Analysis of the U.S. Department of Energy's Energy Efficiency & Renewable Energy and Fossil Energy SBIR Programs

Final Report

February, 2019

Prepared by: Sabrina T. Howell, New York University Stern School of Business

Prepared for: Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy

Acknowledgements

The research underlying this report was possible because of the determined efforts of Yaw Agyeman, Ken Alston, Jeff Dowd, Matthew Dunne, Carla Frisch, Carl Hebron, Tina Kaarsberg, Teryn Norris, and Jamie Vernon, all currently or formerly at the Department of Energy. In particular, Tina Kaarsberg provided tireless support and SBIR information. Matt Dunne overcame substantial administrative hurdles in the final phases of data access. Carl Hebron took a great deal of time from his SBIR office responsibilities to provide the data itself. Charles Wessner (Georgetown University), Irwin Feller (Penn State), Albert Link (University of North Carolina at Greensboro), Lori Lewis (Analytical Evaluation Consultants, LLC.), and Robin Gaster (Innovation Ecologies Inc.) provided review of an earlier 2015 analysis of SBIR patenting that is the basis for part of this report. Finally, the author is grateful to her PhD advisors: Joseph Aldy (Harvard Kennedy School), Raj Chetty (Harvard University), David Scharfstein (Harvard Business School), and especially Josh Lerner (Harvard Business School), whose ideas and support have been central to this project from its inception.

The Census analysis was undertaken jointly with J. David Brown of the U.S. Census Bureau. I am very grateful to him, to Shawn Klimek, and to others at the Census Bureau who made this part of the analysis possible.

Contents

List of Tables

List of Figures

Executive Summary

This report analyzes the impact of the Department of Energy's (DOE) Small Business Innovation Research (SBIR) program, which awards research and development (R&D) grants to small firms. Specifically, it examines two applied programs at DOE, Energy Efficiency and Renewable Energy (EERE) and Fossil Energy (FE). With detailed application and ranking data between 1995 and 2013, it uses a regression discontinuity design to establish a causal relationship between the grants and recipient outcomes.

Background

Although governments around the world use grants to subsidize private sector R&D, there is relatively little rigorous quantitative evidence about these subsidies' effectiveness. This study focuses on the Congressionally-mandated SBIR program, which provides grants across the U.S. government to small businesses that are developing new commercially viable technologies. The SBIR program consists of two phases. Firms first apply for a (up to $150,000 in 2013) Phase 1 award, and then Phase 1 winners are eligible to apply for a (up to $1 million in 2013) Phase 2 award.

Study Purpose and Scope

The purpose of this report is to summarize the author's analysis of DOE's EERE and FE SBIR programs, which was conducted to evaluate the impact of the grants on recipients within these programs. The report focuses on specific firm outcomes: innovation (measured using patents), employees, payroll, revenue, and wages. Howell (2017) also examines the grant effects on applicants' subsequent venture capital financing. One goal of the SBIR program is commercialization. While revenue is used here as a proxy for commercial activity, data on technology deployment were not available. Patenting activity, in addition to providing a measure of innovation, is also viewed by the DOE as an intermediate outcome towards achieving the commercialization objective. For an evaluation of the DOE SBIR program's operation, the reader is directed to NAS (2016).

Methods

This report uses data on over 4,500 firms that applied to all competitions for SBIR awards in the two applied R&D offices at DOE, EERE and FE, between 1995 and 2013. The applicant firms are matched to patent data from the U.S. Patent and Trademark Office and to firm growth data from the U.S. Census Bureau. This report draws from analysis in Howell (2017) and from forthcoming work with J. David Brown of the U.S. Census Bureau's Center for Economic Studies.

The estimation strategy for SBIR Phase 1 data, called a regression discontinuity design, makes use of the ranks that DOE assigns to firms within competitions. It compares firms ranked immediately above and below the award cutoff. These just-winners and just-non-winners are shown to be ex-ante similar, such that comparing them is equivalent to a local random experiment. In this way, the regression discontinuity design enables causal effects to be established. As there are few Phase 2 applicants per competition, the Phase 2 analytical approach is a conventional regression - essentially a difference of means, with no control for rank. If non-winning firms are on average lower quality than winning firms, this will bias the Phase 2 results towards finding positive effects.

