NBER WORKING PAPER SERIES

OPENING UP MILITARY INNOVATION:
CAUSAL EFFECTS OF REFORMS TO U.S. DEFENSE RESEARCH

Sabrina T. Howell
Jason Rathje
John Van Reenen
Jun Wong

Working Paper 28700
http://www.nber.org/papers/w28700

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2021, Revised March 2023

Opening up Military Innovation: Causal Eects of Reforms to U.S. Defense Research
Sabrina T. Howell, Jason Rathje, John Van Reenen, and Jun Wong
NBER Working Paper No. 28700
April 2021, Revised March 2023
JEL No. H56,H57,O31,O32,O33,O36,O38

ABSTRACT

How should governments procure innovation? One choice facing policymakers is whether to tightly specify the innovations they seek (a "Conventional" approach) or to allow firms to suggest ideas (an "Open" approach). We study a natural experiment in the U.S. Air Force Small Business Innovation Research (SBIR) program where Open and Conventional competitions were held simultaneously. We compare them using within-competition regression discontinuity designs on administrative data. Open awards increase desired outcomes; they lead to more adoption of new technologies, measured by (non-SBIR) defense contracts, and more commercial innovation, measured by VC funding and patenting. In contrast, Conventional awards have no effects on these outcomes but do create lock-in through increasing the chances of winning a future SBIR award. The Open program succeeded in its aim of attracting new types of firms, but we demonstrate that openness has a differential impact beyond inducing selection: (i) comparing specific and non-specific Conventional topics; (ii) examining firms that applied to both Open and Conventional programs; and (iii) comparing Open with two other reform programs that attracted similar types of firms to Open but used specific topics. Overall, the results point to benefits from open approaches to innovation procurement.

Sabrina T. Howell
NYU Stern School of Business
KMC 9-93
44 West 4th Street
New York, NY 10012
and NBER
showell@stern.nyu.edu

Jason Rathje
U.S. Air Force
jason.rathje@gmail.com

John Van Reenen
Department of Economics
London School of Economics
Houghton Street
London, WC2A 2AE
United Kingdom
and NBER
j.vanreenen@lse.ac.uk

Jun Wong
University of Chicago
junwong@uchicago.edu

1 Introduction

Amid concern about declining productivity growth, the role of innovation policy has become ever more important (Decker et al. 2016, Syverson 2017, Goolsbee and Jones 2022). Although the economics literature has paid much attention to the government's role as a major funder of R&D, there is much less study of how the public sector should *design* innovation procurement. A key decision is whether to take a centralized approach where the desired innovation is tightly specified or to take a more open, decentralized approach where applicants are given leeway to suggest solutions. There are trade-offs. The open approach may result in too many suggestions that are not useful to the funder, whereas the centralized approach may work poorly if there is uncertainty about what opportunities exist and may result in insularity if a small group of firms specializes in the specified projects.

The ideal experiment to compare these two approaches would take an existing centralized program—the most common form of public sector procurement—and decentralize a portion of competitions while keeping all other design parameters constant. Fortunately, such a quasi-experiment occurred in 2018 in the Small Business Innovation Research (SBIR) program of the U.S. Air Force. SBIR is among the world's largest and most influential small business R&D grant programs, spending $3.11 billion across 11 Federal agencies in 2018. Of this, the Department of Defense (DoD) accounted for $1.32 billion, and the Air Force had the largest program among the military services. The Air Force holds multiple competitions about every four months in which firms apply to develop military technologies, first through a White Paper (Phase 1) and then a prototype (Phase 2). The "Conventional" approach is to hold competitions with highly specific topics such as "Affordable, Durable, Electrically Conductive Coating or Material Solution for Silver Paint Replacement on Advanced Aircraft." After 2018, the Air Force also included an "Open" competition which ran alongside the Conventional model, where firms could propose what they thought the Air Force might need. This reform was partly in response to a growing concern that American military innovation was in decline, something that we confirm to be the case.

We use rich administrative data between 2003 and 2019 on applications and their evaluation scores to assess the causal impact of the Air Force SBIR programs. We have 19,500 proposals from 6,500 unique firms (and 7,300 proposals across 3,200 firms in the baseline 2017-19 sample when Open and Conventional were run simultaneously). An advantage of our setting is that the Air Force has clear and quantifiable objectives. They want firms to not only do the

R&D—which is what SBIR is for—but also to deploy new technologies in both the private and defense sectors. It is unusual for researchers to directly observe this technology adoption. We can measure this accurately through observing future (non-SBIR) U.S. defense contracts with SBIR applicant firms. Moreover, we measure private sector adoption through firms' subsequent venture capital (VC) funding as well as the number of quality of patents. Our discussions with senior Air Force officials confirm the explicit "dual-use" mandate, where they desire technologies that they fund to also be used in the private sector. We observe all outcomes for at least three years after the last award period ends (through January 2023).

We find that the Open program reached new types of firms, which was one policymaker objective. Compared to firms applying to Conventional topics, Open topic applicants are about half as old, half as large, less likely to have previous defense contracts, and more likely to be located in high-tech entrepreneurial hubs. Next, we identify causal effects of winning with a sharp regression discontinuity design (RDD) that compares winning and losing applicants around a cutoff for the award. Our research approach exploits the facts that (a) the review and selection process was the same across the Open and Conventional programs, making them comparable; and (b) the RDD can be implemented using the rank that determines the award decision as a running variable. This ranking is constructed by the forced ordering of independent scores from three evaluators. The cutoff is independent of the evaluation process, making the manipulation of any firm around the cutoff extremely unlikely.[1]

Our main result is that winning an Open topic competition has positive and significant effects on three outcomes indicative of dual-use benefits desired by the program administrators. Winning increases (i) the chances of the military adopting the new technology (as measured by subsequent non-SBIR DoD contracts) by 9.8 percentage points (49% of the mean); (ii) the probability of subsequent VC investment by 6 percentage points (68% of the mean); and (iii) patenting and patent originality. By contrast, winning a Conventional award has no measurable effects on any of these outcomes. Furthermore, there were no causal impacts of winning a Conventional award between 2003 and 2017—before the Open program was introduced—so it is not the case that Conventional projects require a longer time horizon or were crowded out by Open ones. Where Conventional does have an effect is on the chances of winning a future SBIR award, while there is no effect on this outcome in the Open program. This is considered

---

[1]We document a smooth density around the cutoff and continuity in baseline covariates. One downside of RDD is that the results are necessarily local. However, in our case, because the results are generally similar when we use the whole sample or a narrow bandwidth, and with or without rank controls, they seem likely to apply more broadly.

a rather undesirable feature of the Conventional program from a policymaker perspective, as it creates lock-in and insularity.

Although reaching new firms was a key aim of the reform and successfully raised adoption and innovation, it raises the question of whether Open was more successful due to this kind of selection. This would limit the implications of our results, because in many settings it is hard to change the applicant types. We use three complementary research designs to show that the positive effects of openness go beyond inducing a different composition of applicant firms.

First, we implement a direct test for the role of openness by using machine learning techniques on application abstract texts to characterize the degree of specificity for each topic. This measures 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 approach—winning an award in that topic *does* increase innovation. Second, there are more positive effects of Open even among the firms that applied to *both* the Open and Conventional programs. Third, we examine two other DoD SBIR reforms in 2018 that also had specific topics like the Conventional program, but due to other features such as faster contracting and outreach to startup hubs, attracted firms similar in characteristics to those in the Open program. We find that the Open program had larger effects than these other programs. All three designs balance the characteristics of applicant firms in different ways, but all three imply that the success of openness is in part due to the way it incentivizes greater innovation from broadly the same pool of firms.

In sum, the Open program seems to work in part because it provides firms with an avenue to identify technological opportunities about which the government is not yet fully aware. The Open contracts may represent an entry point to much larger public sector contracts, which helps to explain their large effect. 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. It is worth noting that there are no classified (i.e. "secret") SBIR projects. In being applied and not classified, SBIR is representative of the vast majority of defense R&D.[2]

Since our results point to benefits from open innovation, we build a simple model describing the welfare trade-off between the Open and Conventional programs and

---

[2]Although military R&D has the reputation of being highly secretive and therefore less likely to generate civilian benefits than non-military R&D, many papers find significant spillovers from defense innovation to the private sector (e.g. see the discussion in Moretti et al. (2020)).

implement a quantitative cost-benefit analysis from the perspective of the Air Force to running a marginal competition, comparing Conventional to Open topics. The net benefits of an Open competition are around $700,000 more than those of an equivalent Conventional competition. Consistent with our findings, the revealed preference of the Air Force is to pursue Open. After the period of experimentation, they settled for an interim period on an 80-20 budgetary split between Open and Conventional.[3]

The results of our study are of general interest. Governments around the world, including in the EU, UK, and other U.S. agencies such as DARPA, have sourced innovative ideas from firms via open solicitations.[4] For example, the National Institutes of Health funds both "investigator-initiated competitions," similar to Open, as well as more specific "requests for applications," similar to Conventional (Myers 2020). Our setting is also inherently important; the DoD is the largest single investor in R&D in the world and comprises about 60% of total U.S. federal government R&D (CRS 2018).[5] DoD has historically been an important financier and early customer for technology, both transformational and incremental (Mowery and Rosenberg 1991, Mazzucato and Semieniuk 2017, Gross and Sampat 2020). Its investments often have dual-use properties, generating opportunities for large private sector spillovers. Russia's invasion of Ukraine prompted announcements of large increases in Western military budgets, so understanding how to get the best value for this additional money is crucial.[6]

We make two contributions to the literature: (i) we show that decentralization and openness are relevant for public R&D procurement; and (ii) we offer the first causal analysis of a defense

---

[3]See Defense Comptroller 2021 Budget Justification.

[4]The largest EU small business innovation funding program, the Horizon SME Instrument, has shifted to open solicitations over the past decade. The call for applications explains: "If research and development activities are not translated into innovation, i.e. value creation from novelty, SMEs cannot take profit from those activities. Great ideas need to be picked up by users, customers and the market, before they can really transform society....The instrument is now fully bottom-up, meaning that call respondents do not have to worry about answering to specific topics." See https://ec.europa.eu/programmes/horizon2020/en/area/smes. Other examples 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).

[5]Also see here.

[6]Private sector companies are also increasingly using open innovation, especially in R&D-intensive industries (Chesbrough 2003, de Villemeur and Versaevel 2019). 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

R&D program. Regarding (i), our paper joins work on how to motivate or procure innovation, such as Manso (2011), Azoulay et al. (2011), Nanda et al. (2014), Halac et al. (2017) and Krieger et al. (2018). In related work, Che, Iossa and Rey (2021) theoretically compare cash prizes and follow-on contracts to motivate innovation. They argue that in the absence of perfect information, contract rights are the optimal mechanism, and furthermore that bundled approaches in which the innovating firm receives the follow-on contract are ideal for unsolicited proposals, akin to our Open setting. The authors point specifically to DoD SBIR as an important setting in which follow-on contracting is used to incentivize innovation.

More broadly, we contribute to the literature on subsidizing innovation (e.g. Goolsbee 1998, Atkeson and Burstein 2019, Bloom et al. 2019, Pless 2019, Rathje and Katila 2020, Akcigit et al. 2021), especially through direct R&D grants (Lach 2002, Jacob and Lefgren 2011, Azoulay et al. 2019). A subset of this literature concerns the SBIR program, such as Lerner (1999), Wallsten (2000), Howell (2017), and Lanahan and Feldman (2018). Bhattacharya (2021) studies Phase 2 of the U.S. Navy's SBIR program, focusing on competition among Phase 1 winners, who are eligible to apply for Phase 2. He finds that increasing the number of applicants and forced sharing of ex-post patents would be socially beneficial, but would not necessarily benefit the DoD. In his structural model and analysis, the only outcome is Navy contracts. Our paper is complementary, as we examine the causal effect of winning the Phase 1 competition and consider a broad array of outcomes, including all DoD contracts and commercial innovation.

We also contribute to work on defense R&D. Defense is unique because the 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.—almost unlimited buying power in the event of success. The defense setting enables us to study the government as a customer rather than a regulator or 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 even more so in many other countries. The literature has also used military spending as an exogenous shock to demand (Ramey 2011, Barro and Redlick 2011, Nakamura and Steinsson 2014), and has studied the the crowd-in effects of defense R&D (Lichtenberg 1984; 1988; 1995, Middleton et al. 2006, Draca 2013, Moretti et al. 2020, Belenzon and Cioaca 2021).

2 Defense R&D and SBIR Reforms

The addition of an Open program to the Air Force SBIR was motivated by concerns among policymakers about declining innovativeness among prime defense contractors and the SBIR program's failure to generate useful technologies for military and commercial purposes. As we could find no existing quantitative studies showing that, in fact, defense innovation has declined, we describe the economic context for U.S. military R&D and document innovation trends in Appendix A. The results of this exercise reveal dramatic consolidation among prime contractors in recent decades, accompanied by a decline in innovation quality relative to the private sector (see panel A of Figure 1 and Figure A.1). This has occurred despite a substantial increase in prime contractors' profits and assets (Figure A.2 panels C and D). These stylized facts set the stage for the reform we study.

SBIR Context and Challenges. The SBIR is one of DoD's principal tools for bringing new technology and firms into the defense ecosystem, thus reducing insularity among the prime contractors. The Secretary of Defense said in 2021 that the SBIR "program helps fuel American firms to pursue R&D tailored to the department's unique tech requirements" (Garamone 2021). Small contracts in the SBIR program feed into the broader defense industrial base when an SBIR awardee becomes a major contractor, as in the case of Progeny or Qualcomm, or is acquired by a prime contractor (SBIR.gov 2011).

Congress first authorized the SBIR program in 1982 to strengthen the U.S. high technology sector and support small firms. SBIR has two Phases: the smaller Phase 1 award funds proof-of-concept work, after which a firm may apply for a larger Phase 2 award to support later stage demonstration.[7] SBIR applicant firms are typically small and high-tech, a firm type that is crucial to job creation and innovation, especially those that receive VC backing (Kortum and Lerner 2000, Foster et al. 2008, Haltiwanger et al. 2013, Arora et al. 2018, Howell et al. 2020). SBIR at DoD funds applied R&D, as opposed to basic research.

Policymakers have expressed concern about lock-in at the SBIR program, with repeat contracts awarded to firms that are interested neither in commercializing innovation nor in seeking scale in the defense market (Edwards 2020). This concern may relate to the difference between our findings and the strong positive effects of Department of Energy (DoE) SBIR

---

[7]The Small Business Technology Transfer (STTR) program is an add-on to the SBIR program and requires small businesses 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."

grants in Howell (2017)). Indeed, we show that this is the case in Figure A.3 Panel B (see Appendix A.2 for details). The greater lock-in at DoD might reflect 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 research firm. Policymakers' concern appears well-founded because there has been a decline in relative innovation since the 1990s among winners in the Conventional SBIR program (Figure 1 Panel B), paralleling the decline among the prime contractors (see Appendix A).[8]

SBIR Process at the Air Force. The Conventional and reform programs that we study have a common administrative process. 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 scores on three criteria: Technology, Team, and Commercialization.[9] The commercialization sub-score reflects the potential to sell any derived product or service within and outside the government. Firms' proposed cost is not a factor in the evaluation if 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, no single evaluator can manipulate a firm's position around the cutoff

---

[8]Since firms must be small to participate, concentration is not a primary concern in the SBIR. However, Figure A.3 panel A uses the Herfindahl-Hirschman Index (HHI) to show that the DoD SBIR program has become more concentrated over time, with more firms winning many awards in a single year.

[9]The official description for the Conventional program of these criteria is: "(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."

because each evaluator independently scores the proposal. 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.

Overall, the review process, metrics, and selection mechanism are the same across both programs, and indeed sometimes the evaluators are the same Air Force Science and Technology personnel. This makes them well-suited to be compared to each other.

SBIR Reforms: Open vs. Conventional. Conventional topics are highly specific. One example is: "Develop Capability to Measure the Health of High Impedance Resistive Materials" (more examples are in Appendix B). First conducted in May 2018, Open topic solicitations contain no direction regarding the technology that the applicant may propose.[10] 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 its 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). That is, there are three Open topics a year with the same language. 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

---

[10]The SBIR reforms have taken place within a new organization called Air Force Ventures. This is the business division of AFWERX, an office that seeks to foster innovation within the Air Force. Conventional topics are sourced primarily from the Air Force Research Laboratory (AFRL).

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). Therefore, the fact that we identify larger causal effects in Open than in Conventional cannot be explained by the difference in award amounts since Open is less financially generous than Conventional. The Phase 2 awards of $300,000 to $2 million are intended to last 12-24 months and fund a prototype (this paper focuses on Phase 1, so we minimize the discussion of further awards and provide more details on Phase 2 in Appendix B). In the later Open topics, the Air Force 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 Appendix B, we provide further institutional context and details about the contracting process.

3   Data and Summary Statistics

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). In some analyses, we restrict the sample to the years 2017-2019 so that market factors are similar across the sample. We also assess the Conventional program using all available years.

Figure A.4 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 6,654 Phase 1 proposals from 3,058 unique firms. Our main analysis does not consider Phase 2 awards because the sample is small and the Open Phase 2s are too recent to observe outcomes. Summary statistics for Phase 2 are in Table A.2.

Table 1 Panel A describes company competition characteristics at the proposal level. 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 1,600 applicants per topic. While the

number of applicants and share of winners is different in Open, this stems from the fact that Open represents a single topic per solicitation rather than the many specific topics in each solicitation for Conventional. As mentioned above, the review process is the same for Open and Conventional. Table A.3 shows similar statistics for the whole Conventional sample.

We collect outcome data spanning our whole data sample period and at least 37 months after the last award (as of January 2023). This provides sufficient time to observe effects. The outcomes are summarized in Table 1 Panel B. Subsequent VC investment and technology adoption as measured with DoD non-SBIR contracts correspond to the two key metrics of success from the program administrators' perspective. The Air Force leadership views commercial innovation as evidence of initial success, based on the idea that a strong U.S. industrial base will ultimately enable strong defense, especially if its research has early-stage ties to DoD (Williams 2020). From an economic perspective, VC 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 VC deals from Pitchbook, CB Insights, SDC VentureXpert, and Crunchbase. The majority of deals come from Pitchbook. Non-SBIR DoD contracts, which we gather from the Federal Procurement Data System, represent 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."[11] It is equivalent to technological adoption, something that is often hard to observe in empirical studies.

Two ancillary outcomes are related to innovation and lock-in dynamics. The first is patents with application dates after the SBIR award date that 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 all SBIR awardee technology abstracts are publicly available, and no projects are classified. Therefore, secrecy orders on patent applications are unlikely to affect our results. Using the granted patents, we build measures

---

[11]We cannot easily link these non-SBIR contracts to particular SBIR awards because summaries of the contracts' contents are not available. However, we can manually identify a number of examples. One from the Open program is the firm 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 Ascendant Engineering Solutions. After winning a $149,000 Conventional Phase 1 award in September 2016 to work on gimbals, Ascendant Engineering Solutions was awarded a $7.5 million Air Force contract for its tactical gimbals in February 2018. (Note: A gimbal is a pivoting support that permits an object to rotate on a single axis.)

of patent quality, including originality, forward citations, and generality. The final outcome is subsequent Air Force SBIR awards, (results are similar using all-DoD or all-agency SBIR awards). We obtain these from the Small Business Administration.

Table 1 indicates that Conventional applicants have higher average base rates of patenting and contracting. This is equally true across winners and losers, which will be consistent with the absence of a causal effect in the RDD analysis. The intuition of the RDD is to assess a causal effect of winning conditional on selecting into applying. If there is no effect, we conclude that firms would have similarly high base rates of the outcomes in the absence of the competition, suggesting that in an immediate sense, the program is likely a poor use of taxpayer dollars.

Finally, we make use of the text in proposal abstracts to assign proposals to technology clusters. Employing a machine learning algorithm called k-means clustering (Forgy 1965, Bonhomme and Manresa 2015), we classify each abstract based on its word embedding.[12] Applications are clustered 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.5 and A.6, with the word's size scaled to reflect its prevalence in the cluster.

One aim of the Open reform is to attract new types of firms to the defense market. Here, we assess whether it was successful in doing so. 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, half as old (10 vs. 18 years old) and half the size (27 vs. 61 employees) of Conventional applicants. They are more likely to be in the VC hubs of the San Francisco Bay area, greater Boston, and New York City and less likely to be in a county where there is an Air Force base.[13] The technology clusters by program type and winner status are shown in Figure A.7 (Open applicants are more likely to have software technologies). In terms of outcome variables (Panel B of Table 1), Open

---

[12]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

[13]To describe their geographic diversity, we map the location of applicants in Figure A.14, 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.15 and documents similar patterns.

applicants are more likely to have previous VC financing and are less than half as likely to have previous DoD contracts or SBIR awards.

To describe these differences and assess statistical significance, Figure 2 Panel A shows how the Open applicants (blue circles) differ from the Conventional applicants (red dashed line). This figure reports coefficients in the form of relative risk ratios from a multinomial logit model that estimates whether ex-ante firm characteristics predict applying to a particular program along a number of dimensions. Open applicants differ substantially from Conventional applicants. (The other programs, NSIN and Pitch Day, are introduced below for a supplementary test.) Finally, selection across all our observables is visually apparent in Figures A.8-A.11, which we use to validate our empirical design.

In sum, the Open program seems to have attracted a good deal of new entry by new types of firms into defense R&D procurement. In Section 6, we explore whether this selection alone explains our results, or whether there is also an additional role for openness even conditional on firm type.

4   Empirical Design

Since the application and scoring processes for Open and Conventional topics are very similar (see Section 2), we can use the same regression discontinuity design (RDD) for both programs. The RDD approximates the ideal experiment of randomly allocating awards among applicants. The intuition is either a discontinuity at the cutoff (Hahn et al. 2001) or local randomization around the cutoff (Lee 2008). It is relevant in settings where treatment assignment is based on an applicant's location around a cutoff in a rating variable. Our setting permits a sharp RDD because the running variable perfectly predicts award in all topics in both Open and Conventional. This is shown for four representative topics in Figure 3; the probability of treatment jumps from zero to one at a cutoff.

