The commercialization of DoD-SBIR patents: A counterfactual analysis

Carlo Bottai[a], **Gaétan de Rassenfosse**[b], **and Emilio Raiteri**[a,1]

[a]Eindhoven University of Technology, School of Innovation Sciences, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands; [b]École polytechnique fédérale de Lausanne, College of Management of Technology, ODY 201.1, Station 5, 1015, Lausanne, Switzerland

This manuscript was compiled on July 15, 2022

The paper proposes a novel, web-based approach to innovation policy evaluation. The approach overcomes several limitations affecting established evaluation methods used in the literature. We implement it to study the impact of the U.S. DoD-SBIR program on technology commercialization. We start by identifying the universe of USPTO patents that acknowledge support by the SBIR program. We then track whether these patents are mentioned in relation to commercial products in a virtual patent marking page available on the recipient's website. We interpret the latter event as signal of commercialization. Finally, we create a group of control patents and we compare the commercialization probability of SBIR-funded and control inventions. The results support the view that the SBIR program is quite effective at stimulating the commercialization of federally-funded scientific discoveries.

Commercialization | Government-funded Research | Patents | Policy Evaluation | Web-based Evidence

Scholars have long acknowledged the importance of public procurement, notably defense procurement, for scientific and technological progress (1, 2). In a recent paper, Moretti and colleagues suggest that U.S. defense procurement represents the most important industrial policy to affect the speed and direction of innovation (3). Numerous works highlight the role of defense procurement in developing products that have become major commercial successes. Ruttan describes how the purchasing power of the U.S. Department of Defense (DoD) was instrumental for the arrival of the commercial Internet and the GPS technology (4). Mazzucato stresses that popular consumer products, such as the iPhone or the iPad, and services, such as Siri, benefited from public intervention. She also provides anecdotal evidence of a close link between the Apollo program and products widely adopted today, from the shock-absorbing sneaker soles to medical devices such as pacemakers and defibrillators (5, 6). Mazzucato gives particular praise to the U.S. Small Business Innovation Research (SBIR) program for guiding the commercialization of hundreds of new technologies from the laboratory to the market (5).

The Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programs are two related public funding programs. They seek to encourage U.S. small businesses to engage in federal R&D projects with commercialization potential. The SBIR program was introduced by the Small Business Innovation Development Act of 1982, whose objectives include the increase of private sector commercialization of innovations derived from federal R&D.* The STTR came a decade later, in 1992. The U.S. Small Business Administration (SBA) coordinates the programs, that involve eleven participating agencies. In fiscal year 2019, federal agencies obligated about \$3.8 billion of SBIR/STTR funding. The SBIR/STTR programs have two main phases. Phase I funds initial research to establish the technical merit, feasibility, and commercial potential of an R&D project. Successful Phase I participants may proceed to Phase II, in which they receive larger funding to pursue the research started in Phase I. Phase I awards generally amount to \$50-150,000 for six months or one year. Phase II awards may reach \$1 million and last for two years. The two programs, SBIR and STTR, are similar enough to be considered as a joint funding schem for the purpose of this paper. As such, we will use term 'SBIR' to intend both.

The program is considered to be largely successful, and it is broadly emulated and extensively studied all over the world (7–10). Policymakers and scholars alike have devoted special attention to the impact of the SBIR program, in terms of bringing the fruits of federally-funded research to the final consumer. This issue is particularly relevant for defense-related R&D, which accounts for the vast majority of R&D procurement in the United States,† and for about half the overall budget of the SBIR/STTR program (on average in 1990–2012, computed from the balance sheets provided on the program's

*The explicit goals of the program are to (i) stimulate technological innovation, (ii) use small business to meet federal research and development needs, (iii) foster and encourage participation in innovation and entrepreneurship by women and socially or economically disadvantaged persons, and (iv) increase private-sector commercialization of innovations derived from federal research and development funding. For further details about the program, see the Small Business Act (15 U.S.C. § 638), as well as https://www.sbir.gov/about.

†In FY 2017 DoD contract obligations amount to \$320 billion, equal to 63% of federal contract obligations and 8% of all federal spending. Of these contracts, 8% were for R&D contracts, in line with the average federal spending (11).

Significance Statement

Governments invest massively in research and development activities through various support programs. Assessing the 'real impact' of such programs is challenging notably due to the difficulty of tracing the commercialization of publicly-funded inventions. We propose a novel method to address this challenge that involves searching for online traces of commercialization on companies' websites. We apply the method to (patented) inventions funded by the U.S. SBIR program, which seeks to push inventions from the lab to the market. We find that SBIR-funded inventions are 17 percent more likely to be commercialized than a control group of similar, but privately-funded, inventions. Researchers can use the method to assess other such programs or commercialization outcomes.

Conceptualization (CB, GdR, ER); Data curation (CB, ER); Formal analysis (CB, GdR, ER); Funding acquisition (CB, ER); Project administration (ER); Software (CB); Supervision (GdR, ER); Visualization (CB); Writing (CB, GdR, ER).

The authors declare no competing interest.

[1] To whom correspondence may be addressed. Email: e.raiteritue.nl

website). A few academic studies provide evidence of a positive effect of the SBIR program at DoD on the commercialization of new technologies, as proxied by sales and patent applications [LIST STUDIES]. Since 2000, the National Academies have undertaken a quadrennial assessment of each agency's SBIR/STTR program, often using case studies and survey data. The DoD reports assert the program's positive effect on commercialization. According to these assessments, close to half of Phase II projects are associated with sales from products developed with SBIR funds (8, 12, 13).

Nevertheless, a number of studies also highlight the limitations of the program evaluations conducted so far. A Government Accountability Office report stresses that military departments mostly collect commercialization information about selected success stories and that their evaluation systems are not designed to capture detailed information about projects that did not transition to commercialization (14). A recent study by the National Academies of Sciences, Engineering, and Medicine highlights two fundamental issues affecting evaluations conducted thus far (10). First, program evaluation should go beyond observing program recipients improving their records over time. It requires assessing recipients' progress in light of what would have happened in the absence of the program; in short, compared to a credible counterfactual situation. Most of the academic and government-mandated assessments of the SBIR programs have largely neglected this aspect. Second, and equally important for the present paper, extant evaluations do not capture product market introductions. The DoD considers SBIR-funded projects as having a successful transition to commercialization if supported firms report any positive revenues from a product or service developed in the performance of the project. Yet, these revenues might well originate from non-SBIR contracts awarded by the DoD itself. According to (8, p.61) "nearly 60 percent of Phase II projects with sales reported sales to DoD or DoD primes." Thus, although these projects have successfully transitioned to commercialization, their broader impact on private sector innovation remains unclear.

