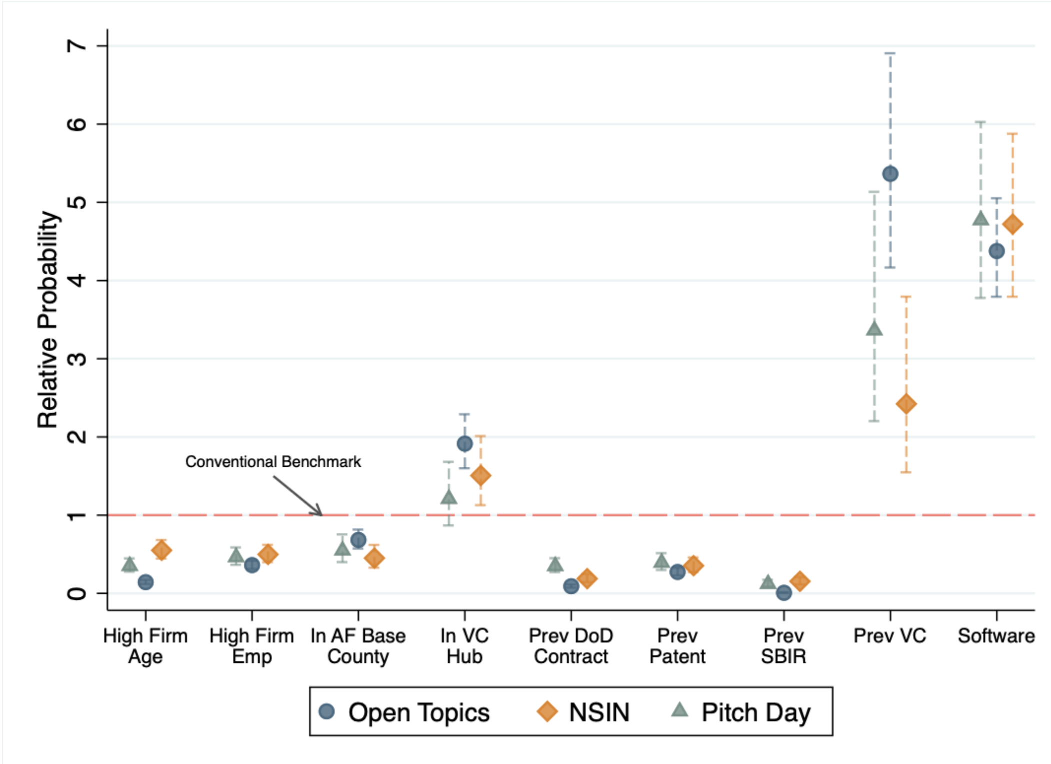


(b) Conventional SBIR Winner Patent Citations



Note: These graphs describe patent quality for the prime defense contractors and their acquisition targets (depicted in Figure A.1). That is, 226 firms are included in 1976, while only six are included in 2019 (as the 226 have merged into these six). Panel A shows the total number of forward citations (solid blue line) and outside non-prime forward citations (dashed orange line) for these firms relative to the average in the same class-year. A value of 1 means the firm's patents have the same number of citations as the average patent in the same class-year. The dashed line makes two changes relative to the blue line. First, it excludes self-citations, where the company cites one of its own previous patents. Second, it excludes any citations from the firms in the figure (prime defense contractors and their acquisition targets). We do not count future cites of a target firm's patents from its future acquirer as self-cites, so the effect is not mechanical from consolidation. Note that the prime and target share of patents in a class year has declined over time, so there are not "fewer outside patents to cite" in a class-year (see Figure 3). Panel B repeats this exercise but for Air Force SBIR winner firms. In this case, the dashed orange line excludes self-citations citations from other AF SBIR winner firms. The measures in both figures are smoothed using kernel-weighted polynomial regressions. The gray band around the relative citations represents the 95% CI. Data are sourced from the USPTO.

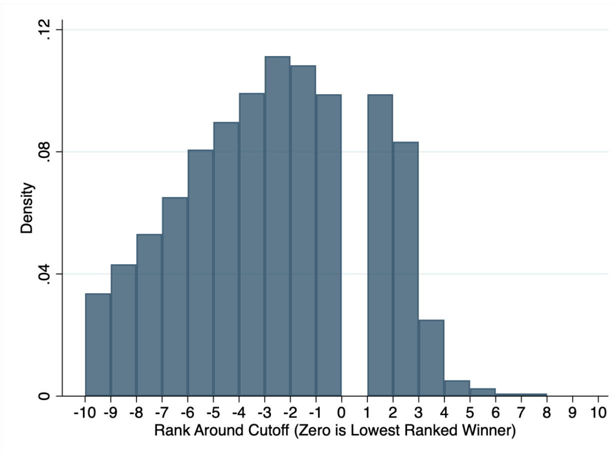
Figure 2: Selection into Programs



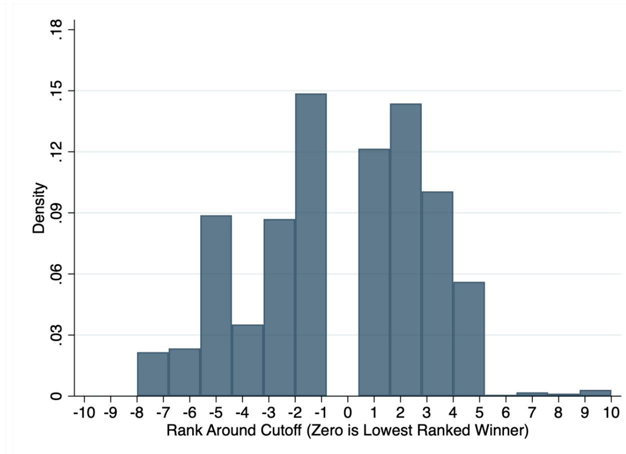
Note: This figure shows how firm characteristics predict selecting into reform programs compared to the Conventional program. The points represent coefficients from a multinomial logistic regression with categories for four programs: the Conventional program, which serves as the base group and is represented by the red dashed line, and the three reform programs (Open, NSIN, and Pitch Day). Data restricted to 2017-2019. The dashed lines around each coefficient point indicate the 90% confidence intervals.

Figure 3: Regression Discontinuity Density Manipulation Test

(a) Conventional (2017-19)



(b) Open

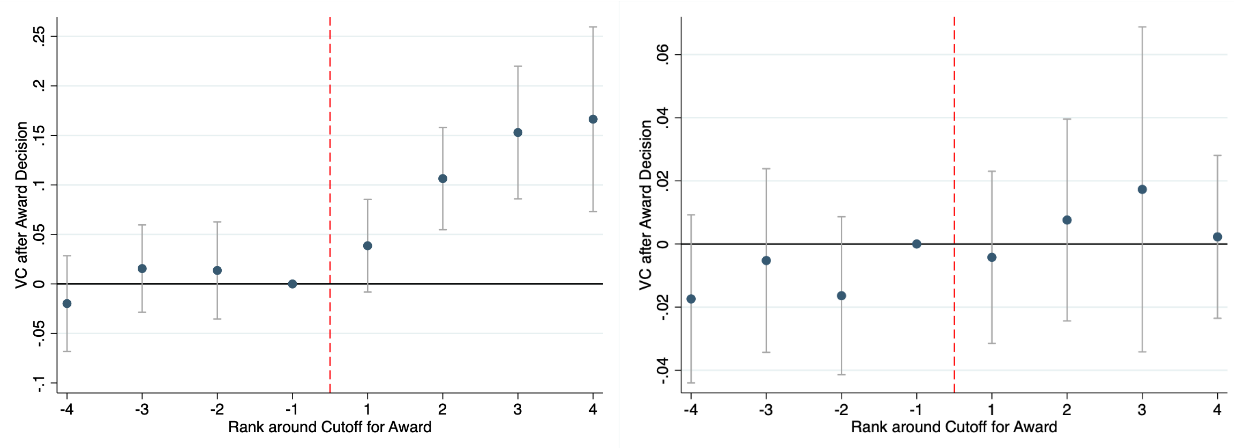


Note: This figure plots the density of applicants by rank around the cutoff using Phase 1 applicants to the Conventional (left graph labeled (a)) and Open (right graph labeled (b)) programs, to test for bunching near the cutoff. There is more density overall to the left of the cutoff because there are more losers than winners.

Figure 4: Probability of Venture Capital by Rank Around Cutoff

(a) Open

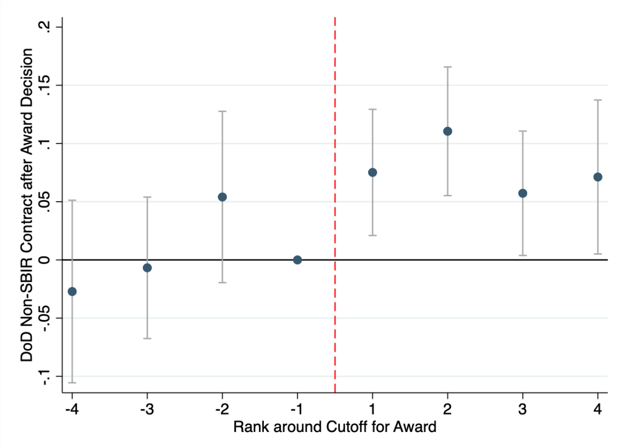
(b) Conventional (2017-19)



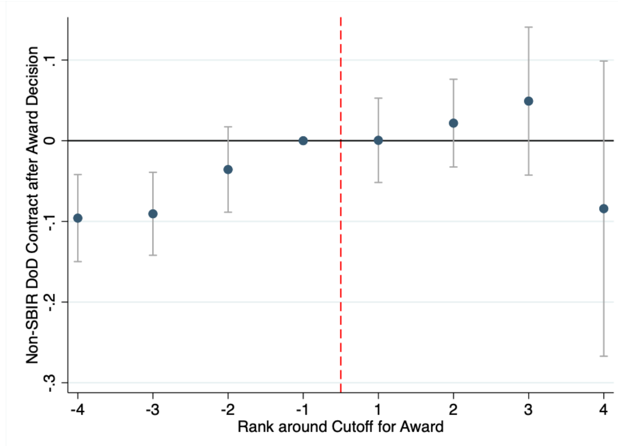
Note: These figures show the probability that an applicant firm raised venture capital investment (VC) within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. We plot the points and 95% confidence intervals from a regression of the outcome on a full complement of dummy variables representing each rank, as well as fixed effects for the topic. The omitted group is rank=-1. We include first applications from 2017-19.

Figure 5: Probability of DoD non-SBIR Contract by Rank Around Cutoff

(a) Open



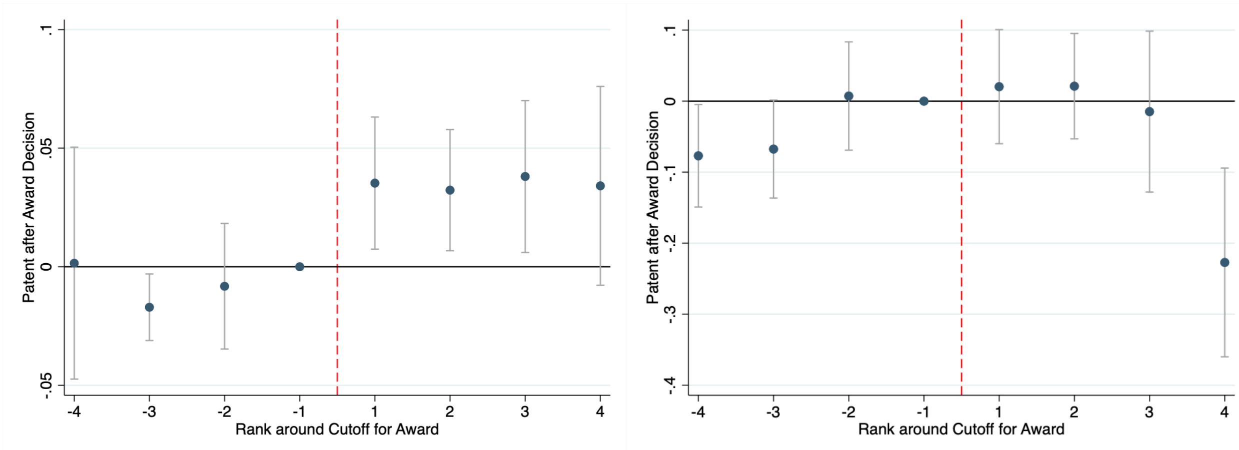
(b) Conventional (2017-19)



Note: These figures show the probability that an applicant firm had any non-SBIR DoD contracts valued at more than \$50,000 within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. We plot the points and 95% confidence intervals from a regression of the outcome on a full complement of dummy variables representing each rank, as well as fixed effects for the topic. The omitted group is rank=-1. We include first applications from 2017-19.

Figure 6: Probability of Patents by Rank Around Cutoff

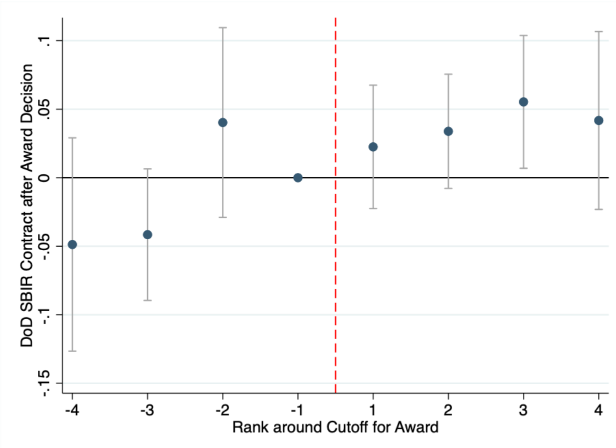
(a) Open (b) Conventional (2017-19)



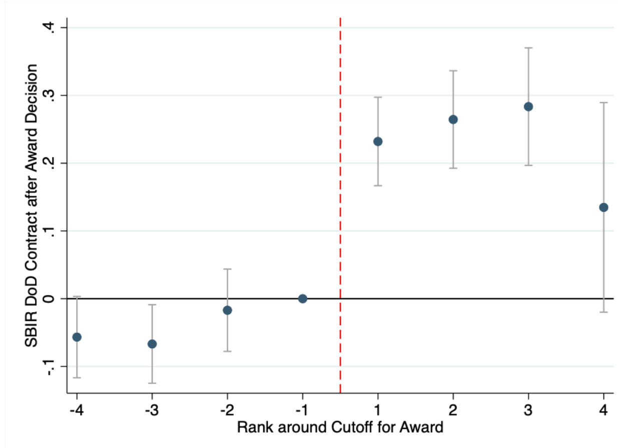
Note: These figures show the probability that an applicant firm had any ultimately granted patent applications within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. We plot the points and 95% confidence intervals from a regression of the outcome on a full complement of dummy variables representing each rank, as well as fixed effects for the topic. The omitted group is rank=-1.We include first applications from 2017-19.

Figure 7: Probability of Air Force SBIR Contract by Rank Around Cutoff

(a) Open



(b) Conventional



Note: These figures show the probability that an applicant firm had any Air Force SBIR contracts within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. We plot the points and 95% confidence intervals from a regression of the outcome on a full complement of dummy variables representing each rank, as well as fixed effects for the topic. The omitted group is rank=-1. We include all data for Conventional rather than only 2017-19 because the effect is only observed in the whole sample, as the sample in 2017-19 is overwhelmingly repeat-firms.

Table 1: Summary Statistics for Main Estimation Sample

Panel A: Competition and Company Summary

	Open Topic				Conventional			
	N	Mean	Median	SD	N	Mean	Median	SD
Competition Summary								
Num Proposals per Topic	1,659	379.327	375	156.453	4,995	19.808	15	17.131
Num Winners per Topic	1,659	212.842	297	115.083	4,995	3.090	2	3.606
Topic Non-Specificity	1,648	3.907	4	0.515	4,974	1.000	1	0.540
Award Amount	269	$49,569	$50,000	$14,636	876	$147,235	$152,718	$25,296
Company Characteristics								
Age	1,659	9.794	5	10.981	4,995	18.166	16	13.133
Number of Employees	1,659	26.885	8	60.687	4,995	60.774	20	90.802
$\mathbb{1}$(in VC Hub)	1,659	0.197		0.397	4,995	0.148		0.355
$\mathbb{1}$(in County with AF Base)	1,659	0.192		0.394	4,995	0.275		0.446
$\mathbb{1}$(Minority Owned)	1,659	0.121		0.326	4,993	0.127		0.333
$\mathbb{1}$(Woman owned)	1,659	0.111		0.314	4,993	0.155		0.362
$\mathbb{1}$(Proposal is Hardware)	1,659	0.240		0.427	4,995	0.514		0.500

Panel B: Pre-Award Outcome Summary

	Open Topic				Conventional			
	N	Mean	Median	SD	N	Mean	Median	SD
$\mathbb{1}$(VC)	1,659	0.114		0.318	4,995	0.060		0.238
Avg VC Amt (Mill)	154	$6.859	$1.925	$14.683	204	$3.643	$0.700	$6.633
$\mathbb{1}$(DoD Non-SBIR Contract)	1,659	0.253		0.435	4,995	0.601		0.490
# DoD Non-SBIR Contracts	420	12.310	4	33.286	3,000	20.174	9	29.737
Avg DoD Non-SBIR Contract Amt (Mill)	420	$1.631	$0.697	$2.772	3,000	$1.757	$0.868	$4.189
$\mathbb{1}$(Patent)	1,659	0.250		0.433	4,995	0.473		0.499
# Patents	415	12.313	3	39.420	2,364	26.678	10	45.638
# Patent Application if Any	515	10.996	3	36.026	2,554	25.691	9	45.461
$\mathbb{1}$(AF SBIR Contract)	1,659	0.189		0.391	4,995	0.593		0.491
# AF SBIR Contracts	313	21.856	8	40.649	2,960	50.405	18	76.366
$\mathbb{1}$(Never Awarded SBIR)	1,659	0.691		0.462	4,995	0.283		0.451

Note: Panel A of this table shows summary statistics about the Phase 1 competitions, as well as select company characteristics as of the application date. Panel B shows summary statistics of variables used as outcomes in the analysis, all calculated for the period before the award decision to facilitate evaluating selection into applying for the different programs. The data are restricted to all applications in our main analysis sample (the Open and Conventional programs from 2017-19).