A valid sharp RDD has four conditions (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 of 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 and, as mentioned above, is illustrated in Figure 3. 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 4, 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 the 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.8-A.11 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. In Table A.4 we formally confirm that there are no significant differences around the cutoff among the baseline covariates.

We conduct a third test for manipulation to address concern 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.12 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

14

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.[14] 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:

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

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 24 months after the award decision. 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 narrower or wider array of controls, including pre-award outcome variables. 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. (We do not consider the award amount because it is co-linear with winning.) 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 fully interacting the right-hand side variables with the vector $\mathbf{Program'_T}$:

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

In the main tables, $\mathbf{Program'_T}$ takes one of two values for Open and Conventional. In Section

---

[14]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).

6, we add two other program types: NSIN and Pitch Day. 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.

5   Effects of the Open and Conventional Programs

This section describes the effects of winning Open and Conventional competitions on innovation and contracting outcomes.

Venture Capital (VC) Investment.   We first examine the effect of winning an award on receiving any VC after the award decision.[15] In addition to being a goal of the program, VC represents high-growth innovation potential and leads to spillovers (Gornall and Strebulaev 2015, Lerner and Nanda 2020). Figure 5 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.[16]

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. Column (1) of Panel A shows that winning an Open award increases the probability of VC by six percentage points, which is 68% of the mean (8.7%) among Open applicants. This compares to a near-zero effect for Conventional whether in the same period 2017-19 (column (2)) or the longer period 2003-19 (column (3)). Column (4) pools the Open and Conventional applicants and estimates Equation 2. The positive coefficient on the interaction between winning and Open implies a significantly larger effect on VC funding of winning an Open relative to a Conventional award in all time periods (columns (4)-(5)) as well as when using all applications

---

[15]In unreported models, we find similar results using the level and log amount of VC funding.

[16]The probability of VC before and after the award decision increases in rank for the Open program (Figure A.10 Panel B and Figure 5 Panel A), pointing to the need to control for rank in the analysis.

instead of only a firm's first proposal (column (6)).[17] Note that since we cluster errors by topic in columns (4)-(6), the significantly larger effect of the Open program does not reflect any across-firm within-topic correlation in the standard errors. Also, the difference between the programs does not reflect insufficient time to observe Conventional results, because we find no effects in this program using the whole 2003-2019 sample (columns (3), (5), and (6)).

As explained in Section 2, 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. Why would small 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 can bring evidence of large defense customers to VCs, which appears to improve their odds of raising funds. The expectation of a Phase 2 award, which averages about $830,000 (see Table A.2 Panel A), may also help to explain the effect if VCs are responsive to Phase 1 because they expect it to lead to substantially more non-dilutive cash. We find no effects of winning a Phase 2 award (Appendix E). 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.

Technology Adoption: Non-SBIR Defense Contracts. One goal of the SBIR reforms is to enable more firms to transition technologies out of the SBIR program to operational

---

[17]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.

programs of record. We therefore consider the effect of winning an award on an indicator for subsequent technology adoption in the form of non-SBIR DoD contracts. The visual results are in Figure 6 where we observe that there is a level shift upwards to the right of the cutoff in Open (Panel A) but not in Conventional (Panel B). In Panel B of Table 2 we replicate the specifications of Panel A, except we use technology adoption as an outcome. The effect of winning in an Open topic is 9.8 percentage points, or 49% of the mean of 19.9% (column (1)). By contrast, there is no significant effect of winning a Conventional award (columns (2) and (3)). 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. Although weaker than the VC results, winning an Open topic has a positive causal effect on a firm subsequently creating something of practical value for the DoD.

For both Open and Conventional, about half of the subsequent technology adoption, or non-SBIR DoD contracts that winners obtain, is from services besides the Air Force. In unreported analysis, we find roughly similar effects of Open vis-a-vis Conventional in both Air Force and non-Air Force contracts. Since all services view any DoD contract as a successful transition of an SBIR firm into the defense pipeline, it is important to consider contracts beyond the sponsoring service.[18]

Patents. We next turn to patenting as an alternative measure of innovation. The visual analysis in Figure 7 shows a 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 increase among Conventional applicants. Panel A of Table 3 confirms the effect of winning an Open award on subsequent patenting, with an effect of 9 percentage points (column (1)), which is 77% of the mean. Column (2) suggests that there may be a positive effect in the Conventional program, although it is smaller, insignificant, and does not appear in the longer time period. 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.

---

[18]For example, the Navy's SBIR webpage states that "The Navy's SBIR/STTR Programs are primarily mission oriented, providing companies the opportunity to become part of the national technology base that can feed both the military and private sectors of the nation. To that end, the Navy incorporates into its Phase II component, the emphasis on the small business' need to market its technology to both military and private sectors." See https://www.navysbir.com/.

Columns (3) and (6) of Table 3 suggest a small negative effect of winning a Conventional award over the 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 the intent to commercialize an invention.

We examine 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)), but no effect on any of the measures, including citations, in Conventional. These results, together with the visual analysis (Figure 7 Panel B), lead us to conclude that Open has a meaningfully larger effect on patenting.

Future SBIR contracts. Last, we consider the chances of subsequent Air Force SBIR contracts. The RDD helps us to overcome the difficulty of separating state dependence (the causal impact of the lagged dependent variable) from unobserved heterogeneity (e.g., if the best firms keep winning contracts). The effect on this outcome is rather different than the previous three. Figure 8 shows a large jump at the threshold for winning a Conventional award, but no change for Open. 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). In contrast, columns (2) and (3) indicate a 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 14 percentage points is 56% of the mean, while over the full time period in column (3) the effect is 140% of the mean. The interaction with Open in the last three columns is consistently negative and significant.

The Open program may have avoided locked-in contractors only because it is new. To assess this, we use 2020 application data, which we do not use in the main analysis in order to have enough time to observe outcomes. If Open awards will also suffer from lock-in, we expect to see some evidence of it in the third year. Appendix Figure A.16 contains a histogram of the number of Open and Conventional applicants in categories defined by the number of Air Force SBIR awards in the past three years, with Open applicants from 2020 and Conventional applicants from 2019.[19] Conventional applicants are far more likely to have many Air Force SBIR awards in the past three years. Nearly 100% of 2020 Open applicants are entirely new

---

[19]We do not observe Conventional application data in 2020, and this approach also aligns the sample with that of our main analysis (2017-19, where 2017 only contains Conventional applicants).

to the program, while only about 60% of Conventional applicants have no SBIR award in the previous three years. Conversely, there is a long tail of Conventional applicants with many Air Force SBIR awards in the previous three years.

In sum, whether through reputation, dedicated staff, or some other channel, the traditional SBIR contract gives birth to recurring SBIR-winners. By contrast, Open topics have avoided this lock-in effect.

Robustness Tests This subsection describes 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.[20] 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.

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 B we use two ranks (to keep the sample size reasonably large) above and below the threshold (±1,±2). The results remain robust. This narrow bandwidth test also helps us understand the degree to which the results apply only in the region around the cutoff. Because the key results are similar when we use the whole sample or a narrow bandwidth, and are similar with and without rank controls, we believe they seem likely to apply more broadly. In Panels C and D 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.

The zero effects for Conventional in the 2017-19 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.4). 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 E 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.

---

[20]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.

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, using alternative vectors of controls for baseline characteristics, or using no controls at all. We also assess whether the differential effect of Open reflects a larger number of applicants (winners) by interacting winning with the topic's number of applicants (winners) being above median (or above the 75th or 90th percentile). The interactions are small and insignificant. In other words, the effects are not larger in large Conventional topics. Finally, we find similar results scaling both the grants and the contracts or VC deals by dollar amounts. The results are stronger in favor of Open because the Open grants are smaller.

6   Does Selection Explain the Larger Effects of Open?

Our conclusions from Section 5 are clear. Winning an Open award increases the chances of future VC funding, patenting and supplying the DoD with future technologies. By contrast, the only effect of winning a Conventional SBIR contract is to increase the firm's chances of winning *another* SBIR contract in the future. From the perspective of whether these programs are useful to DoD, it is irrelevant whether Conventional applicants have higher base rates of patenting and contracting across both winners and losers; the absence of a causal effect means firms would have these higher rates even if they had not won a Conventional award.

We now explore whether the success of Open is due to the composition of applicants (e.g. more startups from high tech hubs) or whether Open is more effective even for observationally identical firms. Although this distinction may be irrelevant from DoD's perspective as Open's success could come from either mechanism, it is an important distinction when considering the implications of our findings in other settings.

We consider three core strategies to achieve "balancing on the unobservables". First, we take the sample of Conventional-only firms but distinguish between competitions that focused on topics that were less specific (more like Open) vs. those that were more specific. Second, we look at two other new Air Force SBIR programs which attracted broadly the same type of firms that applied to Open, but were centralized and top-down like Conventional. Third, we look at firms that applied to *both* Open and Conventional. In all three cases, it is the open nature of a competition that has positive treatment effects regardless of the characteristics of the pool of companies that apply to the respective competitions. Finally, as a further check, we document an absence of systemic treatment effect heterogeneity along the lines of Open

firm characteristics.

More "Open" topics in the Conventional Program. Some Conventional topics are more specific than others in identifying the technology DoD wishes to procure. If openness is important, there should be larger effects of winning a Conventional award when the topic is more technology-neutral, encouraging a broader range of ideas.

To develop a measure of topic specificity, we employ the machine learning algorithm for proposal abstract text introduced in Section 3 (described in detail in Appendix C.3). After summarizing the text as a vector of word embeddings, we measure the topic's specificity based on the distribution of its applications. Specifically, we take the distance between each application and the average (the centroid) for the topic. The non-specificity index for the topic is then the average of these distances within the topic, analogous to variance in the scalar case. Intuitively, if the content across proposals is more diverse, then the topic is less specific (i.e. more "open") and the non-specificity index takes a higher value. As a useful validation exercise, we find that Open topics are much less specific than Conventional topics (Table 1 Panel A). Nevertheless, there is considerable heterogeneity in the degree of specificity within the pool of Conventional topics, and we exploit this in the following designs.

First, we show that subsetting on topic specificity does not induce selection within Conventional. In Figure 2 Panel B, we compare selection into the programs, where we divide the Conventional topics into above and below median specificity. Here, the baseline group, denoted by the horizontal red line, is specific Conventional topics. The orange squares represent the non-specific Conventional topics. Especially relative to the reform programs, the composition of applicants in non-specific and specific Conventional topics is very similar.

Next, we we restrict the sample to Conventional topics and interact winning with a dummy for topics with above median non-specificity in Table 5 Panel A.[21] All columns include topic fixed effects, which absorb the linear specificity measure. We consider three patent-based outcomes because there was some evidence of a possible effect on patents for Conventional (Section 5). Columns (1) and (2) show that for both first and all proposals, there is a positive effect on patenting of the interaction between non-specificity and winning an award, indicating that more "open-style" Conventional topics yield a relatively larger positive effect of winning on patenting within the Conventional sample. In particular, column (1) implies that a non-specific award is 6.6 percentage points more likely to generate a patent than a specific award

---

[21]Results are similar if we use the continuous measure of the non-specificity index.

(which has an insignificant effect), over a mean of 16.6%. We also find significantly larger effects for above-median forward citations and originality in columns (3)-(6).

If openness matters, we also expect that the significantly larger effect of Open relative to Conventional in the baseline results will attenuate when Conventional topics are less specific (i.e. more open). In Table 5 Panel B, we repeat the models from Table 3 Panel A columns (5)-(6), but restrict the Conventional sample according to specificity. Columns (1)-(4) restrict the Conventional sample to topics that are in the bottom tercile of the non-specificity distribution. Here, we see a strong positive effect of Open that is even more striking than in the full-sample models of Table 3. In contrast, in columns (5)-(8) we restrict the sample to topics with top-tercile of non-specificity (which are more like Open topics), and here see no significant differential effect of Open.

Non-specificity is only one characteristic that distinguishes topics from one another, so we checked whether the results in Table 5 were robust to including other observable characteristics of competitions.[22] 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).[23]

To summarize, the effect of winning a Conventional topic on patent activity is significantly larger for less specific—i.e. more open—topics. This suggests that the Conventional program is more impactful when it takes a more open approach, consistent with openness being important independently from selection or other characteristics of the Open reform program.

Other Reform Programs: Pitch Day and NSIN.   We next pit openness against selection using two other reform SBIR programs: Air Force Pitch Day and National Security Innovation Network (NSIN) topics. These programs were instituted at roughly the same time as the Open program. They attract firms with characteristics that are very similar to the Open applicants, but the topics are more specific. By comparing the effects across the three programs, we can assess whether openness appears independently important.

---

[22]We obtain similar patterns of coefficients when we use our other outcomes as the dependent variable in specifications like those in Table 5, although the results are generally insignificant.

[23]For example, we re-ran column (1) of Panel A in Table 5, 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 standard error of 0.033).

Pitch Days were held in VC hubs such as Boston, New York City, and Austin. They seek to bring mission programs with large procurement needs 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, except 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. Summary statistics for these other reform programs are in Table A.3.

Figure 2 Panel A indicates that firms applying to the three reform programs have markedly different characteristics relative to firms applying to the Conventional topics. For example, applicants in Open, NSIN, and Pitch Day are all much less likely to have previous patents or above-median employment. The coefficients for the three reform programs tend to be clustered tightly together with a large difference from Conventional. These common differences in 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 A.5 using Equation 2, with VC as the dependent variable. 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 find a negative 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. We similarly find no effects of these two reform programs on the other main outcomes. A caveat is that the samples for the two additional reform topics are small. However, recall that we observe a significant effect in Open topics on VC even with a narrow bandwidth, which has a similar sample size as the Pitch Day and NSIN samples (Table 4 Panel B). By showing null results in a setting without openness but with firm characteristics similar to Open, these results offer further evidence that selection does not fully explain the effects of the Open reform.

Firms that Applied to Both Programs. The previous two analyses examined treatment effects on firms with Conventional characteristics (the specificity test) and Open characteristics (the other reform programs). We next examine firms who apply to both programs, whose unobservable characteristics should be more tightly matched. Specifically, we repeat the analysis in column (4) from Tables 2-3 but restrict the sample to those firms which at some point in our sample applied to both programs (usually in different years). There are 727 unique firms in this sample of 1,395 applications.[24]

Table A.6 shows that there are significantly larger effects on VC and subsequent DoD contracts of winning an Open competition in this subsample. The interaction coefficient for patents is large and positive. This is remarkable because firms that select into applying to both programs appear more likely to be SBIR specialists who apply to as many topics as possible. Indeed, Figure A.13 shows that along all dimensions except previous VC, firms that apply to both programs are more aligned with those that apply only to Conventional; for example, they are larger and are more likely to have previous DoD and SBIR contracts than those that apply only to Conventional.

Heterogeneity of Treatment Effects: Financial Constraints and other Firm Characteristics. If the causal effects of Open reflect a higher share of firms with startup characteristics, perhaps because they are more financially constrained, we expect systematically larger effects among such firms. Therefore, we employ our main specification with an interaction for above-median age, above-median employment, and a software-based technology, characteristics that are standard metrics for identifying both startups and financially constrained firms in the literature (Fazzari et al. 1988, Clementi and Hopenhayn 2006, Álvarez and Crespi 2015).

The results are in Appendix Table A.7. The first three columns of each panel limit the sample to Open, and the last three columns to Conventional. Of the 18 possible interactions, 16 are insignificant and one is significant at the 10% level.[25] We also did not find meaningful heterogeneity in the award amount, which we would expect under a financial constraint mechanism.[26]   Overall, this exercise indicates a notable absence of interaction effects,

---

[24]The sample is less than two times the number of firms because a few firms applied once to Conventional before this sample begins in 2016.

[25]The significant result is for high age for VC in Open competitions in column (2) of Panel A. This could suggest that entrants are important for the effect on VC, as old firms do not typically attract VC investors.

[26]Using one Open round in which the size of the award was increased to $75,000 from $50,000, we assess whether the effects differ by award amount. We do not find meaningful or significant differences (not

implying that Open's success is not because of the larger treatment effects of financially constrained firms.

A final firm characteristic we consider is quality, as measured by the evaluator ranks. It may be, for example, that non-marginal conventional winners were much better than the non-marginal Open winners. To test this, we interact winning with being in the right-tail of the rank distribution, defined as the top 15% of winner scores within topic. In addition to addressing selection on the quality distribution, this test offers a robustness test of our main results and supports our argument that our findings are not strictly limited to the region around the cutoff. The results, reported in Appendix Table A.8, show that there are no significantly different effects in the right tail. We observe similar results using thresholds besides 15%.

Summary. This section has shown that firm selection does not explain Open's positive effects. In other words, if Conventional were just to restrict itself to firms with "startup-like" characteristics, it would be unlikely to have the same large positive effects as Open.

7 Implications for Innovation Procurement

The findings from the previous section point to benefits from the move toward the Open program. In this section, we first use our estimates to quantify the benefits of program reform with a simple model, before exploring some of the broader qualitative conclusions for counterfactual public procurement policies.

7.1 Quantifying the Benefits of Program Reform

Motivating Set-Up. Consider the Air Force's marginal decision about whether the next SBIR competition should (i) be an Open topic, (ii) be a Conventional topic or, (iii) not be run at all. For simplicity, we model the competition as having five contestants and one winner.[27] We assume that the Air Force has the following objective function, V, from running this hypothetical competition:

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

reported).

[27]This is without loss of generality as having more winners and contestants essentially just scales the aggregate losses up or down without changing the relative benefits between the options.

Here, M represents the benefits from military activities, N represents the benefits from non-military activities, and μ^M and μ^N represent the Air Force's utility weights on military and non-military activities, respectively. C is the costs of the competition, composed of the Phase 1 and 2 award amounts and administrative costs. Note that Air Force places value on non-military industrial development, so $\mu^N > 0$.[28]

We model the military benefits $M(D, u^M)$ as composed of an observable, D, the dollar value of future technology adoption (non-SBIR DoD contracts) and a set of unobservables u^M; with $\partial M/\partial D \geq 0, \partial M/\partial u^M \geq 0$. This is motivated by the fact that a key goal for DoD of the SBIR is for SBIR awardees to transition into a "program of record," i.e. a future non-SBIR contract with the DoD, 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 (linked more to economy-wide benefits), and an unobservable term, u^N; with $\partial N/\partial VC \geq 0, \partial N/\partial u^N \geq 0$.

Cost-Benefit Results. We assign quantitative values to these objects in Table 6, Panel A. To simplify our baseline analysis, we assume $M(D, u^M) = D + u^M$ and $N(VC, u^N) = VC + u^N$. Panel A of Table 6 reports our baseline calculations. The first two columns have Open outcomes for VC investments (column (1)) and technology adoption (column (2)). The next two columns have Conventional outcomes for VC (column (3)) and technology adoption (column (4)). We take our preferred estimates of the treatment effects in Row 1 (columns (1) and (2) of Table 2). We assume that winning an award changes the winner's probability of obtaining VC or technology adoption 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.[29]

We multiply this treatment effect by the average amount of subsequent VC investment (or technology adoption as measured by non-SBIR DoD 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 (about \$10 million) or contracted (about

---

[28]An alternative interpretation of Equation 3 is that V is society's welfare in which case the Social Planner may have a relatively higher weight, μ^N, on non-military benefits.

[29]This is a strong assumption as we are assuming that the local average treatment effect identified by the RDD is the same across all winners of an SBIR contract. It could be that the infra-marginal winners have different treatment effects. While we cannot definitively rule it out, in general, we find that the results are generally similar when we use the whole sample or a narrow bandwidth, and similar with and without rank controls, suggesting they are likely to apply to the whole sample.

$16 million) across all applicants in the 24 months post-award for the 2017-19 main sample in Row 2. We hold these constant across Open and Conventional awards so that differences will be driven by the estimated causal effect, rather than mean contract values. However, the conclusions are robust to using the program-specific means (Panel B, Row 2). 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 \$750,000 for Open and -\$6,000 for Conventional (Row 5).

Next, we consider costs (C), which are composed of the award amount and administration costs.[30] Like the rest of this analysis, the costs are based on the 2017-19 sample and are estimated by Air Force program administrators. A Phase 1 Open award is \$50,000, while a Phase 1 Conventional award is \$150,000. 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.[31] 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. Finally, the Air Force estimates the cost for administering a Phase 2 competition as about three times those of Phase 1, which are \$3,000 per winner and \$1,500 per loser. Using these inputs, Row 6 shows that the total cost in thousands of dollars for Open is $C^O = 50 + 3 + 0.39\,(4.5 + 832) = 379$, while the total cost for Conventional is $C^C = 153 + 3 + 0.22\,(4.5 + 814) = 333$.

Finally, in Row 7, we calculate the net benefit V to be \$371,000 for Open and -\$339,000 for Conventional. Therefore, Open has a \$710,000 higher benefit than Conventional, reflecting its much larger causal impact with only slightly higher costs (Row 8). In Appendix F, we discuss three extensions. First, we conduct a range of robustness tests of Table 6, such as altering the welfare weights, using alternative treatment effects, etc. Second, we allow for *dynamic* effects as winning a Conventional award causally increases the chances of winning a future SBIR award. Given the smaller treatment effects for such incumbents shown in Table 5, this further

---

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

[31] The average Open Phase 2 is \$832,000, while the average Conventional Phase 2 is \$814,000 (see Table A.2). 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 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 Phase 1 Open winners, 39% won Phase 2 competitions compared to only 22% amongst Conventional Phase 1 winners. We use these percentages in our cost analysis.

reduces the benefits of Conventional, although when we calibrate this to our estimates, the magnitude of this correction is small. Third, we address the problem that the $M(.)$ function may not be monotonic in D, as the price of a non-SBIR DoD contract reflects more than just the value of the innovation.

In sum, although there are many caveats around this coarse quantification, it seems that there are much larger benefits to the Air Force from running a marginal competition as an Open topic rather than as a Conventional topic.