Overcoming these issues requires the development of new methods and metrics of commercialization. We propose a novel, web-based approach to evaluate the impact of the SBIR program on commercialization. Specifically, we first link SBIR contracts to patented inventions arising from these contracts, and then connect these patents to the products and services they protect. We can then compare the commercialization rate of SBIR-related patents to a comparable set of patents that did not receive SBIR funding—that is, the counterfactual outcome. The identification of the patent-product connection builds on the work of de Rassenfosse (15). We search for the presence of specific web pages or product information brochures that clearly signal a patent-product link on the SBIR recipient's website. The next section and the Appendix illustrate our approach in detail, but one of its key features is the focus on actual patented inventions. This allows us to exploit the universe of patented inventions generated by SBIR contracts awarded by the DoD, and not exclusively to inventions owned by companies that agreed to respond to a survey or reached more advanced stages of product development.[‡]

We find that SBIR-funded patents are 17 percent more likely to be commercialized compared to control patents. This effect is particularly pronounced for applied and development R&D contracts as well as for Phase II contracts. We also find that SBIR awards signed after the year 2000's 'Phase II Plus' policy were more likely to be commercialized, suggesting that the reform has served its purpose. Finally, an analysis focusing on green inventions does not provide conclusive evidence that public support helped—or hindered—commercialization.

The rest of the paper is organized as follows. The next section, and the appendix, explain the details of the approach and illustrate the key features of the database. The section following presents the results and the last section concludes.

Data and methods

Our evaluation of the SBIR program entails three steps. First, we link patented inventions generated by DoD-SBIR contracts to actual commercial products and services using a novel web-based approach. This first step produces a unique database composed of three main elements: SBIR awards data, patent data, and web pages. Second, we identify a set of suitable patented inventions that form a control group to contrast the impact of SBIR funding. We similarly search for online traces of commercialization for these patents. Third, we perform regression analyses to assess the differences in the probability of commercialization between SBIR-funded and control inventions.

Constructing the database. To construct the database, we first identify patented inventions developed with the support of DoD-SBIR contracts. We exploit the Bayh–Dole Act of 1980 and the U.S. Federal Acquisition Regulation (FAR) as in (16). Under the Bayh–Dole Act, private entities must acknowledge federal funding and rights to an invention in the written specification of the invention for all U.S. patent applications. Furthermore, the FAR requires the applicant to disclose in the patent application the specific governmental agency and the contract number connected with the invention. These requirements allow us to identify the patented inventions produced under a government contract and the related contract information. To connect patents to specific awards, we extract the contract identification number from the patent documents and link them to federal databases providing detailed contract information.[§] Contract-level information allows us to identify patents specifically associated with contracts awarded by the DoD in the context of the SBIR program.

Next, we link patents to commercialized products. To do so, we adopt a web-based approach inspired by (15) and search for the existence of virtual patent marking (VPM) pages on the websites of the owner of the SBIR-related patents. VPM pages list a company's commercial products that are patent-protected. Companies set up VPM pages to provide public notice that a product is patented, allowing them to claim higher damages in case of infringement as per the marking statute in U.S. patent law (35 U.S.C. § 287(a)). Since we are not specifically interested in VPM pages but, more broadly, in any indication of patent protection of commercial products, we look beyond VPM and search for any web page identifying a

[‡] Explain that this is the sampling methodologies for most DoD evaluations.

[§] A detailed explanation of the procedure adopted to extract the contract identifiers—the Procurement Instrument Identifiers (PIID)—is reported in the appendix. Data about the *government interest statement* of a patent is from PatentsView (17). Data about the awards comes from the Defense Contract Action Data System (DCADS), for the years 1984–2001, and from USAspending.gov, for the years 2001–2018.

clear link between a patent and a product as a sign of invention commercialization. For instance, besides 'traditional' VPM pages, product brochures are a valuable source of information for our purpose. Even though product brochures may not strictly comply with 35 U.S.C. § 287(a), these documents often highlight the existence of one or more patents covering the product advertised.

Concretely, we start by identifying the potential website(s) of each patent assignee in the sample. We search for the assignee legal name on Google.com, Bloomberg.com, and the SBIR program's website and extract domain names from the results of each search. We then search for the patent identification number of the SBIR-related patents in each of the identified websites. This process leads to multiple web pages, from the assignee's website, containing a string of characters that matches one of the patent numbers of interest. At this stage, the string of characters may correspond, say, to a phone number or a patent. If it is a patent, it may not link to a product (e.g., patent numbers reported in SEC forms). To ensure the goodness of the patent-product link, each page has been classified as a *true* or *false positive* either by an automatic classifier developed ad hoc, or via human inspection.[¶]

The approach described so far connects a SBIR contract to a potential VPM-like page. To capture the commercialization potential of a patented invention in a more comprehensive manner, we consider two paths leading to a product, as illustrated by Fig. 1. A *direct path* occurs when a patent acknowledging SBIR support protects a product as identified on a VPM-like page belonging to the patent assignee. The top part of Fig. 1 illustrates this case with an autonomous home floor mopper. The company commercializing the product lists the patents protecting it on its VPM page. One of these patented inventions was first developed in the performance of a SBIR contract awarded by the Army Aviation and Missile Command.

An *indirect path* occurs when the SBIR-funded patent is cited by a subsequent patented invention connected to a commercialized product through a VPM-like page. Given the technical function of patent citations as signals of existing prior knowledge relevant for the new invention (18), we also consider this second case as providing evidence of a link between SBIR funding and the introduction of a final product on the market. The bottom part of Fig. 1 reports the example of a set of noise-canceling headphones. One of the key patents protecting the noise-canceling technology embedded in these headphones lists as relevant prior-art a patented invention realized with the support of an Army SBIR contract awarded in 1993.

Descriptive statistics. Following the approach described above, we first identify the universe of DoD-SBIR-funded patents and then establish if they are directly or indirectly connected to one or more products. The final dataset consists of 2,896 granted patent, assigned to 1,062 distinct companies, and with priority years ranging from 1977 to 2019.[‖] We now turn to presenting some descriptive statistics about patents in the sample.

The patents acknowledge 2,092 different procurement contracts, with 15 percent of the patents reporting the support of multiple awards. About eight percent of the patents are linked to a VPM-like page through a *direct path*, and 17.2 percent through an *indirect path*. Considering the two paths together, 21.5 percent of the patents connect to a VPM-like page.