Table 2: Effect of Winning on Main Outcomes

Panel A: Any Subsequent Venture Capital Investment

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.054**	-0.005	0.010	-0.005	0.010	0.005
	(0.025)	(0.019)	(0.010)	(0.019)	(0.010)	(0.003)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				0.059**	0.045*	0.046*
				(0.027)	(0.023)	(0.025)
Observations	1385	2608	7384	3993	8769	21432
Program	Open	Conv.	Conv.	Both	Both	Both
Proposal	First	First	First	First	First	All
Time Period	2017-19	2017-19	2003-19	2017-19	2003-19	2003-19
Outcome Mean	0.079	0.019	0.017	0.040	0.027	0.017

Panel B: Any Subsequent Non-SBIR DoD Contracts

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.075**	0.033	0.015	0.033	0.015	-0.022
	(0.035)	(0.052)	(0.031)	(0.051)	(0.031)	(0.013)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				0.042	0.060	0.109***
				(0.067)	(0.055)	(0.036)
Observations	1385	2608	7384	3993	8769	21432
Program	Open	Conv.	Conv.	Both	Both	Both
Proposal	First	First	First	First	First	All
Time Period	2017-19	2017-19	2003-19	2017-19	2003-19	2003-19
Outcome Mean	0.148	0.324	0.230	0.263	0.217	0.421

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on the probability of any VC investment (Panel A) and any non-SBIR DoD contract valued at more than $50,000 (Panel B) within 24 months after the award decision for Open topics and Conventional topics. Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. In all cases, we control for previous Air Force SBIR awards. Columns (1) to (3) estimate the effect of winning separately for each program. Column (1) contains estimates of the effect of winning an Open topic. Columns (2) and (3) contain estimates of the effect of winning a Conventional topic for years between 2017-19 and 2003-19, respectively. Columns (4) through (6) pool both Open and Conventional programs and interact winning an award with an indicator that is equal to one if a proposal is in an Open topic (and zero otherwise). Column (4) restricts the sample to the years between 2017-2019 and columns (5) and (6) include all years 2003-19. Columns (1) through (5) restrict the sample to the firm's first application within the sample time period whereas column (6) uses all proposals. All columns include topic fixed effects. Standard errors are below coefficients (in parentheses) and are clustered by firm in columns (1) through (3) and by topic in columns (4) through (6). ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Effect of Winning on Ancillary Outcomes

Panel A: Any Subsequent Patents

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.051***	0.077	-0.022	0.077	-0.022	-0.022*
	(0.017)	(0.051)	(0.025)	(0.050)	(0.025)	(0.013)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				-0.026	0.073*	0.069**
				(0.057)	(0.037)	(0.030)
Observations	1385	2608	7384	3993	8769	21432
Program	Open	Conv.	Conv.	Both	Both	Both
Proposal	First	First	First	First	First	All
Time Period	2017-19	2017-19	2003-19	2017-19	2003-19	2003-19
Outcome Mean	0.027	0.146	0.158	0.105	0.137	0.235

Panel B: Any Subsequent SBIR Contracts

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.040	0.172***	0.288***	0.172***	0.288***	0.123***
	(0.028)	(0.048)	(0.028)	(0.047)	(0.027)	(0.011)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				-0.133	-0.248***	-0.114**
				(0.082)	(0.076)	(0.056)
Observations	1385	2608	7384	3993	8769	21432
Program	Open	Conv.	Conv.	Both	Both	Both
Proposal	First	First	First	First	First	All
Time Period	2017-19	2017-19	2003-19	2017-19	2003-19	2003-19
Outcome Mean	0.105	0.312	0.229	0.240	0.210	0.462

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on the probability of any ultimately granted patent applications (Panel A) and any SBIR DoD contracts (Panel B) within 24 months after the award decision for Open topics and Conventional topics. Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. In all cases, we control for previous Air Force SBIR awards. Columns (1) through (3) contain estimates of the effect of winning separately for each program. Column (1) contains estimates of the effect of winning an Open topic. Columns (2) and (3) estimate the effect of winning a Conventional topic for years between 2017-19 and 2003-19, respectively. Columns (4) through (6) pool both Open and Conventional programs and interact winning an award with an indicator that is equal to one if a proposal is in an Open topic (and zero otherwise). Column (4) restricts the sample to the years between 2017-2019 and columns (5) and (6) include all years 2003-19. Columns (1) through (5) restricts the sample to the firm's first application within the sample time period; whereas column (6) uses all proposals. All columns include topic fixed effects. Standard errors are below coefficients (in parentheses) and are clustered by firm in columns (1) through (3) and by topic in columns (4) through (6). ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4: Robustness Tests

Panel A: Controls

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
𝟙(Award)	0.051**	-0.006	0.049***	0.064	0.057*	0.028	0.040	0.189***
	(0.025)	(0.017)	(0.017)	(0.042)	(0.034)	(0.049)	(0.028)	(0.043)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.019	0.027	0.146	0.148	0.324	0.105	0.299

Panel B: No Controls

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
𝟙(Award)	0.060**	-0.005	0.049***	0.081*	0.045	0.040	0.021	0.204***
	(0.025)	(0.018)	(0.016)	(0.045)	(0.037)	(0.054)	(0.029)	(0.050)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.019	0.027	0.146	0.148	0.324	0.105	0.299

Panel C: Narrow Bandwidth

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
𝟙(Award)	0.062***	0.006	0.037***	0.015	0.072***	0.050	0.022	0.028
	(0.020)	(0.011)	(0.013)	(0.031)	(0.027)	(0.034)	(0.021)	(0.031)
Observations	671	902	671	902	671	902	671	902
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.059	0.018	0.021	0.216	0.153	0.434	0.089	0.550

Panel D: All Proposals

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
𝟙(Award)	0.051**	-0.000	0.047***	0.065**	0.088***	0.006	0.009	0.097***
	(0.024)	(0.012)	(0.015)	(0.031)	(0.033)	(0.034)	(0.029)	(0.035)
Observations	1659	4995	1659	4995	1659	4995	1659	4995
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.014	0.028	0.142	0.160	0.467	0.113	0.442

Panel E: Ever-After Outcomes

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(Award)	0.054**	-0.012	0.051***	0.064	0.075**	0.027	0.040	0.172***
	(0.025)	(0.022)	(0.017)	(0.044)	(0.035)	(0.047)	(0.028)	(0.045)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.025	0.027	0.151	0.148	0.362	0.105	0.312

Panel F: Conventional 2003-2017

Dep Var:	Any VC	Any Patents	Any DoD Contracts	Any SBIR
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	0.011	-0.018	0.022	0.332***
	(0.008)	(0.023)	(0.025)	(0.024)
Observations	6628	6628	6628	6628
Outcome Mean	0.016	0.171	0.239	0.206

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on the probability of any venture capital investment (columns (1) and (2)), any ultimately granted patent applications (columns (3) and (4)), any subsequent DoD non-SBIR contract valued at over $50,000 (columns (5) and (6)), and any DoD SBIR contract (columns (7) and (8)) within 24 months after the award decision for Open and Conventional Topics. In Panel A, we add a full suite of controls for whether the firm had any previous patents, previous VC, previous non-SBIR DoD contract, and whether the firm is located in a VC hub city or a county with an Air Force base, whether it has a software rather than hardware-based technology, as well as a continuous variable for firm age, in addition to any previous Air Force SBIR award. In Panel B, we do not include any controls. In Panel C, we restrict the bandwidth to include only two applicants on each side of the cutoff. In Panel D, we include all proposals in the 2017-19 period rather than only a firm's first (so a firm may appear twice). In Panel E, we allow outcomes to be ever-after rather than within two years. In Panel F, we consider Conventional effects before the Open program was implemented (2003-2017). In all panels, the sample is restricted to first-time applicants only, and for Panels A-E, to award years 2017-19. All columns include topic fixed effects. Standard errors (in parentheses) are below coefficients and are clustered by firm for all panels. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Heterogeneity in Effect of Winning on Venture Capital Investment

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.063**	0.083***	0.043*	0.000	-0.008	0.009
	(0.026)	(0.028)	(0.026)	(0.026)	(0.024)	(0.021)
$\mathbb{1}$(Award) × $\mathbb{1}$(Prev. SBIR)	-0.072***			-0.008		
	(0.027)			(0.018)		
$\mathbb{1}$(Award) × $\mathbb{1}$(High Age)		-0.098***			0.004	
		(0.026)			(0.017)	
$\mathbb{1}$(Award) × $\mathbb{1}$(Hardware)			0.058			-0.022
			(0.037)			(0.015)
Observations	1385	1385	1385	2608	2608	2608
Program	Open	Open	Open	Conv.	Conv.	Conv.
Time Period	2017-19	2017-19	2017-19	2017-19	2017-19	2017-19
Outcome Mean	0.079	0.079	0.079	0.019	0.019	0.019

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on the probability of any VC investment within 24 months after the award decision for Open topics (columns (1) through (3)) and Conventional topics (columns (4) through(6)). Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. In all cases, we control for previous Air Force SBIR awards. Columns (1) and (4) interact winning an award with an indicator for having won previous Air Force SBIR awards. Columns (2) and (5) interact winning an award with an indicator for a young firm defined as firm's age below the median of the distribution of firm age (and zero otherwise). Columns (3) and (6) interact winning an award with an indicator for whether the proposal was classified as hardware-related by the k-means clustering algorithm. All columns include topic fixed effects. Standard errors (in parentheses) are below coefficients and are clustered by firm. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6: Effect on Venture Capital Investment of Open, Conventional, and Other Reform Programs Relative to Each Other

Sample:	Open, Conv & NSIN	Open, Conv & Pitch Day	All	NSIN Topics	Pitch Day Topics
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topics)	0.059**	0.059**	0.059**		
	(0.027)	(0.027)	(0.027)		
$\mathbb{1}$(Award) × $\mathbb{1}$(NSIN)	-0.014		-0.014		
	(0.066)		(0.066)		
$\mathbb{1}$(Award) × $\mathbb{1}$(Pitch Day)		-0.106	-0.106		
		(0.060)	(0.060)		
$\mathbb{1}$(Award)	-0.005	-0.005	-0.005	-0.019	-0.111
	(0.019)	(0.019)	(0.018)	(0.063)	(0.067)
Observations	4416	4317	4740	423	324
Outcome Mean	0.043	0.042	0.044	0.071	0.062

Note: This table compares the effect of winning an award on the probability of any VC investment within 24 months after the award decision for Open, Conventional, and two other "reform" topics, relative to conventional topics. The other reform topics are Pitch Day and NSIN (discussed in the text). Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. We fully interact all right-hand side variables with the coefficient of interest (e.g., rank on either side of the cutoff is interacted with the indicator for Open in column (1)). In columns (1) through (3), the base group is Conventional proposals. We assess whether the effect of winning Open and NSIN is significantly different from Conventional in column (1). Column (2) is similar, but considers Pitch Day instead of NSIN. Column (3) includes all four programs. In columns (4) and (5), we consider the effect of winning within Pitch Day and NSIN topics as separate samples, respectively. We use our main sample of first proposals from 2017-19 in all columns. Note the topic fixed effects control for the independent effects of Open, NSIN, and Pitch Day. Standard errors are clustered by topic in columns (1) through (3), and by firm in columns (4) and (5). ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 7: Effect of Award on Patenting in Conventional Program by Topic Specificity

Panel A: The Role of Topic Specificity in Conventional Topics

	Any Patent		Any High Citation Patent		Any High Originality Patent	
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	-0.022	-0.007	-0.036*	-0.013	-0.028	-0.011
	(0.024)	(0.015)	(0.019)	(0.011)	(0.020)	(0.012)
$\mathbb{1}$(Award) $\times$ Non-specificity	0.055**	0.040**	0.038**	0.045***	0.046**	0.023*
	(0.025)	(0.017)	(0.018)	(0.014)	(0.022)	(0.014)
Observations	7384	17500	7384	17500	7384	17500
Proposals	First	All	First	All	First	All
Outcome Mean	0.158	0.253	0.076	0.118	0.103	0.177

Panel B: The Role of Topic Specificity in Conventional and Open Topics

	Any Patent		Any High Originality Patent	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	-0.022	-0.007	-0.028	-0.011
	(0.025)	(0.015)	(0.020)	(0.013)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open)	-0.045	-0.047	-0.028	-0.007
	(0.079)	(0.052)	(0.066)	(0.043)
$\mathbb{1}$(Award) $\times$ Non-specificity	0.046*	0.039**	0.036*	0.021
	(0.026)	(0.016)	(0.022)	(0.014)
Observations	8769	19159	8769	19159
Proposals	First	All	First	All
Outcome Mean	0.137	0.235	0.090	0.165

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on patent-based variables within 24 months after the award decision in the Conventional program, where the effect of winning is modulated by the index of topic specificity. Specifically, we interact winning with our demeaned "non-specificity" index. A higher value of non-specificity means the topic is more bottom-up based on the diversity of proposals it attracted (see Section 6.3 for details). In Panel A, the dependent variable in columns (1) and (2) is an indicator for whether a firm was granted a patent that was applied for after the award decision. The other two outcomes are quality measures of these patents. Columns (3) and (4) consider whether a firm obtained a patent with above sample median future citations (defined among the applications in our sample). This is a measure of patent quality that is informative about the impact of a patent on future research. Columns (5) and (6) consider whether the patent had above median originality, which measures whether the patent cites previous patents in a wide range of fields. These outcomes are described in detail in Appendix C.2. In Panel B, we include the Open applications to assess whether there is still a robust interaction of award and Open after controlling for the effect of winning in non-specific Conventional topics. We do not include the citations outcome because there is not enough time after the Open awards for citations to accrue. The following statements apply to both panels: Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. For each outcome, we report one model (odd columns) restricted to first-time applicants and one using all applications (even columns). All columns include topic fixed effects, which absorb the measure of specificity. We include all years (2003-19). Standard errors are under coefficients (in parentheses) and clustered by firm. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Appendix

(For Online Publication)

A Slowing Innovation in the US Defense Industry

Despite widespread concern over several decades among policymakers, we know of no studies of the evolution of prime defense contractors' innovation.[31] Here, we document innovation trends focusing on the top eight contractors over the past two decades: Boeing, Raytheon, Lockheed Martin, Northrop Grumman, General Dynamics, United Technologies Corp, Harris, and L-3. We researched all of their acquisitions since 1976 of companies that were also defense contractors and linked the eight primes and all their acquisition targets to the NBER/USPTO patent database and Compustat.

Figure A.1 shows that between 1976 and 2019, 225 companies consolidated into just six, with L-3 and Harris merging in 2018, and Raytheon and United Technologies merged in 2020. Remarkably, the dollar share of total defense contracts that these firms have won, shown in the grey area, has stayed fairly constant over the years at roughly 35%.[32] The value (in 2019 dollars) of these contracts increased from around $70 billion spread across 225 companies in the late 1970s to $115 billion awarded to just six companies in 2019. The number of firms responsible for the remaining roughly 65% of contract value not represented in the graph declined slightly from 25,339 unique contractors in 1976 to 24,656 in 2018. To confirm that the remaining contracts have not become more dispersed, we present the Herfindahl-Hirschman Index (HHI) of concentration for all non-SBIR DoD contracts, though this measure is not very insightful because the defense market is composed of myriad small markets for items ranging from food supplies at a particular base to a fleet of fighter jets. Nonetheless, the dashed orange line in Figure A.4 Panel A shows that overall concentration has remained relatively stable, albeit volatile.

The dramatic consolidation among the primes has been accompanied by a decline in innovation quality as measured by patent citations, which shed light on private sector spillovers. Figure 1 shows patent activity for the firms in Figure A.1, weighted by future citations. Patent activity is only one proxy for innovativeness, but it is relevant to DoD-funded innovation. While a patent involves some disclosure, there are often trade secrets that prevent a competitor from copying the invention even once the patent is public, and a patent can coexist with classified aspects of the research that do not appear in the patent itself.

In 1976, the figure includes patents from all 225 companies, and in 2019 we are considering patents from the six companies. Citations are normalized by the average number of citations for all patents in the same CPC3 Technology class by year cohort, so that a number above one indicates the patent is more impactful than the average patent in its class-year.[33] The

[31] Carril and Duggan (2020) show that the substantial consolidation among major defense contractors in the mid-1990s reduced competition and led to a shift to cost-plus contracts in which cost escalations are uncapped.

[32] We exclude DoD contracts to Humana (health insurance provider) and universities.

[33] We use a kernel-weighted polynomial to smooth the lines (the results are very similar with a binscatter approach).

solid blue line includes all forward citations, and we see a secular decline across the unit threshold, so that defense patents changed from being relatively more innovative to relatively less innovative within their narrow technology areas. This pattern is even starker when we include only outside citations to patents from firms that are not featured in the graph. That is, we exclude citations from firms outside the prime contractor universe. These citations are shown in the dashed green line. They decline from having 17% more citations from outside defense than the average patent in the class year in 1976 to 60% fewer citations in 2019. These trends suggest a prime contractor base that has become markedly more insular over time.

To assess whether firms are innovating in new areas that could have novel defense applications (e.g. software, clean energy), we also calculate a firm's share of "explorative" patents in any given year, following Manso (2011). An explorative patent is a patent filed in technology classes previously unknown to the firm in a given year. Figure A.2 shows the average share of exploratory patents relative to other firms with similar in age, size, and year. As above, all firms from Figure A.1 are included. Age is defined as the year from the firm's first observed patent and size is defined as the firm's patent stock in a given year. As firms merge, they acquire new areas of expertise, and we expect this should lead to increasing exploration since the assignee after the acquisition is usually the acquiring parent firm. This seems to be true to some extent for the big mergers of the 1990s but is not true subsequently. Instead, in Figure A.2 we see a marked decline over time, indicating that the defense contractors are not patenting in new technology areas even as they acquire each other. By 2019, the share of explorative patents was 60% lower than firms with similar patent stocks and age since the first patent.

Figure A.3 shows other variables relevant to prime contractor innovation. In Panel A, we compare the growth in the number of patents for the primes to growth among all other U.S. assignees in the USPTO. Until the early 1990s, the defense contractors were patenting at similar rates as the overall universe, but we see a subsequent divergence, with defense contractors patenting at a lower rate.[34] The subsequent three panels use Compustat data and compare primes to other firms in the same three-digit NAICS industry.[35] Panel B shows that before the mid-1990s, the primes had a higher ratio of profits to R&D than peer firms, but by 2019, they earned $8 for each R&D dollar compared to $5.50 in the comparison group. Panel C shows that the level of profits has increased much more for primes than for other firms and Panel D shows that R&D has grown since 1976, but more slowly than revenue and assets.

In short, there has been a big increase in concentration among prime defense contractors. Although their profits and assets have increased substantially, this has been accompanied by a fall in the primes' relative innovation whether measured by citations, patenting or R&D intensity. The key transition appears to have occurred after the Cold War ended, during the period of lower defense budgets and consolidation during the 1990s but continued into the period of higher spending following 9/11 and the Iraq War.

B Institutional Details on Open Reform and Comparison to Conventional Topics

The SBIR reforms have taken place within a new organization called Air Force Ventures (AFVentures), a business division of AFWERX.[36] AFVenture's stated goals are to leverage private capital to deploy new innovations for the military, to expand the industrial base interested in defense, and to grow the U.S. economy. That is, they hope to address the

[34] This coincides with a major merger wave in the mid-1990s when, among others, Northrup merged with Grumman, McDonnell Douglas merged with Boeing, and Lockheed merged with Martin Marietta.

[35] Since many acquisitions were of unlisted firms, the figures only include the acquisition targets after acquisition, so must be treated with more caution.

[36] https://www.afwerx.af.mil/

challenges facing military procurement identified in Section 2.2. The idea is that if the bottom-up approach is successful in this context, it might be applied to the larger acquisition programs with the hope of garnering interest in the defense market among the large tech firms in areas such as cybersecurity and artificial intelligence.[37] Senior leaders perceive commercial innovation metrics as measures of successful Air Force R&D investment, with the idea that an innovative U.S. industrial base will, in the long term, enable military supremacy, especially if the research has early-stage ties to the defense market.

AFWERX and AFVentures are one of a number of initiatives that the Defense Department has instituted, since about 2015, aiming to reduce barriers between defense field missions and commercially focused companies that are not traditionally defense contractors.[38] Many of these programs make use of Congressional authorization for increased spending through "Other Transaction Authorities" (OTA), which do not require adherence to the arduous regulations and competition requirements that govern most contracts. Congress noted when making these authorizations in 2016 that "We believe that expanded use of OTAs will support Department of Defense efforts to access new source[s] of technical innovation, such as Silicon Valley startup companies and small commercial firms."[39]

More broadly, AFWERX is representative of many institutions established in the 2010s around the world reflecting a realization that the traditional defense sector is no longer at the cutting edge of innovation. Instead, the private sector, especially nimble startups and the venture capitalists who fund and guide them are perceived to be at the frontier of innovation in many areas. Important features of this entrepreneurial ecosystem are a willingness to experiment and access, through both co-location as well as pecuniary and non-pecuniary benefits, to high-skill human capital. Militaries around the world are therefore focusing energies on funding and working with high-tech, small businesses that possess "dual-use" technologies with commercial as well as defense applications.

DoD SBIR awards are in the form of contracts. This contrasts with some agencies, such as the DoE or the NIH, which deliver SBIR awards in the form of grants. With a grant, the application defines the scope of work, payment is entirely up-front, and the government has little recourse in the event that the firm does not use the money as intended. Conversely, contracts represent a binding agreement between the government and the firm to deliver a good or service. Payment only comes after the firm has accomplished some pre-established milestone. Therefore, risk and liquidity are allocated differently across the two instruments. Grants offer the firm money upfront, and the government takes the risk that the project (or the firm) will fail. Contracts allocate more risk to the firm and require the firm to finance the investment upfront. In the context of financially constrained startups, this may present a challenge.

First conducted in May 2018, Open topics are the centerpiece of AFWERX's reformed SBIR program. Open topics are "bottom-up" in that the solicitation contains no direction regarding the technology – including software – that the applicant may propose. With an explicit reference to seeking "unknown unknowns" in the solicitation, Open topics are designed to let the private sector do the work of identifying military applications for its technology. The solicitation explains:

> "The objective of this topic is to explore Innovative Defense-Related Dual-Purpose Technologies that may not be covered by any other specific SBIR topic and thus

[37] With the notable exception of Palantir, the large West Coast-based tech firms do minimal work with the DoD.

[38] Some of the new initiatives include SOFWERX (part of the Special Operations Command), the Defense Innovation Unit (DIU), the Defense Innovation Board, and the National Security Innovation Network (NSIN), the Army Venture Capital Initiative, and the Capital Factory in Austin, an incubator "tech hub" that houses offices of AFWERX, Army Applications Lab, and DIU.

[39] U.S. Congress, House Committee on Armed Services, National Defense Authorization Act for Fiscal Year 2016, committee print, Legislative Text and Joint Explanatory Statement to accompany S. 1356, P.L. 114-92, 114th Cong., 1st sess., November 2015, pp. 700-701.