7.2   Some Implications for Innovation Procurement Reform

At a high level, our results suggests that there are benefits to moving towards more open and less tightly specified public procurement of innovation. Open is an extreme example of a radical reform. There are three obvious more incremental means. First, government could use more Requests For Information (RFI) in advance of Requests For Procurement (RFPs). This might improve the information flow. Second, we have shown that the benefits of technology adoption spill over to other parts of the DoD from the Air Force (as well as to the private sector), even though it is the Air Force who solely make the decisions over project funding. One reform might be to bring in other parts of the DoD in making evaluation decisions (although this might increase administrative burden) to internalize some of these spillovers. In the context of Equation 3, this would increase the weight on non-Air Force utility. Third, if the problem was essentially the wrong mix of firms (e.g. too few start-ups) then rather than the radical move to Open, a simpler option would be more set-asides/subsidies for start-ups, etc. The evidence from Section 6, however, showed that this would be insufficient, as the benefits of Open are not only due to a different composition of applicant firms.

Finally, there are policy implications that emerge from the contrast between our average null results for the Conventional program and the positive effects of U.S. Department of Energy (DoE) SBIR grants in Howell (2017). Specifically, there are two possible reasons for the difference that are relevant to innovation procurement design. First, the DoE SBIR program's topics appear to be substantially broader than the Air Force's Conventional topics.[32] In other words, DoE's SBIR program seems to compare more closely to the non-specific Conventional

---

[32]While it is not possible to verify this systematically, our perusal of the two programs suggests it is the case. Examples of DoE topics from 2022 include "Solar Hardware and Software Technologies: Affordability, Reliability, Performance, and Manufacturing", "Advanced Subsurface Energy Technologies" (i.e. geothermal), and "Rare Earth Elements and Critical Minerals" (i.e., making advanced metal alloys for a wide range of uses, such as in batteries and catalysts).

topics where we do find positive effects. Second, Howell (2017) finds that the positive effects of DoE's SBIR program are driven by new entrants. Above, we showed that DoD's SBIR program has more repeat awardees than the DoE program (Figure A.3 Panel B). This likely relates to DoD's massive procurement capability and much larger SBIR program, which allow firms to focus solely on the defense market.

8 Conclusion

U.S. defense R&D is often regarded as an exemplar of how to stimulate innovation through mission-driven research, but (as we have shown) the luster has faded in recent decades, with the prime defense contractors becoming less innovative than the rest of the U.S. economy on several dimensions. One response of the U.S. Air Force to this concern was to introduce an Open innovation dimension to its SBIR program, departing from the tightly specified topics of its Conventional program. Although though the SBIR program differs from mainline procurement, we have documented parallel problems of declining innovation and lock-in of repeat contractors.

We show that the Open program succeeded in its objective of attracting more high-tech startups to the defense market. Our primary outcomes are proxies for civilian innovation benefits (VC investment and patenting) and military benefits of technology adoption (winning future non-SBIR DoD contracts). Using a regression discontinuity design, we find that winning an Open topic award has positive effects on these outcomes, 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. Open's success was *not* simply due to attracting a different composition of firms. Using three alternative designs, we find openness *per se* matters: (i) among the conventional competitions, those topics which were less specific generated more innovation; (ii) although two alternative reformed SBIR programs also attracted applicants with similar characteristics to those applying to Open competitions, they had no causal effects; and (iii) among the sub-set of firms who applied to both Open and Conventional competitions, we again find beneficial effects only for the Open competitions.

The success of Open relates to the broad question of how to procure defense innovation. The main mechanisms are government contracting with private organizations, design competitions, and in-house R&D at government laboratories. In recent decades, the U.S. has emphasized the

30

first channel, but before World War II and in certain parts of the defense establishment such as SBIR, the second channel of design competitions has been important (Mansfield 1971, Jacobsen 2015). 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. Competitions can enable more government flexibility and encourage contractor risk-taking (Lichtenberg 1984). However, a downside 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 in the public sector.

Beyond these points, our context is important because the U.S. DoD funds more R&D than any other single entity in the world. 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. Even beyond the public sector, our findings relate to efforts at open innovation in the private sector, ranging from the large pharmaceutical companies sourcing innovation from biotech startups (Schuhmacher et al. 2013) to LEGO Ideas, which has led to 30 LEGO model kits based on externally submitted ideas.[33] An important avenue for future work is whether causal evaluations of other open R&D programs reveal similar patterns to those found here.

---

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

References

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

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde, "Back to basics: Basic research spillovers, innovation policy, and growth," *The Review of Economic Studies*, 2021, *88* (1), 1–43.

Álvarez, Roberto and Gustavo A Crespi, "Heterogeneous effects of financial constraints on innovation: Evidence from Chile," *Science and Public Policy*, 2015, *42* (5), 711–724.

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

Atkeson, Andrew and Ariel Burstein, "Aggregate implications of innovation policy," *Journal of Political Economy*, 2019, *127* (6), 2625–2683.

Azoulay, Pierre, 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," *The Quarterly Journal of Economics, forthcoming*, 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," *Econometrica*, 2021, *89* (5), 2189–2224.

Bloom, Nicholas, 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.

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.

Che, Yeon-Koo, Elisabetta Iossa, and Patrick Rey, "Prizes versus contracts as incentives for innovation," *The Review of Economic Studies*, 2021, *88* (5), 2149–2178.

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

Clementi, Gian Luca and Hugo A Hopenhayn, "A theory of financing constraints and firm dynamics," *The Quarterly Journal of Economics*, 2006, *121* (1), 229–265.

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.

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," *The Journal of Finance, forthcoming*, 2020.

Dial, Jay and Kevin J Murphy, "Incentives, downsizing, and value creation at General Dynamics," *Journal of Financial Economics*, 1995, *37* (3), 261–314.

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.

Fazzari, Steven M, R Glenn Hubbard, Bruce C Petersen et al., "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, 1988, *19* (1), 141–206.

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

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.

Garamone, Jim, "Austin Tells Reagan Forum How U.S. Will Take on Challenge of China," Technical Report, Department of Defense 2021.

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.

__ **and Ben Jones**, *Innovation and Public Policy*, Univiversity of Chicago, 2022.

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.

Halac, Marina, Navin Kartik, and Qingmin Liu, "Contests for experimentation," *Journal of Political Economy*, 2017, *125* (5), 1523–1569.

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.

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.

Jensen, Michael C, "The modern industrial revolution, exit, and the failure of internal control systems," *the Journal of Finance*, 1993, *48* (3), 831–880.

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.

Lundquist, Jerrold T, "Shrinking fast and smart in the defense industry.," *Harvard Business Review*, 1992, *70* (6), 74–85.

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

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

Mazzucato, Mariana 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, 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.

Myers, Kyle, "The elasticity of science," *American Economic Journal: Applied Economics*, 2020, *12* (4), 103–34.

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.

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.

Rathje, Jason 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.

SBIR.gov, "Qualcomm Inducted into SBIR Hall of Fame," Technical Report, https://www.sbir.gov/success-story/qualcomm-inducted-sbir-hall-fame 2011.

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.

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
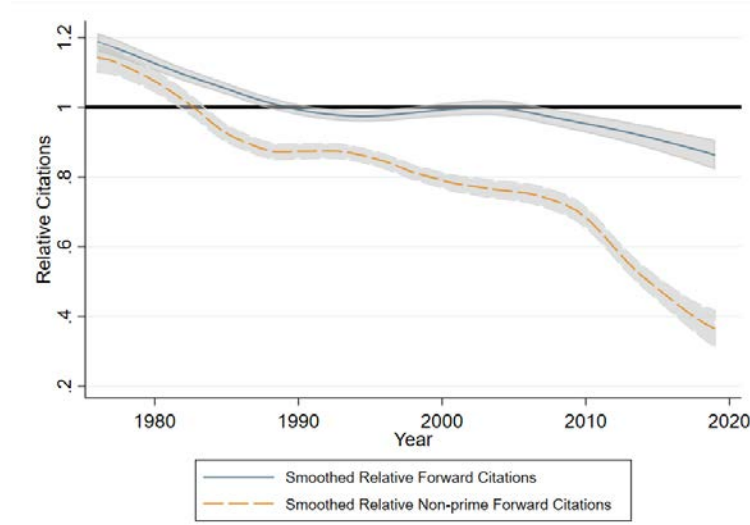

(b) Conventional SBIR Winner Patent Citations
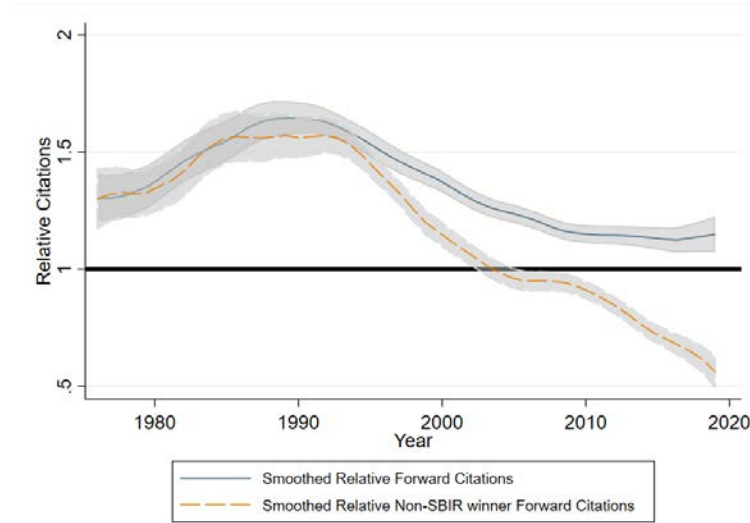
Note: These graphs describe patent quality for the prime defense contractors and their acquisition targets (depicted in Figure A.1). That is, 226 firms are included in 1976, while only six are included in 2019 (as the 226 have merged into these six). Panel A shows the total number of forward citations (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

(a) Selection into Programs
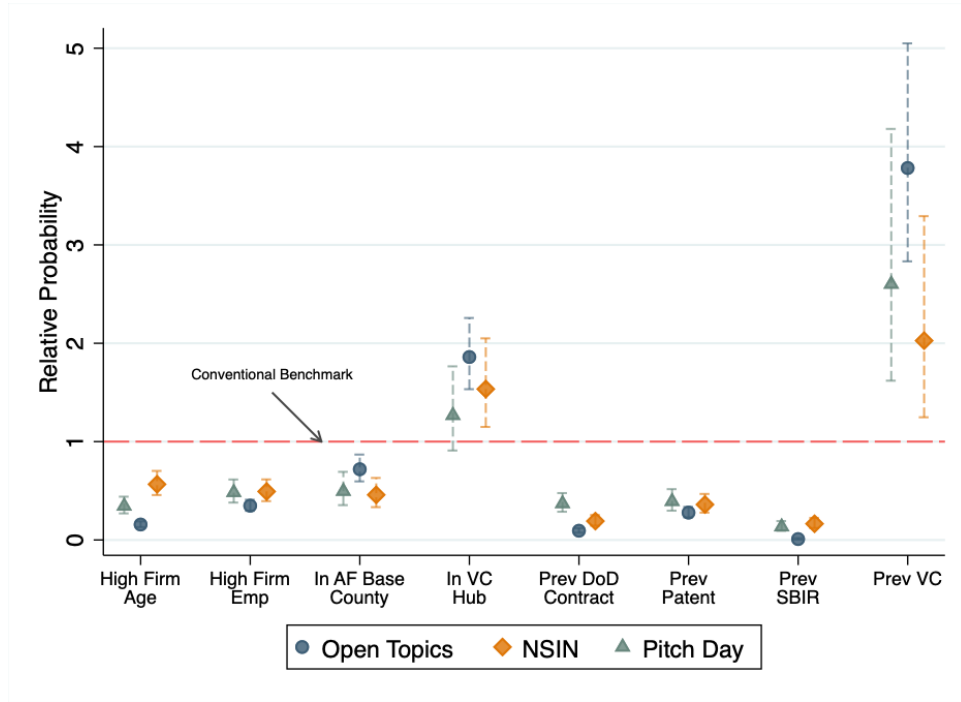
(b) Conventional Divided by Specificity
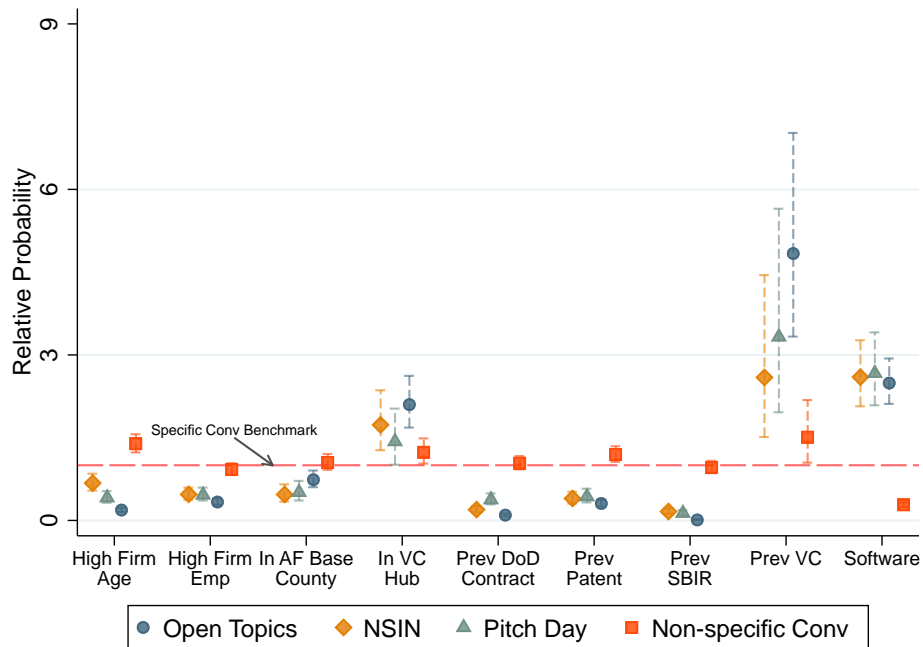
Note: Panel A in this figure shows how firm characteristics predict selecting into reform programs compared to the Conventional program. The points represent coefficients from a multinomial logistic regression with categories for four programs: the Conventional program, which serves as the base group and is represented by the red dashed line, and the three reform programs (Open, NSIN, and Pitch Day). Panel B in this figure repeats Panel A but sets the baseline as applicants to Conventional topics with above-median specificity, and then adds applicants to below-median specificity topics as an additional group. Data restricted to 2017-2019. The dashed lines around each coefficient point indicate the 90% confidence intervals.

Figure 3: Raw Scores and Award Probability in Four Representative Topics
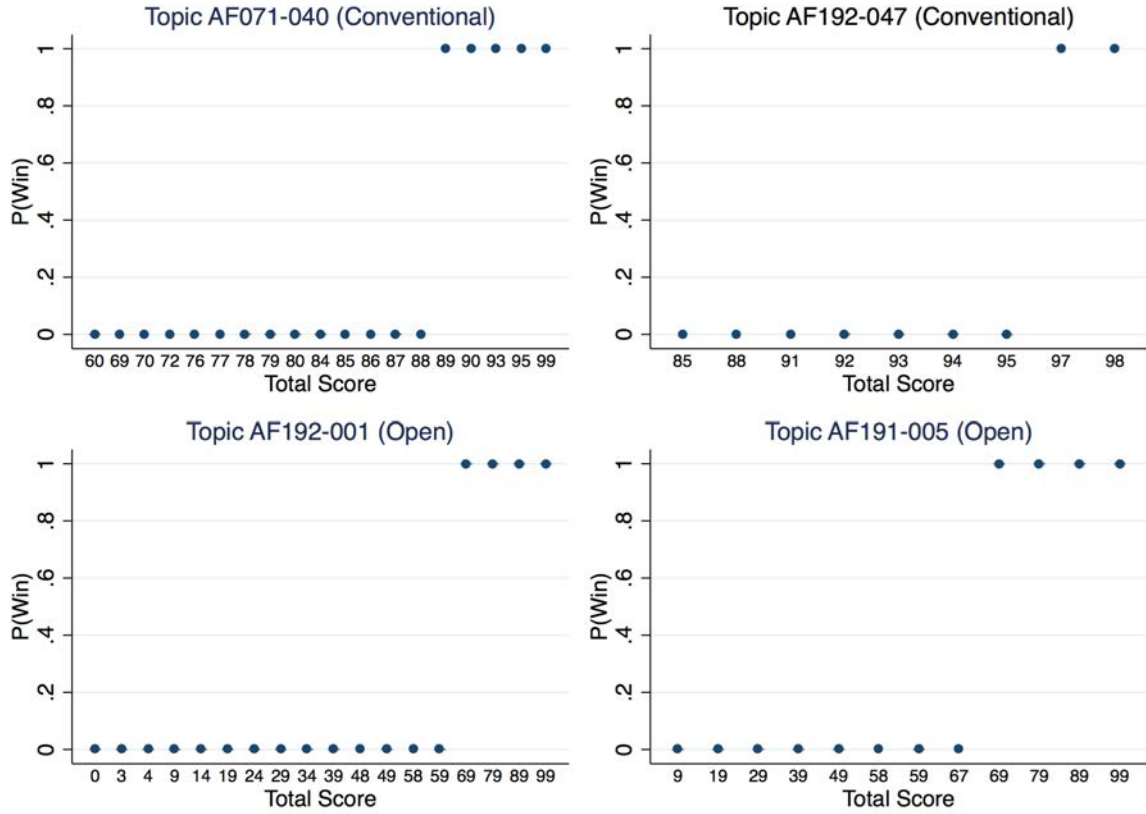
Note: These plots document the sharp RDD in each topic by showing the probability of winning by raw score. The score perfectly predicts award except occasionally when an awardee is declined in the contracting process because some ineligibility was identified (these instances are dropped in analysis). Note that the range of scores differs across topics, which is we construct a rank normalization for combined analysis.

Figure 4: Regression Discontinuity Density Manipulation Test
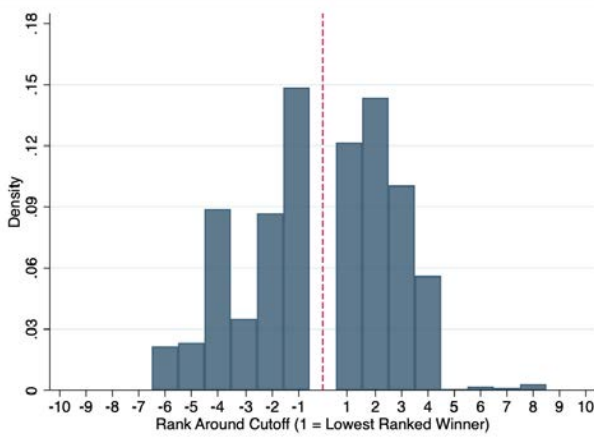
(a) Open   (b) Conventional (2017-19)
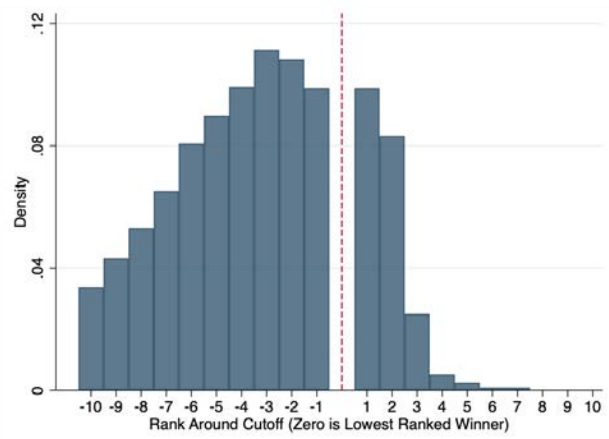
Note: This figure plots the density of applicants by rank around the cutoff using Phase 1 applicants to the Open (left graph labeled (a)) and Conventional (right graph labeled (b)) programs, to test for bunching near the cutoff. There is more density overall to the left of the cutoff because there are more losers than winners.