We augment the base data with contract level information from the Federal Procurement Database System (FPDS). Unsurprisingly, all SBIR contracts connected to the patent in our sample are awarded to perform R&D activity. FPDS data allow us to distinguish between three different stages of R&D efforts, from more fundamental research to developmental activities that are supposed to be closer to technology commercialization. Among our patents, 1,036 acknowledge at least one *basic research* contract; 932 an *applied research* contract; and 568 a *development* contract.

A key characteristic of a SBIR contract is whether it is awarded for a Phase I or a Phase II project. Of the patents in our sample, 1,486 (51.3%) acknowledge at least one Phase I contract, and 595 patents acknowledge exclusively Phase I contracts. A total of 1,723 patents (59.5%) acknowledge instead one or more Phase II contracts. For patents linked to Phase I contracts only, we also determine if the project never reached Phase II or if a Phase II contract exists but the patent simply did not mention it (see the appendix for further explanation). Accounting for Phase I contracts later extended to a Phase II contract not acknowledged in the patent document, we find that 2,374 patents (82.0%) are connected to Phase II funding.

Fig. 2a illustrates that most patents acknowledging support from the DoD SBIR program concern recent years, with the median patent being applied to the USPTO in 2007. In particular, the chart shows a significant increase in patenting activity by DoD-SBIR recipients from 1997 onwards. This pattern partly reflects the growth in overall patenting activity, already noted in scholarly work (19). The temporal distribution of the DoD-SBIR-funded patents also reflects the fact that the law was implemented in 1980, such that we can expect a lower compliance rate in the earlier years of the time window.[**] The commercialization of DoD-SBIR-funded technologies, either *directly* or *indirectly*, appears particularly strong in the central period of the time-window. In the years 1994–2002, about 35-40 percent of funded patents are linked to a product. This percentage is about 25 percent in the preceding period (1986–1993) and in the subsequent one (2003–2011). This temporal trend is not surprising, at least concerning earlier years. Besides lower compliance rate, the Web searches will miss older pages as they are being removed from the Web. Hence, we should interpret these figures cautiously. However, it is worth noting that the counterfactual analysis will compare the commercialization of patents from the same age cohorts.

As Fig. 2b illustrates, the DoD-SBIR-funded patents are concentrated in a few technological fields, reflecting the DoD's R&D needs. A total of 22.6 percent of the patents relate to electrical and electronic technologies; 21.2 percent to the domain of computers and communications; 18.3 percent to chemical; and 17.8 percent to mechanical fields. The proportion of commercialized patents is surprisingly similar across the technological categories (from 17.8 percent to 24.8 percent),

[¶] The classification process is described in more detail in the appendix. The automatic classifier identifies pages that unequivocally link patents and products, such as well-structured VPM pages and product brochures. We manually assess the web pages whose classification is automatically marked as uncertain.

[‖] The data are available at ...

[**] This hypothesis is strengthened by the fact that the reporting of the PIID was made mandatory only later, through the FAR.