Results

The $150,000 Phase 1 award has powerful effects on innovation and firm growth. Receiving a Phase 1 grant increases a firm's subsequent patents weighted by future citations by about 250 percent, relative to an average of 12 cite-weighted patents. Note that it is not possible to tie a firm's patents to the specific research that the firm conducted with its SBIR grant. We do not observe how firms use the grant, nor do we observe the technologies underlying the patents. The estimation strategy permits us to conclude that the grant causes the difference in patenting between winning and non-winning applicants.

The effect of the Phase 1 grant on patenting is larger among young firms and among firms that have no or few existing patents. It also declines in the number of previous SBIR awards a firm has won; for each additional previous SBIR award, the effect of an award on cite-weighted patents declines by 20 percent.

The $1 million Phase 2 grant doubles a firm's cite-weighted patents, relative to a mean of 20. This effect is much smaller than the Phase 1 effect on a per-grant-dollar basis. Also, the effect of Phase 2 falls and loses significance when the sample includes previous winners.

There are also multiple positive Phase 1 effects related to firm growth, including wages, employment, and payroll. Using a subset of the applicant firms that were matched to U.S. Census Bureau data, the Phase 1 grant is shown to lead firms to have about 19 percent more employees relative to the year of application than they would have had otherwise. The similar result for payroll is 29 percent, and the effect on revenue is 15 percent. The grant also increases firm wages by about 11 percent. Like the effects on innovation, the Phase

1 grant effect is larger for younger firms. It is also larger for smaller firms. The Phase 2 grant was not found to have any statistically significant effects on firm employment, payroll, revenue, or wage growth. Note that as with the patent data, it is not possible to tie specific employees or portions of revenue to the SBIR grant. Again, the estimation strategy permits us to identify the effects on firm growth as being caused (perhaps indirectly) by the SBIR grant.

1 Introduction

This section first discusses the economic rationale for evaluating the DOE SBIR program, and then states the primary research questions. Section 1.2 provides background on the SBIR program at DOE, and Section 1.3 describes the data that are used. Section 1.4 analyzes how the applicant firms cluster geographically.

1.1 Economic Rationale for Analysis

Governments worldwide subsidize new, high-tech ventures in an effort to spur innovation. The largest single grant program for high-tech entrepreneurs in the U.S. is the SBIR program, which disbursed around $2.5 billion in 2017.[1] In addition to the 11-agency federal SBIR, many U.S. states have similar programs, including New Mexico, Ohio, and Hawaii. Parallels overseas include the UK's Innovation Investment Fund, China's Innofund, Israel's Chief Scientist incubator program, Germany's Mikromezzaninfonds and ZIM, Finland's Tekes, Russia's Skolkovo Foundation, and Chile's InnovaChile. Many of these programs deliberately mimic the U.S. SBIR program.

One rationale for such subsidies is that the private sector does not internalize the social benefits of innovation.[2] Another is that financial frictions lead to underinvestment in early stage R&D (Hall and Lerner 2010). Grants might increase investment if given to startups that face excessively costly external finance. For example, it is extremely difficult for potential investors to conduct due diligence on new energy technologies. Early stage startups also often do not have collateralizable assets with which to raise debt. Yet critics contend that government R&D subsidies may be ineffective because they displace private investment that would have occurred in the absence of the subsidy, or because they allocate funds inefficiently (Wallsten 2000, Lerner 2009).

Despite opposing theoretical arguments, there is relatively little empirical evidence about the effectiveness of direct R&D subsidies.[3] Existing research has mainly studied non-U.S.

[1] See https://www.sbir.gov/reports/state-summary?year%5B%5D=2017.

[2] For evidence that startups contribute disproportionately to economic growth, see Akcigit and Kerr (2011), Haltiwanger et al. (2013), and Audretsch, Keilbach and Lehmann (2006).