> to explore options for solutions that may fall outside the Air Force's current fields of focus but that may be useful to the U.S. Air Force. An additional objective of this topic is to grow the industrial base of the U.S. Air Force."

The firm's objective is to demonstrate the feasibility of developing a product or service with an Air Force partner interested in potentially procuring the firm's technology. The Phase 1 deliverable is a white paper, or report describing the outcomes of research. The Open topics are aimed at firms already developing a technology aimed at commercial use, even if it is in the very early stages

In contrast, Conventional topics tend to fund R&D projects nominally geared towards a particular military use. Conventional topics are sourced primarily from the Air Force Research Laboratory (AFRL). They are highly specific; some examples of topics are:

- "Affordable, Durable, Electrically Conductive Coating or Material Solution for Silver Paint Replacement on Advanced Aircraft"
- "Safe, Large-Format Lithium-ion (Li-ion) Batteries for ICBMs"
- "Develop Capability to Measure the Health of High Impedance Resistive Materials"
- "Standalone Non-Invasive Sensing of Cyber Intrusions in FADEC for Critical Aircraft System Protection"
- "Hypersonic Vehicle Electrical Power Generation through Efficient Thermionic Conversion Devices"
- "Cyber Attack model using game theory"

Each year, there are usually three solicitations, each of which has many Conventional topics but only one Open topic since 2018. For example, in the second solicitation of 2019, there was one Open topic and 61 Conventional topics. All Open topics are the same; there are multiple topics because they are issued at different points in time (i.e., in different solicitations). The pool of competitors a given applicant faces in the Open topic depends on when it applies, as scoring and ranking are within-topic. This creates a different distributional structure in Open topics relative to Conventional, as there are many more applicants but also far more winners. The difference in topic structure should not bias the results towards favoring a stronger effect in Open because we estimate the effect of winning within each program, and the cutoff point for winning is lower in the score distribution for Open.

Open topic awards are also smaller than Conventional ($50,000 vs. $150,000) and have shorter time frames (3 vs. 9 months). AFWERX's belief that offering many very small awards can be useful was in part informed by existing research finding strong positive effects on VC and patenting from small, early-stage Phase 1 awards (Howell 2017). Note that the budget for each of the hundreds of topics is determined before the competition, and depends on factors such as the overall funding settlement for U.S. Air Force's SBIR program, military priorities, etc. Hence, the precise threshold will be competition-specific and depend on the number and quality of the applicants for each solicitation.

This paper focuses on Phase 1, so we minimize the discussion of further awards. The Phase 2 awards of $300,000 to $2 million are intended to last 12-24 months and fund a prototype, software, or dataset. For all but the first two of its Open SBIR topics, AFWERX sought to encourage Phase 1 winners to access outside funding from either private or or government sources with a matching provision in Phase 2. Below, we evaluate the impact of match availability separately from openness. Figure A.5 shows clustering of awards, particularly in Phase 1, around the maximum amount. Some firms apply for less than the maximum, apparently because firms must apply for the amount of money required to do the work they are proposing.[40] Phase 2 contracts are much more detailed, bespoke, with higher and more varied amounts than Phase 1 (see Figure A.5).

C Data Sources, Sample Construction, and Key Variables

C.1 Data Sources and Sample Construction

Our starting point is a dataset of applications and awards to the Air Force SBIR program between 2003 and 2019. All awards are publicly available, most easily from the SBA's website www.sbir.gov. Our causal analysis, however, requires applications and evaluations; that is, knowing which firms applied and lost as well as the internal scores that determine award status for all applications. We observe complete evaluation data for all topics between 2017 and 2019, and further evaluation data for Conventional topics in select earlier years: 2003-2007, 2015, and part of 2016. The remaining years' data were, unfortunately, inadvertently destroyed in 2016. The application and evaluation materials are not public information.

We analyze the effect of winning a Conventional award using all our data, but the main focus of this paper is to compare Open and Conventional. To do this, we restrict the sample to the three years of 2017-2019, so that the relevant economic environment and defense procurement factors are similar across the sample. In 2017, all applicants are Conventional. In 2019, four-fifths of applicants are Open. Figure A.6 shows the number of awards by program and year, and Table A.1 describes counts of topics, firms, and proposals for all programs. In the 2017-19 sample, there are 7,229 Phase 1 proposals from 3,170 unique firms.

Conventional topics average 20 applicants and three winners (i.e., awardees). Open topics have a very different model, leading to many applicants and winners per topic (on average, 380 and 213, respectively). Table A.2 shows similar statistics for the whole Conventional sample. Summary statistics for Phase 2 are in Table A.3.

C.2 Outcome Variables

The two main outcomes of interest, in Table 1 Panel B, are subsequent VC investment and DoD non-SBIR contracts, which correspond to the two key metrics of success from AFWERX's perspective. The current Air Force leadership views commercial innovation as evidence of initial success, based on the idea that a strong U.S. industrial base (especially if its research has early-stage ties to DoD) will ultimately enable strong defense (Williams 2020). From an economic perspective, VC investment is a useful proxy for high-growth innovation potential. Although VC-backed startups make up only 0.11% of new firms, over 44% of public company R&D is performed by formerly VC-backed startups (Puri and Zarutskie 2012, Gornall and Strebulaev 2015).

We obtain unique private financing deals from Pitchbook, CB Insights, SDC VentureXpert, and Crunchbase. The majority of deals come from Pitchbook, which we observe through January 10, 2021. While there are likely VC investments that do not appear in these databases, they are the industry state-of-the-art and are widely used (Lerner and Nanda 2020, Gornall and Strebulaev 2020). We match firms to these datasets on name and state, and then check manually for false positives. Clearly there may be some errors, in particular as firms can change names. However, there is no reason the error rate should be systematically different across Open and Conventional in the 2017-19 time frame.

[40] There is also apparently some misconception that cost will be a key factor in evaluation, despite explicit information in the solicitation that it will not.

The second outcome is non-SBIR DoD contracts, representing defense procurement success in the sense that the research has led to a practical application for the military; in the DoD jargon this is often termed "transition to programs of record." An example of a successful Open applicant from the perspective of "transitioning" to Air Force operations is Alabama-based Aevum, which designs drone-launched rockets in a former textile mill. After winning a $50,000 Open Phase 1 award in July 2019, Aevum was awarded a $4.9 million Air Force launch contract in September 2019. An example in the Conventional program is Cornerstone Research, which won a $1.9 million SBIR Phase III contract from AFRL (the sponsor of its SBIR award) for further research and development in biotechnology.

To construct this outcome, we use complete data from the Federal Procurement Data System (FPDS) through July 2020. The FPDS dataset is a single, comprehensive dataset of all federal contracts. We remove all non-DoD contracts and then match to our data on firm DUNS number. For the portion that do not match on DUNS, we use firm name and state, and manually check for false positives. Among the matched contracts, 64% were matched on DUNS. While it is possible that we are missing some contracts, the error rate should be small. We restrict to contracts worth at least $50,000 so that we do not capture very small add-on type awards or minor purchases. Among the matched contracts, 42% of contracts by volume and 99% by value are over this threshold. The results are similar using all matched contracts.

We also consider two ancillary outcomes that are related to innovation and lock-in dynamics, respectively. First, we consider patents from the USPTO through December 2020 to assess technical innovation with potential commercial applications. We match SBIR applicants to patent assignees on firm name and state. Our primary outcome is an indicator for the firm having any granted patents that were applied for after the award date. That is, we use the application date (as opposed to the award date), but we restrict to granted patents.

For supplementary analysis, we construct four other patent-based variables. First, we consider the number of forward citations, which we normalize by patent class and by year to adjust for the systematic differences across classes and years.[41] Forward citations are informative about the impact of a patent on future research. The second measure is generality. A high generality score indicates that the patent influenced subsequent innovations in a variety of fields (Trajtenberg et al. 1997).[42] The third measure is originality. The originality score will be low if a patent cites previous patents in a narrow set of technologies, whereas citing patents in a wide range of fields leads to a high score.[43] The last measure is the number of patent applications, which could represent innovation effort and is less truncated due to the lag between application and award. We obtained application data courtesy of Liat Belenzon, and merged these data to the SBIR data on firm name and state.

The second ancillary measures is subsequent Air Force SBIR awards, using data from the Small Business Administration (results are similar using all-DoD or any-agency SBIR awards). We examine whether winning one SBIR award causally increases the probability of winning a future one, to assess lock-in to the SBIR program.

[41] The citations data are from the USPTO.

[42] Generality for patent i is defined as $1-\sum_j s_{ij}^2$, where s_{ij} is the percentage of citations received by patent i that belong to patent class j. Thus, if a patent is cited by subsequent patents that belong to a wide range of fields the measure will be high, whereas if most citations are concentrated in a few fields it will be low (close to zero).

[43] Originality for patent i is defined as $1-\sum_j c_{ij}^2$, where c_{ij} is the percentage of citations that patent i makes that belong to patent class j.

C.3 Machine learning approach to classifying applications and topics

Here we give further details on how we measured characteristics of applications from their text, which is used both to assign firms to technology areas and to identify the non-specificity (i.e. openness) of topics. As noted in the main text, the raw applications data is not classified by industry or technology. As a way of classifying application types, we make use of the abstracts in the application proposals. We employ a machine learning algorithm called "k-means clustering" (see Forgy (1965) in the statistics literature or Bonhomme and Manresa (2015) in the econometrics literature) to classify each abstract based on its word "embedding."

We first map each word of the abstract into vector space using a pre-trained model that, based on corpuses of text, is able to identify words that are conceptually similar. For example, the vectors for words such as "happy" and "joy" would be close in distance, while vectors for words such as "happy" and "toolbox" would be quite distant from each other. Specifically, we use the SpaCy pipeline in Python, whose model is trained on OntoNotes with GloVe (Global Vectors for Word Representation) vectors trained on Common Crawl.[44] Each word embedding vector consists of 300 elements where an element is a value between -1 and +1. We then estimate the abstract embedding as the average of the word embeddings that make up the abstract. In this way, we can capture how similar abstracts are to one another using the average embedding. Next, we reduce the dimensionality of the abstract embeddings from three hundred dimensions to two. We do so nonlinearly using isometric mapping, following the framework in Tenenbaum et al. (2000).

Next, we cluster these abstract embeddings using the k-means clustering algorithm, whose objective is to minimize the total within-cluster variance. Note that this is unlike traditional topic modeling methods such as Latent Dirichlet Allocation, which focuses on the co-occurrence of words within topics and within the corpus of the given text but does not take into account the semantics and context of the words (i.e. the relationship between words themselves). The number of clusters is pre-set by the researcher, but the "elbow method" provides a criterion for finding the optimal number of clusters, which in our case is five (see Figure A.7). We present the five- and two-cluster model, as the latter yields a clear dichotomy between software- and hardware-based technologies. The word clouds for the clusters are in Figures A.8 and A.9. They show the keywords that form a topic cluster, with the word's size scaled to reflect its prevalence in the topic. We assign a name to the cluster based on the nature of the words. For the two-cluster model we have a cluster over what could describe as "Training/Software" and one which we could describe as "Hardware". For the five-cluster model we describe these as "Software/security," "Network/Data," "Simulation/Engineering," "Optical/Measurement," and "Materials/Manufacturing". We remove the most prevalent 75 words across all topics from the word clouds for clarity, as these are mostly filler words.

To calculate the "non-specificity" of a topic, we calculate the (squared Euclidean) distance of each application's embedding to the centroid of each topic and calculate the topic's average distance. This is akin to a measure of variance. When the average distance is low, this means that applications tend to be close to each other in the words they use, indicating that the topic is likely to have been relatively tightly specified. When the average distance in high, this suggests the words used to describe the applications are quite different from each other so the topic is "non-specific". In general, we would expect Open topics to be more non-specific than Conventional topics as there is little restriction on what can be proposed. This is indeed what we see in Table 1, where the index of non-specificity of Open is four times that of Conventional at the mean and median.

[44] OntoNotes is a large corpus consisting of various texts from news, conversational telephone speech, blogs, broadcasts, and talkshows. It is available at catalog.ldc.upenn.edu/LDC2013T19. GloVe is an unsupervised learning algorithm where training is done on global word-word co-occurrence statistics from a corpus (Pennington et al., 2014). Common Crawl is an open repository of web crawl data, available at www.commoncrawl.org.

We validate our approach to measuring topic non-specificity by manually examining the top and bottom 1% of topic titles. Among the top 1% of topics by non-specificity are "Wearable Device to Characterize Chemical Hazards for Total Exposure Health" and "Extended Weather Measurements in Support of Remotely Piloted Aircraft." Among the bottom 1% (most specific) are "Landing Gear Fatigue Model K Modification" and "Mitigation of Scintillation and Speckle for Tracking Moving Targets." This quick validation exercise gives us confidence that the non-specificity measure indeed reflects topic specificity.

D Further Results

D.1 Further Patent-Based Measures

In our RDD analysis, we also examined several other patent-related variables. The first is the number of applications. In unreported analysis, we find no effects, suggesting that the positive effect on granted patents does not simply reflect different levels of effort to apply for patents.

The second outcome is originality, which increases in the number of classes a patent cites. A more original patent is less siloed in a particular field and is more basic (Jaffe and Trajtenberg 2002). Given that our data are so recent, this measure has the advantage of being backward-looking. Using the measure from Jaffe and Trajtenberg (2002), Table A.5 shows that winning Open has a positive effect on producing an above-median originality patent (defined among all the applicants in our sample) while winning Conventional has no effect.

The third outcome is the number of patent citations, which reflect patent quality. We do not find any effects of Open on citations, which likely reflects insufficient time for them to accrue; indeed, there are no citations recorded whatsoever for the Open patents as they were granted mostly just before we stopped collecting data. For the 2003-19 period in Conventional, there is sufficient time for citations to accrue. However, in unreported analysis we find no effects of winning a Conventional award on either the number of citations. We also find no effects on the level of citations or in a negative binomial model for zero-inflated counts.[45]

D.2 Role of the Matching Program in the VC results

In Section 5.1 we found a large effect of winning an Open topic contract on VC and argued that one reason appears to be the potential of these contracts to serve as a gateway to much larger contracts at the Air Force beyond the SBIR program, which will support technology development and ultimately lead to off-the-shelf procurement in concert with commercial sales. There is also a second possible reason: the SBIR Phase 2 matching program. As explained in Section 2.3.3, an additional reform in the Open topics was to offer matching in Phase 2. Phase 2 applicants could request additional funds to match private or government money that they secured during the Phase 1 period. While the matching reform makes it more difficult to establish a pure treatment effect of openness, it also offers to our knowledge the first opportunity to evaluate a VC matching program. Researchers have long been interested in whether government programs that match VC solve information problems for the government agency or crowd out private capital (Lerner 2012).

Several features of the program's implementation facilitate evaluation. First, we can redefine the VC outcome to exclude VC investments that were matched in the Phase 2 stage.

[45] This contrasts with Howell (2017), where there is a large effect of DoE SBIR grants on patent citations. Above, we showed greater firm lock-in at DoD than at DoE. The greater focus on the defense market among DoD SBIR winners could reduce incentives to patent in the Conventional program or reduce limitations on patenting among non-winners of a topic. The Open program, by reaching firms that are already oriented towards the civilian market, appears to have a more positive effect on granted patents though it is too soon to identify effects on patent quality.

Second, the matching was not available at all for the first Open topic, and for the second topic, it was made available only just before firms submitted their Phase 2 applications. We can therefore assess whether the effect of winning an Open topic Phase 1 is concentrated in the later topics, where matching could have affected selection into applying for Phase 1. Third, we can assess whether the causal effect of Phase 1 on VC is driven by firms that apply for a Phase 2 match.

At the firm level, the fraction of Phase 1 winners that raised VC but never had a private match is 6.1%. The fraction that raised VC and also obtained a private match from the Phase 2 program is 1.3%. Table A.6 provides summary statistics on the matching program within the sample of firms that applied to Phase 2. The average confirmed private funding amount – that is, the event for which a matching contract was awarded for up to $750,000 – is $1.3 million.[46] Among Phase 2 applicants, 20.6% applied for a private match and 14.2% both won Phase 2 and received a matching award. Private funds are categorized as either VC, which means the matching came from an institutional VC fund or any other private source. Almost 40% of the private matches are from VC.

We are interested in whether the matching program was successful in driving subsequent VC, and also whether there are effects of winning an Open Phase 1 award on VC independently of whether the firm ultimately received a Phase 2 matching contract. In Table A.7, we repeat the main specification from Table 2 Panel A column (1) but make certain adjustments. In column (1) we redefine the outcome variable to be an indicator for subsequent VC if the firm did not receive a Phase 2 VC match. That is, the outcome of VC is zero if the firm did receive VC and got it matched in Phase 2. The effect is 4 percentage points, significant at the .1 level. This is 59% of the mean. Comparing it with the main result from Table 2 of 5.7 percentage points (71% of the mean) suggests that while matching may increase the effect, the majority of the Open Phase 1 effect cannot be explained by subsequent matching. In column (2), we consider the complement. The dependent variable is redefined to be zero for firms that got VC but had no private match. As we would expect, the effect is larger relative to the mean, at 1.5 percentage points relative to a mean of 2.7%.

Even if it does not lead to differential effects of winning, potential matching could affect selection into Phase 1 and perhaps VC decision-making. However, this is not possible for the topics that did not offer VC matching. We split the samples into topics that offer VC matching (column (3)) and topics that do not offer VC matching (column (4)). There is not a statistically significant effect in column (3). The effect in column (4), topics with no matching offered, sees a large but not statistically significant effect of 7.4 percentage points (double the mean). Finally, we interact winning with an indicator for the topic having no match, and exclude topic fixed effects, in column (5). The coefficient on the interaction is small and insignificant, reflecting the fact that the effect in topics without matching is very similar to the effect in topics with matching.

E Analysis of Phase 2

In this Appendix, we consider the effect of Phase 2 awards which, as noted above, are more generously-funded, larger-scale follow-ups to Phase 1 awards (a Phase 1 award is a necessary condition for a Phase 2). We must be cautious, however, in interpreting the results because the main models using data from 2017-19 Phase 1 awards have only a very short time frame for evaluating Phase 2. Furthermore, the sample is quite small, making it impossible to perform an analysis on Open, so we limit ourselves to Conventional Phase 2 competitions. An interesting aspect of Phase 2 is that it enables considering the amount of award, as unlike Phase 1, there is substantial variation in the Phase 2 award amounts (Figure A.5).

[46] It is also possible to have an outside government match (as the table shows, 13% of Phase 2 applicants had matching government funds). We find no relationship between the government match and VC.

Table A.8 show RDD estimates of the effect of winning a Phase 2 award on all four outcomes of interest. We find no effects of the Conventional topic Phase 2 on any outcome, even over the long term (the even columns of each panel), which is consistent with Howell (2017), where Phase 2 grants also have no effects, in part because firms with successful innovation tend to go to the private sector for funding rather than come back to the government for research grants.

If Phase 2 is important for VC but only through a dynamic channel – via its implications for the Phase 1 treatment effect – this would help explain both why the small Phase 1 award is so impactful for Open and why there is no observable Phase 2 effect on VC. The expectation of a Phase 2 award, which averages about \$830,000 (see Table A.3 Panel A), may help to explain the large Phase 1 treatment effect on VC. VCs may believe that if they invest, the chances of a Phase 2 award are very high. In practice, about half of applicants to Phase 2 win, but this rises to all among Phase 2 applicants that raised VC in the 12 months after the Phase 1 award (12 months is roughly the period between the two phases). VC after Phase 1 may affect the Phase 2 decision firstly because VC is one measure the evaluators use to gauge commercialization, and secondly because the VC can provide support in the Phase 2 process. For these reasons, VCs may be responsive to Phase 1 because they expect it to be associated with substantially more non-dilutive cash. Under this hypothesis, there would be little marginal effect of winning Phase 2 because it has been, in a sense, "priced in" to the Phase 1 effect.

F Cost-Benefit Analysis

In this Appendix, we try to shed some light on welfare effects from the DoD's perspective by conducting a back-of-the-envelope cost-benefit analysis informed by our estimates.

F.1 Set-Up

Consider the marginal decision over whether or not to run an SBIR competition. This can be run this as (i) Open topic, (ii) Conventional topic or (iii) not run at all. For simplicity, we think of the competition as having just one winner. We assume that the DoD has the following objective function from running this hypothetical competition:

$$V = \mu^M M + \mu^N N - C. \tag{3}$$

Here, M represents the benefits from military activities, N represents the benefits from non-military activities, and μ^M and μ^N represent the utility weights on military and non-military activities, respectively. C represents the costs of the competition. Note that DoD claims to place value on non-military industrial development (see Sections 2.3.2 and 3.1), so $\mu^N > 0$. An alternative interpretation of Equation 3 is that V is society's welfare in which case the Social Planner would have a relatively higher weight,μ^N , on non-military benefits..