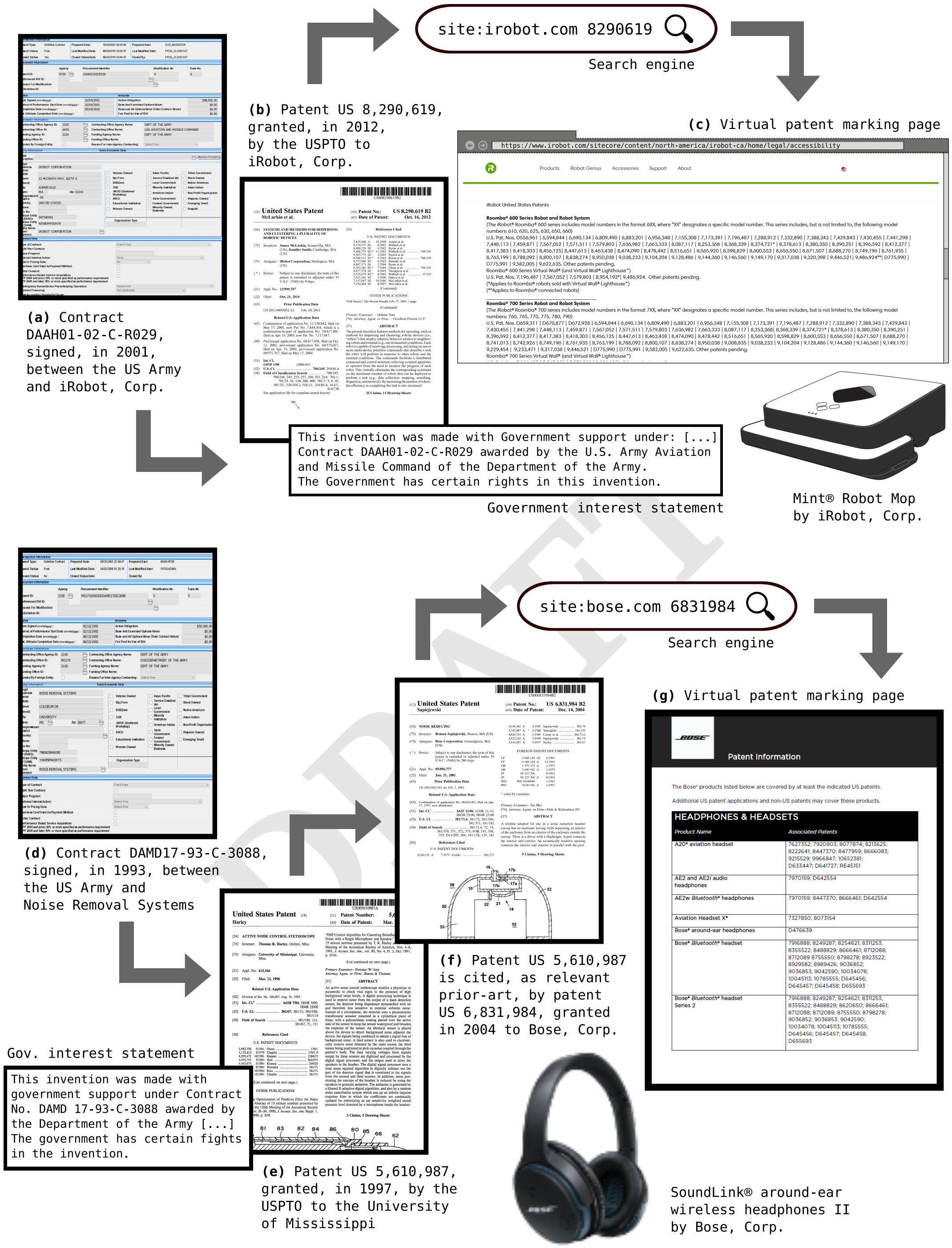


Fig. 1. Two illustrative examples of paths covered by the paper. On top, a *direct path* where, in 2001, the US Army signed the contract No. DAAH01-02-C-R029 with iRobot, Corp. (a). The company applied for a patent, granted in 2012 as US 8,290,619 (b), acknowledging the government's support for this invention. As declared by iRobot on its website (c), this same patent is protecting the company's Mint® Robot Mop, Mint Plus® Robot Mop, and Braava® Robot Mop products. The bottom figure illustrates an *indirect path*. In this case, the contract DAMD17-93-C-3088, signed between the US Army and Noise Removal Systems in 1993 (d), is acknowledged in patent US 5,610,987, granted by the USPTO in 1997 (e). This patent is cited, as relevant prior-art, by patent US 6,831,984, granted to Bose, Corp. in 2004 (f). Bose informs us, through its website (g), that this last patent is protecting products like its SoundLink® around-ear wireless headphones II and A20® aviation headset.

suggesting little technology-specific effects.

Lastly, turning to the spatial distribution, Fig. 2c illustrates that SBIR-funded patents are unevenly concentrated in a few metropolitan areas (MSA) around the United States. This observation is consistent with the geography of innovation literature (e.g., 20). Fig. 2d depicts the commercialization rate of SBIR-funded patents. Looking at the two maps combined suggest no correlation between the capacity of an MSA to attract public funding and its ability to commercialize the technology (Pearson's correlation coefficient of −0.01).

Econometric approach. As mentioned above, the second step in the evaluation of the DoD-SBIR program involves constructing a control set of patents with similar characteristics to the SBIR-funded patents in the sample. For each treated patent, we select up to three controls from a pool of patents assigned to a private company classified as small business by the USPTO and applied between 1984 and 2018.[††] Each of the selected control patents shares the main USPC technological class and the priority year of its respective treated patent. Our final control set consists of 4,622 granted patents, assigned to 3,895 distinct companies. By design, they have priority years distributed within the same time frame as the SBIR-funded ones. Of these control patents, 6.0 percent are *directly* linked to a VPM-like page, while 15.1 percent of them are linked to one of these web documents only *indirectly*. All in all, 18.5 percent of the control patents are linked to a VPM-like page, either directly or indirectly.

The third step involves comparing the commercialization performance of the treated and control patents using standard regression analyses. More specifically, we estimate the following linear probability model:

$$\Pi_i = \beta_0 + \beta_1 \cdot \text{SBIR}_i + \mathbf{X_i} \cdot \beta + \gamma_i + \delta_i + \varepsilon_i. \qquad [1]$$

Π_i is the main outcome variable. It takes the value 1 if patent i is linked to a product through a VPM-like page, and 0 otherwise. We construct three different versions of Π_i, based on the commercialization path: direct, indirect, or any of the two. The variable SBIR_i is the variable of interest. It takes the value 1 if patent i acknowledges funding from the DoD SBIR program, and 0 otherwise. The vector $\mathbf{X_i}$ includes patent-level control variables that might correlate with the commercialization outcome. In particular, following the extant patent literature, we control for (the log of): the number of independent claims in the patent (`claims`); the number of citations made to other patents (`bwd_cit`) and to the non-patent literature (`npl_cit`); the number of citations received by patent i in the first three years after its application date (`fwd_cit`); and the geographical family size of patent i, i.e., the number of countries in which patent protection is sought (`geo_fam`). Lastly, the model includes the year of first priority, γ_i, and USPC patent class, δ_i, fixed effects, to control for time- and technology-dependent factors.

Fig. 3c reports descriptive statistics for control variables. On average, control and treated patents appear to have fairly similar values.

[††] The data providing the information about the type of entity comes from the USPTO's Patent Examination Research Dataset (PatEx) database (21); see also https://www.uspto.gov/ip-policy/economic-research/research-datasets/patent-examination-research-dataset-public-pair. The assignee is classified as a small business based on the type of maintenance fee paid. Small enterprises pay a reduced fee. Patents assigned to an assignee whose name recur also between the SBIR-funded patents, or within the list of SBIR recipients, have been excluded.

In addition to the baseline regression described above, we exploit the contract-level information to analyze whether specific characteristics of a SBIR contract disproportionately affect the probability of commercialization of the inventions arising from that contract. In particular, we focus on the stage of the R&D work procured by DoD (basic, applied, or developmental research stage) and on the phase of the contract (Phase I or Phase II). Finally, for some robustness analysis, we also collect additional information about the commercialization timing, by proxying the commercialization year of a final product as the earliest creation date of each VPM-like page. This information will offer insights on the time-lag that it takes for an invention to reach the consumer market.

Results

The top part of Fig. 3a depicts the results of the baseline regression model for the three outcome variables, focusing of the coefficient β_1. As regression results (1a)–(1c) show, an invention introduced with the support of a DoD-SBIR contract has a higher likelihood of commercialization than a control invention. The effect appears to be sizable: SBIR support increases the probability of a commercial product introduction by about 17 percent (any path). We find a similar effect if we consider only *direct* (1b) or only *indirect paths* (1c). As discussed in section Data and methods, we observe an *indirect path* when a patented invention connected to a product cites one of the focal patents as relevant prior art. One might argue that a positive effect of SBIR support on *indirect paths* provides only weak evidence of a decisive impact of public support on commercialization. However, a more careful look at the data suggests a different interpretation. We find that for about 40 percent of the patents that are linked to a product indirectly, the connecting citation is a self-citation, i.e., it comes from a patent applied for by the same assignee as the focal patent. Accordingly, we run the baseline model on two distinct sets of focal patents: patents that did receive at least one self-citation from a subsequent patent and patents that did not receive any self-citation. Interestingly, the effect of SBIR support on commercialization disappears—and even turns negative—when we consider patents with no ensuing self-citations. By contrast, the results are in line with baseline model (1c) when we consider exclusively patents with self-citations, with a 3.1 percentage points higher probability of commercialization for SBIR supported patents (see the appendix for an in-depth description of this analysis). This finding suggests that the long-term, indirect effect on commercialization is achieved only if the company that received SBIR support is actively involved with further technological developments and, hence, only if the *indirect path* is closely connected to the SBIR funding. This finding is consistent with an 'input additionality' effect of the SBIR program.