Figure 5: Probability of Venture Capital by Rank Around Cutoff

(a) Open

(b) Conventional (2017-19)
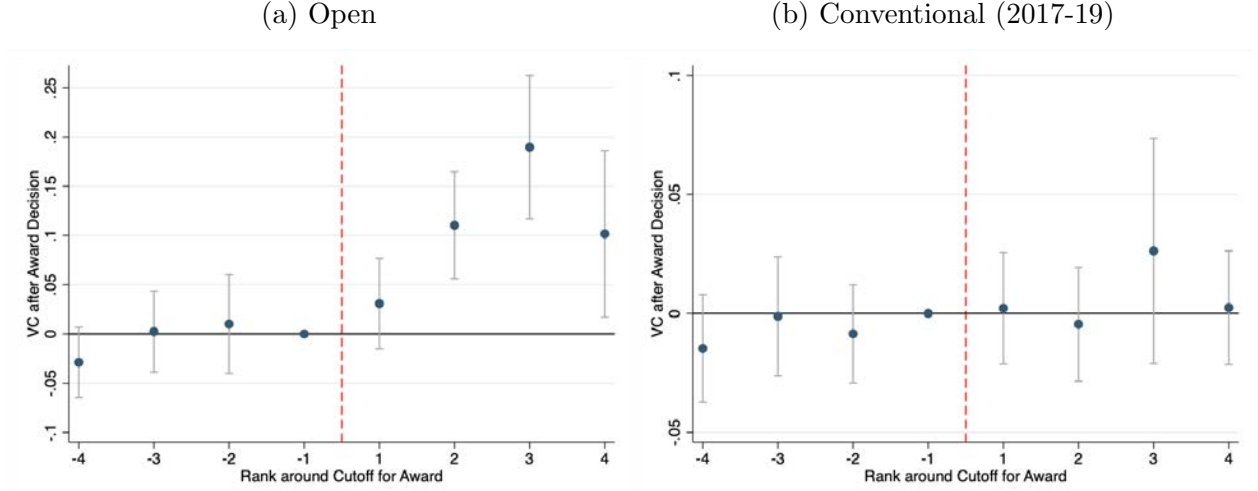
Note: These figures show the probability that an applicant firm raised venture capital investment (VC) within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, 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 6: Probability of Technology Adoption (DoD non-SBIR Contract) by Rank Around Cutoff

(a) Open

(b) Conventional (2017-19)
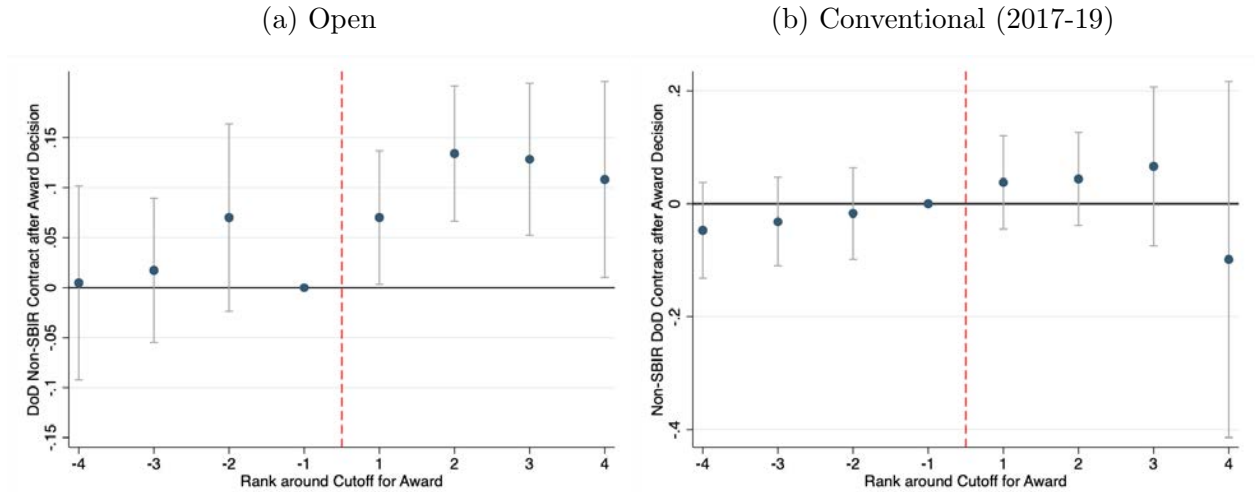
Note: These figures show the probability that an applicant firm had any non-SBIR DoD contracts valued at more than $50,000 within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the 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 7: Probability of Patents by Rank Around Cutoff

(a) Open

(b) Conventional (2017-19)
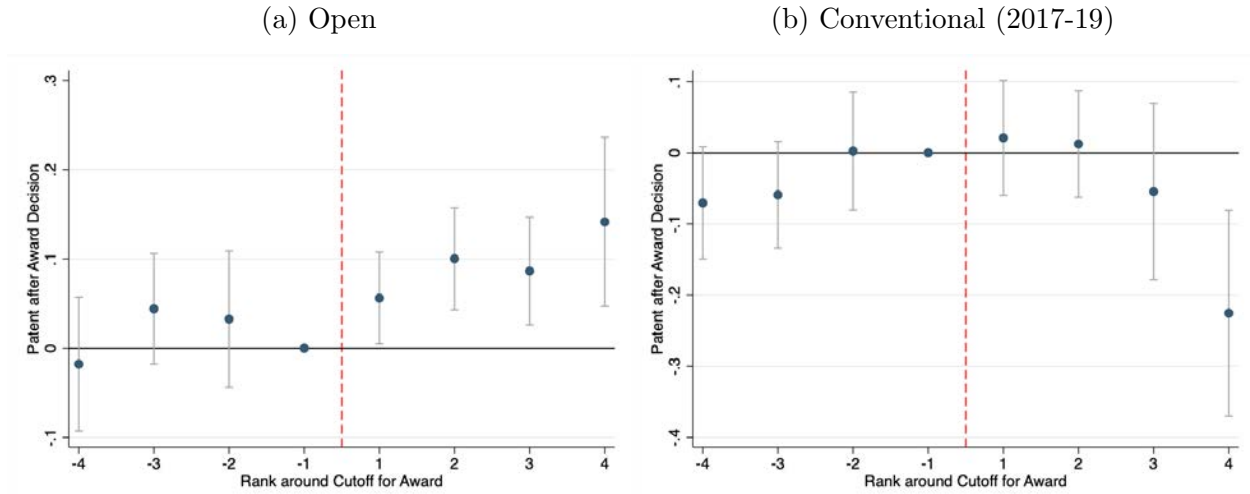
Note: These figures show the probability that an applicant firm had any ultimately granted patent applications within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score 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 8: Probability of Air Force SBIR Contract by Rank Around Cutoff

(a) Open

(b) Conventional
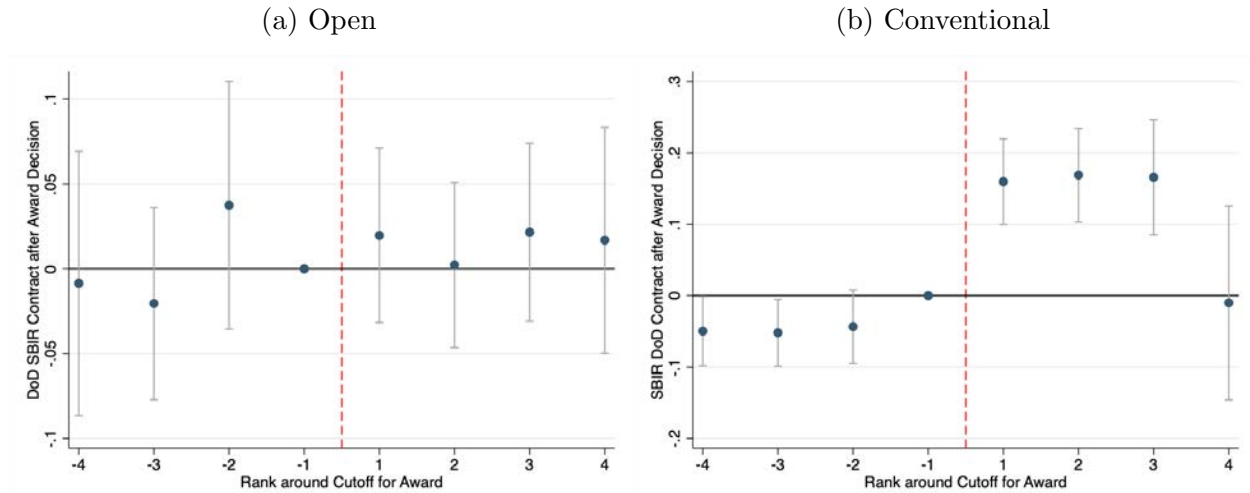
Note: These figures show the probability that an applicant firm had any Air Force SBIR contracts within 24 months after the award decision. In both panels, the x-axis shows the applicant's rank around the cutoff for an award. A rank of 1 indicates that the applicant had the lowest score among winners, 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.

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	$ 147,237	$ 152,718	$ 25,297
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

Panel B: Pre-Award Outcome Summary

	Open Topic				Conventional			
	N	Mean	Median	SD	N	Mean	Median	SD
$\mathbb{1}$(VC)	1,659	0.101		0.302	4,995	0.060		0.237
Avg VC Amt (Mill)	168	$ 5.357	$ .5594	$ 9.624	299	$ 2.234	$ .0277	$ 6.017
$\mathbb{1}$(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	20.177	9	29.736
Avg DoD Non-SBIR Contract Amt (Mill)	420	$ 1.286	$ .698	$ 1.598	3,000	$ 1.351	$ .8652	$ 1.580
$\mathbb{1}$(Patent)	1,659	0.250		0.433	4,995	0.473		0.499
# Patents	415	12.313	3	39.420	2,364	26.685	10	45.638
# Patent Application if Any	515	10.996	3	36.026	2,554	25.701	9	45.462
$\mathbb{1}$(AF SBIR Contract)	1,659	0.188		0.391	4,995	0.584		0.493
# AF SBIR Contracts	312	16.679	7	28.961	2,915	37.837	14	55.299
$\mathbb{1}$(Never Awarded SBIR)	1,659	0.573		0.495	4,995	0.277		0.448

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)	0.060**	-0.004	0.009	-0.004	0.009	0.005
	(0.024)	(0.020)	(0.010)	(0.020)	(0.010)	(0.003)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				0.065*	0.051*	0.052*
				(0.034)	(0.030)	(0.027)
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	0.087	0.016	0.016	0.041	0.027	0.017

Panel B: Any Subsequent Non-SBIR DoD Contracts

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	0.098***	0.029	0.046	0.029	0.046	-0.014
	(0.037)	(0.053)	(0.031)	(0.051)	(0.031)	(0.014)
$\mathbb{1}$(Award) × $\mathbb{1}$(Open Topic)				0.069	0.052	0.111**
				(0.070)	(0.059)	(0.045)
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	0.199	0.329	0.231	0.284	0.226	0.427

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 technology adoption measured by 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)	0.090***	0.057	-0.004	0.057	-0.004	-0.020
	(0.029)	(0.055)	(0.026)	(0.054)	(0.025)	(0.014)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open Topic)				0.033	0.094**	0.128***
				(0.067)	(0.048)	(0.038)
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	0.117	0.194	0.166	0.167	0.159	0.274

Panel B: Any Subsequent SBIR Contracts

	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$(Award)	-0.003	0.140***	0.194***	0.140***	0.194***	0.050***
	(0.029)	(0.050)	(0.028)	(0.049)	(0.028)	(0.012)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open Topic)				-0.143***	-0.197***	-0.065***
				(0.050)	(0.030)	(0.015)
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	0.098	0.250	0.138	0.197	0.132	0.370

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) 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 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)
$\mathbb{1}$(Award)	0.060**	-0.005	0.080***	0.041	0.081**	0.024	-0.008	0.131***
	(0.024)	(0.018)	(0.028)	(0.042)	(0.036)	(0.049)	(0.028)	(0.046)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.087	0.016	0.117	0.194	0.199	0.329	0.098	0.250

Panel B: Narrow Bandwidth

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(Award)	0.064***	0.003	0.037***	0.015	0.072***	0.050	-0.003	0.076**
	(0.020)	(0.010)	(0.013)	(0.031)	(0.027)	(0.034)	(0.016)	(0.032)
Observations	671	902	671	902	671	902	671	902
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.064	0.017	0.021	0.216	0.153	0.434	0.062	0.409

Panel C: All Proposals

Dep Var:	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(Award)	0.056**	0.001	0.108***	0.033	0.097***	0.003	-0.016	0.088**
	(0.023)	(0.011)	(0.027)	(0.038)	(0.035)	(0.035)	(0.029)	(0.038)
Observations	1659	4995	1659	4995	1659	4995	1659	4995
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.087	0.012	0.125	0.270	0.216	0.476	0.113	0.416

Panel D: 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)	0.144***	0.002	0.087***	0.065	0.116***	-0.017	-0.006	0.082
	(0.031)	(0.026)	(0.030)	(0.047)	(0.041)	(0.048)	(0.029)	(0.050)
Observations	1385	2608	1385	2608	1385	2608	1385	2608
Program	Open	Conv	Open	Conv	Open	Conv	Open	Conv
Outcome Mean	0.152	0.039	0.123	0.229	0.245	0.404	0.102	0.285

Panel E: Conventional 2003-2017

Dep Var:	Any VC	Any Patents	Any DoD Contracts	Any SBIR
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	0.012	-0.018	0.022	0.117***
	(0.008)	(0.023)	(0.025)	(0.022)
Observations	6628	6628	6628	6628
Outcome Mean	0.015	0.171	0.239	0.167

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 restrict the bandwidth to include only two applicants on each side of the cutoff. In Panel C, we include all proposals in the 2017-19 period rather than only a firm's first (so a firm may appear twice). In Panel D, we allow outcomes to be ever-after rather than within two years. In Panel E, 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-D, 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: 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)	-0.039	-0.028	-0.059**	-0.044**	-0.046**	-0.037***
	(0.027)	(0.020)	(0.024)	(0.018)	(0.022)	(0.014)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$($> P50$ Non-specificity)	0.066**	0.043*	0.067**	0.047**	0.056**	0.042***
	(0.032)	(0.023)	(0.030)	(0.020)	(0.028)	(0.016)
Observations	7384	17500	7384	17500	7384	17500
Proposals	First	All	First	All	First	All
Outcome Mean	0.166	0.287	0.131	0.202	0.109	0.198

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

	Bottom Tercile Conventional Nonspecificity				Top Tercile Nonspecificity			
	Any Patent		Any High Originality Patent		Any Patent		Any High Originality Patent	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(Award)	-0.038	-0.028	-0.045	-0.045**	0.081*	0.060*	0.051	0.033
	(0.049)	(0.028)	(0.035)	(0.023)	(0.049)	(0.031)	(0.044)	(0.028)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open)	0.129**	0.136***	0.101**	0.107***	0.009	0.048	0.005	0.029
	(0.064)	(0.046)	(0.044)	(0.031)	(0.063)	(0.047)	(0.052)	(0.035)
Observations	4038	7722	4038	7722	3357	6197	3357	6197
Proposals	First	All	First	All	First	All	First	All
Outcome Mean	0.131	0.225	0.079	0.142	0.162	0.264	0.106	0.185

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, where the effect of winning is modulated by the index of topic non-specificity. A higher value of non-specificity means the topic is more open based on the diversity of proposals it attracted (see Section 6 for details). In Panel A, we interact winning with the demeaned "non-specificity" index within the Conventional program sample. 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 compare the effects of winning in Open vs. Conventional, where the independent coefficient on Award represents the effect of Conventional. Columns (1)-(4) restrict the Conventional sample to topics with bottom-tercile non-specificity, while columns (5)-(8) restrict the sample to topics with top-tercile non-specificity (which are more like Open 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 in Panel A. We include all years (2003-19). Standard errors are under coefficients (in parentheses) and clustered by topic. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6: Cost-Benefit Analysis

All $ in thousands

A. Baseline Results

Program: Outcome:	Open VC (1)	Open Tech Adoption (2)	Conv VC (3)	Conv Tech Adoption (4)
1. Treatment Effect	0.060	0.098	-0.004	0.029
2. Average Contract Size	$9,921	$1,577	$9,921	$1,577
3. Implied Benefit	$595	$155	-$40	$46
4. Utility Weight	1	1	1	1
5. Benefit Sum	$750		-$6	
6. Cost	$379		$333	
7. Net Benefit (Benefit - Cost)	$371		-$339	
8. Net Benefit Difference (Open - Conv)	$710			

B. Robustness

	Net Benefit Difference (1)	Open Net Benefit (2)	Conv Net Benefit (3)
1. Baseline	$710	$371	-$339
2. Pre-award Average Contract Sizes	$429	$114	-$315
3. Zero Weight on Non-Military	$63	-$224	-$287
4. Zero Coefficient on VC for Conv	$666	$379	-$287
5. Conv Coefficients from 2003-19	$550	$379	-$171
6. Separate Contract Size Averages for Open and Conv	$806	$484	-$322

Note: This table shows the cost benefit analysis using the regression discontinuity (RD) estimates of the treatment effect in Table 2. Panel A shows the baseline scenario, where we consider monetary benefits within 24 months after the award decision, specifically the average VC funding amount and technology adoption measured as the non-SBIR DoD contract size. The cost of the programs consists 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.

Appendix

(For Online Publication)

A Slowing Innovation in the US Defense Industry

In this appendix, we describe some of the economic context for U.S. military R&D and explain the concerns among policymakers about declining innovativeness, which motivated the SBIR reform we study. Despite these concerns, there is no public evidence about the decline. Therefore, in the second part of the appendix, we document the evolution of prime defense contractors' innovation.

A.1 Economic Context for U.S. Defense R&D

Military R&D has shaped technological advances since antiquity, both "pushing" and "pulling" civilian innovation.[34] 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 (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. 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

---

[34]For example, many historians (e.g. Polybius' *Histories)* credit Archimedes with 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).

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.

Much more so than other Western countries, the U.S. procures 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 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. 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. 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.[35]

[35]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.

A.2 Documenting Declining Innovation

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%.[36] 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.3 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

---

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

one indicates the patent is more impactful than the average patent in its class-year.[37] The 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.17 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.17 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.2 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.[38] The subsequent three panels use Compustat data and compare primes to other firms in the same three-digit NAICS industry.[39] Panel B shows that

---

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

[38]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.

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

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. While these changes clearly increased efficiency, they may also help to explain the decline in innovation we observe in the defense sector and, more directly related to our thesis, have left a remaining cadre of smaller, less innovative, and more locked-in defense contractors.

Finally, our results are consistent with case study evidence. Dial and Murphy (1995) document how one prime, General Dynamics, generated substantial wealth for shareholders despite facing declining demand in the post-Cold War period. After tying executive pay to stock price increases, the firm dramatically increased stock returns through downsizing, including cutting R&D spending in half, and by shifting resources out of the defense industry (p. 262-3, 277). Lundquist (1992) also explains how the defense industry more broadly created value for shareholders by consolidating and reducing overall research investment. Jensen (1993) specifically explains how these acquisitions transferred large sums to target-firm shareholders so that they could reinvest in more productive sectors, outside of defense.

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.

Finally, we conduct a similar analysis for the SBIR program. Since firms must be small to participate, concentration is not a primary concern. 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 Panel B shows that among winners in the Conventional SBIR program, there has been a decline in relative innovation since the 1990s, similar to that for prime contractors (see Appendix A).[40]

This decline may be related to the difference between the findings of this paper and that

---

[40]Furthermore, Figure A.3 panel A uses the Herfindahl-Hirschman Index (HHI) to show that the DoD SBIR program has become more concentrated over time, with more firms winning many awards in a single year.

in Howell (2017) for the Department of Energy (DoE), where there are large positive effects on innovation of winning a Phase 1 grant. One explanation is that there is more severe lock-in of the SBIR firm base at DoD than at DoE. Indeed, we show that this is the case in Figure A.3 Panel B. Each line 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. The greater lock-in at DoD might reflect 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 research firm. The higher incidence of repeat contracts in the SBIR offers a parallel to the consolidation among prime contractors in the larger acquisitions programs.

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.[41] 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 challenges facing military procurement identified in Section **??**. The idea is that if the open 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.[42] 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

---

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

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

commercially focused companies that are not traditionally defense contractors.[43] 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."[44]

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. Military forces 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

---

[43]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.

[44]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.

SBIR program. Open topic solicitations contain no direction regarding the technology 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 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). The fact we find larger causal effects in Open than in Conventional cannot be explained by the difference in award amounts since Open is less financially generous than Conventional. 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. 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.18 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.[45] Phase 2 contracts are much more detailed, bespoke, with higher and more varied amounts than Phase 1 (see Figure A.18).

---

[45]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.

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.4 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.3 shows similar statistics for the whole Conventional sample. Summary statistics for Phase 2 are in Table A.2.

C.2 Outcome Variables

The two main outcomes of interest, in Table 1 Panel B, are subsequent VC investment and technology adoption measured by 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.

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 from the Conventional program is Ascendant Engineering Solutions. After winning a $149,000 Conventional Phase 1 award in September 2016, Ascendant Engineering Solutions has been awarded with a $7.5 million Air Force contract for its tactical gimbals in February 2018.

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.[46] 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).[47] 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.[48] 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.

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-

---

[46]The citations data are from the USPTO.

[47]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).

[48]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.

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.[49] 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.19). 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.5 and A.6. 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

---

[49]OntoNotes is a large corpus consisting of various texts from news, conversational telephone speech, blogs, broadcasts, and talk shows. 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.

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.

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.9 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.[50]

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

In Section 5 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, 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. 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.

---

[50]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.

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.10 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.[51] 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.11, 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. This is 52% of the mean. Comparing it with the main result from Table 2 of 5.2 percentage points (60% 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

---

[51]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.

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.18).

Table A.12 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.2 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 examine the robustness of the welfare calculation sketched in Section 7 in more detail. First, we examine a dynamic extension. Second, we consider how the cost of a defense contract may be non-monotonic in the military benefit function. Third, we consider alternative assumptions of the magnitudes of the elements underlying Table 6. We show that our finding of the higher net benefits for Open is robust across these different extensions.

F.1 A dynamic extension to the welfare effects

We consider a dynamic extension to the model to incorporate negative effects of lock-in and an insular SBIR base. Write the present value of military welfare as:

$$V_t^Z = \pi_t^Z + \beta V_{t+1}^Z \tag{4}$$

where superscript Z denotes whether a competition was Open (O) or Conventional (C), $\beta \leq 1$ is the discount factor and π_t^Z is the Air Force's flow utility net of costs as in Equation (3). The subscript t denotes the ordering of the competitions (which are about every four months).

In the empirical results, we have shown that winning a Conventional (but not an Open) competition increases the probability, p, of winning a future SBIR competition.[52]

Consequently, the value of running a marginal Open competition does not change from the static analysis of Section 7, i.e. $V_t^O = \pi_t^O$. However, since $p > 0$ for Conventional topics, V_{t+1}^C may be empirically relevant for assessing the value of today's decision to run a Conventional competition. In Table A.7 we found that (i) the treatment effects on VC (for Open awards) are significantly smaller for past SBIR winners (Panel A column (1)) and (ii) the treatment effects on technology adoption (for Conventional awards) are significantly smaller for past SBIR winners (Panel B column (4)) . This implies that the value of running a marginal Conventional competition is:[53]

---

[52]We could not statistically reject that the RDD effect of winning a past Conventional competition was different on the chances of winning a future Open vs. a future Conventional competition. Thus, there is no need for superscript Z on the probability, p.