We model the military benefits $M(D, u^M)$ as being composed of an observable, D, the dollar value of future non-SBIR DoD contracts and a set of unobservables u^M. As argued in the main text, a major motivation for DoD is whether the SBIR contract transfers into a "program of record", i.e. a future non-SBIR contract with the military, typically to further develop and implement the technology funded at the SBIR stage. Similarly, we model the non-military benefits $N(VC, u^N)$, as a function of the dollar value of future VC investments, which is linked more to economy-wide benefits, and an unobservable term, u^N.

F.2 Results

Panel A of Table A.9 reports our baseline calculations. The first two columns have Open outcomes for VC investments (column (1)) and non-SBIR contracts (column (2)). The next two columns have Conventional outcomes for VC (column (3)) and non-SBIR contracts (column

(4)). We take our preferred estimates of the treatment effects in Row 1 (columns (1) and (2) of Table 2). Our assumption is that winning an award changes the winner's probability of obtaining VC or non-SBIR contract by the point estimate of the RDD effect. That is, the treatment effects we estimate in our main analysis give the percentage point increase in the probability of, for example, any VC.

We multiply this treatment effect by the average amount of subsequent VC investment (or non-SBIR contract value) among all applicants in the post-award period conditional on having any VC (or contracts) to arrive at a dollar amount of the benefit. We report the average amounts raised or contracted across all applicants in the 24 months post-award for the 2017-19 main sample in Row 2. These are similar in magnitude (\$11 million to \$12 million).[47] We keep these the same across Open and Conventional awards as we want differences to be driven by the estimated causal impact of the program, rather than mean contract values. The qualitative conclusions are robust to using the program-specific means as shown in robustness tests in Panel B. The implied benefit in Row 3 is the product of the previous two rows. Our baseline assumption on the welfare weights is $\mu^M = \mu^N = 1$ (Row 4). The sum of the benefits, representing $\mu^M M + \mu^N N$, is \$1.48 million for Open and \$0.33 million for Conventional (Row 5).

Next, we consider costs (C), which are composed of the award amount and administration costs. Like the rest of this analysis, the costs are based on the 2017-19 sample and are estimated by program administrators in the Air Force. A Phase 1 Open award is \$50,000, while a Phase 1 Conventional award is \$150,000. It is also important to consider Phase 2. As explained in Appendix E, although we find no RDD effects of winning Phase 2, the expected Phase 2 money may be an important element in the success of Phase 1. Therefore, the Phase 2 award should be considered as an additional cost of Phase 1. The average Open Phase 2 is \$832,000, while the average Conventional Phase 2 is \$814,000 (see Table A.3). Winning Phase 1 is a condition for applying to Phase 2, but of course Phase 1 winners need not apply for Phase 2. Conditional on applying, about about 50% of Phase 2 applicants win in both programs. However, the rates of selecting into applying to Phase 2 differ across the two programs. Among all Open(Conventional) winners, 39%(22%) win Phase 2. We use these percentages in our cost analysis.[48]

When it comes to administrative costs, the program administrators calculate that the cost of evaluating and contracting a Phase 1 award in DoD employee-hours is about \$1,000. The cost for a non-winner (which has no contracting dimension) is about \$500. We assume there are an average of four losers per winner, though the results are very similar under alternative assumptions. Finally, we calculate the cost for administering a Phase 2 as triple those of Phase 1, which are \$3,000 per winner and \$1,500 per loser. Using these inputs described above, the total cost in thousands of dollars for Open is $C = 50 + 3 + 0.39\,(4.5 + 832) = 379$, while the total cost for Open is $C = 153 + 3 + 0.22\,(4.5 + 814) = 333$.

Finally, in Row 7 of Table A.9, we calculate the net benefit V to be \$1,104,000 for Open and \$1,000 for Conventional. Therefore, Open has a \$1.1 million higher benefit than Conventional, reflecting its much larger causal impact with only slightly higher costs (Row 8).

F.3 Robustness and Caveats

Panel B of Table A.9 contains some robustness tests around these baseline calculations. Column (1) presents the difference between the net benefit from Open (Column (2)) and

[47]They are different from the statistics in Table 1 Panel B because those are the pre-award amounts, which are smaller.

[48]If we were to focus on society's utility, since the awards are a transfer, the social cost is likely to be less than the expenditure. It will be the deadweight cost of public funds required to raise the tax to finance the government expenditure. In this case, we would also want to consider the cost of applying for unsuccessful firms.

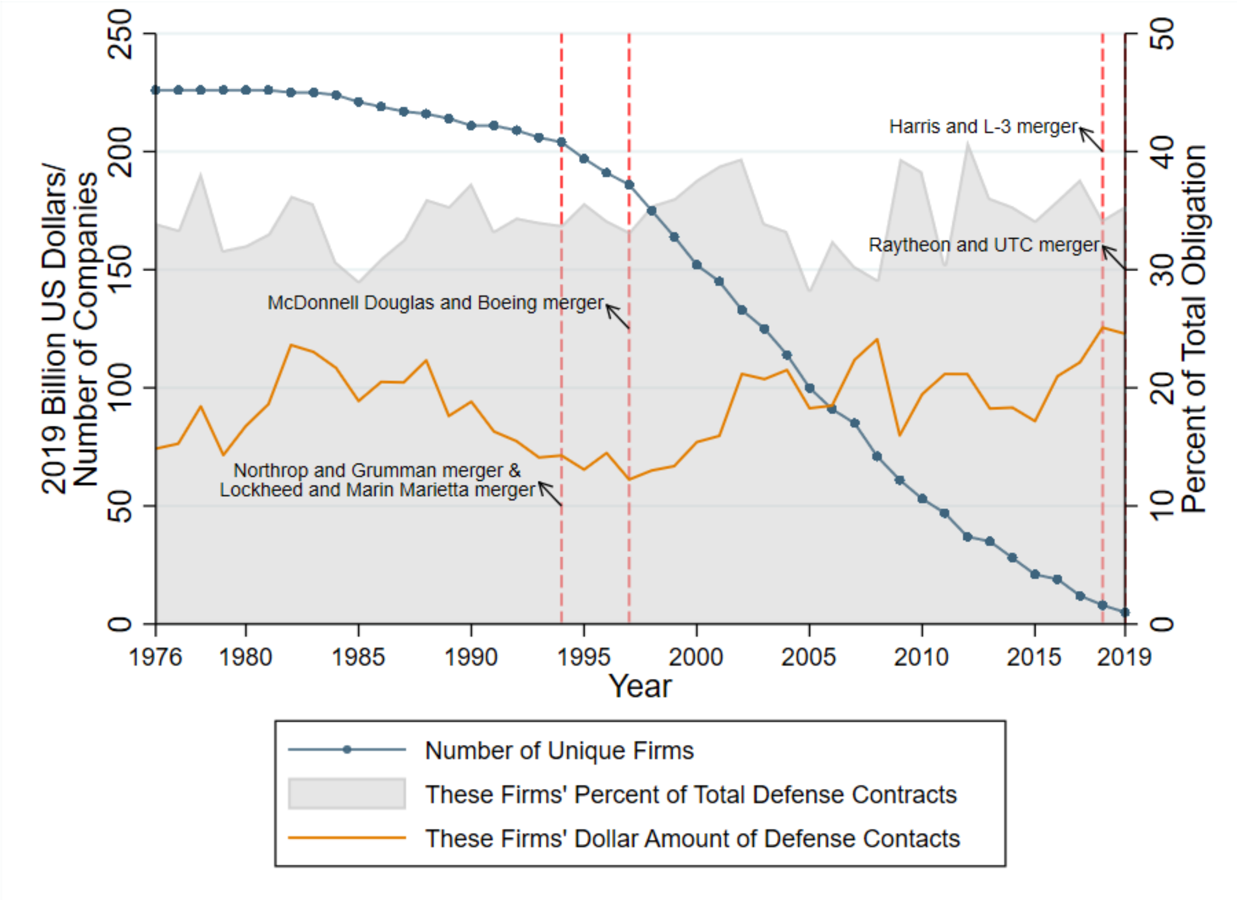
Conventional (column (3)) competitions. Row 1 contains the baseline estimates from Panel A. Row 2 uses the mean in the two years prior to (rather than after) the award, as the amount of the award could also be influenced by the shift to Open. The net benefit difference falls to \$0.82 million.

Recall that we are interpreting Equation 3 as DoD's objective function, not the Social Planner, so an extreme assumption is that no weight is placed on non-military objectives, i.e. $\mu^N = 0$. Row 3 implements this idea, which effectively switches off the VC effects. Even here, however, Open has a \$0.45 million advantage over Conventional. In Row 4 we assume that the treatment effect on VC for Conventional is zero instead of the estimated -0.005 since it is statistically insignificant. This slightly reduces the Open advantage to \$1.05 million. Row 5 uses the 2003-2019 treatment effects for Conventional instead of the results from 2017-2019 that have been used so far. The advantage of Open is larger with the net benefit difference rising to \$1.27 million. Finally, in Row 6 we calculate mean VC and non-SBIR DoD contracts separately for Open and Conventional, rather than together as in Panel A Row 2. As Conventional contracts are on average larger even though they are less frequent, the Open advantage falls to \$0.42 million.

The calculations in Table A.9 are, of course extremely coarse and only intended to give some crude insight into what the numbers may mean in terms of policy. As noted in main text of the paper, there may be many other unobserved benefits that we do not measure (captured by u^M and u^N). For example, obtaining the tightly specified Conventional results might have benefits for non-award winners. However, this may also be true for Open awards, so these unmeasured benefits and costs can go either way. Also, we have assumed that VC is a positive for society, but it is possible that Open winners crowd out other useful VC projects. That said, they might just as easily crowd in other projects (as was found more generally for defense R&D in Moretti et al. (2020)). We have also focused only on the shorter-run benefits from Open and abstracted away from the costs of putting together an application for firms.

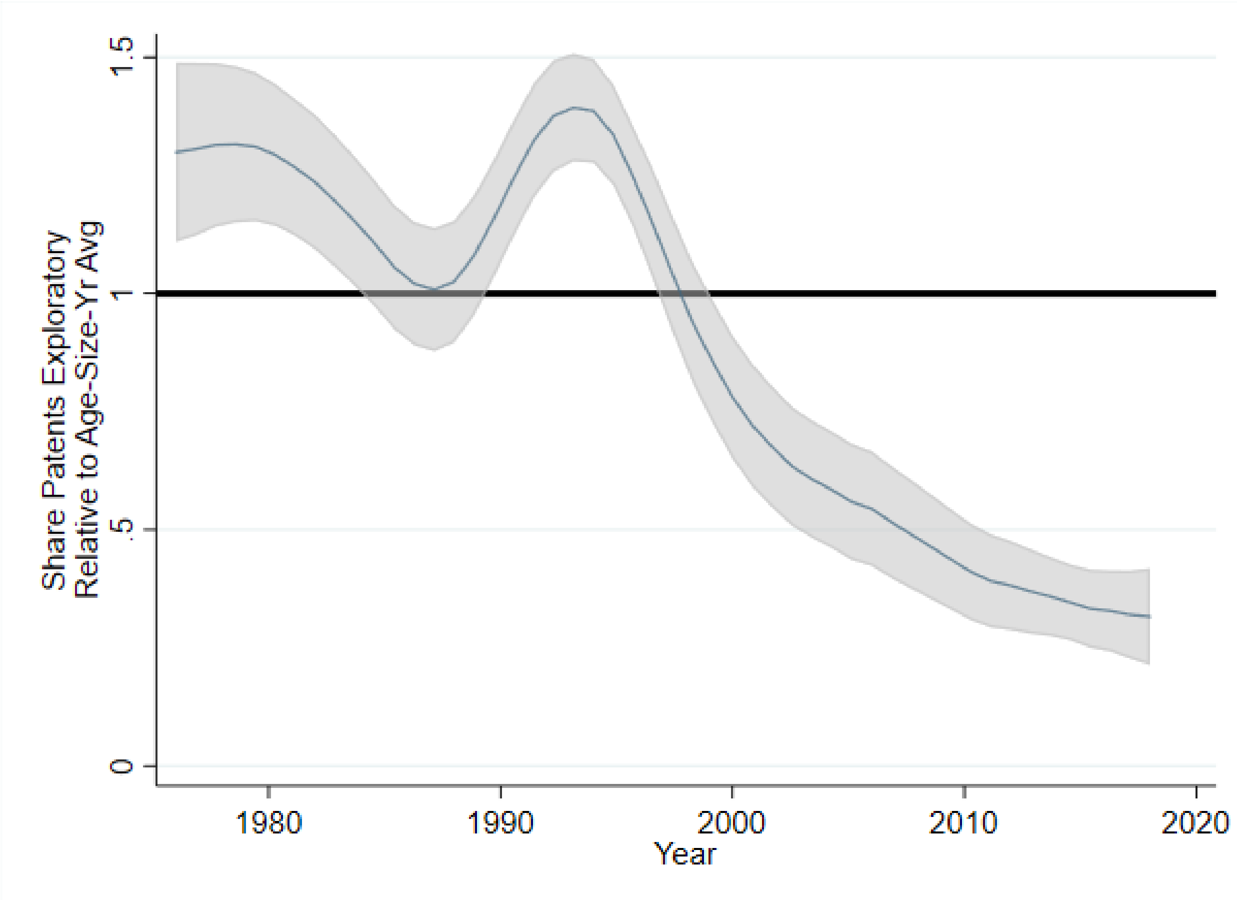
In summary, although the exact net benefits depend on assumptions, it is clear that throughout this analysis the Planner's marginal decision rule would be to run an Open competition rather than a Conventional one.

Figure A.1: Consolidation of Prime Defense Contractors



Note: This figure shows the trend of defense contractors' consolidation since the 1980s. We first define prime defense contractors as the top contractors between 2000 and 2020: Boeing, Raytheon, Lockheed Martin, Northrop Grumman, General Dynamics, United Technologies Corp, Harris, and L-3. We then identify all their acquisitions of other defense contractors starting in 1976. The blue line shows the number of unique firms in each year, from 226 in 1976 to just six in 2020. The gray area shows the share of all DoD contracts (in nominal dollars) that are awarded to the top eight prime defense contractors and their acquisition targets. The total value of these contracts (in 2019 dollars) is shown in the orange line. For example, the 226 firms accounted for about $70 billion or 33% of the total defense contract value, in 1976. They consolidated to six companies by 2019, at which point those six accounted for $115 billion, or 35% of the total defense contract value. Data are sourced from the Federal Procurement Data System (FPDS) and Defense Contract Action Data System (DCADS).

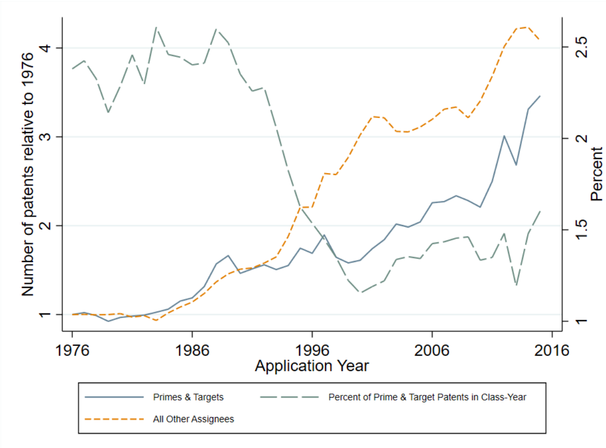
Figure A.2: Exploratory Patents from Prime Defense Contractors



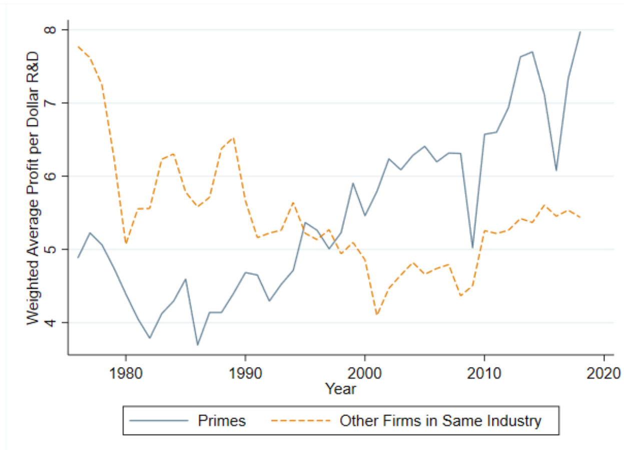
Note: This figure describes the trend of exploratory innovation by the prime defense contractors and their acquisition targets over time. The firms are the same set from Figure A.1. That is, 226 firms are included in 1976, while only six are included in 2019 (as the 226 have merged into these six). The graph shows these firms' average share of exploratory patents relative to other firms with similar in age, size, and year. An exploratory patent is a patent filed in a technology class previously unknown to the firm in a given year. Age is defined as the year from the firm's first observed patent and size is defined as the firm's patent stock in a given year. The measures in both figures are smoothed using kernel-weighted polynomial regressions. The gray band around the relative citations represents the 95% CI. Data are sourced from the USPTO.

Figure A.3: Historical Dynamics of Prime Defense Contractors

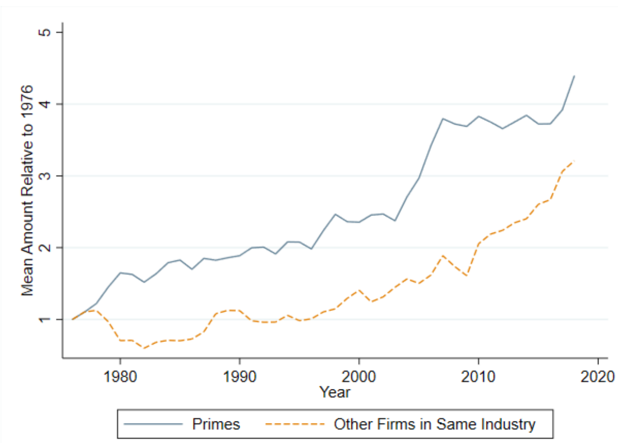
(a) Number of Patents



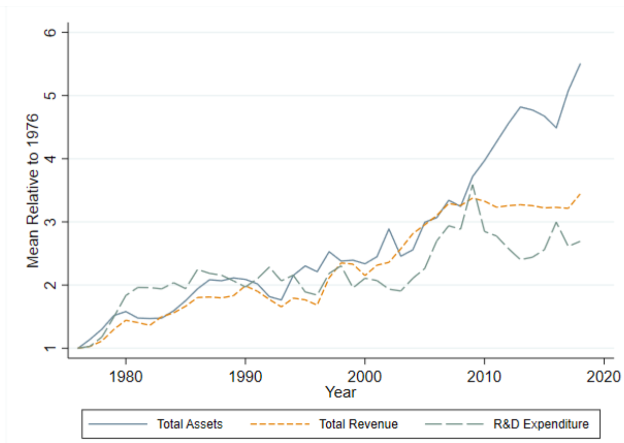
(b) Profit per Dollar R&D



(c) Profits



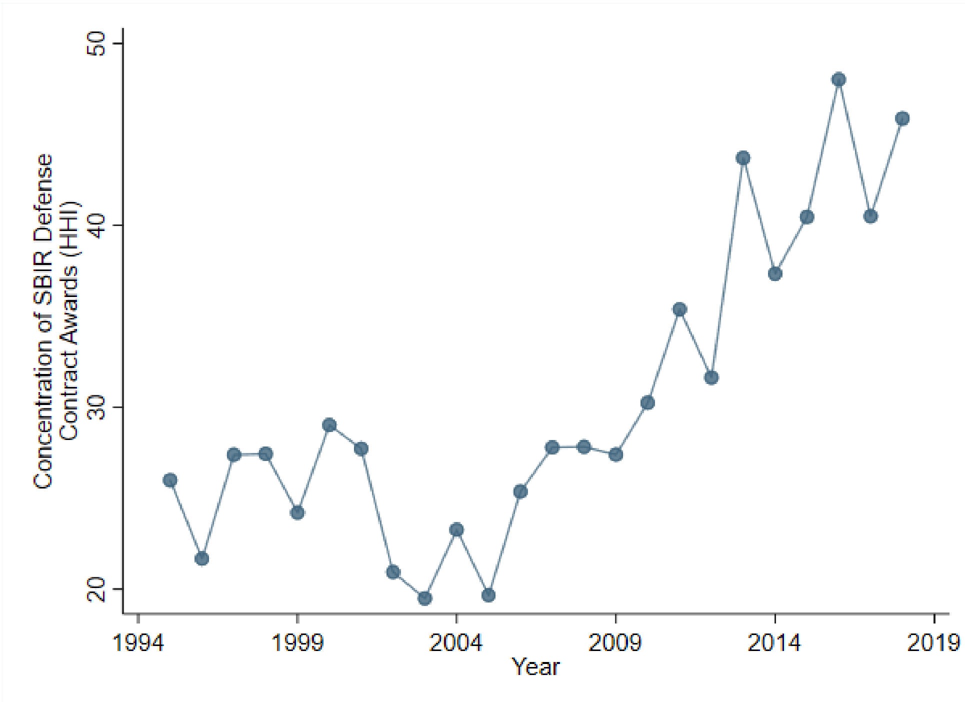
(d) Total Assets, Revenue, and R&D Expenditure