[3] There is substantial work on R&D tax credits, including Hall (1993), Mamuneas & Nadiri (1996), Hall & Van Reenen (2000), Bloom, Griffith and Van Reenen (2002), Clausen (2009), Rao (2016), Wu (2008), Wilson (2009), Dechezleprêtre, Einiö, Martin, Nguyen & Van Reenen (2016), and Balsmeier, Kurakina & Fleming (2018).

programs and come to disparate conclusions, such as Lerner (2000), Wallsten (2000), Lach (2002), Takalo, Tanayama and Toivanen (2013), and Almus and Czarnitzki (2003).[4] The challenge has been that in the absence of a randomized experiment, it is difficult to identify a counterfactual: What would happen to the firms if they didn't get the grants? This report details an approach that approximates a random experiment: comparing applicant firms on either side of the award cutoff, while controlling for the rank that determines their award status.

1.2 Background on the SBIR Program at DOE

The SBIR program has two "Phases." Phase 1 grants fund proof-of-concept work intended to last nine months, and are up to \$150,000 (in 2013 the amount has increased stepwise from \$50,000 in 1983 to \$200,000 in 2019). Firms must demonstrate progress on their Phase 1 projects to win the up to \$1 million Phase 2 grants in 2013. The Phase 2 grant size has also increased stepwise from \$500,000 to \$1,100,000 in 2019 over the life of the program. Phase 2 funds more extensive or later stage demonstrations, and the money is awarded over two years.[5]

Congress first authorized the SBIR program in 1982 to strengthen the U.S. high technology sector and support small firms.[6] As of 2017, 11 federal agencies must allocate 3.2 percent of their extramural R&D budgets (i.e., R&D carried out by non-federal scientists with federal funds) to the SBIR program. There is no required private cost sharing in the SBIR program. Also, the government neither takes equity in the firm nor assumes intellectual property (IP) rights.[7] Eligible applicants are for-profit, U.S.-based, and at least 51 percent American-owned firms with fewer than 500 employees. Although the SBIR grant is non-dilutive, it is

[4]Evaluations of R&D subsidies include Czarnitzki and Lopes-Bento (2012), Serrano-Velarde (2008), Busom (2000), Duguet (2004), González et al. (2005), González and Pazó (2008), Blasio, Fantino and Pellegrini (2014), and Henningsen et al. (2014). In the U.S., Nemet and Kammen (2007) find little evidence of crowding out in federal energy R&D, but Popp and Newell (2009) do. Link and Scott (2010) use SBIR Phase 2 awardee survey data to analyze the likelihood of commercialization. To my knowledge, only the working papers by Zhao and Ziedonis (2013) and Bronzini and Iachini (2011) use data on applicants to R&D incentive programs. The former evaluates a Michigan loan program (N=104), and the latter grants to large firms in Northern Italy (N=171). Both programs have private cost sharing, which SBIR does not. Other researchers have used RD to evaluate grants to university researchers, such as Jacob and Lefgren (2011) and Benavente et al. (2012).

[5]Phase 3 is commercialization of the technology. It is ineligible for SBIR funds except when agencies are purchasing the technology, which does not occur at DOE but is common at the Department of Defense.

[6]For more background on the origin of SBIR see: https://www.sbir.gov/birth-and-history-of-the-sbir-program.

[7]However, If the private company does not explicitly protect its inventions, the government can exert march-in rights (35 U.S.C. §203 in Bayh-Dole Act, P.L. 96-517).

not costless. In a few interviews conducted by the author, investors and startup founders described the application and reporting process as onerous. Applying for an SBIR grant can require two months of 1-2 employees working full time.

The DOE assigns SBIR responsibility to program offices responsible for a broad technology area. Two of the applied R&D offices in DOE are Fossil Energy (FE) and Energy Efficiency and Renewable Energy (EERE).[8] Each year, DOE officials in these technology-specific program offices develop a series of competitions. For example, the Solar Energy Technologies office, which is responsible for all solar power research, including very large grants to national laboratories and universities, is also responsible for developing topics and then evaluating SBIR applicants to those topics.