[53]It is theoretically possible that the reason why "success breeds success" in Conventional is due to improved learning of what the military really needs, hence the higher win probability (so $\delta < 0$). However, it is more likely that $\delta > 0$ because (as noted) our heterogeneity analysis in Table A.7 found that previous SBIR winners had *lower* treatment effects.

$$V_t^C = \pi_t^C - \mu^M \delta^M M^C \sum_{k=1}^{\infty} (\beta p)^k - \mu^N \delta^N N^O \sum_{k=1}^{\infty} (\beta p)^k. \tag{5}$$

The term δ^M is the reduction in the value of an incumbent winning a military contract rather than a new entrant winning it, so that the future utility flow is reduced by this term multiplied by the probability that an incumbent will win (p). Hence, in the next $(t+1)$ competition the flow utility is shaded by $p\delta^M$. In competition $t+2$, there is a p^2 probability that a winner in period t will win, so military welfare suffers a further loss of $p^2 \delta^M \mu^M M^C$. This cumulates giving the total utility loss of $\mu^M \delta^M M^C \sum_{k=1}^{\infty} (\beta p)^k$ in Equation (5) for observable military benefits. Similarly, the reduction in observable non-military benefits (Venture Capital) from a past incumbent winning on an Open competition (N^O), is denoted δ^N, generating the last term in Equation (5).Since we assume stationarity, we can simplify Equation (5) to:

$$V_t^C = \pi_t^C - \frac{\beta p}{1 - \beta p} \left(\delta^M \mu^M M^C + \delta^N \mu^N N^O \right). \tag{6}$$

This discussion illustrates a dynamic reduction of the value to the military of running a Conventional topic instead of an Open one. Instead of the full net static benefit of $\pi^C = \mu^M M^C + \mu^N N^C - C^C$, the $\mu^M M^C$ is shaded by $\frac{\beta p \delta^M}{1 - \beta p} < 1$, as well as reducing the value of future Open competitions $(\mu^N N^O)$ by $\frac{\beta p \delta^N}{1 - \beta p} < 1$. Another way of looking at this is to note that although on the observables of VC and technology adoption, Open seems to outperform Conventional, it may be that there are unobservables which still make Conventional competitions favorable on static criteria. Equation (6) shows that even if this were the case (i.e. $\pi^C > \pi^O$), it could still be than Open is preferred by the Planner (i.e. $V^O > V^C$) because of the costs of lock-in as reflected by the second term on the right hand side of Equation (6).

To quantify this potential loss, consider our baseline calibration values where $\mu^M = \mu^N = 1$ and conservatively assume that $\beta = 1$.[54]

$$V_t^C = \pi_t^C - \frac{p}{1 - p} \left(\delta^M M^C + \delta^N N^0 \right) \tag{7}$$

From our econometric estimates, we found that $p \simeq 0.17$, $\delta^N \simeq 0.06$ and $\delta^M \approx 0.05$.[55] It can

---

[54]Making $\beta < 1$, would strengthen the argument below, making the dynamic adjustment even smaller.

[55]These are taken from Table 3 Panel B column (2), Table A.7 Panel A column (1) and Table A.7 Panel B column (4), respectively.

immediately be seen that these dynamic losses are small in magnitude: of the order of about 1% (= 0.05*(0.17/(1 - 0.17)) of the sum of M^C and N^0. This is because the shading of the future losses are only in the order of 5% to 6% and the persistence term is closer to zero (17%) than to unity. Hence, incorporating these calculations into the welfare implications would not materially change the qualitative or quantitative conclusions of Table 6.

F.2   Innovation Value and Production Costs

In the application we have assumed the military value of SBIR is given by the dollar value of the contract:

$$V = \mu^M M + \mu^N N - C = \mu^M (D + u^M) + \mu^N (VC + u^N) - C. \tag{8}$$

This is obviously a crude assumption, and there are several problems with this linearization. First, we may be *over-estimating* military welfare because we are ignoring the real resource cost of production. This is part of the "purchase price" and not a surplus to the Air Force. Second, we may be *underestimating* welfare because the military (and Venture Capitalists) may be willing to pay more than they actually offer. Indeed, by revealed preference their value cannot be less than the purchase price.

To see this, consider Bhattacharya (2021)'s model of the Navy SBIR program. Although his empirical setting is like our Conventional competitions, his structural approach is quite different from ours. For example, (i) In his baseline model he does not consider non-DoD benefits (the value we proxy through VC)[56] and (ii) he considers only the military benefits flowing through the particular SBIR contents from Phase 1, 2 and 3 (the latter being a specific type of the non-SBIR DoD contracts we consider). By contrast, we also allow a Phase 1 win to have a causal effect on (a) using the technological idea to obtain other non-Air Force DoD contracts (b) contracts outside the specific technology supported in Phase 1 (e.g. a firm may win a Phase 1 award and use the resources, reputation and ideas in a complementary project elsewhere[57]).

A strength of the approach in Bhattacharya (2021), however, is that he has an explicit

---

[56]In any extension, he considers a model where there are external benefits. He uses Phase 1 values to help fit these, but calculates that these benefits are small compared to the direct military benefits. By contrast we find that VC funding is in a similar range to military contracts, suggesting large non-military benefits.

[57]This could occur without even without going to a Phase 2 competition, let alone a Phase 3

structural model which enables him to quantify welfare and counterfactual design reform more rigorously. To see this, we follow his baseline model and assume $\mu^N N = 0$ (no external non-military benefits).[58] He assumes the dollar value of the DoD production contract is $D = T_3 = c + \eta(v - c - s)$, where c is the SBIR innovation's cost of production in the military contract, v is the value of the innovation, $0 < \eta < 1$ is the bargaining parameter (the share of the incremental surplus the Air Force allows the firm to keep) and s is $(v - c)^+$, the surplus of the contract winner's closest rival.[59] This illustrates the fact that if production costs were zero ($c = 0$), the price paid by the Air Force (T_3) would be an underestimate of this welfare.

Bhattacharya (2021) focuses on cases where there is only one realistic contender so $s = 0$. In this case the surplus in Open compared to Conventional competitions is:

$$\mu^M[D^0 - D^C] = \mu^M[\eta(v^O - c^O) - (v^C - c^C)] = \mu^M \eta[v^O - v^C - (c^O - c^C)]$$

Hence, if the costs of Open are equal or less to those of Conventional ($c^O \leq c^C$) we will *underestimate* the benefit of Open relative to Conventional using our approach. However, if the cost of Open is greater than Conventional $c^O > c^C$ we may (for sufficiently large η) *overestimate* the benefits of Open. Since we have no reason to believe Open projects are more costly, a natural benchmark is $c^O = c^C$. Given that Bhattacharya (2021) estimates $\eta \approx 0.5$, this means that we should roughly double the difference in the size of military contracts between Open and Conventional, which would *strengthen* our conclusions over the relative benefits of Open. Nonetheless, this discussion highlights some of the issues with using the value of contracts as a measure of benefits rather than also explicitly modeling the production costs of delivery.

F.3 Robustness and Caveats

Panel B of Table 6 contains some robustness tests around these baseline calculations. Column (1) presents the difference between the net benefit from Open (Column (2)) and 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

---

[58] He also assumes that the DoD transfers in Phases 1 and 2 equal the firms' R&D costs (so there is no net R&D research costs to the firms). Hence, administrative costs aside, $C = 0$

[59] Note that there is a hold-up problem because the full value of the surplus is not collected by the firm, so there is under-provision of R&D effort.

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 the Air Force'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 6 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.

F.4   Summary on Welfare

In summary, although the exact net benefits depend on assumptions, this cost-benefit analysis clearly indicates that the military 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
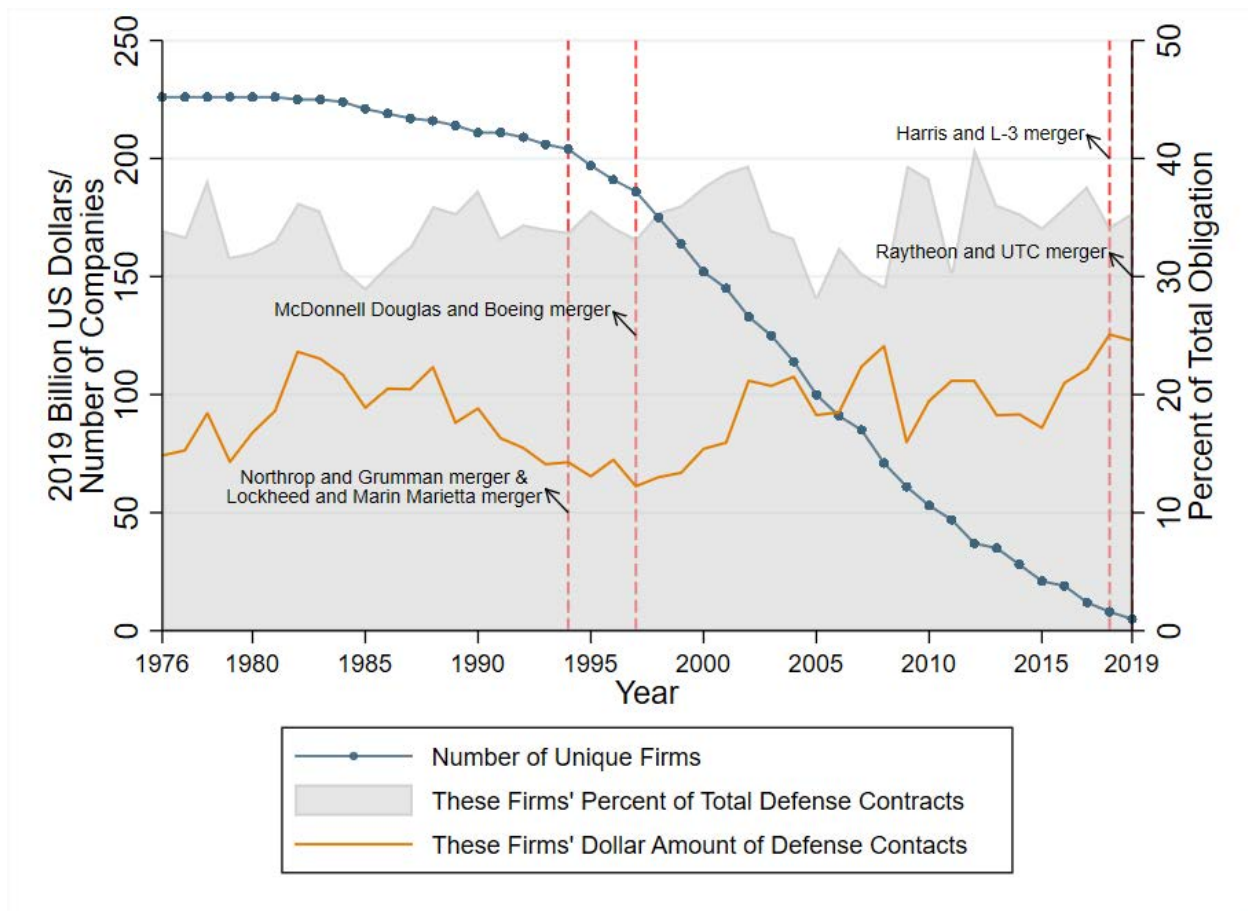
Note: This figure shows the trend of defense contractors' consolidation since the 1980s. We first define prime defense contractors as the top contractors between 2000 and 2020: Boeing, Raytheon, Lockheed Martin, Northrop Grumman, General Dynamics, United Technologies Corp, Harris, and L-3. We then identify 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).

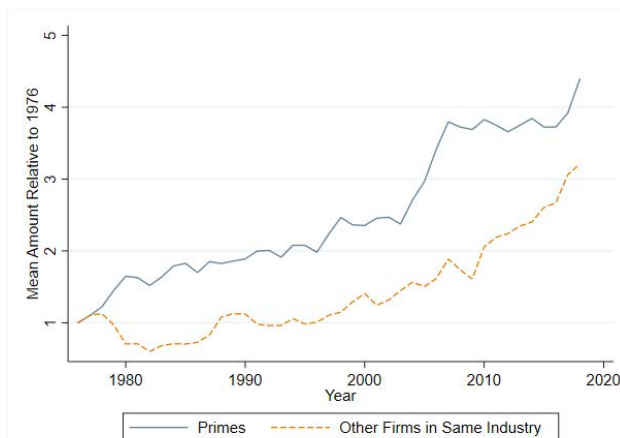## Figure A.2: Historical Dynamics of Prime Defense Contractors
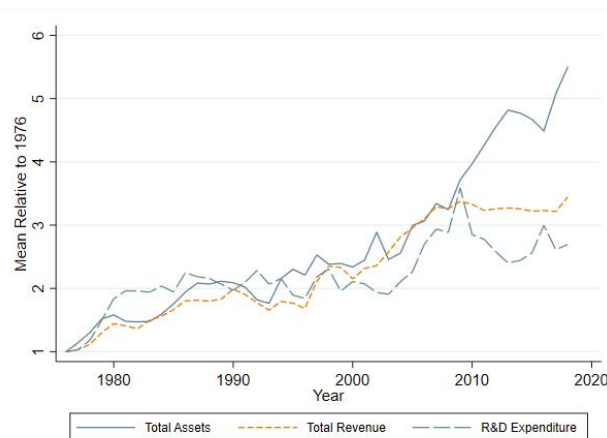
(a) Number of Patents
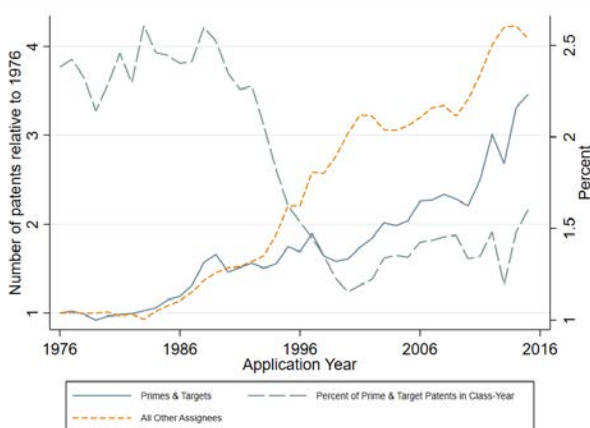
(b) Profit per Dollar R&D

(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.3: Concentration of Federal Contracts

(a) Concentration of Department of Defense SBIR and Non-SBIR Contracts
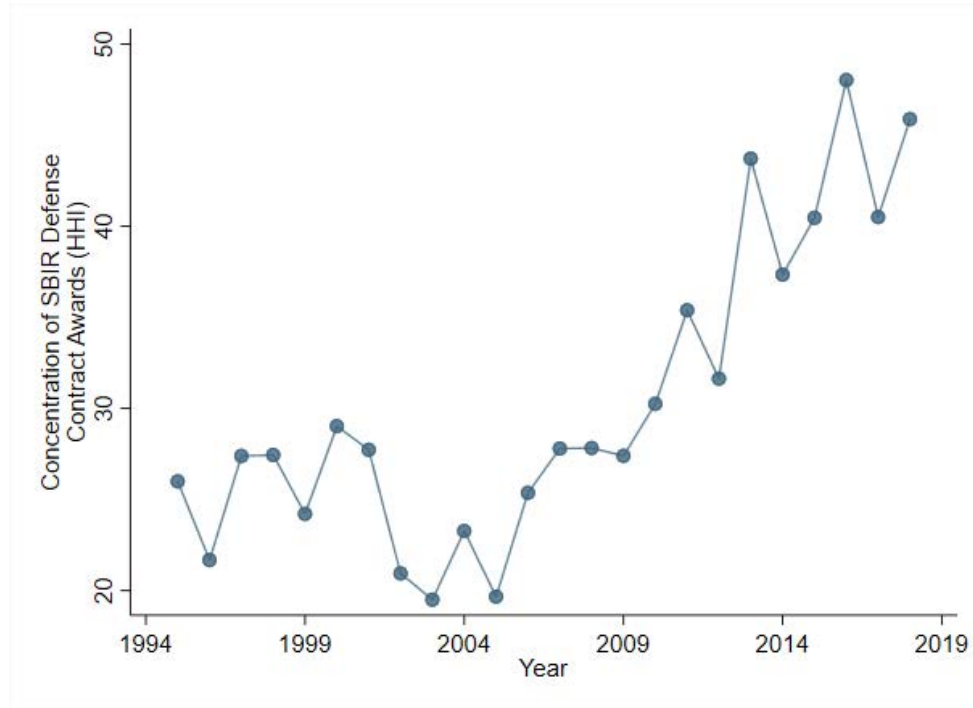
(b) Share of Firms without Recent Repeat Contracts in Two SBIR Programs
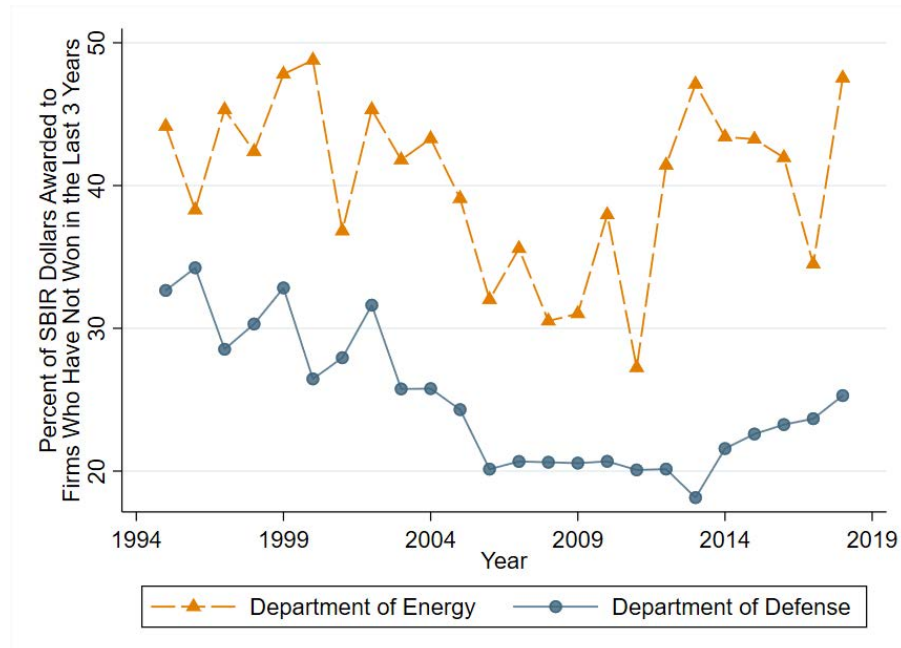
Note: Panel A in this figure shows the Herfindahl–Hirschman Index (0-10,000) for Non-SBIR Department of Defense contracts from 1990 to 2018. Panel B shows the share of "new" firms winning awards from the SBIR programs at the Department of Defense (DoD) and the Department of Energy (DoE). Each line plots 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.4: Number of Applications and Awards Over Time by Topic Type (Analysis Sample)

(a) Applications
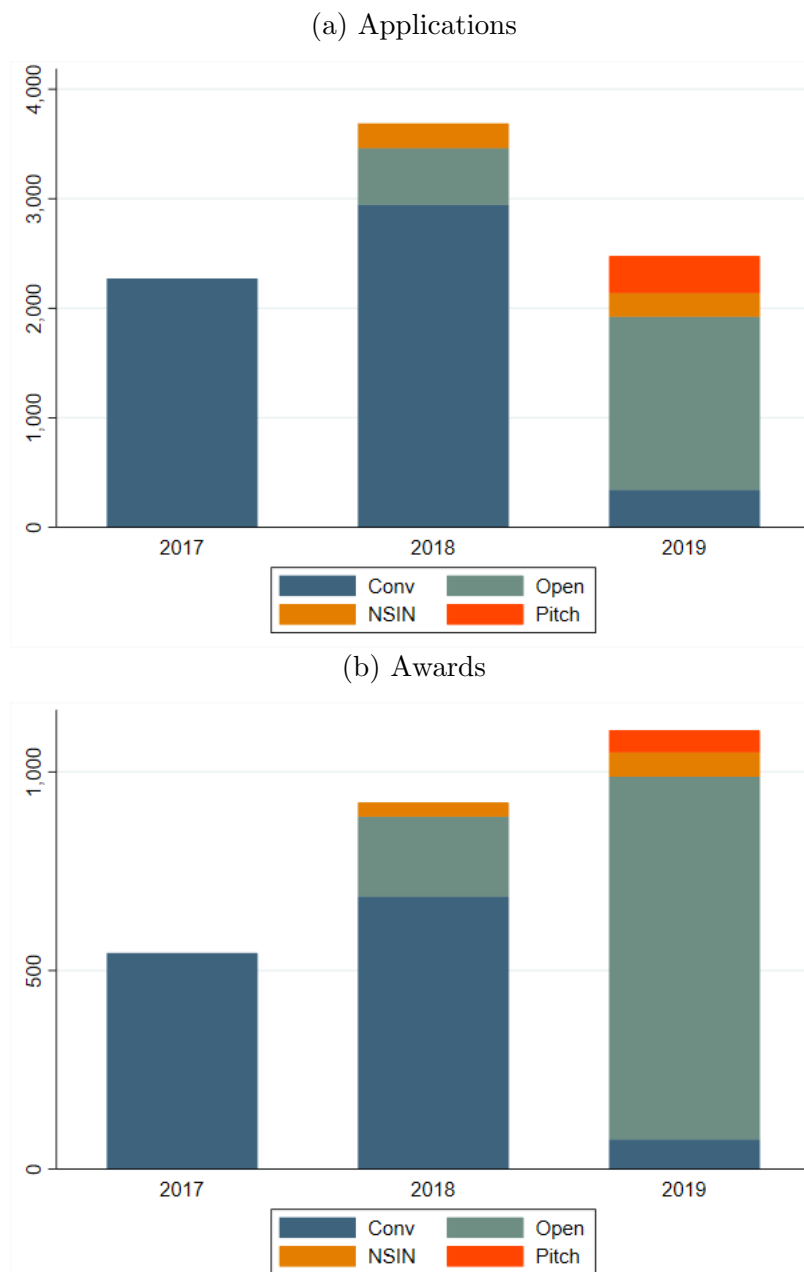
(b) Awards

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.5: Keywords for K-means 2 Cluster Model Scaled by Importance to Topic

(a) Hardware
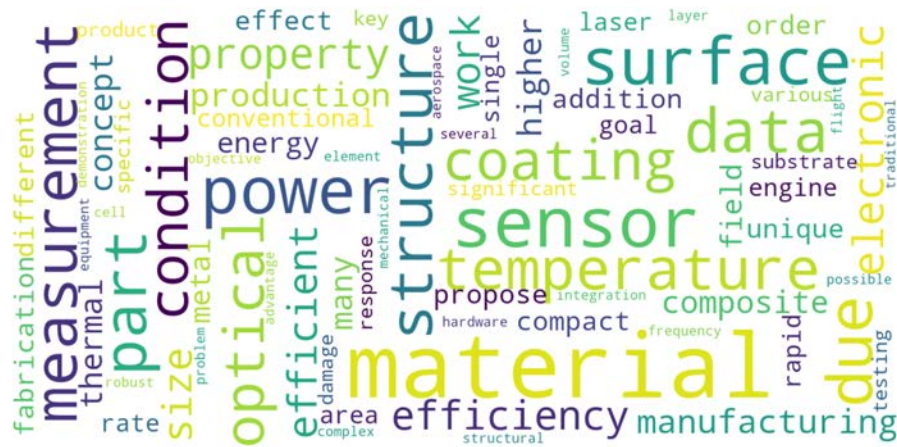
(b) Training/Software

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.6: 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.19). The word's size reflects its prevalence in the topic.