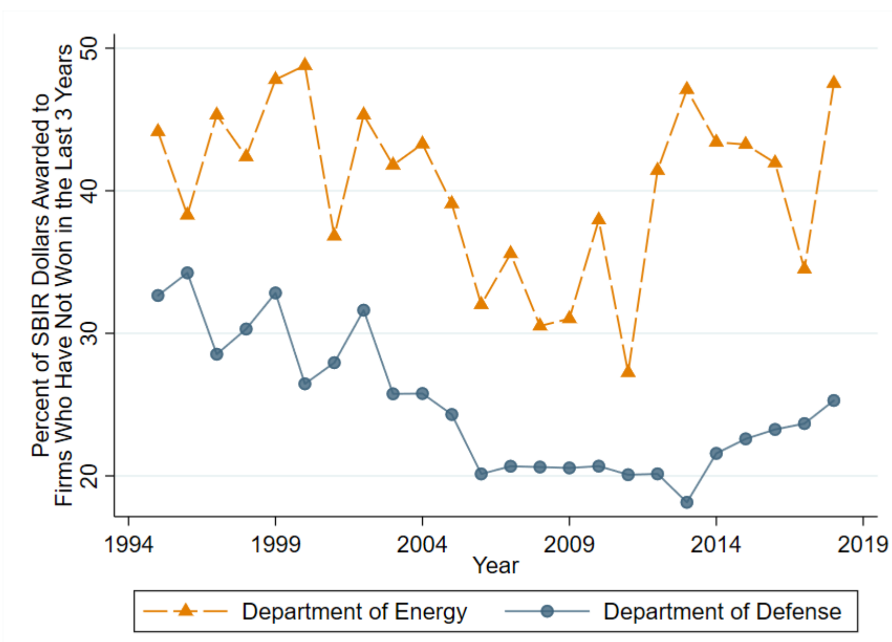
Note: This figure shows the dynamics of prime defense contractors. Panel A shows growth in the number of granted patents for prime defense contractors and their acquisition targets (blue line) and the number of granted patents for all other assignees (orange line) from 1976 - 2016, using data from the U.S.PTO. The teal line shows the share of prime defense contractors and their acquisition targets' patents in their class-year. The number of patents is scaled by 1976 levels (1976=1). We exclude 2016–on because there is a 2-3 year time period between application and patent award, so there are far fewer granted patents in the most recent application years. Panel B shows the weighted average profit per dollar of R&D for prime defense contractors compared to other Compustat firms in the same 3-digit SIC code (334 and 336). Panel C shows the growth of profits for prime defense contractors compared to other Compustat firms in the same 3-digit SIC code (334 and 336) relative to 1976 (1976=1) from 1976 to 2019. Panel D shows the growth of total assets, total revenue, and R&D expenditures in constant 2019 U.S. dollars for prime defense contractors, scaled by the 1976 level. Panel A includes the prime defense contractors and their acquisition targets; Panels B, C, and D only include the prime defense contractors and not their acquisition targets.

Figure A.4: Concentration of Federal Contracts

(a) Concentration of Department of Defense SBIR and Non-SBIR Contracts



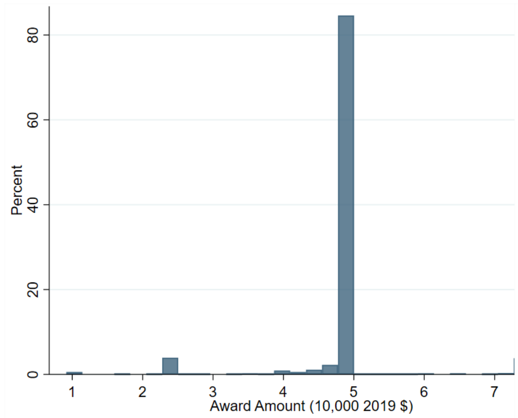
(b) Share of Firms without Recent Repeat Contracts in Two SBIR Programs



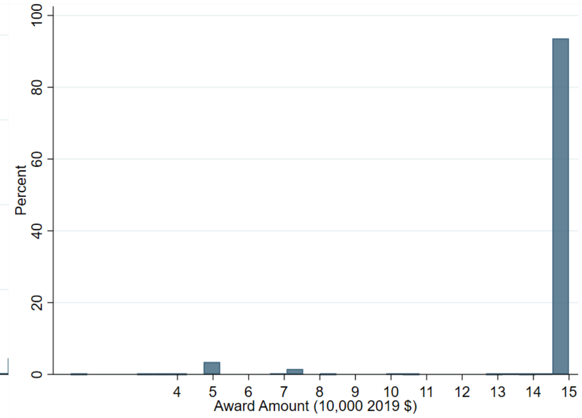
Note: Panel A in this figure shows the Herfindahl–Hirschman Index (0-10,000) for Non-SBIR Department of Defense contracts from 1990 to 2018. Panel B shows the share of "new" firms winning awards from the SBIR programs at the Department of Defense (DoD) and the Department of Energy (DoE). Each line plots the percentage of SBIR contract dollars awarded to firms that have not won a contract in the last three years. At the beginning of the sample in the early 1990s, the share of SBIR awards to firms that have not won in the last three years are relatively similar at the two agencies, but the series subsequently diverge. Data from DCADS, FPDS, and the U.S. Small Business Administration.

Figure A.5: Histograms of Award Amounts by Topic Type and Phase

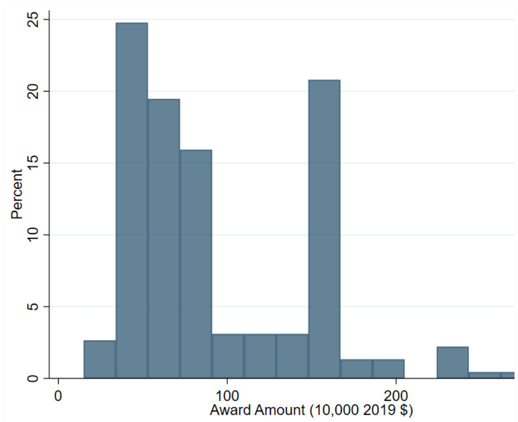
(a) Open Phase 1



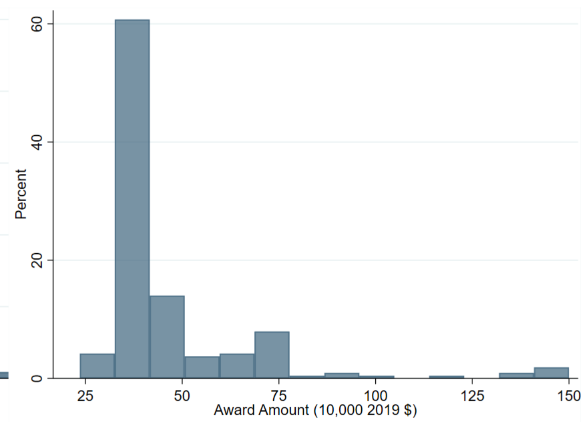
(b) Conv Phase 1 2017-19



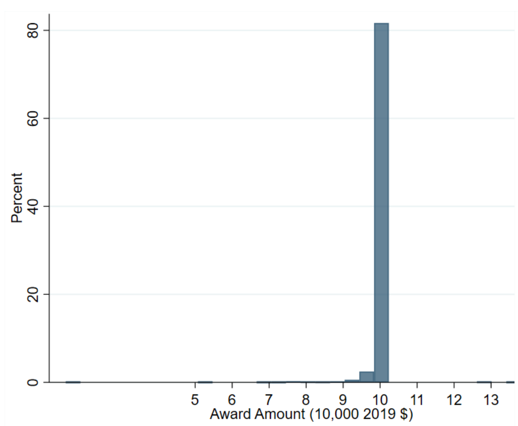
(c) Open Phase 2



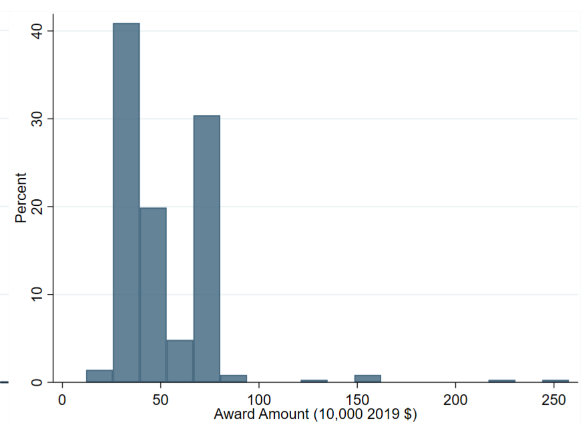
(d) Conv Phase 2 2017-19



(e) Conv Phase 1 $<$ 2017



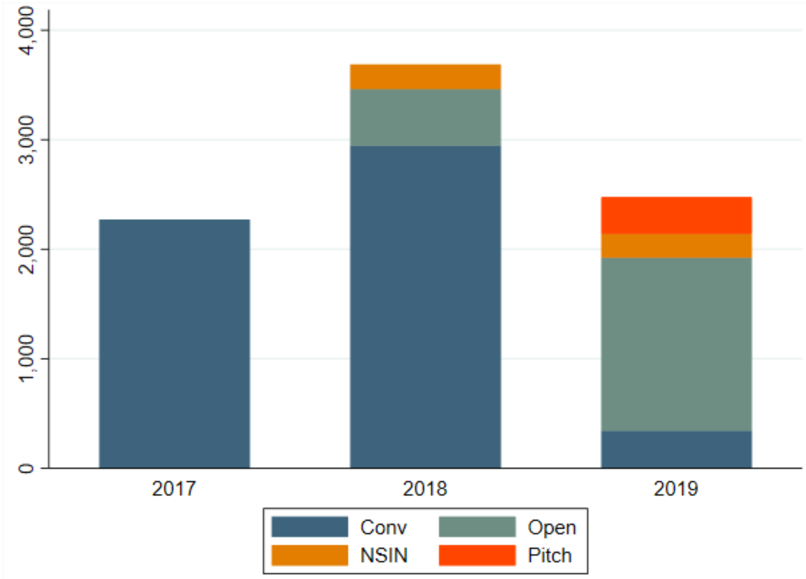
(f) Conv Phase 2 $<$ 2017



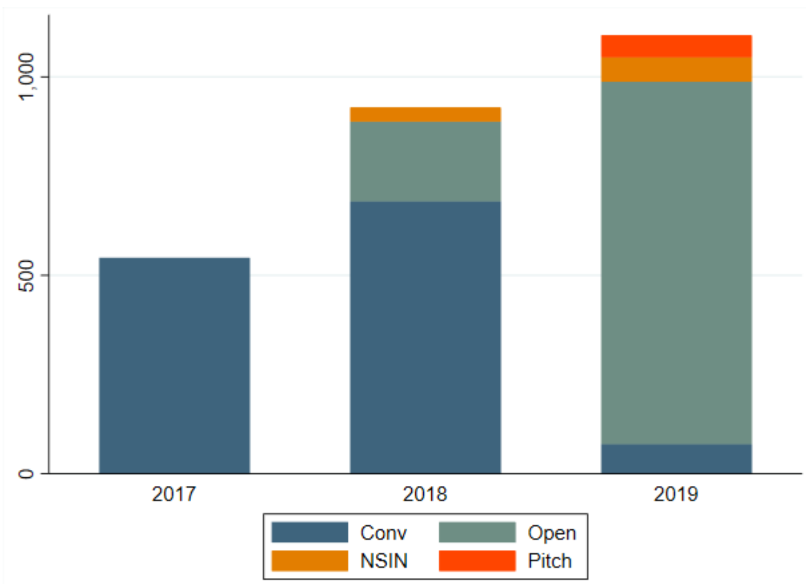
Note: These histograms show the share of awards by amount, in real 2019 dollars. For the bottom right graph (Phase 2 <2017), we omit one outlier $12 mill contract.

Figure A.6: Number of Applications and Awards Over Time by Topic Type (Analysis Sample)

(a) Applications

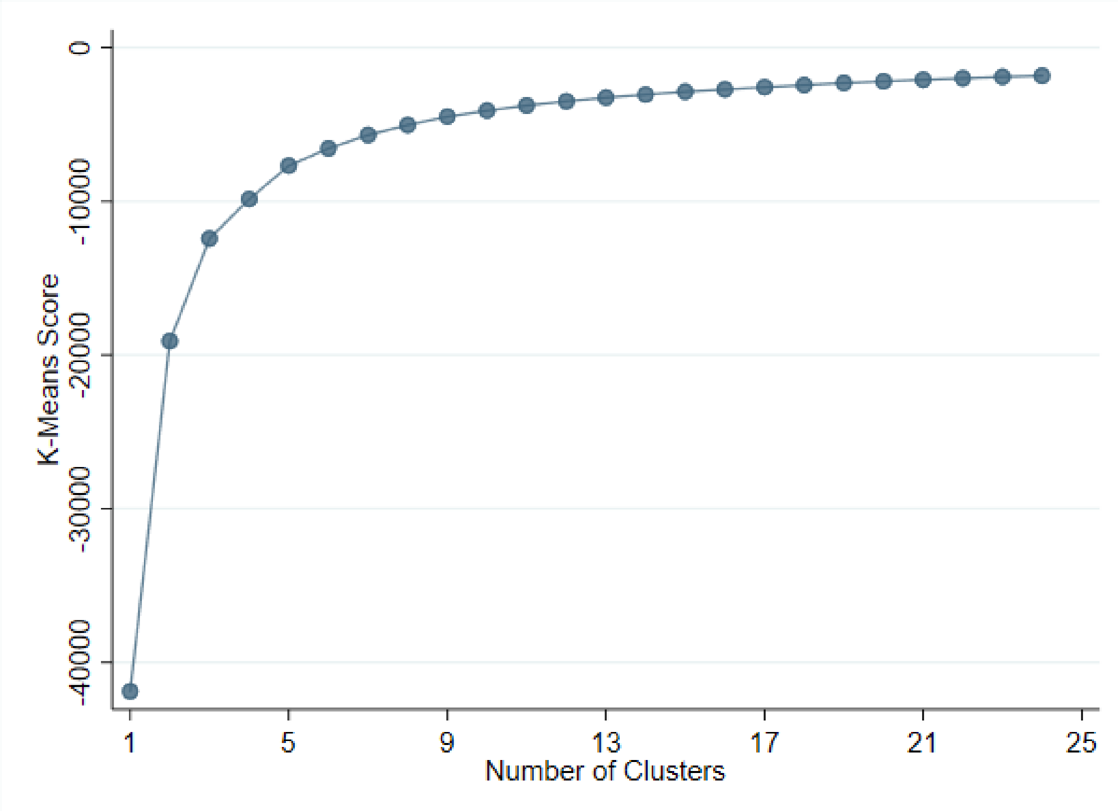


(b) Awards



Note: These figures show the number of applications (top) and awards (bottom) in our "analysis sample" of data from 2017-2019 by topic type.

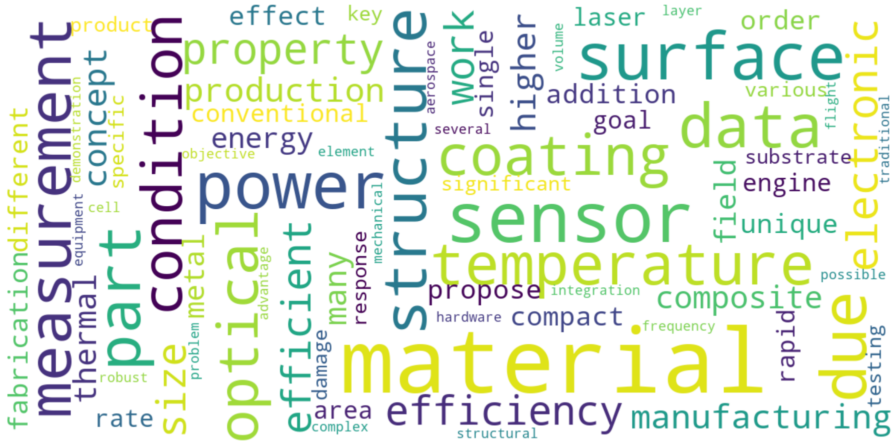
Figure A.7: K-means Optimal Number of Clusters (Elbow Method)



Note: This figure shows the k-means score for 1-25 clusters. The k-means score is defined as the negative total squared error between the cluster members and their centroid. The optimal number of clusters, or the "elbow," is identified for the number of clusters k for which the objective improves significantly and plateaus thereafter. We identify the optimal cluster to be $k = 5$.

Figure A.8: Keywords for K-means 2 Cluster Model Scaled by Importance to Topic

(a) Hardware



(b) Training/Software

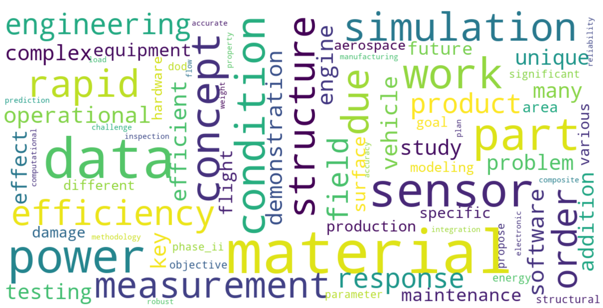
Note: These figures show the keywords that are identified as a topic cluster by the k-means cluster algorithm, where the algorithm has been assigned to find two clusters. The word's size reflects its prevalence in the topic.