An applicant firm must propose a project that fits with a specific competition's scope. Examples of competitions include "Solar Powered Water Desalination," and "Improved Recovery Effectiveness In Tar Sands Reservoirs." The empirical strategy in this study compares firms within competitions. Applications are evaluated according to three criteria: 1) Strength of the scientific/technical approach; 2) Ability to carry out the project in a cost-effective manner; and 3) Commercialization impact (Oliver 2012). Program officials rank applicants within each competition based in part on written expert reviews from scientists at National Labs, universities, and the private sector. Program officials submit the rank ordered lists to an independent, separate DOE SBIR office. The cutoff within each competition is unknown to the program officer when she produces the rankings. The number of awards varies across competitions.

1.3 Data Description and Summary Statistics

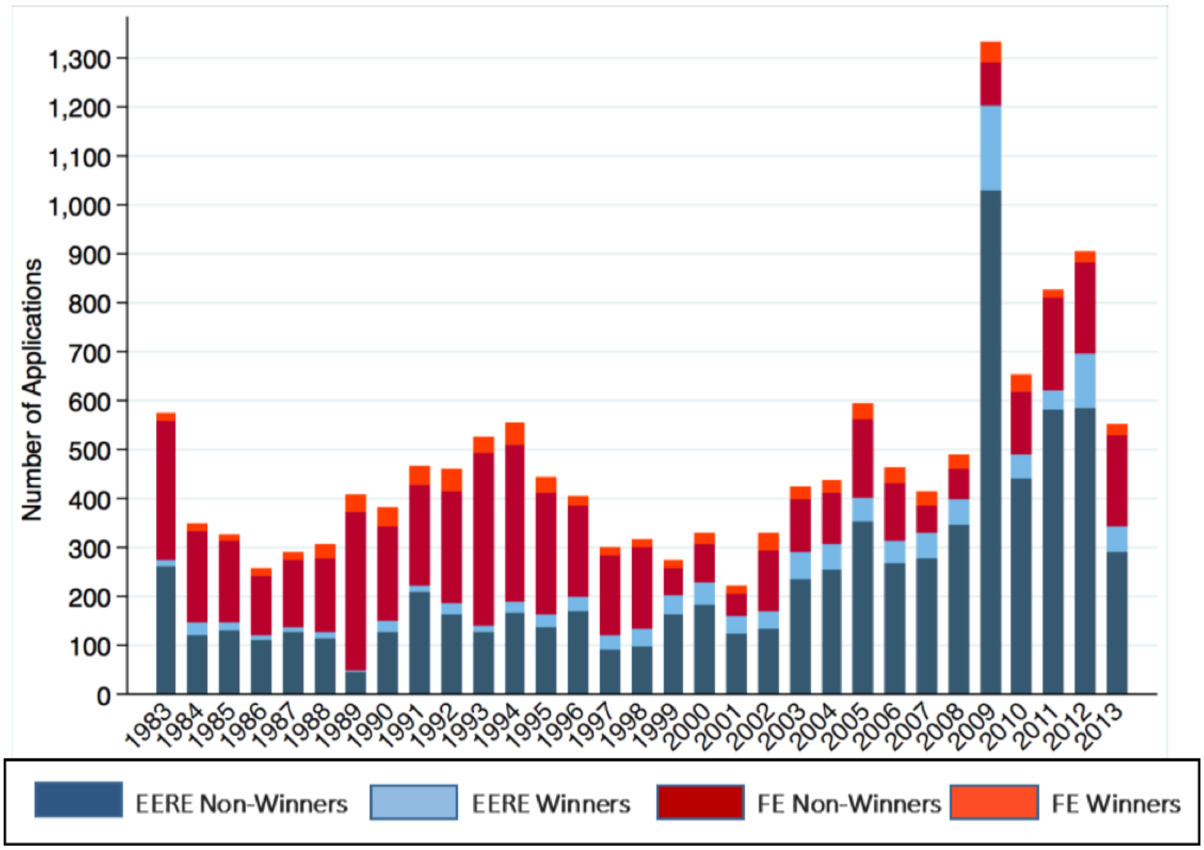
1.3.1 SBIR Data

This study begins with data on all applicants to the EERE and FE offices for the entirety of the DOE SBIR program, from 1983 to 2013. Applicants to the Small Business Technology Transfer (STTR) program are excluded. The data include 7,436 small energy technology