All in all, the results so far confirm a strong and positive effect of SBIR funding on commercialization outcomes. To better understand the nature of this effect, we evaluate the importance of specific contract characteristics. We start by considering the stage of the R&D work for which a contract is awarded. To do so, we split the sample of treated patents in three groups, basic, applied, or developmental R&D, based on the features of the contract connected to each invention. We then couple each of the patents in these groups with its respective control patents and run the baseline model

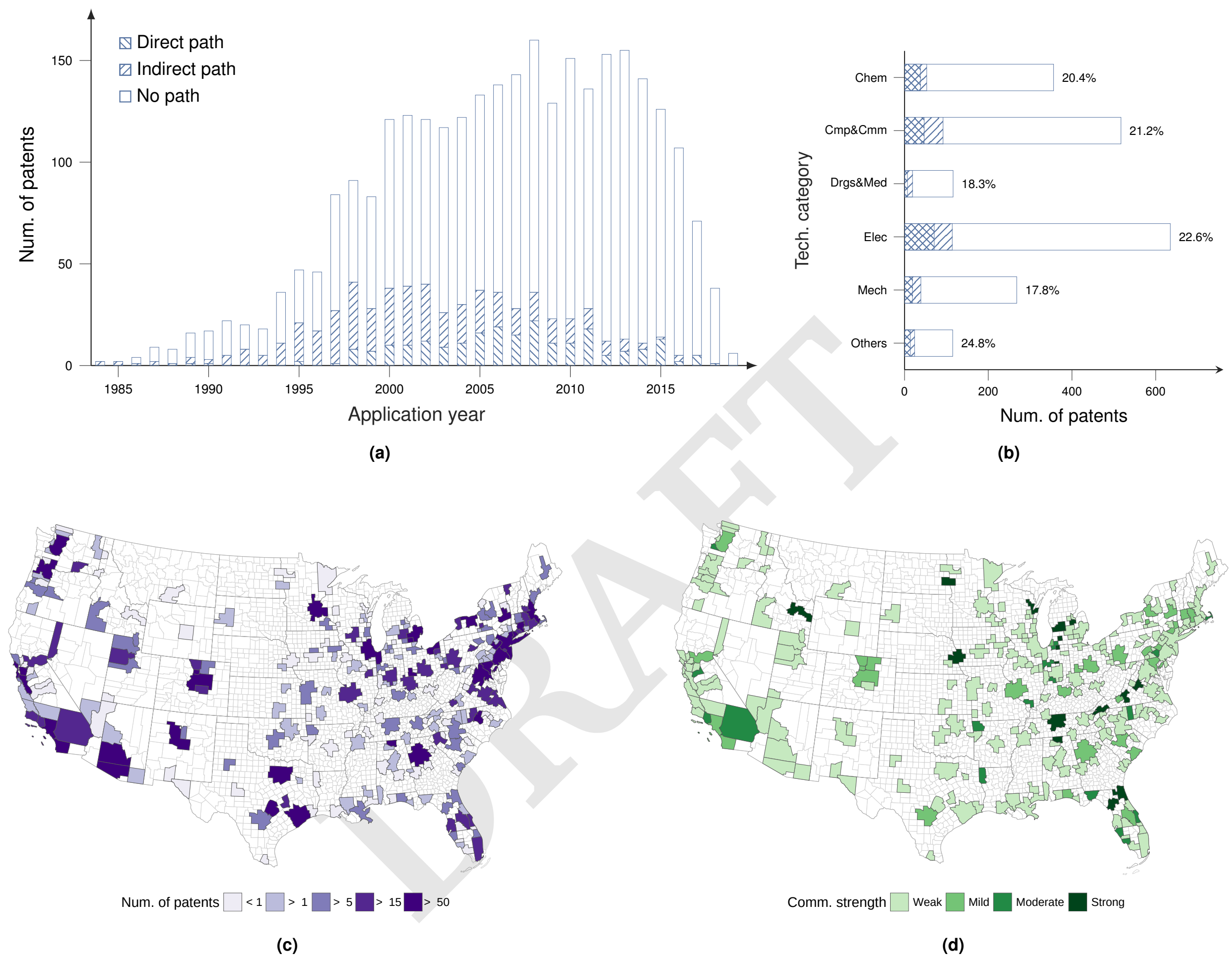


Fig. 2. Descriptive statistics (colored figure online). **(a)** Distribution of SBIR-funded patents by patent's application year. The figure distinguishes between patents for which we did not find any commercialization trace, those directly protecting a product, and those cited by a product-protecting patent. Notice that a patent both directly and indirectly linked to a VPM-like page is counted among the direct paths. A similar figure, for the control patents, can be found in the appendix. **(b)** Distribution of SBIR-funded patents by patent's NBER technological category (Chemical; Computers & Communications; Drugs & Medical; Electrical & Electronic; Mechanical; Others, respectively). The percentage reported represents the fraction of product-protecting patents over the total number of SBIR-funded patents in each technological category. **(c)** Spatial distribution of SBIR-funded patents by U.S. Metropolitan Statistical Area (MSA). **(d)** Spatial distribution of the *commercialization capacity* index (CCI) by MSA. The CCI measures the ability of each metropolitan area to commercialize SBIR-funded science and is defined as $\text{CCI} = (\text{CP}_c/\text{FP}_c)\big/(\sum_c \text{CP}_c / \sum_c \text{FP}_c)$; where $\text{CP} =$ number of patents linked to a product and $\text{FP} =$ number of patents funded by the SBIR program. In the maps, only the conterminous United States is reported; non-metropolitan counties are colored white; and for each patent, a fraction of it has been assigned to a given MSA proportionally to the share of its inventors resident in such metropolitan area. Note that less than 1.5% of the patents included in our data do not belong to any MSA.

on each sample separately. Fig. 3a reports the summary results of these regressions for the three outcome variables. Focusing our attention on patents connected to *basic R&D* contracts, the effect of the SBIR support on direct or indirect commercialization outcomes is never significantly different from zero (models (2a)–(2c)). Receiving an *applied R&D* contract increases the commercialization likelihood (3a), but model (3c) suggests that this effect is driven primarily by *indirect paths*. SBIR-supported inventions have a 4.9 percent higher likelihood to be indirectly connected to a product, whereas the effect on *direct paths* only is not statistically significantly different from zero, see model (3b). Looking at patents connected to *development R&D* contracts, our data show a strong positive effect for both *direct* and *indirect paths* to commercial products (models (4a)–(4c)). Overall, the results of this split sample analysis suggest that the impact of SBIR funding increases with the R&D stages. The more applied the stage of the R&D activity that led to patenting, the higher the impact of public support on the commercialization likelihood of a specific invention.