Figure A.9: Keywords for K-means 5 Cluster Model Scaled by Importance to Topic

(a) Software/Security

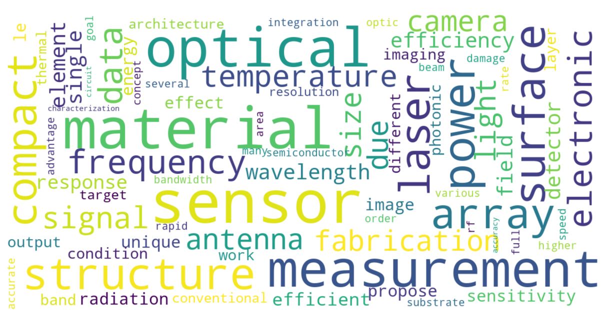
(b) Materials/Manufacturing

(c) Simulation/Engineering



(d) Network/Data

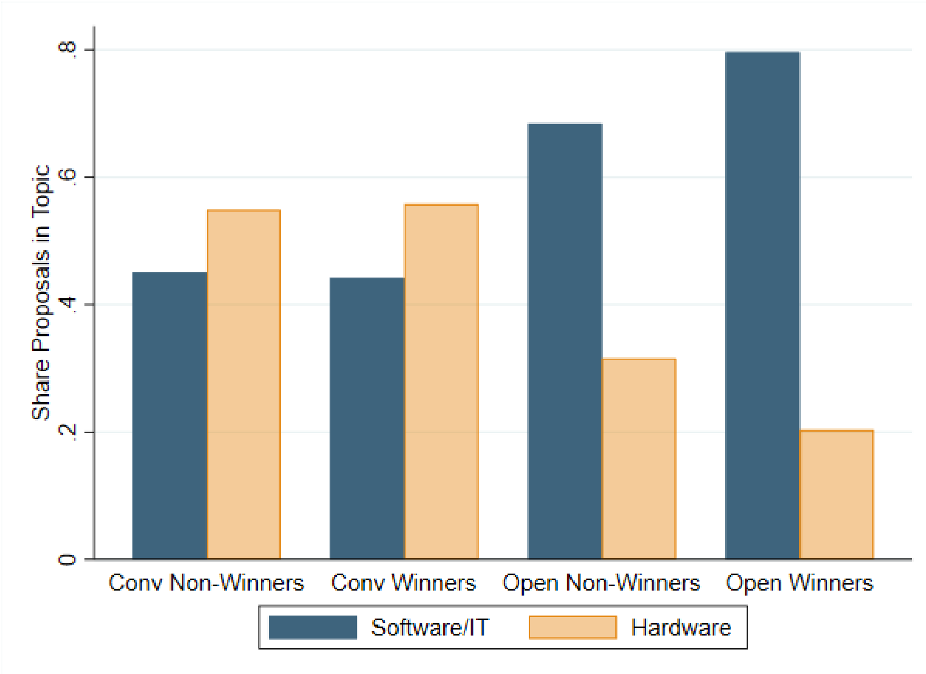
(e) Optical/Measurement



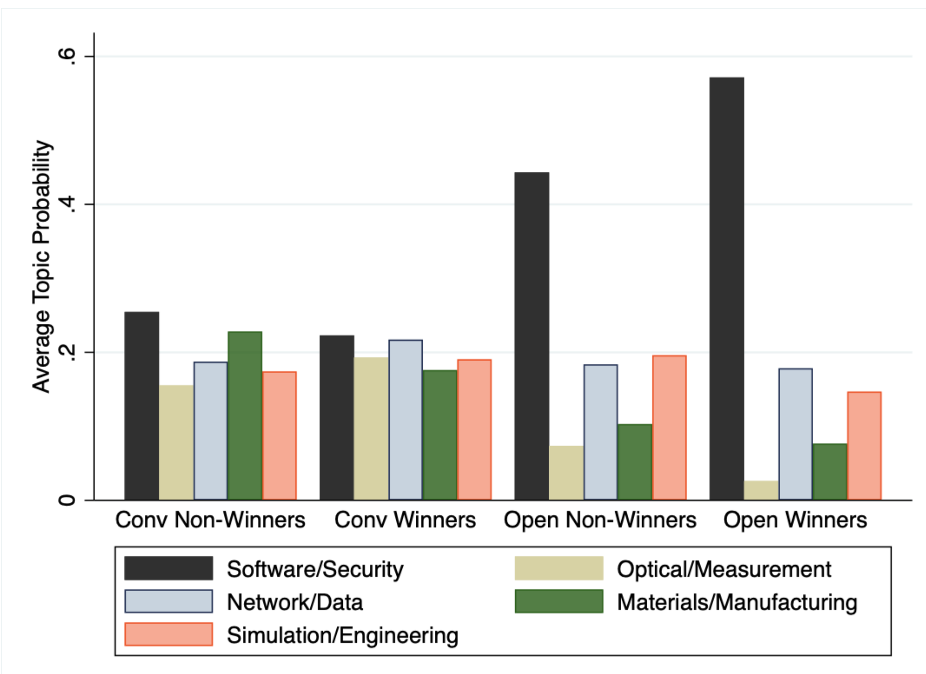
Note: These figures show the keywords that are identified as a topic cluster by the k-means cluster algorithm, where the algorithm has been assigned to find five clusters. Five is the optimal number of clusters according to the elbow method (see Figure A.7). The word's size reflects its prevalence in the topic.

Figure A.10: Applicant Technologies (Based on K-means Clustering of Abstract Text)

(a) Model with Two Clusters

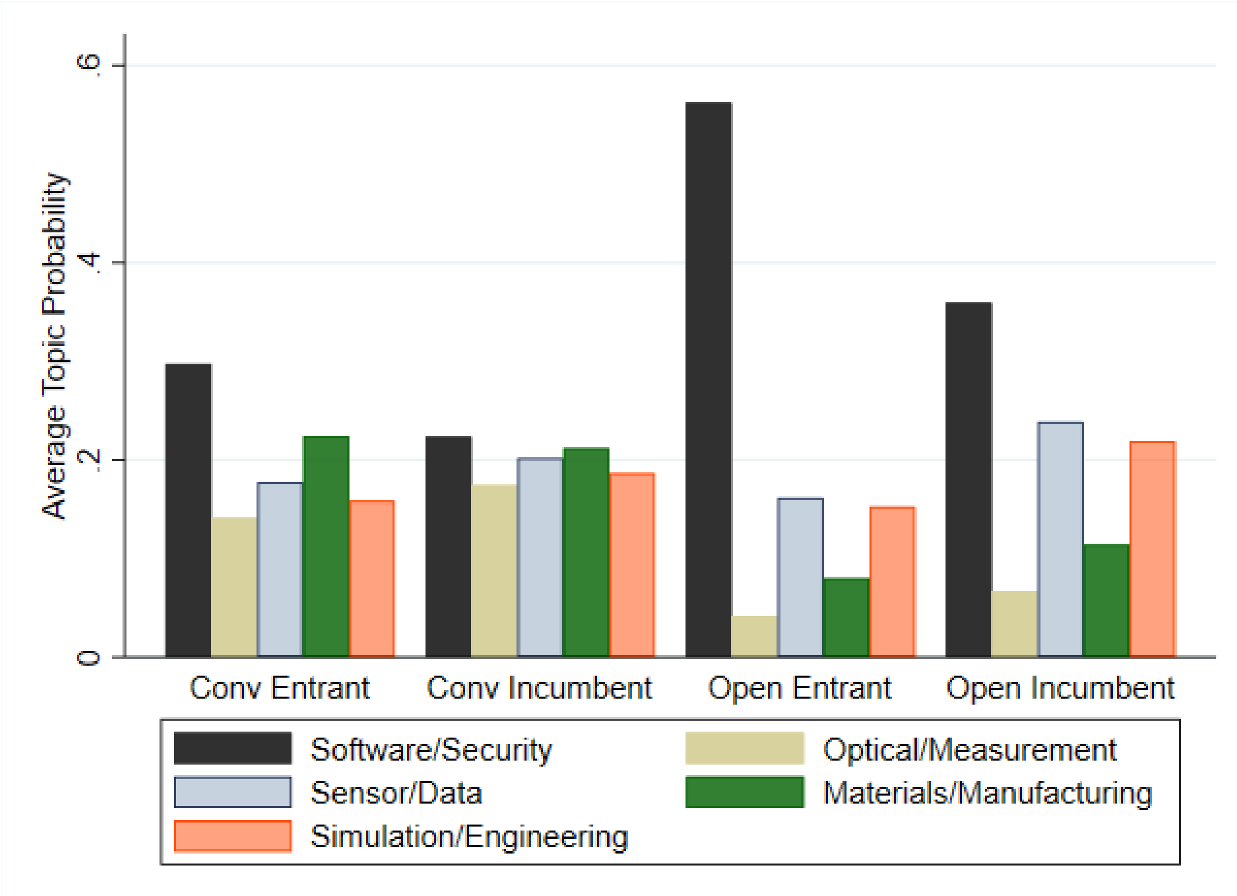


(b) Model with Five Clusters



Note: These figures show the distribution of applications by topic based on a k-means cluster model, described in Section 3.2. Panel (a) shows the model pre-set to produce two clusters. Panel (b) shows the model with the optimal number of clusters, which is the one for five clusters. In each case, all applications are assigned to exactly one cluster. In each panel, we divide the sample into four groups by program type (Conventional and Open) and winner status. We then show the percent of applications within each group by topic type.

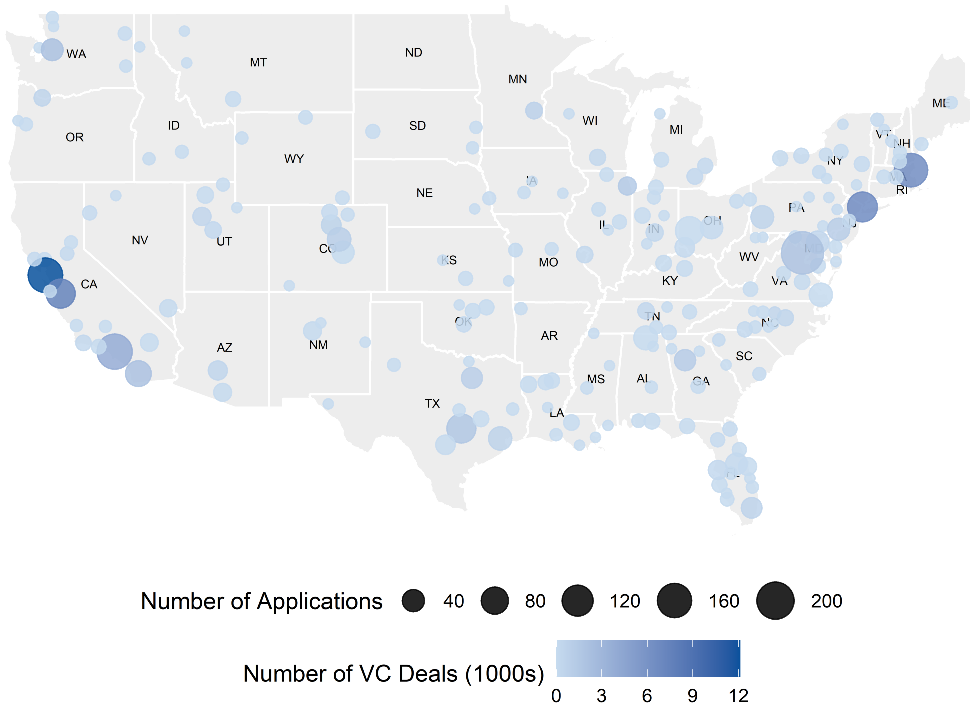
Figure A.11: Applicant Technologies By Entrant Status (Based on K-means Clustering of Abstract Text)



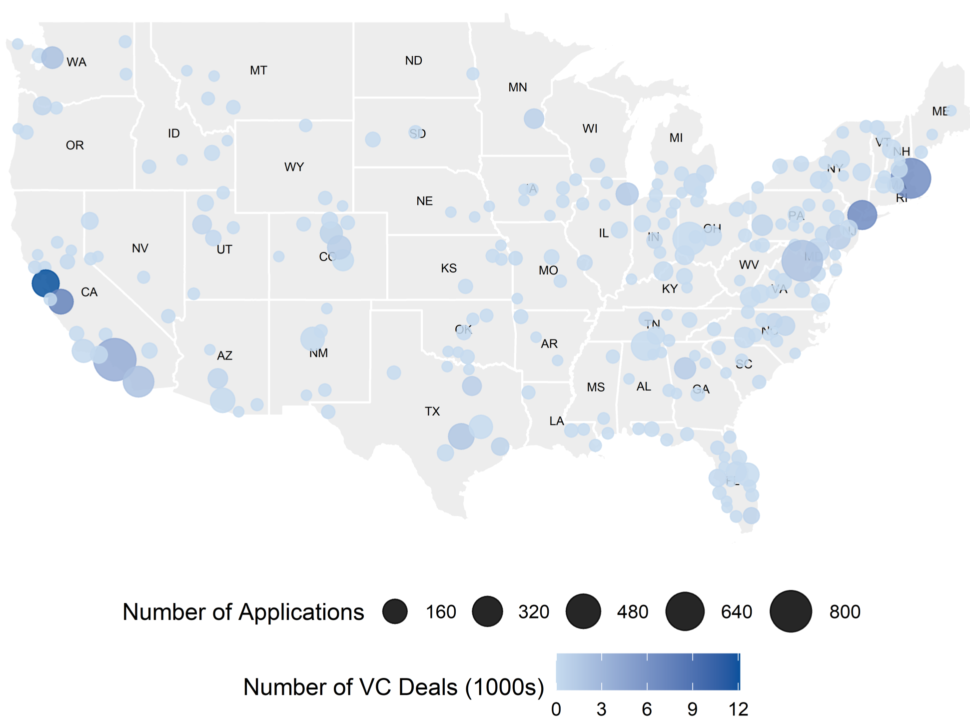
Note: This figure shows the distribution of applications by topic based on a k-means cluster model. All applications are assigned to one of five topics. In each case, we divide the sample into four groups by program type (Conventional and Open) and Entrant status, where an Entrant is defined as a firm with no previous DoD SBIR awards, and an Incumbent is defined as a firm with at least one previous DoD SBIR award. We then show the percent of applications within each group by topic type.

Figure A.12: Geographic Dispersion of Applications (2017-19)

(a) Open Topic Applications and VC Deals



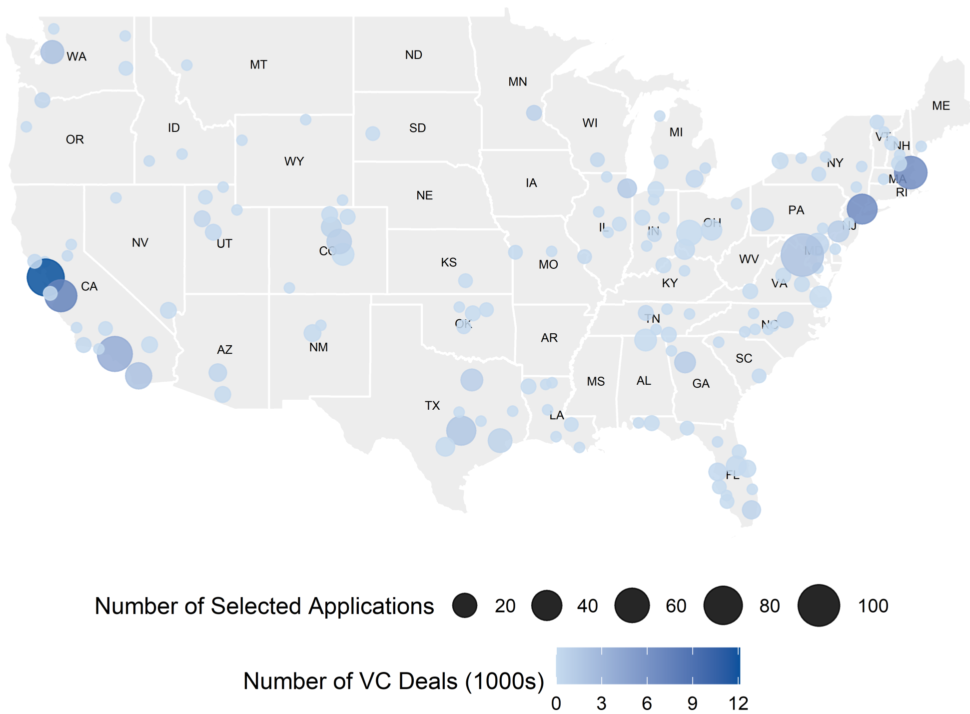
(b) Conventional Topic Applications and VC Deals



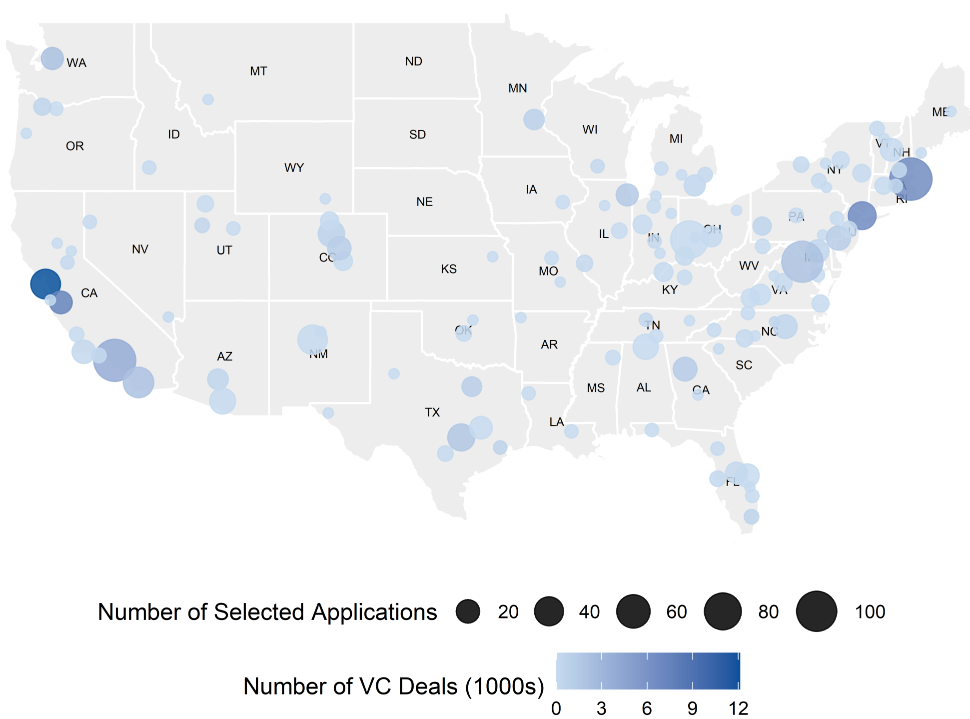
Note: These maps show the number of applications to open (Panel A) and conventional SBIR topics (Panel B) by MSA from 2017 to 2019. The size of the bubble represents the relative number of applications. The color gradient in both maps also show VC activity by MSA.

Figure A.13: Geographic Dispersion of Awards (2017-19)

(a) Open Topic Awards and VC Deals



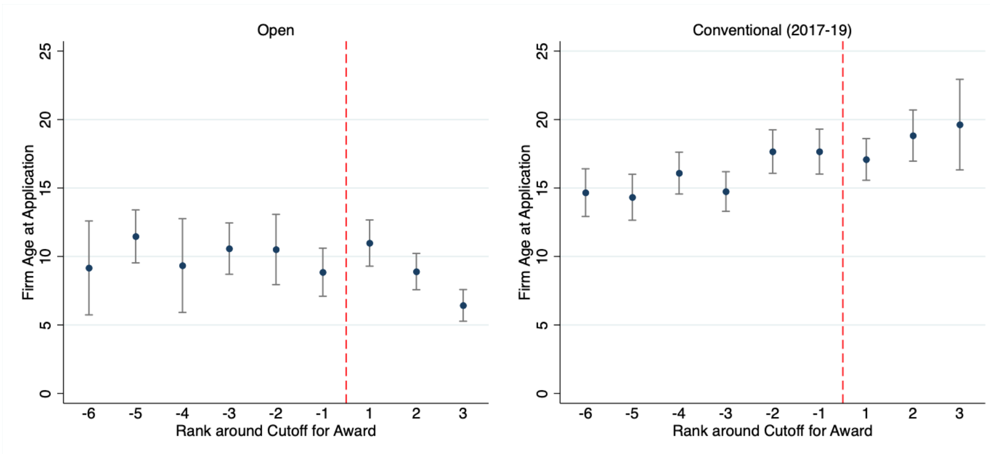
(b) Conventional Topic Awards and VC Deals



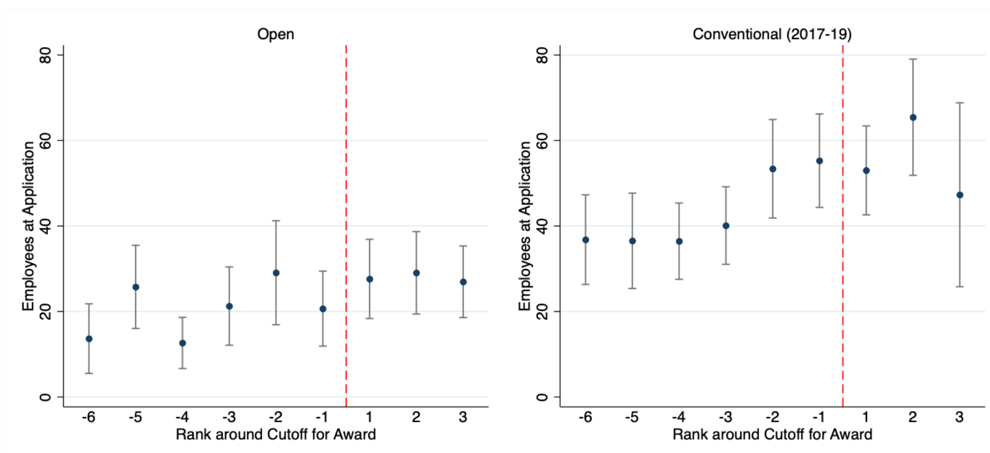
Note: These maps show the number of awards (i.e. contracts) for open (Panel A) and conventional SBIR topics (Panel B) by MSA from 2017 to 2019. The size of the bubble represents the relative number of applications. The color gradient in both maps also show VC activity by MSA.