[8]Besides EERE and FE, the other offices that received DOE-SBIR funding during the period examined were: Office of Science; Nuclear Energy; Environmental Management and Electricity Delivery & Energy Reliability. DOE's ARPA-E—stood up in 2009 - manages its own SBIR program. During the estimation period (1995-2013) there were several major reorganizations and re-namings of various offices including subunits of EERE. Within EERE, the nine program offices are now: Solar Energy Technology, Bioenergy Technologies; Fuel Cell Technologies; Geothermal Technology; Wind Energy Technologies, Water Power Technologies; Vehicle Technology; Building Technology and Advanced Manufacturing.

firms. Figure 1 shows the number of applicants to the EERE & FE Phase 1 SBIR Program and annual Phase 2 applicants are shown in Figure 2. The spike in 2010 is due to the American Recovery and Reinvestment Act (Stimulus). Ranking information is available only from 1995, so estimation starts in that year. The ranks and non-winning applicant identities are confidential and not disclosed to the public.[9]

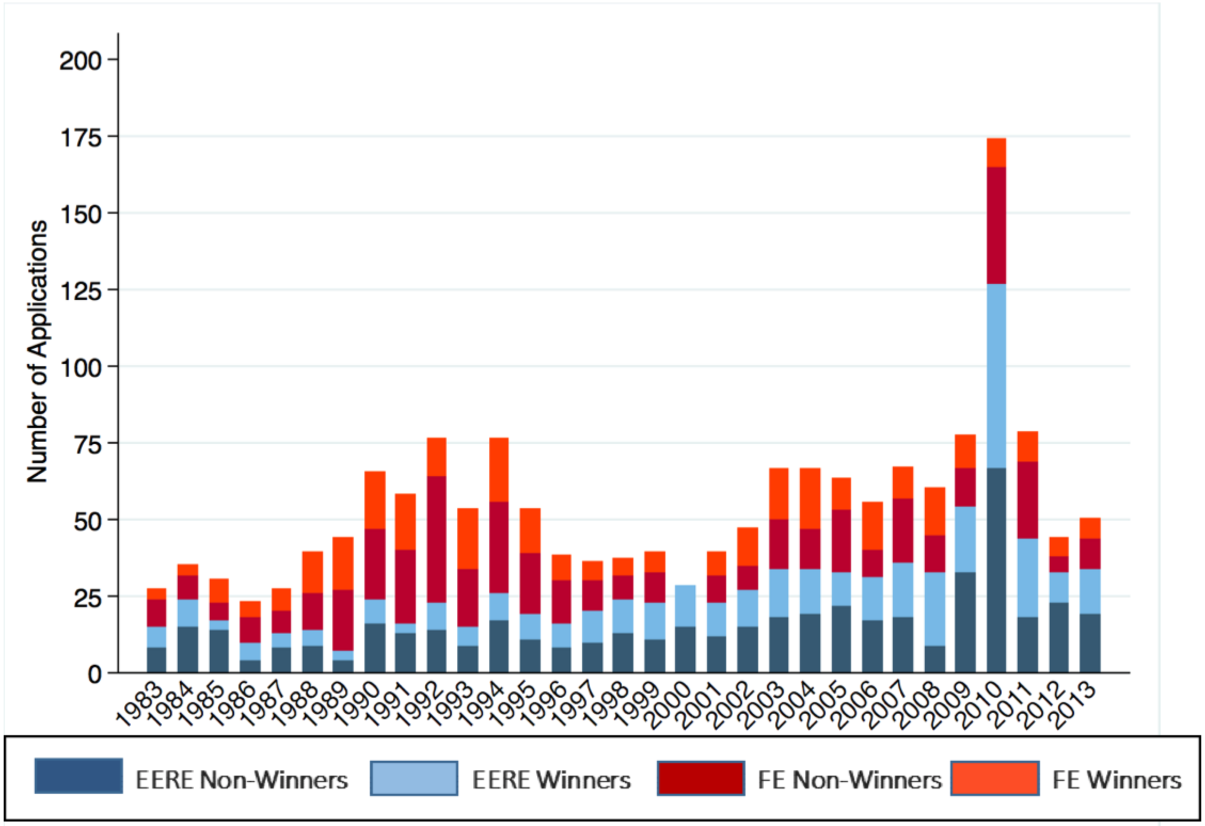
Figure 1: Number of Applicants to EERE & FE Phase 1 SBIR Program