Another key characteristic of SBIR contracts is whether they relate to a Phase I or a Phase II project. As discussed above, only successful and promising Phase I projects have the opportunity to receive Phase II funding. It allows the recipient to further develop the ideas and technologies generated during the initial phase. Therefore, by design, Phase II projects are closer to commercialization. In addition, the bulk of the funding that successful applicants receive arrives in Phase II, where the award size is an order of magnitude larger than in Phase I. If the SBIR program was indeed effective at spurring commercialization, we should expect it to be especially true for Phase II projects. The results of models (5a)–(5c) and (6a)–(6c) in Fig. 3a contrast the impact of the two phases. Focusing on Phase I projects that never reached Phase II, the difference between the treated and the control group is never statistically different from zero. By contrast, the impact is perfectly in line with the baseline models once we consider only the patents linked to projects that obtained Phase II funding. These results seem to confirm the effectiveness of the SBIR program. Phase I projects are awarded to assess both the capacity of an SME to perform R&D and the quality of an innovative idea; therefore, the likelihood for an invention generated by a Phase I project to reach the commercialization stage is not particularly higher than for a comparable but privately-funded invention. However, through this preliminary stage, it seems that DoD agencies acquire enough information to provide adequate support to inventions with higher commercialization potential than non-SBIR comparable inventions. These results are in line with others reported in the appendix. Compare patents acknowledging at least a Phase I contract with these acknowledging at least a Phase II contract, the former group exhibits weaker commercialization potential than the latter.

To shed more light on the mechanism behind the results, we exploit a policy change in the design of SBIR that put greater focus on commercialization. With the Small Business Reauthorization Act of 2000 (§110), the U.S. Congress demanded the Small Business Administration "to provide for the requirement of a succinct commercialization plan with each application for a Phase II award that is moving toward commercialization" (22). Specifically for the DoD, the Act also introduced the Phase II Enhancement policy—also known as Phase II Plus—to further encourage the transition of SBIR research into DoD acquisition programs as well as the private sector (13). Under this policy, a Phase II recipient can receive additional SBIR funds matching private or public financing the company obtains from non-SBIR sources. Both these changes affected the implementation of Phase II, but not Phase I, projects and provided additional emphasis on the commercialization goals of the program. Interestingly, these adjustments had limited impact on the technical merit or the scientific focus of the projects selected for Phase II. We exploit the latter fact to provide tentative evidence on whether the positive impact of the program on commercialization outcomes stems from a pure selection effect, i.e., DoD agencies simply selecting the projects with the highest commercialization potential, or from the support and the explicit push towards commercialization offered by the program.

We adopt a difference-in-differences (DiD) estimator and focus on SBIR-funded patents awarded in the years immediately before and after this policy change (1996–2005). More specifically, we assess whether Phase II-related patents connected to SBIR awards signed after the year 2000 have a higher likelihood to be directly linked to a commercial product than Phase II patents connected to pre-2000 contracts, using Phase I-related patents as the control group. If the results were entirely driven by selection, we should not observe any effect of the policy change on the commercialization likelihood. Tab. 3b reports the results of the DiD analysis. As the table shows, our main variable of interest, the interaction term Phase II $\times$ Post 2000, is positive and significant. In other words, it seems that the additional push towards commercialization introduced in the year 2000 indeed lead to a higher commercialization propensity of the average Phase II-related patent.

Overall, the results support the view that the SBIR program is quite effective at stimulating the commercialization and transfer of new inventions to the final consumers. SBIR-backed patented inventions have a higher likelihood to end up in commercial products than similar inventions developed by the private sector without government support. So far, the results are silent on the timing of commercialization. The government might simply provide more *patient capital* compared to the private sector (23). Hence, the difference in the commercialization rate may come from fully privately-funded projects that are abandoned early because of their lower potential, while similar publicly-funded projects move forward with government money. To explore this possibility, we exploit the data on patents connected to products to look into the *time-to-market* of each invention. As explained in more detail in the appendix, we proxy the commercialization year of a product with the earliest date of creation of any of the VPM-like pages reporting the patent-product link. We then computed the *time-to-market* of each patent as the number of years between the patent filing date and the product commercialization. Even though our proxy for the commercialization timing is likely to be noisy, Fig. 3d offers a preliminary view of the direct and indirect commercialization lag for treated and control patents. The chart shows no striking differences between SBIR-funded and control inventions, in terms of time-to-market. Looking at *direct paths*, for the average SBIR-funded invention it takes about eight years to reach the final consumers, whereas it takes seven years for control inventions. However, as reported in the appendix (p. XX), this difference is not statistically significant.

The picture is very similar for the indirect paths, for which the commercialization path is 14 years long, on average.

In a separate analysis (reported here), we have identified which of the treated and control patents were 'green,' in the sense that they relate to climate change mitigation technologies (Y02 CPC technological sub-class). We found 6.63 percent green treated patents and 8.42 percent green control patent. Overall, the probability of commercialization of green patents is 3.7 to 4.6 percentage points lower than non-green patents. The difficulty in commercializing green inventions is typically seen as one justification for public support (24). However, a regression model that interacts green patents with SBIR support leads to inconclusive results. We do not find clear evidence that public support hindered or helped commercialization of green inventions.

Discussion

We have proposed a novel method for evaluating the performance of the SBIR program by the DoD. The method involves searching the web for traces of commercialization of SBIR-funded patents. This approach is part of a broader trend in the literature of using internet data for economic research (e.g., 25, 26), (26).

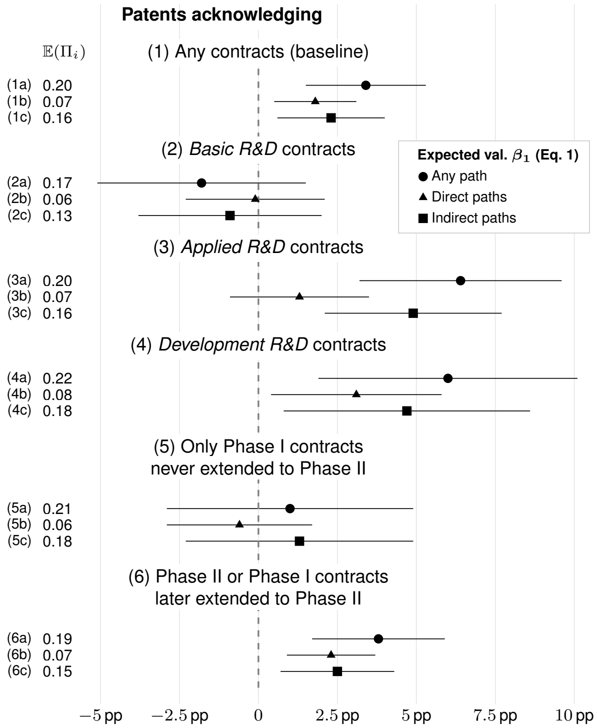
The present work focuses on projects that have led to patents. It does not consider the set of SBIR-funded projects that did not lead to patents. Although such data are directly available from the relevant agencies, performing a counterfactual analysis to evaluate the success rate of SBIR-funded vs. privately-funded projects is particularly challenging, for it requires observing the patent outcome of private projects, for which representative data are notoriously difficult to access.