Figure A.7: Applicant Technologies (Based on K-means Clustering of Abstract Text)

(a) Model with Two Clusters
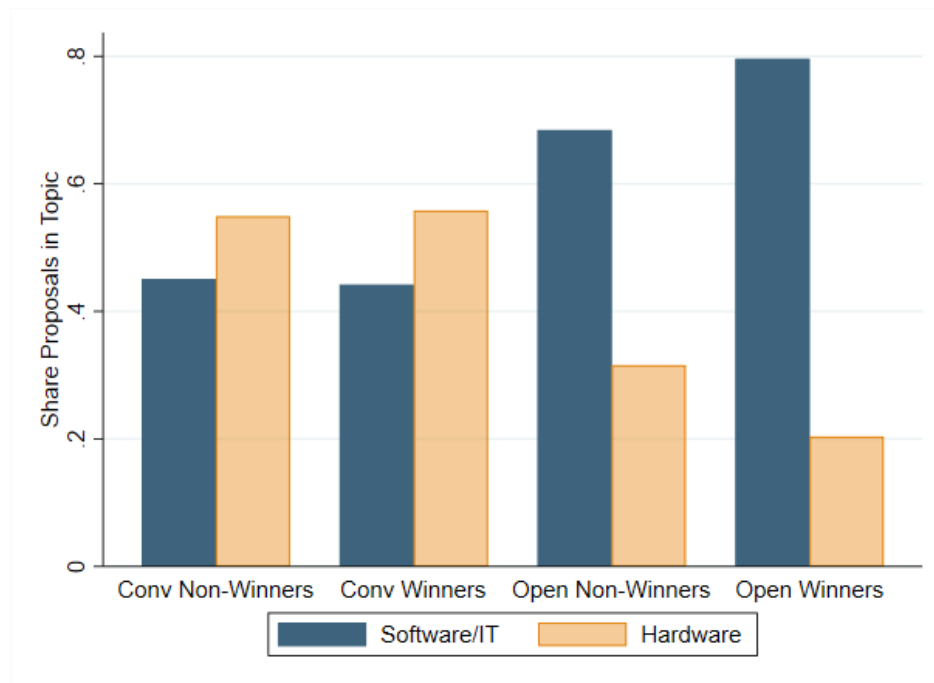
(b) Model with Five Clusters
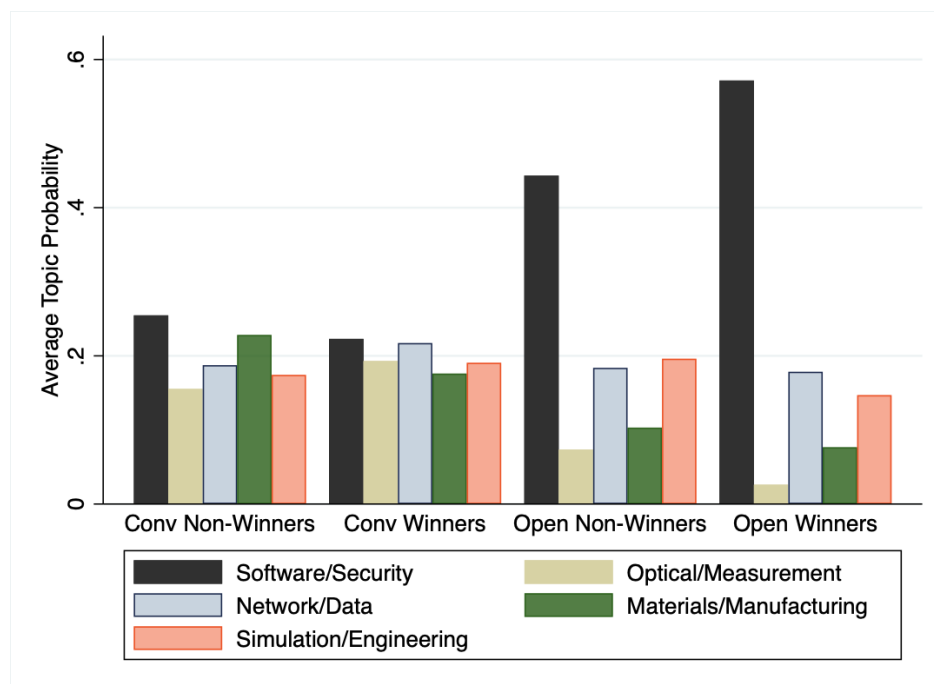
Note: These figures show the distribution of applications by topic based on a k-means cluster model, described in Section **??**. Panel (a) shows the model pre-set to produce two clusters. Panel (b) shows the model with the optimal number of clusters, which is the one for five clusters. In each case, 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.8: Continuity of Baseline Characteristics by Rank around Cutoff (Part 1 of 4)

(a) Firm Age at Application
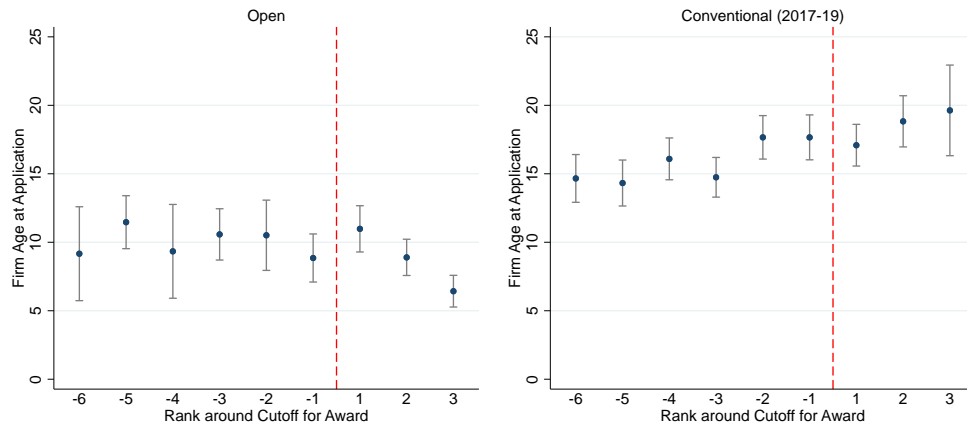
(b) Firm Employment at Application
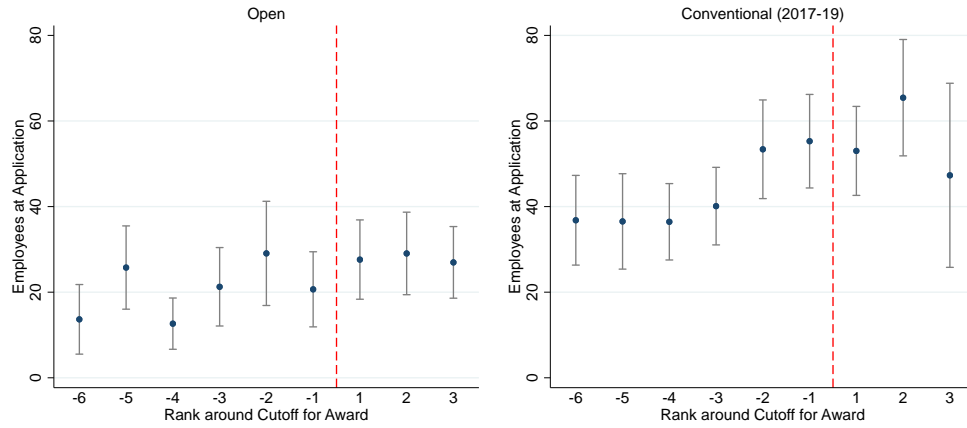
(c) Software vs. Hardware Technology
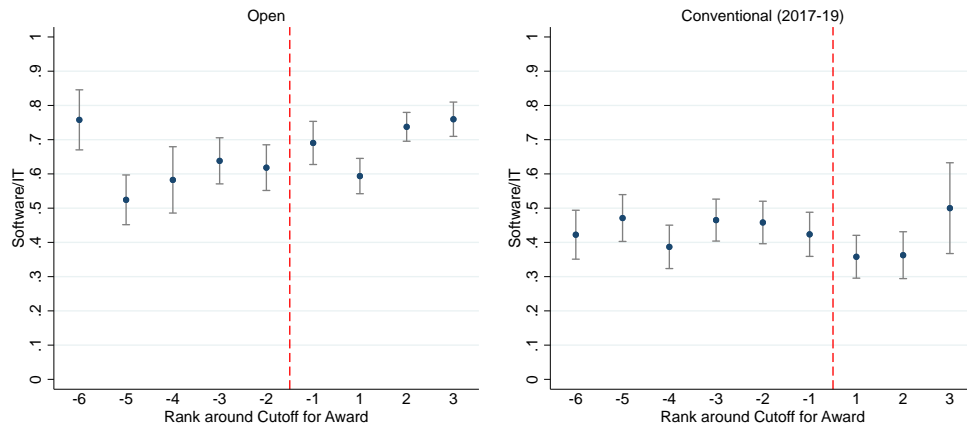
Note: These figures show applicant firm age (top figures), employment (middle figures), and the k-means 2 cluster abstract classification which yields a software and a hardware group (bottom figures) at the time of the application. In all cases, the x-axis shows the applicant's rank around the cutoff 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.

Internet Appendix 30

Figure A.9: Continuity of Baseline Characteristics by Rank around Cutoff (Part 2 of 4)

(a) Probability Firm Located in VC Hub City
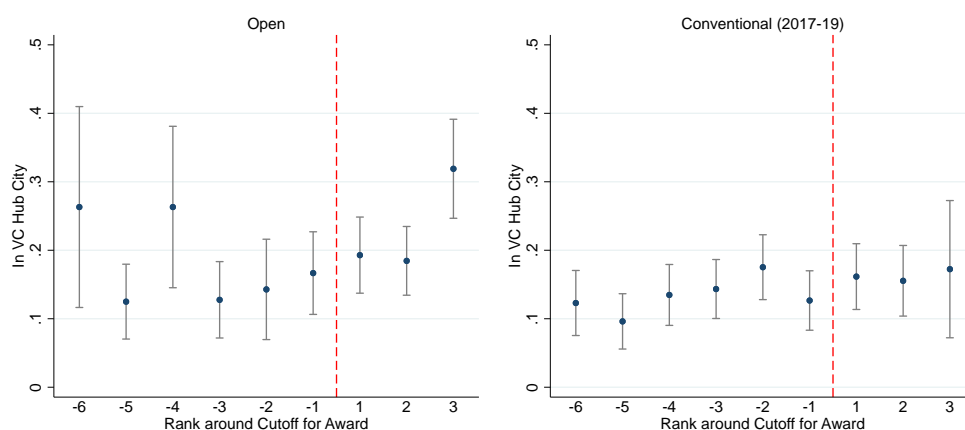
(b) Probability Firm Located in a County with an Air Force Base
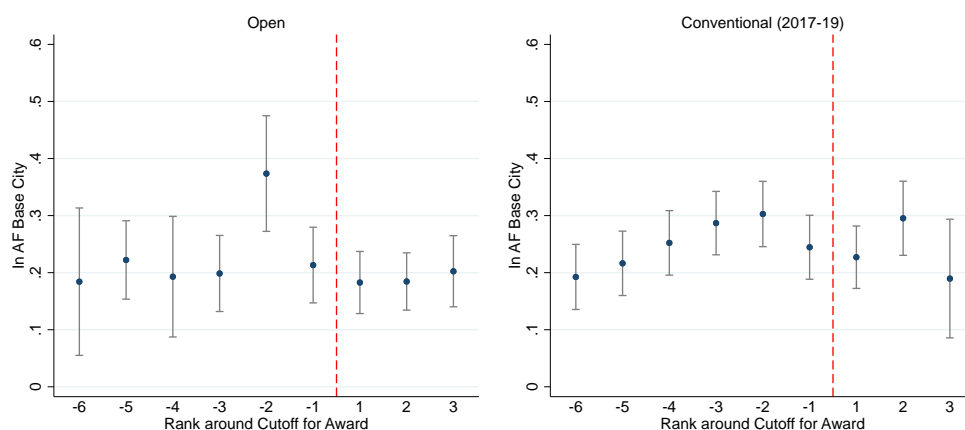
(c) Probability Firm Woman-Owned at Application
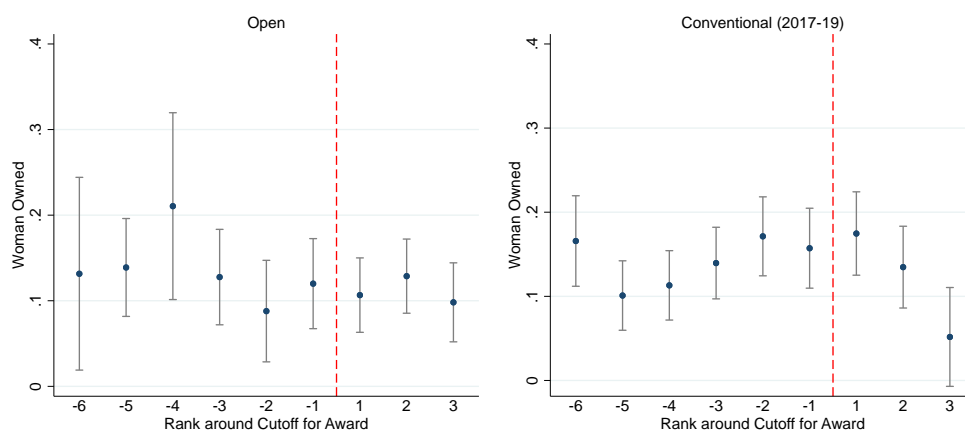
Note: These figures show the probability that an applicant firm is located in either San Francisco/San Jose, Boston, or New York City (top figures), located in a county with a U.S. Air Force base (middle figures), and woman-owned (bottom figures) at the time of the application. In all cases, the x-axis 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.

Internet Appendix 31

Figure A.10: Continuity of Baseline Characteristics by Rank around Cutoff (Part 3 of 4)

(a) Probability Firm Minority-Owned at Application
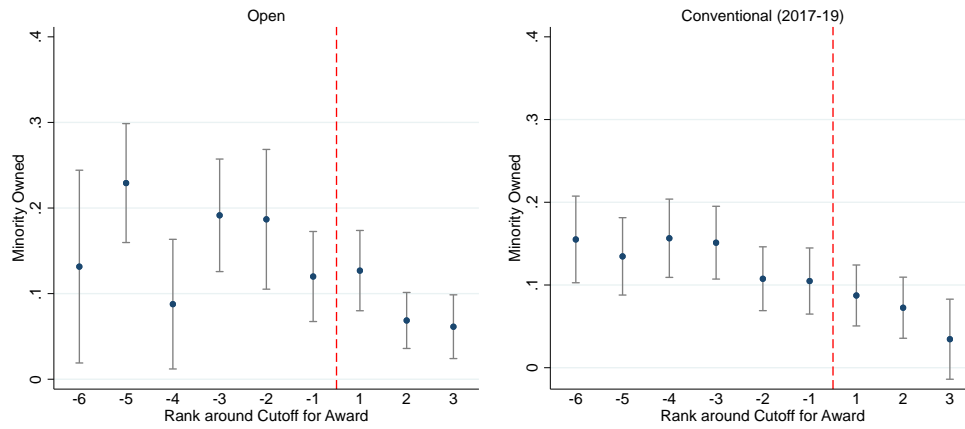
(b) Probability of Venture Capital Before Award Decision
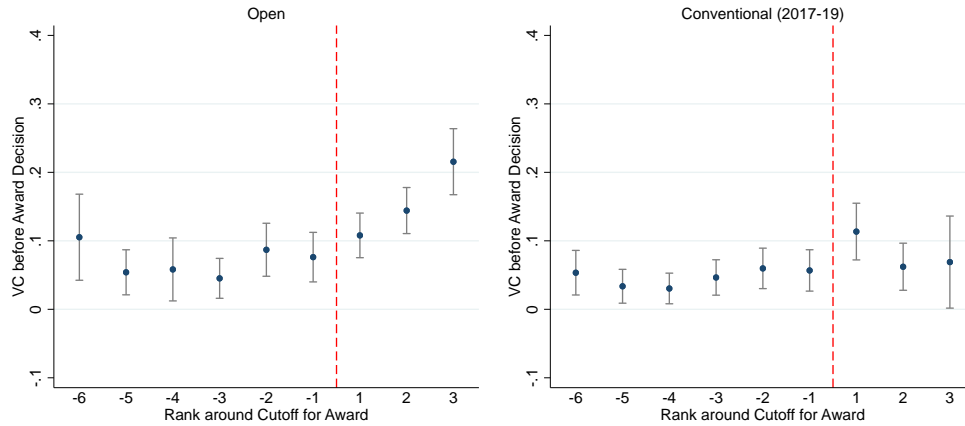
(c) Probability of Patent Before Award Decision
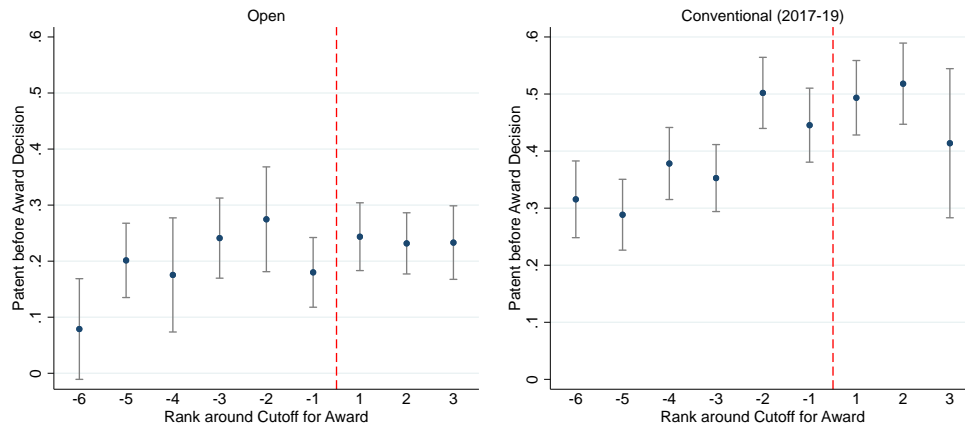
Note: These figures show the probability that an applicant firm is minority-owned (top figures), raised venture capital investment (VC, middle figures), and had any patents at the time of the application (bottom figures) at the time of application. In all cases, the x-axis shows the applicant's rank 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.