Note: This figure shows the number of non-winning and winning Phase 1 grant applicants over time. Note that firms may appear more than once.

[9]It is only in the author's capacity as an unpaid DOE employee that she was able to use this data. Throughout the paper, specific references to companies will only include winners.

Figure 2: Number of Applicants to EERE & FE Phase 2 SBIR Program



Note: This figure shows the number of non-winning and winning Phase 2 grant applicants over time. Note that firms may appear more than once.

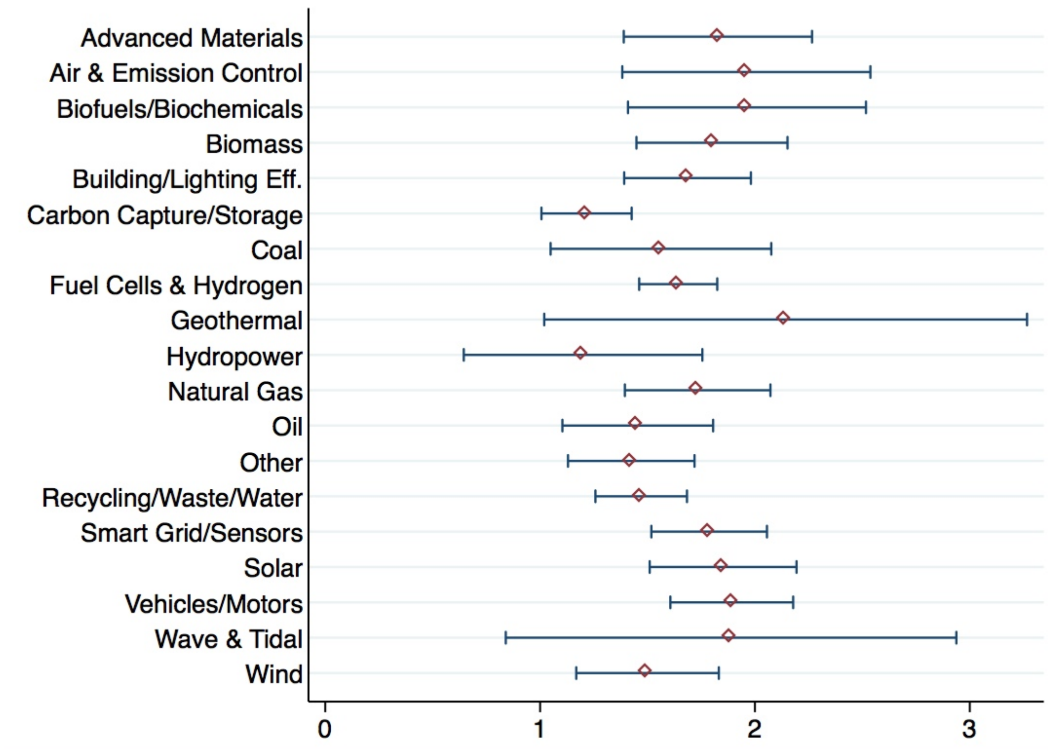
Table 1 contains summary statistics about the applications and competitions for EERE and FE SBIR. Each Phase 1 competition has on average 11 applicants, with a standard deviation of eight. The number of awards also varies by competition; the average is 1.7. The number of awards is determined by program budget constraints, recent funding history, office commitments to projects such as large National Laboratory grants, and the overall number of ranked applicants the central DOE SBIR office receives (the number of applicants deemed "fundable").[10] The cutoff used to separate winners from non-winners is exogenous to the ranking process used by DOE. The number of SBIR awards does not systematically vary by any covariate (such as by program office, technology area, or time).[11] Figure 3 shows that there are no obvious differences among technology areas in the average number of awards.[12]

[10]The number of competitions or subtopics per program office/topic are deliberately designed to scale with the program budget.