Having collected information posted on companies' websites, the analysis could be subject to a reporting bias. Specifically, SBIR recipients could be more likely to publish information online than non-SBIR recipients, for instance, to please the program manager or signal the DoD funding to investors. Although such bias is presumably less severe than in surveys, we cannot guarantee that our estimates do not suffer from it. In a robustness test, we have performed the analyses exclusively using commercialization as observed from 'proper' VPM pages (excluding product brochures and other web pages)—because these web pages do not mention DoD funding. The results remain qualitatively similar.

Finally, although we observe a significant effect of SBIR funding on commercialization, the magnitude of the impact is difficult to assess for a lack of comparable studies. We hope future research will exploit the method to evaluate other such programs or commercialization outcomes.

ACKNOWLEDGMENTS. We thank the EuroTech Universities Alliance for sponsoring this work. C.B. was supported by the European Union's Marie Sklodowska-Curie programme for the project *Insights on the "Real Impact" of Science* (H2020 MSCA-COFUND-2016 Action, Grant Agreement No 754462). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

1. PA Geroski, Procurement policy as a tool of industrial policy. *Int. Rev. Appl. Econ.* **4**, 182–198 (1990).
2. J Lerner, *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed and What to Do About It.* (Princeton University Press), (2009).
3. E Moretti, C Steinwender, J Van Reenen, The intellectual spoils of war? Defense R&D, productivity and international spillovers, (National Bureau of Economic Research), Working Paper 26483 (2019).
4. VW Ruttan, *Is War Necessary for Economic Growth? Military Procurement and Technology Development.* (Oxford University Press, Oxford, UK), (2006).
5. M Mazzucato, *The Entrepreneurial State: Debunking Public vs. Private Sector Myths.* (Anthem Press, London, UK), (2013).
6. M Mazzucato, *Mission Economy: A Moonshot Guide to Changing Capitalism.* (Allen Lane, London, UK), (2021).
7. DB Audretsch, Standing on the shoulders of midgets: The U.S. Small Business Innovation Research program (SBIR). *Small Bus. Econ.* **20**, 129–135 (2003).
8. National Research Council, *SBIR at the Department of Defense.* (The National Academies Press, Washington, DC), (2014).
9. ST Howell, Financing innovation: Evidence from R&D grants. *Am. Econ. Rev.* **107**, 1136–1164 (2017).
10. National Academies of Sciences, Engineering, and Medicine, *Review of the SBIR and STTR Programs at the Department of Energy.* (The National Academies Press, Washington, DC), (2020).
11. M Schwartz, JF Sargent, Jr., CT Mann, Defense acquisitions: How and where DOD spends its contracting dollars, Congressional Research Service R44010 (2018).
12. National Research Council, *An Assessment of the SBIR Program at the Department of Defense.* (The National Academies Press, Washington, DC), (2009).
13. National Research Council, *Revisiting the Department of Defense SBIR Fast Track Initiative.* (The National Academies Press, Washington, DC), (2009).
14. United States Government Accountability Office, Small business innovation research: DOD's program has developed some technologies that support military users, but lacks comprehensive data on transition outcomes, Testimony Before the House Committee on Small Business GAO-14-748T (2014).
15. G de Rassenfosse, Notice failure revisited: Evidence on the use of virtual patent marking, (National Bureau of Economic Research), Working Paper 24288 (2018).
16. G de Rassenfosse, A Jaffe, E Raiteri, The procurement of innovation by the U.S. government. *PLOS ONE* **14**, e0218927 (2019).
17. C Jones, S Madhavan, Patentsview: Government interest extraction and processing – version 2.0, (American Institutes for Research), Mimeo (2020).
18. AB Jaffe, G de Rassenfosse, Patent citation data in social science research: Overview and best practices. *J. Assoc. for Inf. Sci. Technol.* **68**, 1360–1374 (2017).
19. C Fink, M Khan, H Zhou, Exploring the worldwide patent surge. *Econ. Innov. New Technol.* **25**, 114–142 (2016).
20. MP Feldman, DF Kogler, Stylized facts in the geography of innovation in *Handbook of The Economics of Innovation*, Handbook of the Economics of Innovation, eds. BH Hall, N Rosenberg. (North-Holland) Vol. 1, pp. 381–410 (2010).
21. SJH Graham, AC Marco, R Miller, The USPTO patent examination research dataset: A window on the process of patent examination, USPTO Economic Working Paper 2015-4 (2015).
22. U.S. Congress, Small business reauthorization act of 2000 (2000) HR 5667. Pub. L. 106-554, Appendix I.
23. M Mazzucato, Innovation, the State and patient capital. *The Polit. Q.* **86**, 98–118 (2015).
24. R Owen, G Brennan, F Lyon, Enabling investment for the transition to a low carbon economy: Government policy to finance early stage green innovation. *Curr. Opin. Environ. Sustain.* **31**, 137–145 (2018).
25. B Edelman, Using internet data for economic research. *J. Econ. Perspectives* **26**, 189–206 (2012).
26. SK Arora, S Kelley, S Madhavan, Building a sample frame of SMEs using patent, search engine, and website data. *J. Off. Stat.* **37**, 1–30 (2021).