Figure A.14: Continuity of Baseline Characteristics by Rank around Cutoff (Part 1 of 4)

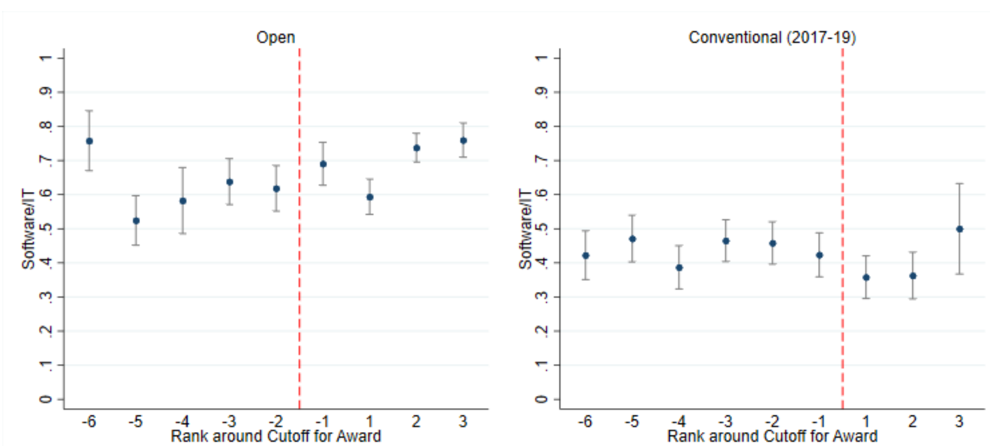
(a) Firm Age at Application



(b) Firm Employment at Application



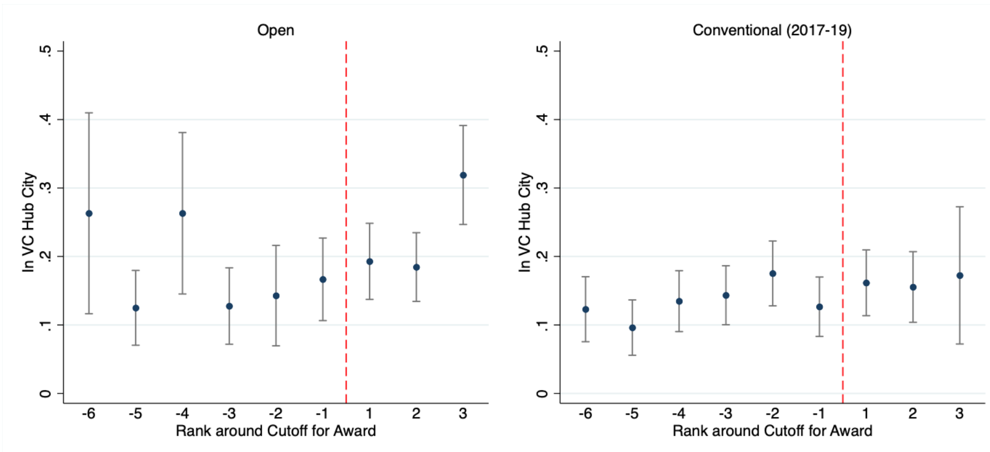
(c) Software vs. Hardware Technology



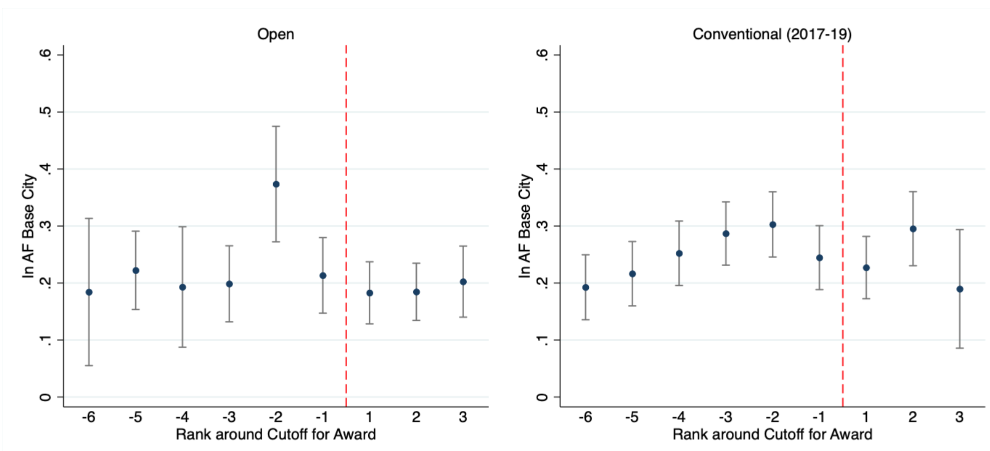
Note: These figures show applicant firm age (top figures), employment (middle figures), and the k-means 2 cluster abstract classification which yields a software and a hardware group (bottom figures) at the time of the application. In all cases, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. The grey capped lines represent 95% confidence intervals.

Figure A.15: Continuity of Baseline Characteristics by Rank around Cutoff (Part 2 of 4)

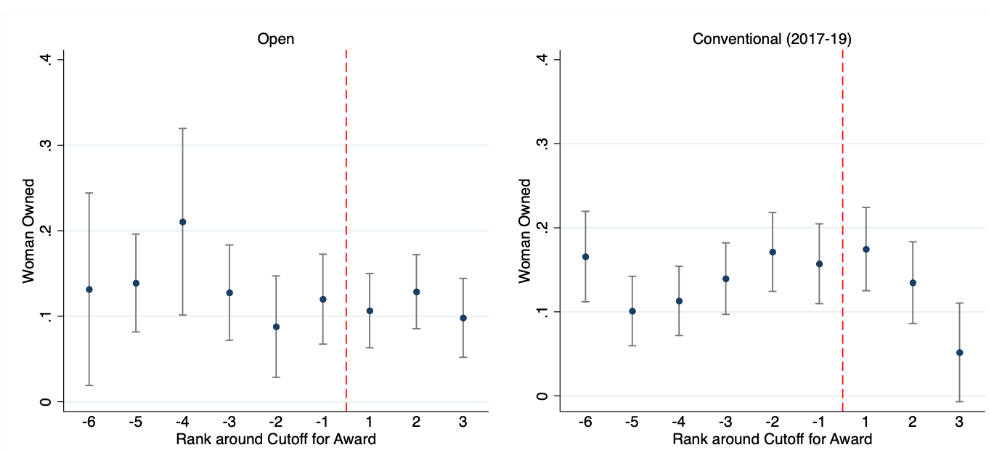
(a) Probability Firm Located in VC Hub City



(b) Probability Firm Located in a County with an Air Force Base



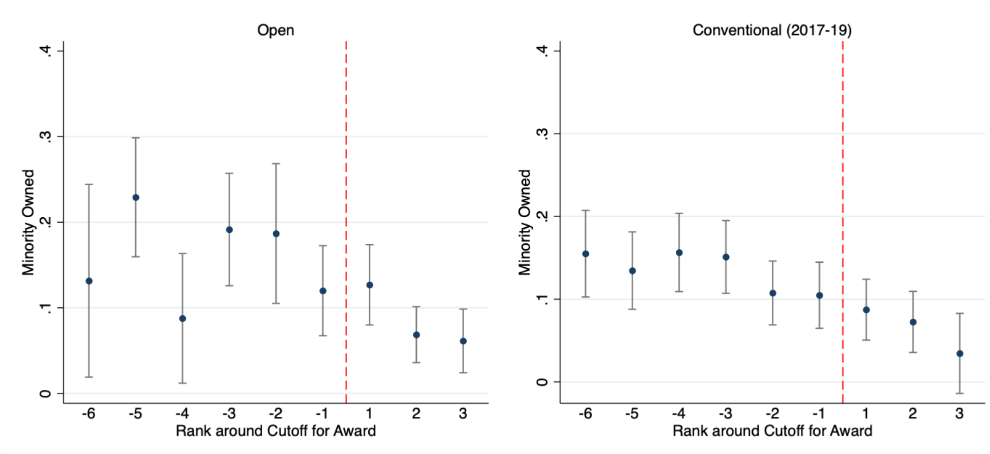
(c) Probability Firm Woman-Owned at Application



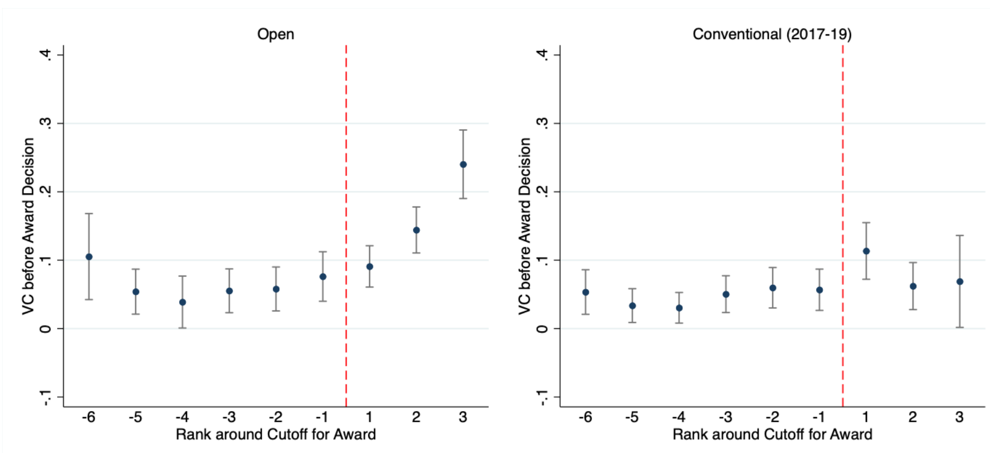
Note: These figures show the probability that an applicant firm is located in either San Francisco/San Jose, Boston, or New York City (top figures), located in a county with a U.S. Air Force base (middle figures), and woman-owned (bottom figures) at the time of the application. In all cases, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. The grey capped lines represent 95% confidence intervals.

Figure A.16: Continuity of Baseline Characteristics by Rank around Cutoff (Part 3 of 4)

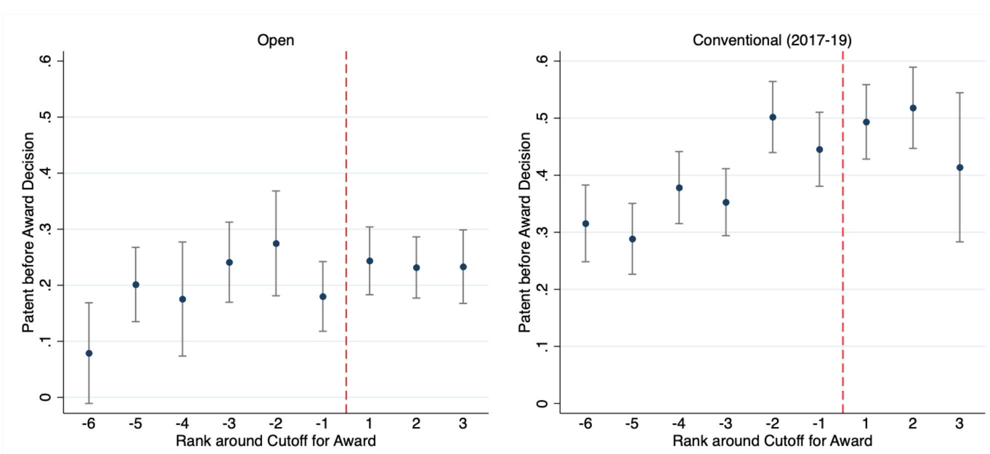
(a) Probability Firm Minority-Owned at Application



(b) Probability of Venture Capital Before Award Decision



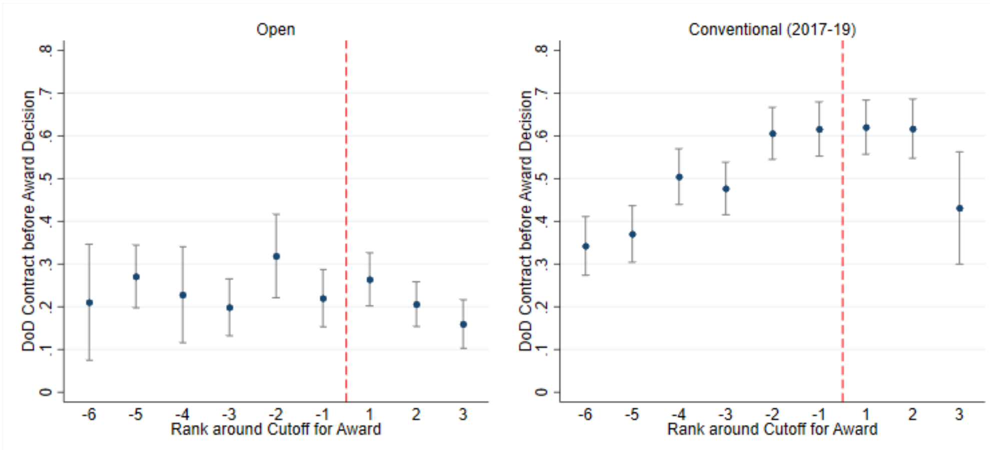
(c) Probability of Patent Before Award Decision



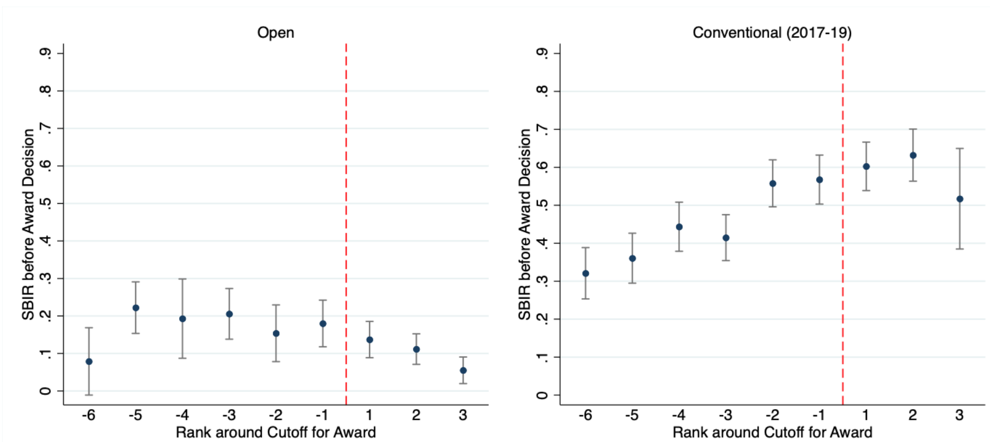
Note: These figures show the probability that an applicant firm is minority-owned (top figures), raised venture capital investment (VC, middle figures), and had any patents at the time of the application (bottom figures) at the time of application. In all cases, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. The grey capped lines represent 95% confidence intervals.

Figure A.17: Continuity of Baseline Characteristics by Rank around Cutoff (Part 4 of 4)

(a) Probability of DoD Non-SBIR Contract Before Award Decision

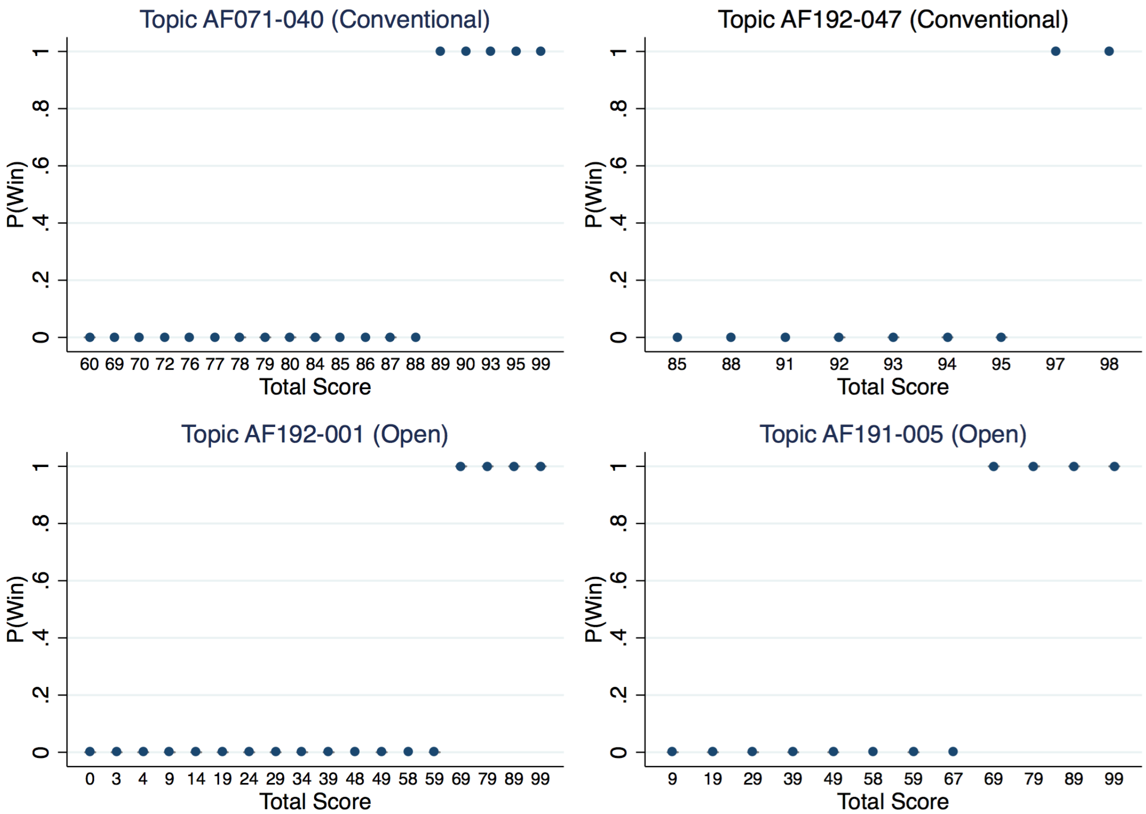


(b) Probability of SBIR Before Award Decision



Note: These figures show the probability that an applicant firm had any SBIR contracts after the award decision (top figures) and had any non-SBIR DoD contracts valued at more than $50,000 at the time of the application (bottom figures). A rank of 1 indicates that the applicant had the lowest score among winners, while a rank of -1 indicates that the applicant had the highest score among losers. The grey capped lines represent 95% confidence intervals.

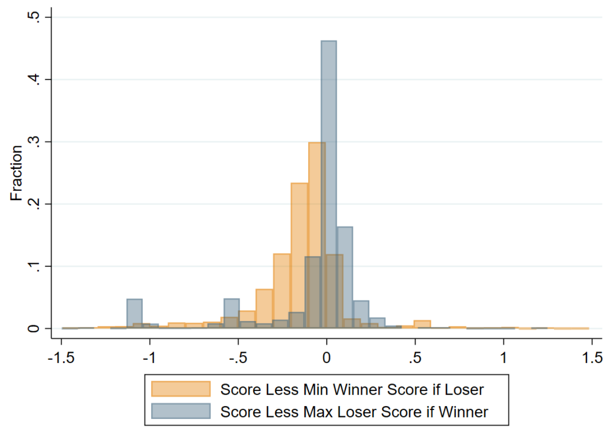
Figure A.18: Raw Scores and Award Probability in Four Representative Topics



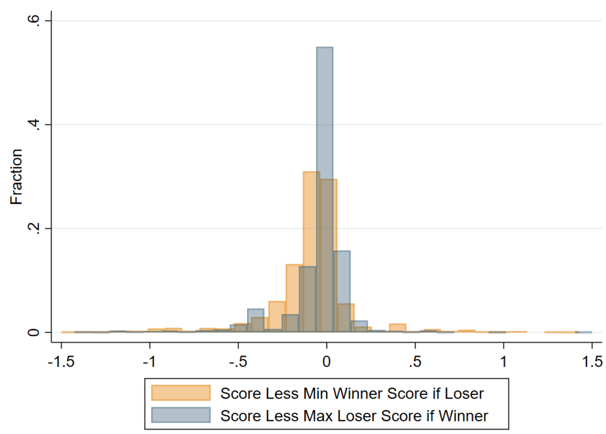
Note: These plots show the probability of winning by raw score. The score perfectly predicts award except occasionally when an awardee is declined in the contracting process because some ineligibility was identified (these instances are dropped in analysis). Note that the range of scores differs across topics, which is we construct a rank normalization for combined analysis.