[11]The author's understanding of the exogeneity of the cutoff to the ranking comes from conversations with stakeholders in the DOE SBIR program, and from historical email records containing rank submissions.

[12]See Howell (2017) for the average number of applicants per competition by program office, the number of awards per office, and the number of awards per competition over time.

Figure 3: Average Number of Awards per Competition by Technology Areas, 1983-2013



Note: This figure shows that within competitions, the average number of Phase 1 awards does not vary systematically across program offices. All DOE EERE & FE competitions from 1995 are included. Capped lines indicate 95% confidence intervals.

Among applicant firms, 71 percent applied only once, and a further 14 percent applied twice. However, a small subset of firms apply and win many times. Seven companies in the data each submitted more than 50 applications. However, the majority of firms in the data apply only once and meet general criteria for startups. The firm median age is six years, and many firms are less than a year old. Among the 23 solar firms that have ever had an IPO, nine appear in the data; SBIR winners include Sunpower, First Solar, and Evergreen Solar (Cleantech Group i3). Although there is no strict definition of "startup," they must be young, small, and have location-unconstrained growth potential. (This is why restaurants, plumbers, and other local small businesses are not startups.) In this context, they are also high-tech.

Table 1: Summary Statistics of DOE's EERE and FE SBIR Applicants

Panel A: Application Data from DOE (EERE and FE)

1983-2013	
# Phase 1 Applications	14,522
# Unique Phase 1 Applicant Firms	7,419
# Competitions	1,633
1995-2013	
# Phase 1 Applications	9,659
# Unique Phase 1 Applicant Firms	4,545
# Phase 1 Applications with ranking data used in RD	5,021
# Phase 1 Competitions used in RD ($\geq$ 1 award)	428
Average # Phase 1 Applicants per Competition	11 (8.3)
Average # Phase 1 Awards per Competition	1.7 (1.1)
# Phase 2 Applications used in RD	919

Panel B: Variables Used in Analysis from Non-DOE Sources

	Type	Mean	Std Dev	Median	N
Pre-award cite-weighted patents	Count	21	122	0	5,021
Pre-award patents	Count	1.9	7.5	0	5,021
Post-award cite-weighted patents $\left(Cites_i^{post}\right)$	Count	12	117	0	5,021
Post-award patents	Count	2	11	0	5,021
Probability in major metro area (top 6)	0-1	0.30	0.46	0	5,021
Age (years)	Count	9.5	11	6	3,427
Probability tech is hardware ($Hardware_i$)	0-1	0.43	0.49	1	2,571
Prob. new sub-sector ($Emerging\ Sector_i$)	0-1	0.58	0.49	1	2,571
Probability minority owned	0-1	0.077	0.27	0	1,722
Probability woman owned	0-1	0.084	0.28	0	1,722
All-gov't SBIR wins ($SBIR_i$)	Count	10	36	0	5,021
Future patents in modal class	Count	9,758	11,809	5,453	1,583
MSA VC investment 2011 ($ mill)	Cont.	851	1,570	0	4,950
MSA median per cap. income 2011 ($ thou)	Cont.	56	14	56	4,603

Notes: This table summarizes the DOE SBIR application data in panel A, and variables used in the regression analysis in panel B. Parentheses in Panel A indicate the standard deviation. Cite-weighted patents refers to the number of citations for all the firm's patents. MSA refers to metropolitan statistical area.

1.3.2 Patent Data

This study uses the number of patents and cite-weighted patents as proxies for innovation.[13] Patenting activity is an imperfect measure of innovation; this is only one way that firms

[13]

Cite-weighted patents refers to the number of citations for all the firm's patents.

protect their intellectual property, and patents have an ambiguous relationship with technological progress (e.g. Arora, Ceccagnoli and Cohen 2008). Nonetheless, they are positively associated with economic value creation and stock market returns (Hall, Jaffe, and Trajtenberg 2005, Eaton and Kortum 1999). They are also the only currently available quantitative innovation metric.