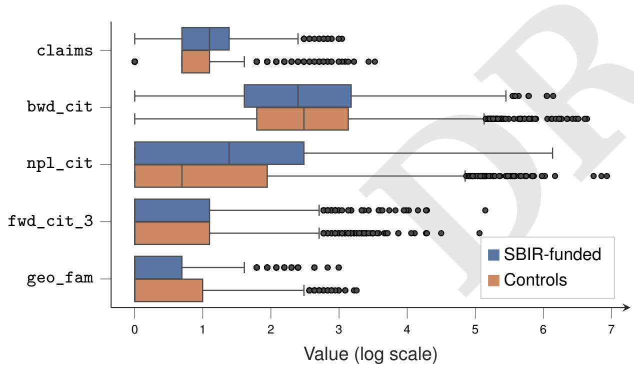


(a)

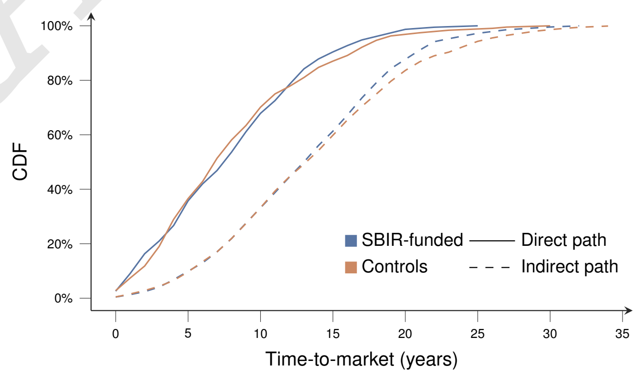
Dep. var.:	OLS		Probit	
Direct path	(1)	(2)	(3)	(4)
Phase II	0.015	-0.034	0.032	-0.053
	(0.021)	(0.031)	(0.035)	(0.050)
Post 2000	0.050*	-0.023	0.060	-0.058
	(0.029)	(0.047)	(0.037)	(0.072)
Phase II × Post 2000		0.083*		0.131*
		(0.046)		(0.071)
$\log$(claims)	0.002	0.002	-0.005	-0.004
	(0.013)	(0.013)	(0.019)	(0.019)
$\log$(bwd_cit)	0.005	0.006	0.006	0.007
	(0.009)	(0.009)	(0.013)	(0.013)
$\log$(npl_cit)	0.012	0.012	0.016	0.017
	(0.008)	(0.008)	(0.010)	(0.010)
$\log$(geo_fam)	0.002	0.003	0.004	0.006
	(0.016)	(0.016)	(0.020)	(0.020)
$\log$(fwd_cit)	0.022*	0.021*	0.030**	0.028*
	(0.011)	(0.011)	(0.015)	(0.014)
Constant	-0.030	0.010		
	(0.113)	(0.113)		
Observations	1422	1422	801	801
R^2	0.239	0.241		
Pseudo R^2			0.147	0.151

Standard errors in parentheses.
$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.010$
$\mathbb{E}(\Pi_i) = 0.10$.
$\mathbb{E}(\Pi_i|\text{Award pre 2000}) = 0.09$.

(b)



(c)



(d)

Fig. 3. Results of the empirical analysis (colored figure online). **(a)** Effect of the SBIR/STTR program on the commercialization likelihood of a patent. The following models are reported: (1) Any patent; (2) Patents acknowledging at least a *basic R&D* contract; (3) Patents acknowledging at least an *applied R&D* contract; (4) Patents acknowledging at least a *development R&D* contract; (5) Patents acknowledging only Phase I contracts never extended to Phase II; (6) Patents acknowledging at least a Phase II contract or a Phase I later extended to Phase II. For each model, (i) the dot point considers any path; (ii) the triangle point only *direct paths*; (iii) the square point only *indirect paths*. Points represent the betas of the treatment variable estimated through a Linear Probability model (β_1), while bars report the corresponding 95% Confidence Intervals. The gray grid is set to 2.5 percent points distance. On the left, the average value of the dependent variable of each model is reported. In the appendix, it is possible to find the corresponding regression tables, in full detail, as well as the results for Probit models, corresponding to each model here discussed. Notice that some patents have been zero-weighted in any of the models except for (1a)–(1c) since they can be linked, solely, to contracts with characteristics other than the one considered by the specific model. Moreover, since a patent can acknowledge more contracts at the same time, the classification in the three R&D kinds or in the two SBIR Phases is not exclusive. **(b)** Table with the results of the policy-change regression. Only SBIR-funded patents, funded by contracts signed in 1996–2005, included. Phase II contracts include also Phase I ones later extended to the second phase of the SBIR/STTR program. For these last contracts, we considered the extending contract date. In note, the average value of the dependent variable is reported, both considering all the patents included in the regression and only these patents acknowledging a procurement contract signed not later than the year 2000. **(c)** Distribution of patents' quality indicators used as *control variables* in the regression exercises below. For each variable, the box-plot on top relates to the SBIR-funded patents, while the other to the control ones. **(d)** Time-to-market. We have been able to date 193 SBIR-funded patents and 216 controls directly linked to a VPM-like page. While for patents indirectly linked to a VPM-like page, we attributed a date to 455 SBIR-funded ones and 641 controls.

Guide to using this template on Overleaf

Please note that whilst this template provides a preview of the typeset manuscript for submission, to help in this preparation, it will not necessarily be the final publication layout. For more detailed information please see the PNAS Information for Authors.

Author Affiliations. Individual authors must link their ORCID account to their PNAS account at www.pnascentral.org. For proper authentication, authors must provide their ORCID at submission and are not permitted to add ORCIDs on proofs.

Submitting Manuscripts. All authors must submit their articles at PNAScentral. If you are using Overleaf to write your article, you can use the "Submit to PNAS" option in the top bar of the editor window.

Manuscript Length. A standard 6-page article is approximately 4,000 words, 50 references, and 4 medium-size graphical elements (i.e., figures and tables). The preferred length of articles remains at 6 pages, but PNAS will allow articles up to a maximum of 12 pages.

Data Archival. PNAS must be able to archive the data essential to a published article. Where such archiving is not possible, deposition of data in public databases, such as GenBank, ArrayExpress, Protein Data Bank, Unidata, and others outlined in the Information for Authors, is acceptable.

Language-Editing Services. Prior to submission, authors who believe their manuscripts would benefit from professional editing are encouraged to use a language-editing service (see list at www.pnas.org/page/authors/language-editing). PNAS does not take responsibility for or endorse these services, and their use has no bearing on acceptance of a manuscript for publication.

Digital Figures. EPS, high-resolution PDF, and PowerPoint are preferred formats for figures that will be used in the main manuscript. Color images must be in RGB (red, green, blue) mode. Include the font files for any text.

Images must be provided at final size, preferably 1 column width (8.7cm). Figures wider than 1 column should be sized to 11.4cm or 17.8cm wide. Numbers, letters, and symbols should be no smaller than 6 points (2mm) and no larger than 12 points (6mm) after reduction and must be consistent.