Figure A.19: Prevalence of Crossover Sub-scores

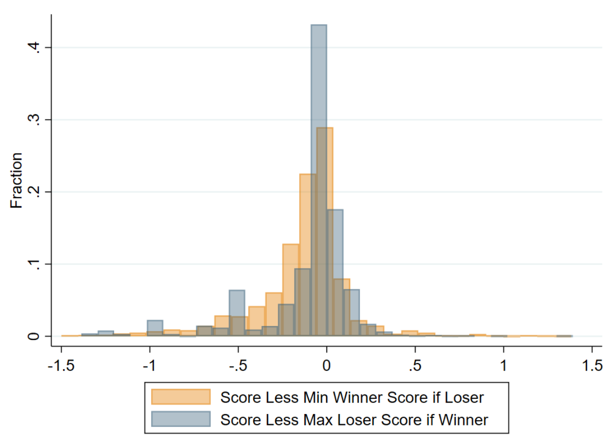
(a) Tech Score



(b) Team Score

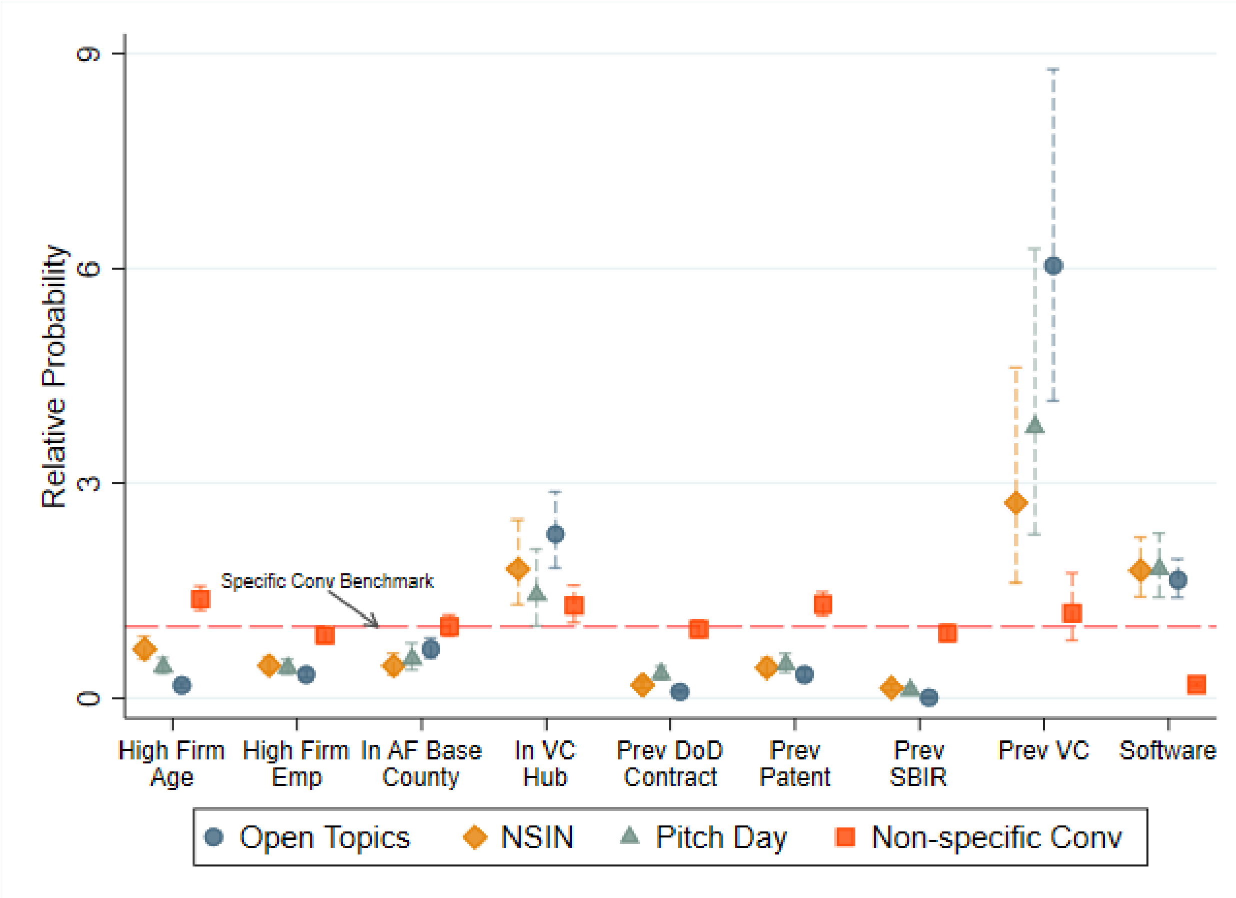


(c) Commercialization Score



Note: These histograms demonstrate the substantial variation in the three sub-scores (tech, team, commercialization) around the cutoff. The red bars to the right side of zero show that many unsuccessful applicants (losers) have a sub-score that exceeds the lowest sub-score among winners. Similarly, the blue bars to the left side of zero show that many winners have sub-scores that are lower than the highest loser sub-score. Altogether, 81% of applicants have at least one sub-score that is a "crossover." All topics 2017-19 are included.

Figure A.20: Selection into Programs



Note: This figure shows how firm characteristics predict selecting into reform programs compared to Conventional topics that are below the median non-specificity score. The points represent coefficients from a multinomial logistic regression with categories for four programs: the specific Conventional topics, which serves as the base group and is represented by the red dashed line, and the three reform programs (Open, NSIN, and Pitch Day) and non-specific Conventional topics. Data restricted to 2017-2019. The dashed lines around each coefficient point indicate the 90% confidence intervals.

Table A.1: Proposal and Firm Counts

Panel A: Open & Conventional (2017-19)

	Both	Open Topic	Conventional
Number of Topics:			
Phase I	512	6	506
Phase II	180	5	175
Number of Proposals:			
Phase I	7229	1656	5573
Phase II	865	444	421
Number of Firms:			
Applied to Type	3170	1408	2409
Exclusively Applied to Type	647	761	1762

Panel B: Full Sample (2003–2019)

	Both	Open Topic	Conventional
Number of Topics:			
Phase I	1796	6	1790
Phase II	661	5	656
Number of Proposals:			
Phase I	19446	1656	17790
Phase II	1684	444	1240
Number of Firms:			
Applied to Type	6485	1419	5724
Exclusively Applied to Type	658	761	5066

Panel C: NSIN and Pitch Day

	Both	NSIN	Pitch Day
Number of Topics:			
Phase I	11	8	3
Phase II	2	1	1
Number of Proposals:			
Phase I	747	423	324
Phase II	28	18	10
Number of Firms:			
Applied to Type	606	361	286
Exclusively Applied to Type	41	320	245

Note: This table shows the counts of topics, proposals (i.e. applications), and unique firms that applied for the Open and Conventional programs in 2017-19 (Panel A), 2003-19 (Panel B), and the NSIN and Pitch Day programs in 2017-19 (Panel C). For example, 1,408 unique firms have applied to Open topics, of which 761 applied exclusively to Open topics. There are 7,229 proposals (note firms can apply multiple times), of which 1,656 are in open topics.

Table A.2: Summary Statistics for Other Reform Programs and Full Conventional Time Period

Panel A: Competition and Company Summary

	Conventional, 2003-19				NSIN & Pitch Day			
	N	Mean	Median	SD	N	Mean	Median	SD
Competition Summary								
Num Proposals per Topic	19,773	18.620	16	12.568	747	78.898	73	27.573
Num Winners per Topic	19,773	3.167	2	3.959	747	14.100	16	5.183
Topic Non-Specificity	19,717	0.977	1	0.582	738	2.257	2	0.985
Award Amount	3,561	$137,120	$130,924	$27,111	83	$119,444	$146,451	$41,841
Company Characteristics								
Age	19,773	15.572	13	12.010	747	12.746	7	12.751
Number of Employees	19,773	48.322	16	76.341	747	34.000	8	76.972
$\mathbb{1}$(in VC Hub)	19,773	0.173		0.379	747	0.190		0.393
$\mathbb{1}$(in County with AF Base)	19,773	0.280		0.449	747	0.170		0.376
$\mathbb{1}$(Minority Owned)	19,773	0.032		0.176	747	0.158		0.365
$\mathbb{1}$(Woman owned)	19,773	0.045		0.207	747	0.123		0.329
$\mathbb{1}$(Proposal is Hardware)	19,773	0.536		0.499	747	0.169		0.375

Panel B: Pre-Award Outcome Summary

	Conventional, 2003-19				NSIN & Pitch Day			
	N	Mean	Median	SD	N	Mean	Median	SD
$\mathbb{1}$(VC)	19,773	0.060		0.237	747	0.072		0.259
Avg VC Amt (Mill)	832	$ 5.825	$1.810	$8.318	40	$6.003	$1.400	$9.892
$\mathbb{1}$(DoD Non-SBIR Contract)	19,773	0.369		0.483	747	0.301		0.459
# DoD Non-SBIR Contracts	7,301	12.837	4	22.681	225	18.338	5	33.293
Avg DoD Non-SBIR Contract Amt (Mill)	7,301	$2.254	$0.720	$6.985	225	$2.049	$ 1.005	$3.213
$\mathbb{1}$(Patent)	19,773	0.440		0.496	747	0.261		0.439
# Patents	8,705	19.434	6	35.112	195	13.200	5	26.916
# Patent Application if Any	8,925	19.275	6	35.230	238	12.105	4	26.668
$\mathbb{1}$(AF SBIR Contract)	19,773	0.514		0.500	747	0.213		0.410
# AF SBIR Contracts	10,173	34.193	12	57.090	159	38.535	10	65.716
$\mathbb{1}$(Never Awarded SBIR)	19,773	0.367		0.482	747	0.660		0.474

Note: This table repeats the summary statistics from Table 1 but for two different data sets. First, on the left side of each panel, is data from the whole Conventional sample spanning 2003-19. Second, on the right side of each panel, is data from the other DoD SBIR reform programs that we consider in Section 6.2.

Table A.3: Phase 2 Competition Summary Statistics

Panel A: Competition and Company Summary

	Open Topic				Conventional			
	N	Mean	Median	SD	N	Mean	Median	SD
Competition Summary								
Num Proposals per Topic	647	137.393	163	57.690	459	8.272	2	12.621
Num Winners per Topic	647	77.811	87	40.737	459	1.793	1	1.955
Topic Non-Specificity	627	2.885	3	0.162	441	0.829	0	1.084
Award Amount	62	\$832,463	\$762,881	\$470,903	74	\$813,940	\$782,165	\$183,199
Company Characteristics								
Age	647	8.622	5	9.821	459	22.986	24	13.509
Number of Employees	645	30.484	10	72.174	459	76.490	35	95.453
$\mathbb{1}$(in VC Hub)	647	0.162		0.369	459	0.155		0.362
$\mathbb{1}$(in County with AF Base)	647	0.088		0.284	459	0.283		0.451
$\mathbb{1}$(Proposal is Hardware)	647	0.133		0.340	459	0.429		0.496

Panel B: Pre-Award Outcome Summary

	Open Topic				Conventional			
	N	Mean	Median	SD	N	Mean	Median	SD
$\mathbb{1}$(VC)	647	0.121		0.326	459	0.076		0.266
Avg VC Amt (Mill)	63	\$6.080	\$2.800	\$8.917	25	\$3.940	\$0.150	\$6.930
$\mathbb{1}$(DoD Non-SBIR Contract)	647	0.195		0.396	459	0.706		0.456
# DoD Non-SBIR Contracts	126	10.063	4	14.773	324	27.219	12.5	32.229
Avg DoD Non-SBIR Contract Amt (Mill)	126	\$1.553	\$0.805	\$2.502	324	\$1.805	\$0.927	\$2.572
$\mathbb{1}$(Patent)	647	0.260		0.439	459	0.580		0.494
# Patents	168	12.065	3	42.972	266	24.850	10	34.960
# Patent Application if Any	214	10.664	3	38.826	282	24.316	10	34.797
$\mathbb{1}$(AF SBIR Contract)	647	0.189		0.391	459	0.728		0.446
# AF SBIR Contracts	122	13.180	2	23.481	334	62.141	20	86.903
$\mathbb{1}$(Never Awarded SBIR)	647	0.742		0.438	459	0.214		0.410

Note: This table repeats the summary statistics from Table 1 Panel A but for the Phase 2 competitions from 2017-19.

Table A.4: Evaluation Sub-scores and Outcomes

Panel A: Total Score

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Score	0.009	0.005	-0.016	0.014	0.007	0.041	0.084***	-0.006
	(0.016)	(0.011)	(0.011)	(0.027)	(0.026)	(0.040)	(0.031)	(0.032)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.019	0.027	0.146	0.148	0.324	0.105	0.299

Panel B: Commercial Score

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Commercial Score	0.081***	0.029	0.001	0.011	0.047**	0.095*	0.044*	0.046
	(0.020)	(0.030)	(0.011)	(0.039)	(0.024)	(0.056)	(0.023)	(0.044)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.019	0.027	0.146	0.148	0.324	0.105	0.299

Panel C: Team Score

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Team Score	-0.115***	0.020	-0.033**	0.028	0.008	-0.021	-0.013	-0.037
	(0.023)	(0.046)	(0.014)	(0.043)	(0.029)	(0.062)	(0.030)	(0.052)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.019	0.027	0.146	0.148	0.324	0.105	0.299

Panel D: Technical Score

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Technical Score	0.004	-0.048	0.009	0.002	-0.064**	0.009	0.059**	-0.053
	(0.023)	(0.059)	(0.012)	(0.042)	(0.031)	(0.064)	(0.027)	(0.054)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.079	0.019	0.027	0.146	0.148	0.324	0.105	0.299

Note: This table shows estimates of the relationship between the proposals' normalized total (Panel A), commercial (Panel B), team (Panel C), and technical sub-scores (Panel D) and the main outcomes within 24 months after the award decision for Open and Conventional Topics. Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. We also control for whether a firm had any previous venture capital investment (columns 1-2), granted patent (columns 3-4), DoD non-SBIR contract over $50,000 (columns 5-6), and DoD SBIR contract (columns 7-8). Standard errors are clustered by firm. All columns include topic fixed effects. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.5: Effect of Winning on Any Patenting Outcomes (originality and generality defined within AF SBIR applicants in year)

Panel A: Any Subsequent High Originality Patent

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.038***	0.020	-0.027	0.020	-0.027	-0.017
	(0.015)	(0.042)	(0.021)	(0.041)	(0.021)	(0.012)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				0.018	0.066**	0.054***
				(0.045)	(0.028)	(0.020)
Observations	1385	2608	7384	3993	8769	21432
Program	Open	Conv.	Conv.	Both	Both	Both
Proposal	First	First	First	First	First	All
Time Period	2017-19	2017-19	2003-19	2017-19	2003-19	2003-19
Outcome Mean	0.018	0.094	0.103	0.068	0.090	0.165

Panel B: Any Subsequent High Citation and Generality Patent

Dep Var:	High Citation		High Generality	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	-0.002	-0.036*	0.015	-0.052***
	(0.003)	(0.019)	(0.018)	(0.019)
Observations	2608	7384	2608	7384
Program	Conv.	Conv.	Conv.	Conv.
Proposal	First	First	First	First
Time Period	2017-19	2003-19	2017-19	2003-19
Outcome Mean	0.001	0.076	0.010	0.085

Note: This table shows effects of winning an award on alternative patent-based metrics. Panel A shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on any subsequent granted high originality patent within 24 months after the award decision, for Open and Conventional topics. The originality score measures whether the patent cites previous patents in a wide range of fields. We define a patent to be highly original if its originality score is above the median in the sample of all applicant patents. Rank within the topic (competition) is controlled separately as a linear function on either side of the cutoff. Panel B shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on any subsequent granted high citation and generality patent within 24 months after the award decision for Conventional topics. These measures are based on citations that accrue after a patent is granted and there is not enough time to observe them for the more recent Open program. Columns 1 and 3 include years 2017-19 and columns 2 and 4 include all years 2003-19. The sample is restricted to first-time applicants only, except in Panel A column 6. In all cases, we control for previous Air Force SBIR awards. All columns in both panels include topic fixed effects. Standard errors are clustered by firm in panel A columns (1)–(3). Standard errors are clustered by topic in columns (4)–(6) in panel A and all columns in panel B. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.6: Phase 2 VC and Government Matching (Open Topics Only) Summary Statistics

	N	Mean	Median	SD
Share Government Match	647	0.131		0.338
Share Private Match	647	0.145		0.353
Confirmed Govt Match Amt	79	$769,446	$600,000	$810,078
Confirmed Private Match Amt	23	$1,273,499	$1,500,000	$468,870
Share Applied Government Match	647	0.182		0.386
Share Applied Private Match	647	0.206		0.404
Applied Govt Match Amt	118	$680,240	$529,619	$538,458
Applied Private Match Amt	133	$1,355,232	$1,500,000	$940,224

Note: This table contains summary statistics about the private and government matching among Open Phase 2 awardees.

Table A.7: Effect of Winning Phase 1 Interacted with Phase 2 Match

Dependent Variable:	VC If No Prvt Match	VC If Prvt Match	Any VC		
Sample:			Match Offered	No Match Offered	
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$(Award)	0.040*	0.015	-0.047	0.074	0.026
	(0.025)	(0.013)	(0.042)	(0.062)	(0.049)
$\mathbb{1}$(Award $\times$ Match Offered in Topic)					0.030
					(0.042)
Observations	1385	1385	1004	381	1385
Outcome Mean	0.068	0.027	0.083	0.071	0.079

Note: This table contains regressions showing the effect of winning a Phase 1 award on measures of VC within 24 months of the award decision interacted with indicators for private and government matching (only available to Open Phase 2 awardees) on subsequent venture capital. In column 1, the dependent variable is redefined to be zero for firms that got a VC match. That is, the dependent variable is zero if a firm got VC and also got a VC match. In column 2, we consider the complement. The dependent variable is redefined to be zero for firms that got VC but had no VC match. That is, the dependent variable is only equal to one for firms that got VC and a VC match and is zero otherwise. Column 3 includes only those topics that offered a match, (19.1, 19.2, and 19.3), while column 4 includes the remaining topics that did not offer a match (18.2 and 18.3). Column 5 shows the interaction. All models include topic fixed effects. The sample is restricted to first-time applicants only. Standard errors are clustered by firm. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.8: Effect of Phase 2 Award and Award Amount on VC and AF Contracts (non-SBIR)

Panel A: Any Subsequent VC and Non-SBIR DoD Contracts

Dep Var:	Any VC		Any DoD Non-SBIR	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	-0.058	-0.015	-0.024	-0.004
	(0.063)	(0.017)	(0.141)	(0.078)
Observations	457	1703	457	1703
Program	Conv	Conv	Conv	Conv
Time Period	2017-19	2003-19	2017-19	2003-19
Outcome Mean	0.013	0.009	0.562	0.524

Panel B: Any Subsequent Patenting and SBIR Contracts

Dep Var:	Any Patents		Any SBIR	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	-0.072	-0.044	0.050	-0.082
	(0.112)	(0.077)	(0.140)	(0.080)
Observations	457	1703	457	1703
Program	Conv	Conv	Conv	Conv
Time Period	2017-19	2003-19	2017-19	2003-19
Outcome Mean	0.100	0.265	0.420	0.579

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 2 award on the main outcomes within 24 months after the award decision. We include both an indicator for award and the award amount in real 2019 dollars. Note the coefficient on the award amount is not shown here. This is possible as the award amount varies, which it does not for Phase 1. Standard errors are clustered by firm. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.9: Cost-Benefit Analysis

All $ in thousands

A. Baseline Results

Program:	Open	Open	Conv	Conv
Outcome:	VC	Non-SBIR	VC	Non-SBIR
	(1)	(2)	(3)	(4)
1. Treatment Effect	0.054	0.075	-0.005	0.033
2. Average Contract Size	$11,085	$11,800	$11,805	$11,800
3. Implied Benefit	$599	$885	-$55	$389
4. Utility Weight	1	1	1	1
5. Benefit Sum	$1,484		$334	
6. Cost	$379		$333	
7. Net Benefit (Benefit - Cost)	$1,104		$1	
8. Net Benefit Difference (Open - Conv)	$1,103			

B. Robustness

	Net Benefit Difference (1)	Open Net Benefit (2)	Conv Net Benefit (3)
1. Baseline	$1,103	$1,104	$1
2. Pre-award Average Contract Sizes	$816	$797	-$19
3. Zero Weight on Non-Military	$449	$506	$56
4. Zero Coefficient on VC for Conv	$1,048	$1,104	$56
5. Conv Coefficients from 2003-19	$1,268	$1,104	-$164
6. Separate Contract Size Averages for Open and Conv	$416	$486	$69

Note: This table shows the cost benefit analysis using the regression discontinuity (RD) estimate of the treatment effect in Table 2. Panel A shows the baseline scenario, where we take the average VC and non-SBIR DoD contract size within 24 months after the award decision. The cost of the programs consist of the average Phase 1 and Phase 2 award amounts and the administrative cost of evaluating a winner. Panel B shows the various robustness scenarios, including using the pre-award contract sizes, considering solely the military (non-SBIR) benefit, assuming zero effects of winning a Conventional program on VC, expanding the analysis to all years (2003-19), and using different average contract sizes for Open and Conventional programs.