Internet Appendix 32

Figure A.11: Continuity of Baseline Characteristics by Rank around Cutoff (Part 4 of 4)

(a) Probability of DoD Non-SBIR Contract Before Award Decision
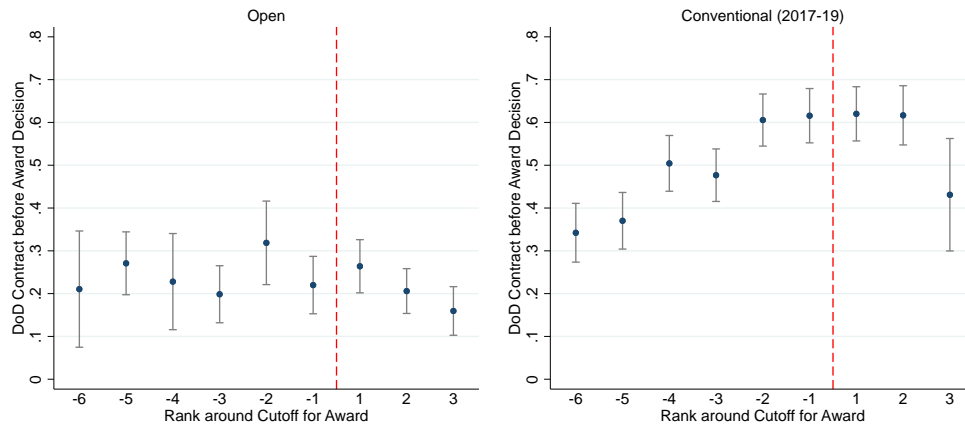
(b) Probability of SBIR Before Award Decision
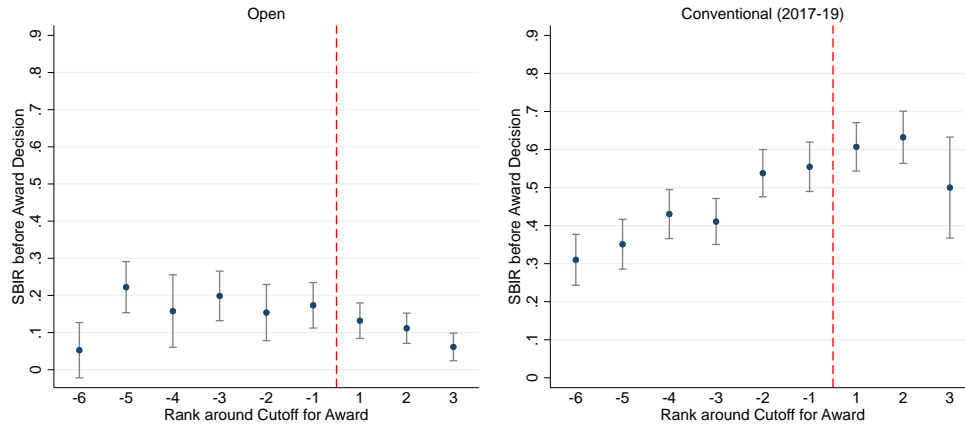
Note: These figures show the probability that an applicant firm had any SBIR contracts after the award decision (top figures) and had any non-SBIR DoD contracts valued at more than $50,000 at the time of the application (bottom figures). A rank of 1 indicates that the applicant had the lowest score among 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.12: Prevalence of Crossover Sub-scores

(a) Tech Score
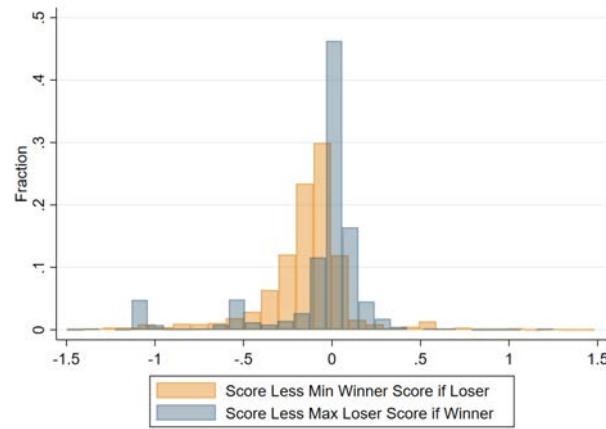
(b) Team Score
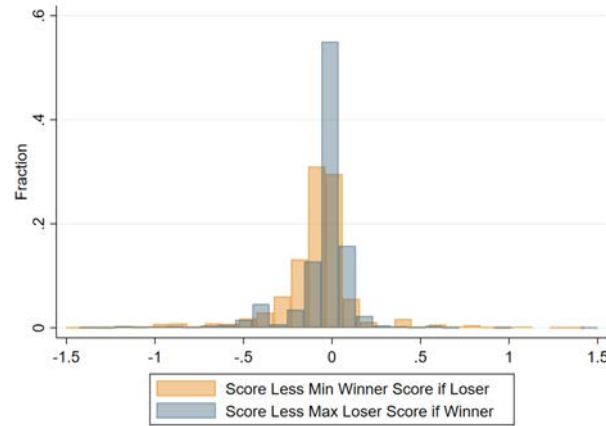
(c) Commercialization Score
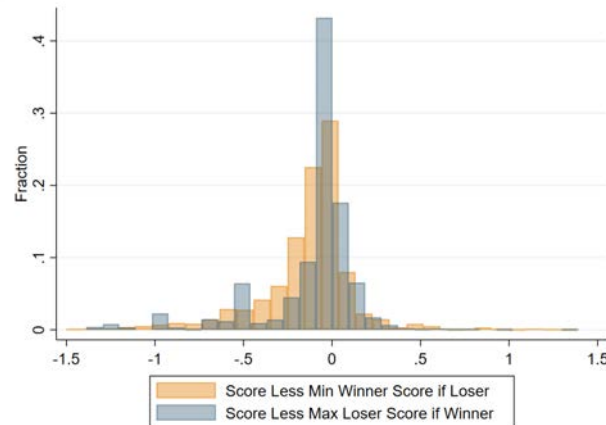
Note: These histograms demonstrate the substantial variation in the three sub-scores (tech, team, commercialization) around the cutoff. The red bars to the right side of zero show that many unsuccessful applicants (losers) have a sub-score that exceeds the lowest sub-score among winners. Similarly, the 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.

Internet Appendix 34

Figure A.13: Selection into Programs; Conventional Divided by Whether Applied to Only Conventional or Both Open and Conventional
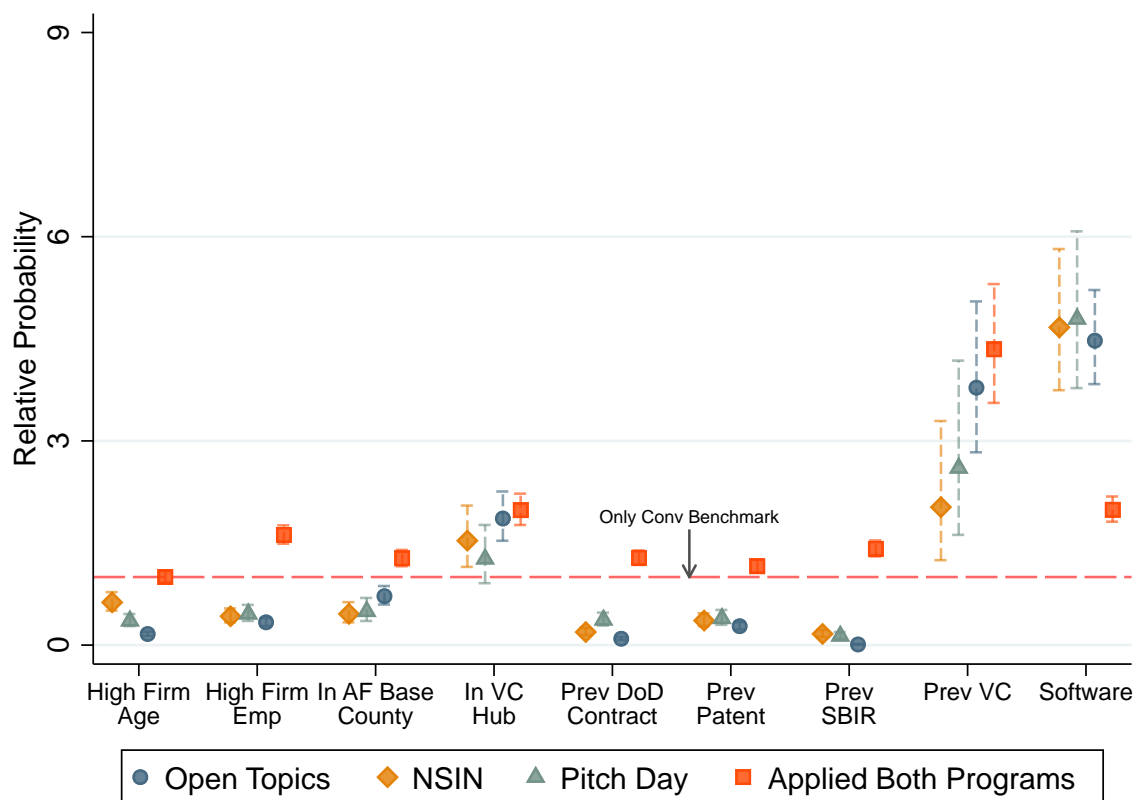
Note: This figure repeats Figure 2 but sets the baseline as applicants who only applied to Conventional topics, and then adds applicants who applied to both Conventional and Open topics as an additional group. The figure shows how firm characteristics predict selecting into reform programs compared to the specific subset of the Conventional program. The points represent coefficients from a multinomial logistic regression with categories for five programs: the applicants to only Conventional program, which serves as the base group and is represented by the red dashed line, the applicants who applied to both Conventional and Open program, and applicants who only applied to the respective 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 A.14: Geographic Dispersion of Applications (2017-19)

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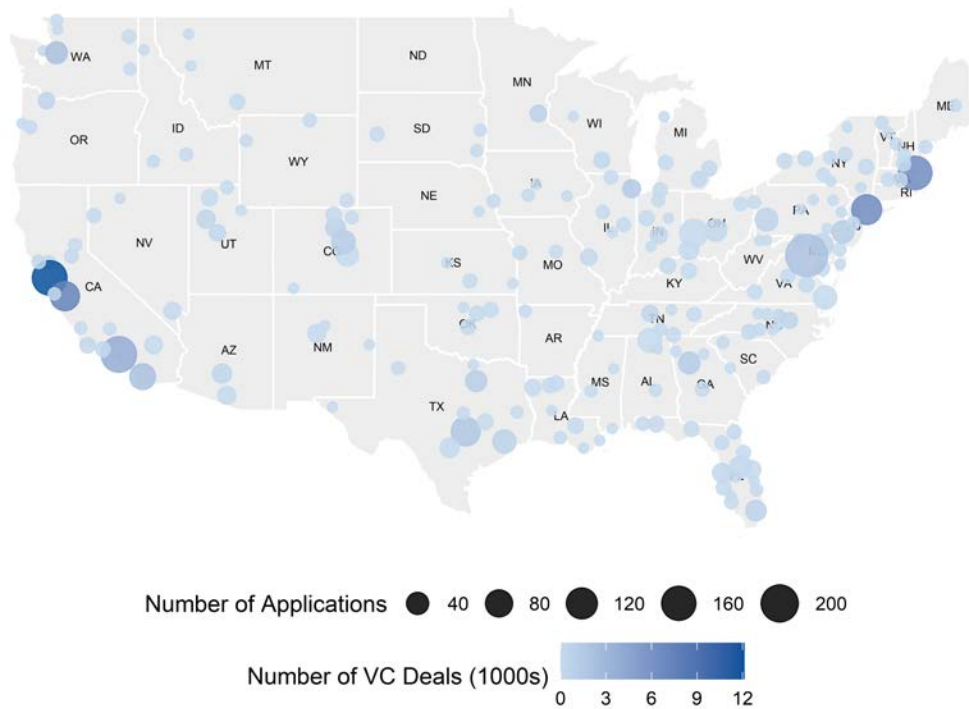
(b) Conventional Topic Applications and VC Deals
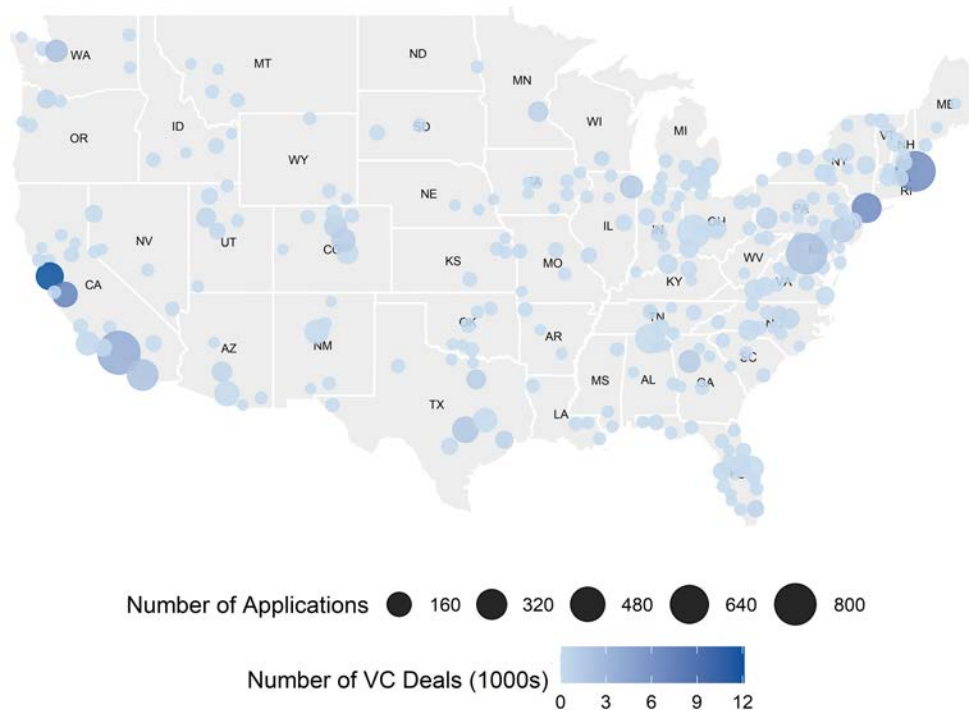
Note: These maps show the number of applications to open (Panel A) and conventional SBIR topics (Panel B) by MSA from 2017 to 2019. The size of the bubble represents the relative number of applications. The color gradient in both maps also show VC activity by MSA.

Figure A.15: Geographic Dispersion of Awards (2017-19)

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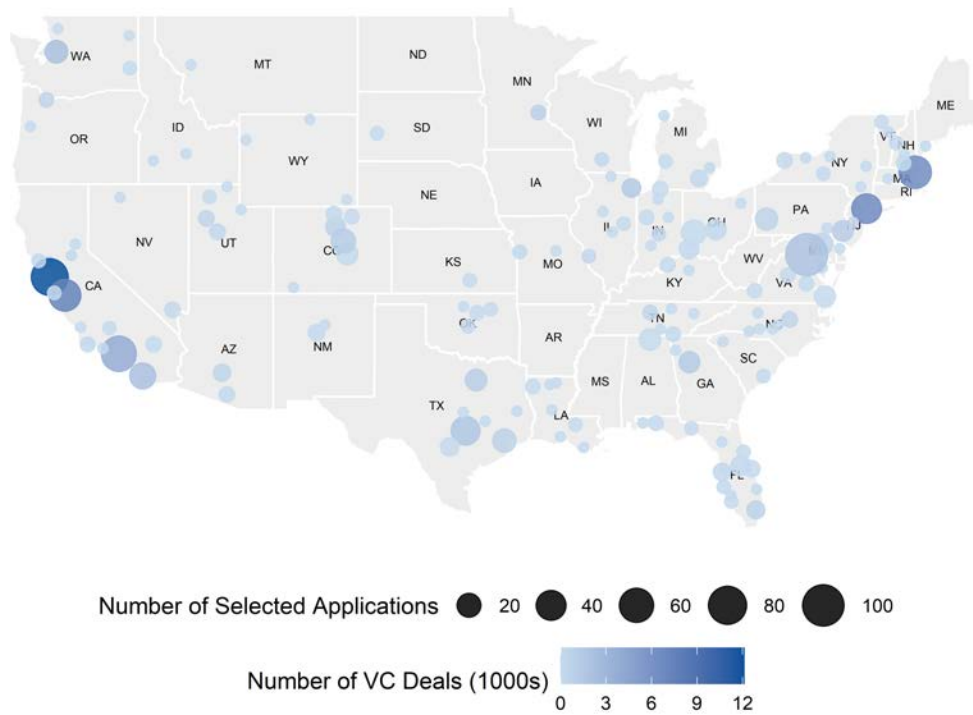
(b) Conventional Topic Awards and VC Deals
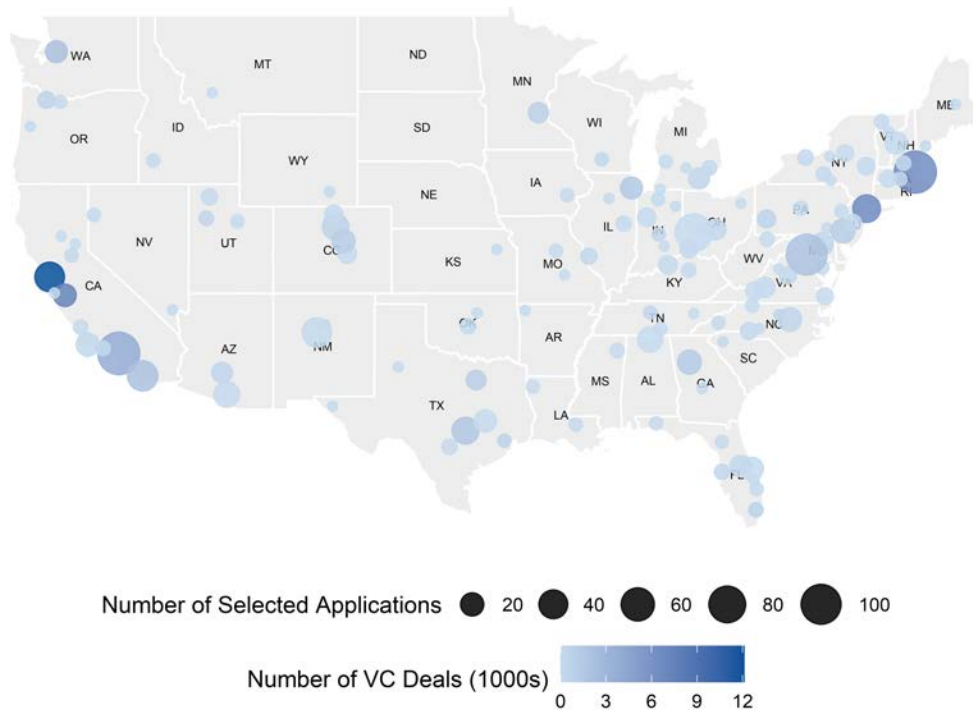
Note: These maps show the number of awards (i.e. contracts) for open (Panel A) and conventional SBIR topics (Panel B) by MSA from 2017 to 2019. The size of the bubble represents the relative number of applications. The color gradient in both maps also show VC activity by MSA.

Figure A.16: Lock-in for Open and Conventional over Three-Year Window
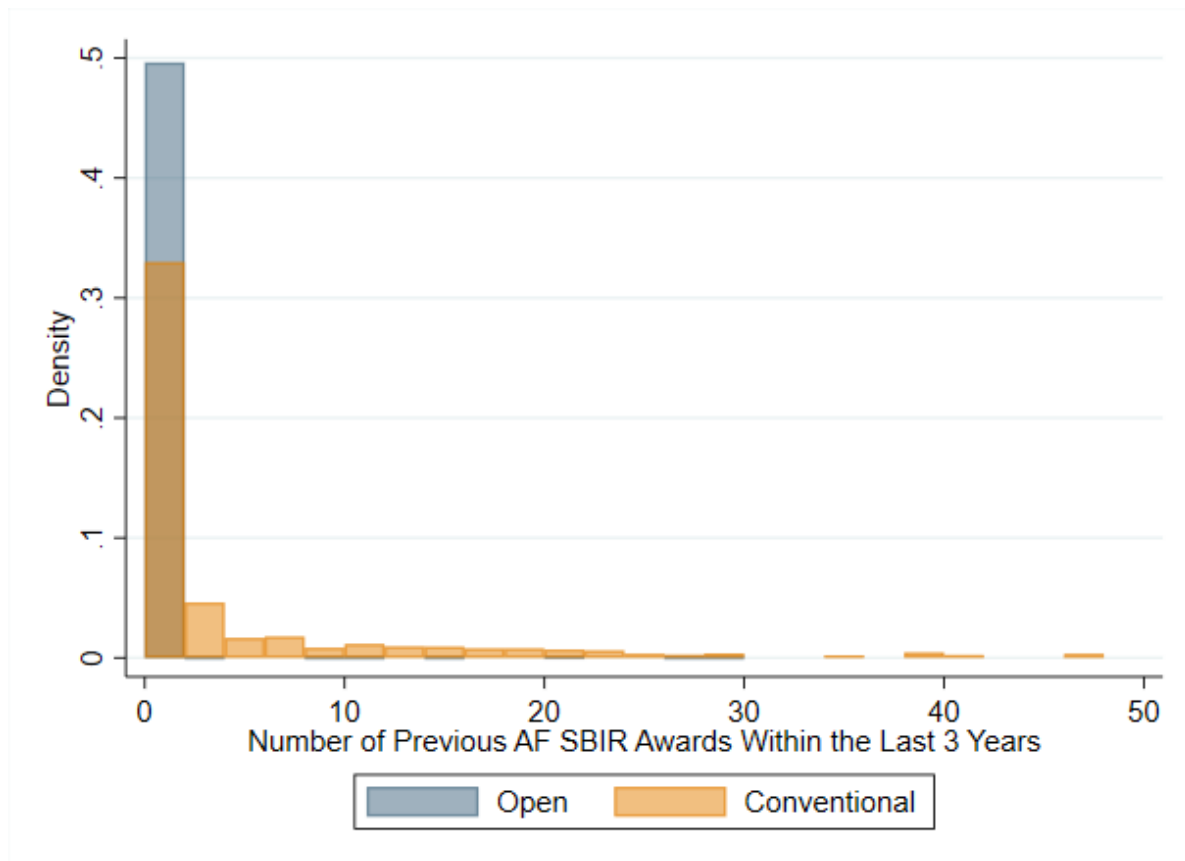
Note: This figure contains a histogram of the number of Open and Conventional applicants in categories defined by the number of Air Force SBIR awards in the past three years. For the Open applicants, we use 2020 data so that we have three years in which to look back for lock-in. For Conventional, we use 2019 data and also look back for three years. We do not observe Conventional application data in 2020, and this approach also aligns the sample with that of our main analysis (2017-19, where 2017 only contains Conventional applicants).

Figure A.17: Exploratory Patents from Prime Defense Contractors
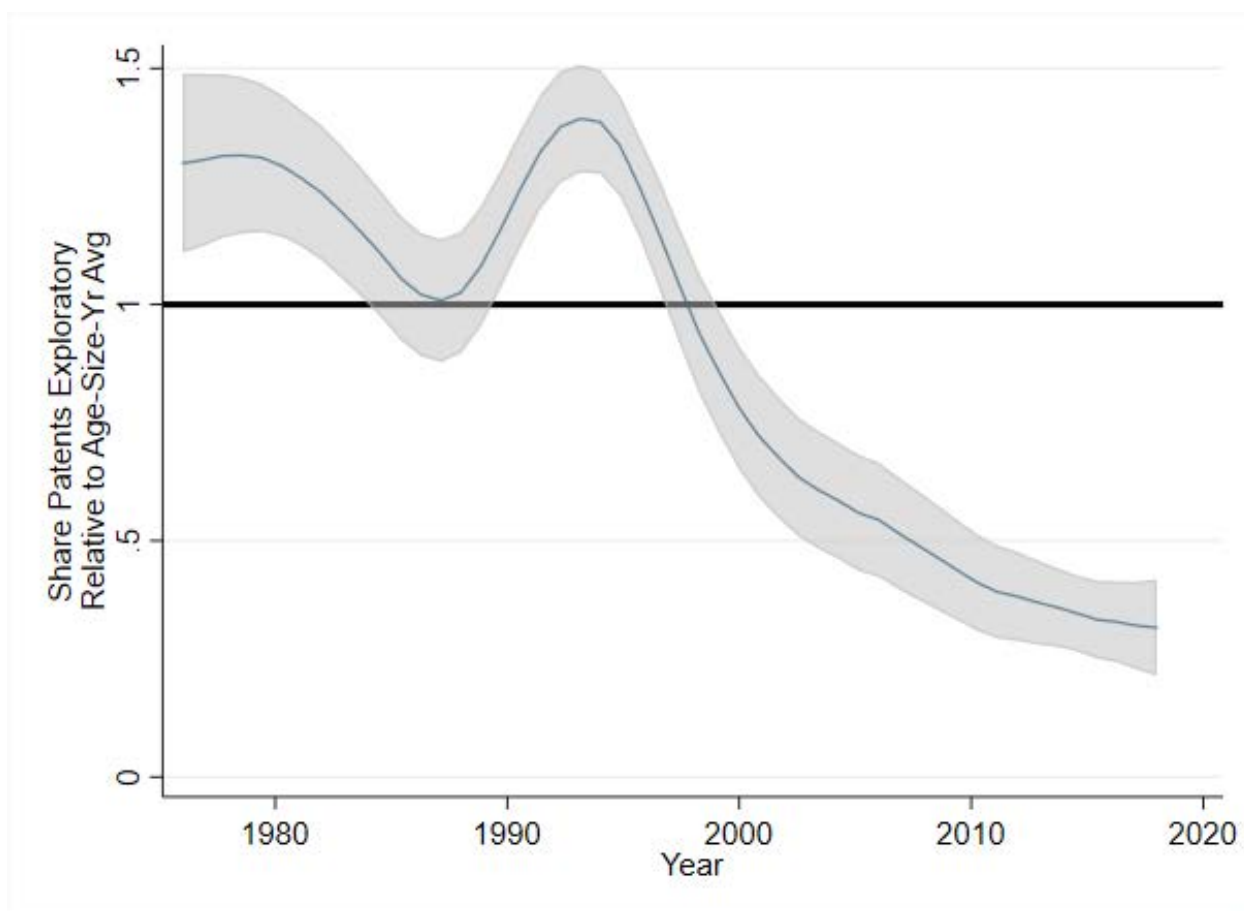
Note: This figure describes the trend of exploratory innovation by the prime defense contractors and their acquisition targets over time. The firms are the same set from Figure A.1. That is, 226 firms are included in 1976, while only six are included in 2019 (as the 226 have merged into these six). The 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.18: 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
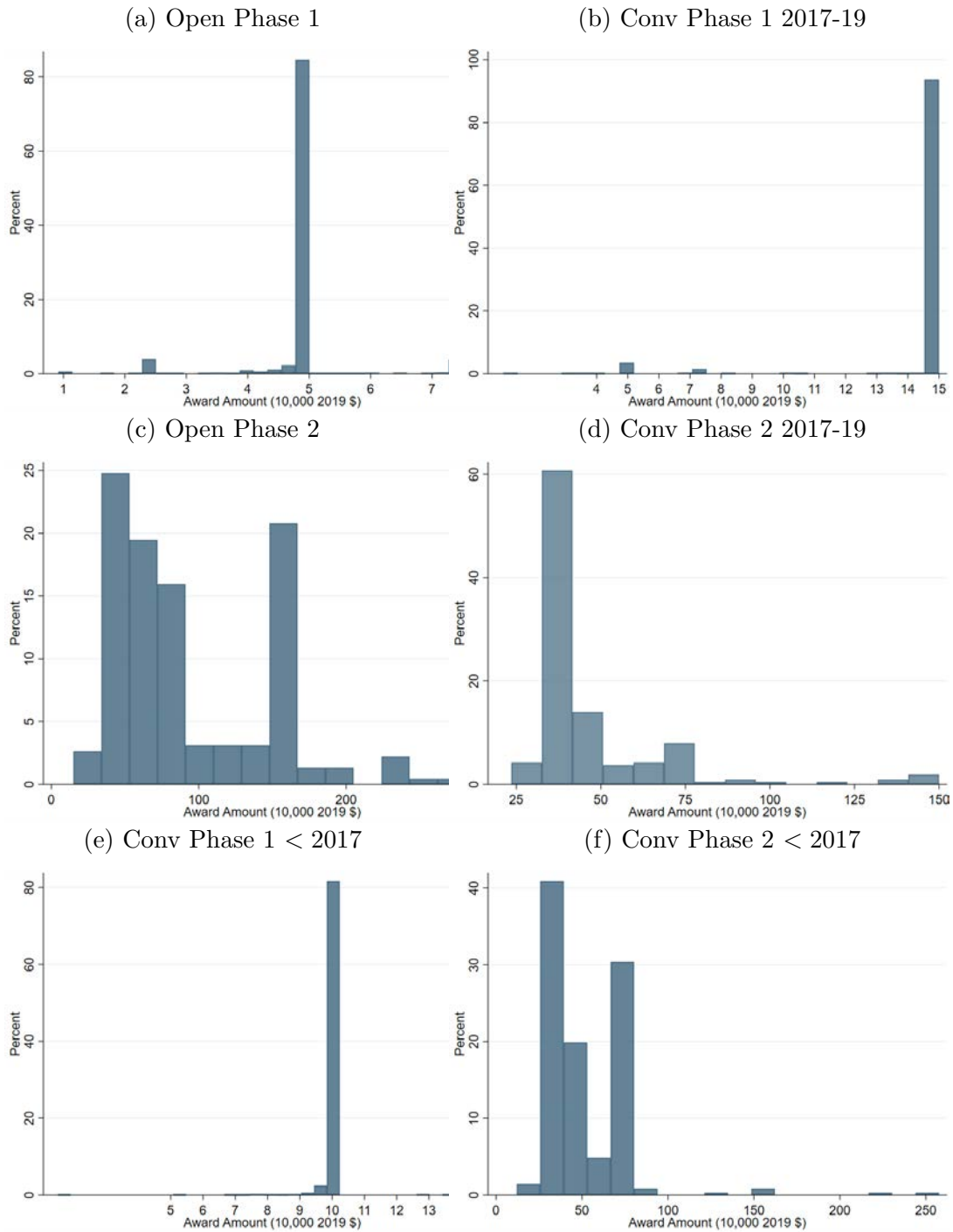
Note: These histograms show the share of awards by amount, in real 2019 dollars. For the bottom right graph (Phase 2 <2017), we omit one outlier $12 mill contract.

Figure A.19: K-means Optimal Number of Clusters (Elbow Method)
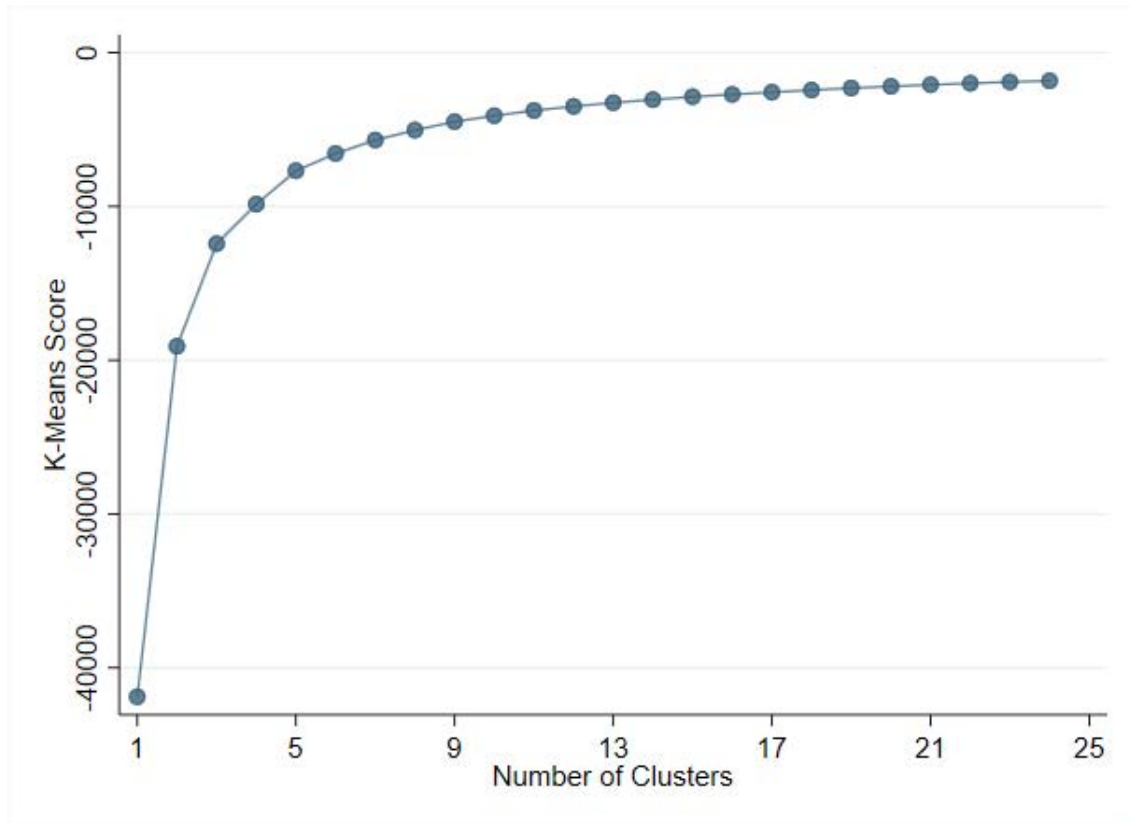
Note: This figure shows the k-means score for 1-25 clusters. The k-means score is defined as the negative total squared error between the cluster members and their centroid. The optimal number of clusters, or the "elbow," is identified for the number of clusters k for which the objective improves significantly and plateaus thereafter. We identify the optimal cluster to be $k = 5$.

Table A.1: Proposal and Firm Counts

Panel A: Open & Conventional (2017-19)

	Both	Open Topic	Conventional
Number of Topics:			
Phase I	428	6	422
Phase II	178	6	172
Number of Proposals:			
First Phase I	1523	1385	2608
Phase I	6654	1659	4995
Phase II	1106	647	459
Number of Firms:			
Applied to Type	3058	1411	2506
Exclusively Applied to Type	859	552	1647

Panel B: Full Sample (2003–2019)

	Both	Open Topic	Conventional
Number of Topics:			
Phase I	1778	6	1772
Phase II	788	6	782
Number of Proposals:			
First Phase I	5010	1385	7384
Phase I	21432	1659	19773
Phase II	2354	647	1707
Number of Firms:			
Applied to Type	6835	1423	6283
Exclusively Applied to Type	871	552	5412

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, 1,412 unique firms have applied to Open topics, of which 651 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: 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	$ 832,467	$ 762,884	$ 470,905	74	$ 813,967	$ 782,191	$ 183,205
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

Panel B: Pre-Award Outcome Summary

		Open Topic				Conventional		
	N	Mean	Median	SD	N	Mean	Median	SD
$\mathbb{1}$(VC)	647	0.159		0.366	459	0.070		0.255
Avg VC Amt (Mill)	103	$ 7.232	$ 1.99	$ 10.325	32	$ 1.719	$ .0277	$ 5.393
$\mathbb{1}$(DoD Non-SBIR Contract)	647	0.257		0.437	459	0.706		0.456
# DoD Non-SBIR Contracts	166	8.783	3	13.907	324	27.281	12.5	32.289
Avg DoD Non-SBIR Contract Amt (Mill)	166	$ 2.036	$ .9221	$ 2.759	324	$ 1.427	$ .9054	$ 1.436
$\mathbb{1}$(Patent)	647	0.284		0.451	459	0.580		0.494
# Patents	184	13.630	4	42.768	266	25.214	12	35.072
# Patent Application if Any	223	13.058	3	40.353	282	24.713	10	34.973
$\mathbb{1}$(AF SBIR Contract)	647	0.206		0.404	459	0.732		0.443
# AF SBIR Contracts	133	9.639	2	16.781	336	47.307	17.5	64.533
$\mathbb{1}$(Never Awarded SBIR)	647	0.623		0.485	459	0.192		0.394

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

Table A.3: 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	$ 130,911	$ 27,110	83	$ 119,445	$ 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

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	0.068		0.252
Avg VC Amt (Mill)	1,182	$ 4.848	$ .812	$ 9.058	51	$ 4.422	$ .1478	$ 8.625
$\mathbb{1}$(DoD Non-SBIR Contract)	19,773	0.369		0.483	747	0.301		0.459
# DoD Non-SBIR Contracts	7,301	12.841	4	22.703	225	18.338	5	33.293
Avg DoD Non-SBIR Contract Amt (Mill)	7,301	$ 1.454	$ .7132	$ 2.200	225	$ 1.616	$ .9903	$ 1.997
$\mathbb{1}$(Patent)	19,773	0.440		0.496	747	0.261		0.439
# Patents	8,705	19.448	6	35.156	195	13.200	5	26.916
# Patent Application if Any	8,925	19.291	6	35.280	238	12.105	4	26.668
$\mathbb{1}$(AF SBIR Contract)	19,773	0.421		0.494	747	0.213		0.410
# AF SBIR Contracts	8,327	21.518	6	41.012	159	30.164	10	51.767
$\mathbb{1}$(Never Awarded SBIR)	19,773	0.373		0.484	747	0.625		0.484

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.

Table A.4: Formal Test for Continuity of Baseline Covariates

| | Open (N= 422) | | | Conventional (N= 960) | | |
| | Rank = -1 | Rank = 1 | | Rank = -1 | Rank = 1 | |
	Mean	Mean	p	Mean	Mean	p
Age	9.571	10.714	0.320	20.196	19.603	0.463
Employees	21.367	29.909	0.169	76.663	77.592	0.882
Firm in VC Hub City	0.172	0.202	0.434	0.168	0.168	0.999
Firm in AF Base County	0.200	0.202	0.950	0.294	0.270	0.393
Woman-owned	0.117	0.095	0.473	0.143	0.176	0.166
Minority-owned	0.117	0.116	0.976	0.092	0.093	0.949
Software	0.298	0.271	0.547	0.586	0.591	0.874
Any Pre-Appplication Patent	0.206	0.252	0.264	0.552	0.590	0.234
Any Pre-Appplication VC	0.067	0.095	0.297	0.059	0.100	0.020
Any Pre-Appplication Non-SBIR Contract	0.239	0.277	0.381	0.734	0.705	0.313
Any Pre-Appplication SBIR Contract	0.211	0.194	0.670	0.695	0.726	0.293

Note: This table shows the t-test results on the marginal winners and losers (Rank = -1 or 1).

Table A.5: 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) $\times$ $\mathbb{1}$(Open Topics)	0.065*	0.065*	0.065*		
	(0.034)	(0.034)	(0.034)		
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(NSIN)	0.013		0.013		
	(0.068)		(0.068)		
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Pitch Day)		-0.154***	-0.154***		
		(0.031)	(0.031)		
$\mathbb{1}$(Award)	-0.004	-0.004	-0.004	0.008	-0.158**
	(0.020)	(0.020)	(0.020)	(0.064)	(0.028)
Observations	4392	4317	4716	399	324
Outcome Mean	0.041	0.041	0.042	0.050	0.052

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 A.6: Effect of Winning Conditional on Firm Applying to Both Programs

Dep Var:	Any VC	Any DoD Contracts	Any Patents	Any SBIR
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	-0.007	-0.078	0.044	0.002
	(0.013)	(0.059)	(0.058)	(0.054)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open Topic)	0.071***	0.152**	0.051	-0.006
	(0.021)	(0.063)	(0.060)	(0.059)
Observations	1492	1492	1492	1492
Outcome Mean	0.054	0.369	0.213	0.284

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award on the four outcomes, all within 24 months after the award decision. We repeat column 4 in Tables 2-3 but restrict the sample to those firms which at some point in our sample applied to both the Open and Conventional programs (usually in different years). 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, firm age, and firm age squared. These latter two variables ensure that the difference between Open and Conventional does not reflect selection on firm age, where a firm applies early in its life to one program and later to another. All columns in both panels include topic fixed effects, and errors are clustered by topic. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.7: Role of Financial Constraints in Effect of Winning: Heterogeneity in Firm Characteristics

Panel A: Any Subsequent Venture Capital Investment

	(1)	(2)	(3)	(4)	(5)	(6)
𝟙(Award)	0.063**	0.090***	0.039	0.005	-0.019	-0.004
	(0.025)	(0.027)	(0.025)	(0.020)	(0.022)	(0.023)
𝟙(Award) × 𝟙(Hardware)	0.019			-0.015		
	(0.034)			(0.013)		
𝟙(Award) × 𝟙(High Age)		-0.104***			0.021	
		(0.026)			(0.013)	
𝟙(Award) × 𝟙(High Emp)			0.032			-0.002
			(0.036)			(0.016)
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	0.087	0.087	0.087	0.016	0.016	0.016

Panel B: Any Subsequent Non-SBIR DoD Contracts

	(1)	(2)	(3)	(4)	(5)	(6)
𝟙(Award)	0.073*	0.097**	0.072**	-0.002	0.003	0.057
	(0.042)	(0.038)	(0.035)	(0.064)	(0.061)	(0.056)
𝟙(Award) × 𝟙(Hardware)	-0.039			0.070		
	(0.052)			(0.058)		
𝟙(Award) × 𝟙(High Age)		-0.013			0.044	
		(0.048)			(0.052)	
𝟙(Award) × 𝟙(High Emp)			-0.054			-0.085*
			(0.052)			(0.049)
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	0.199	0.199	0.199	0.329	0.329	0.329

Panel C: Any Subsequent Patents

	(1)	(2)	(3)	(4)	(5)	(6)
𝟙(Award)	0.076**	0.101***	0.077***	0.058	0.035	0.040
	(0.030)	(0.030)	(0.029)	(0.055)	(0.052)	(0.050)
𝟙(Award) × 𝟙(Hardware)	0.012			0.014		
	(0.045)			(0.050)		
𝟙(Award) × 𝟙(High Age)		-0.033			0.034	
		(0.039)			(0.048)	
𝟙(Award) × 𝟙(High Emp)			-0.015			0.006
			(0.046)			(0.046)
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	0.117	0.117	0.117	0.194	0.194	0.194

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) - (3)) and Conventional topics (columns (4) - (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 hardware-related proposals. Columns (2) and (5) interact winning an award with an indicator for an old firm defined as firm's age above the median of the distribution of firm age (and zero otherwise). Column (3) and (6) interact winning an award with an indicator for a high employment firm defined as firm's number of employees above the median of the distribution of firm employment (and zero otherwise). 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 A.8: Effect of Open vs. Conventional Among the Highest-Quality Applicants

	Any VC		Any Patents		Any DoD Contracts		Any SBIR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$(Award)	-0.008	-0.009	0.029	0.006	0.043	0.034	0.053*	0.045*
	(0.011)	(0.008)	(0.028)	(0.023)	(0.031)	(0.022)	(0.030)	(0.023)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open)	0.082***	0.086***	0.014	0.036	0.002	0.002	-0.082***	-0.083***
	(0.017)	(0.015)	(0.029)	(0.025)	(0.034)	(0.026)	(0.030)	(0.025)
$\mathbb{1}$(High Rank)	0.019	0.006	0.065	-0.006	0.070	0.036	-0.070	-0.114
	(0.059)	(0.035)	(0.098)	(0.083)	(0.106)	(0.082)	(0.109)	(0.088)
$\mathbb{1}$(Open) $\times$ $\mathbb{1}$(High Rank)	-0.051	-0.024	-0.014	0.064	-0.022	0.004	0.099	0.132
	(0.072)	(0.052)	(0.106)	(0.092)	(0.117)	(0.094)	(0.115)	(0.093)
Observations	3993	6654	3993	6654	3993	6654	3993	6654
Program	Both	Both	Both	Both	Both	Both	Both	Both
Proposal	First	All	First	All	First	All	First	All
Time Period	2017-19	2017-19	2017-19	2017-19	2017-19	2017-19	2017-19	2017-19
Outcome Mean	0.041	0.031	0.167	0.234	0.284	0.411	0.197	0.340

Note: This table shows regression discontinuity (RD) estimates of the effect of winning a Phase 1 award. 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), (3), (5), and (7) restrict the sample to the firm's first application within the sample time period while columns (2), (4), (6), (8) include all proposals over the sample time period. We interact winning an award with an indicator that is equal to one if the proposal is in an Open topic (and zero otherwise) and an indicator that is equal to one if the proposal is in the top 15 percentiles of scores among winners within topic (and zero otherwise). 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 A.9: Effect of Winning on Alternative 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)	0.056**	0.013	-0.017	0.013	-0.017	-0.019
	(0.022)	(0.047)	(0.021)	(0.046)	(0.021)	(0.013)
$\mathbb{1}$(Award) $\times$ $\mathbb{1}$(Open Topic)				0.043	0.073**	0.080***
				(0.053)	(0.035)	(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	0.070	0.119	0.109	0.102	0.103	0.189

Panel B: Any Subsequent High Citation and Generality Patent

Dep Var:	High Citation		High Generality	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	0.039	-0.024	-0.007	-0.025
	(0.035)	(0.024)	(0.025)	(0.020)
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	0.039	0.131	0.026	0.096

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.10: 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	$ 529,618	$ 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.11: 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)	0.050**	0.015	0.041	0.096	0.067
	(0.023)	(0.013)	(0.052)	(0.079)	(0.062)
$\mathbb{1}$(Award $\times$ Match Offered in Topic)					0.081
					(0.054)
Observations	1385	1385	1004	381	1385
Outcome Mean	0.075	0.027	0.166	0.113	0.152

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.12: Effect of Phase 2 Award 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)	-0.054	-0.017	0.004	-0.018
	(0.060)	(0.018)	(0.141)	(0.077)
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	0.580	0.531

Panel B: Any Subsequent Patenting and SBIR Contracts

Dep Var:	Any Patents		Any SBIR	
	(1)	(2)	(3)	(4)
$\mathbb{1}$(Award)	0.204	0.033	-0.033	-0.142*
	(0.135)	(0.081)	(0.142)	(0.080)
Observations	457	1703	457	1703
Program	Conv	Conv	Conv	Conv
Time Period	2017-19	2003-19	2017-19	2003-19
Outcome Mean	0.305	0.332	0.436	0.497

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.