This study uses patent data from Berkeley's Fung Institute for Engineering Leadership, which includes all patents filed between 1976 and 2014. Utility patents are matched to applicant firms, and checked by hand. It is assumed that when a firm cannot be matched to the patent data, the firm has no patents. The pre- and post-treatment variables use the patent application date rather than the issue date, as is standard in the literature. This is because we are interested in when the invention occurs, rather than the grant date, which is on average a few years later. To control for patent quality, cite-weighted patents (patents weighted by their future citations, as in Aghion et al. (2013) and Bloom, Schankerman and Van Reenen (2013)) are used. The patent count is not normalized by classification or year because competition fixed effects control for sub-sector and date.

Note that it is not possible to tie a firm's patents to the specific research that the firm conducted with its SBIR grant. We do not observe how firms use the grant, nor do we observe the technologies underlying the patents. The estimation strategy permits us to conclude that the grant causes the difference in patenting between winning and non-winning applicants.

1.3.3 U.S. Census Bureau Data

The second analysis employs outcome measures from U.S. Census Bureau data, which was gathered for research with the U.S. Census Bureau Center for Economic Studies. The Census Bureau maintains an annual Business Register containing all business establishments in the U.S. private non-farm sector with at least one employee. Applicant firms were matched to the Business Register by EIN (when available) or probabilistically on name, address, and zip code. About 70 percent of firms were matched successfully. The matching process erred on the side of including only matches with high confidence to minimize false positives. While it is important to note that the matched sample is not the same as the whole sample, there is no correlation of the matching with technology area or other observables, so there is no reason to believe that bias exists. Firms were also linked to other Census Bureau business datasets, including IRS W-2 data. These permit information about employee wages, ethnicity, and imputed education levels, among other variables.

The Business Register-matched firms were then linked to the Longitudinal Business Database (LBD), which begins in 1976 and ends in 2015. The LBD is the universe of non-farm, non-public administration business establishments with paid employees. This study employs three outcome variables from the LBD. The first is employment, which is observed in the pay period that includes March 12 from 1976 until 2005, when employment is observed for all four quarters of each year. The second is payroll, which is observed quarterly throughout. The third is revenue, which is observed annually starting in 1996. W-2 data on annual earnings for each employee begin in 2005 and end in 2013. Capital gains or other types of income besides wages are not observed. The wage should be thought of as salary income, as most of the jobs in this sample are full-time jobs. Note that as with the patent data, it is not possible to tie specific employees or portions of revenue to the SBIR grant. Again, the estimation strategy permits us to identify the effects as being caused (perhaps indirectly) by the SBIR grant.

The highest paid individual in the first year that the firm appears in the LBD is identified as the "firm founder". This is only a rough approximation, but it is in line with other studies using Census data, which do not identify firm owners or founders (e.g. Kerr and Kerr 2017, Azoulay et al. 2018, Babina and Howell 2018).

The main summary statistics for the matched data are presented in Table 2. There are 2,100 unique applicant firms in 270 competitions with adequate time series data for us to include in the analysis. Some firms apply more than once, so there are 4,300 applications. Wages, payroll, and revenue are in thousands of real 2010 dollars. Variables are summarized at the annual firm panel level as this is the level of analysis. This includes, for example, industries, because industry assignations may change over time within a firm. Industry is based on six-digit NAICS codes. Where a firm has multiple units, and therefore potentially multiple industries, the NAICS associated with the firm's largest employment share is used.

Table 2: Summary Statistics of SBIR data Matched to U.S. Census Data (1995-2013)

Panel A: SBIR Phase 1 competition data (counts)

	N
Unique applicant firms	2100
Applications	4300
Grant award winners	800
Grant award non-winners	3600
Competitions	270

Panel B: Outcome and control variables (firm-year level data)

Outcome variables			
	N	Mean	Std Dev
Payroll ('000 2010 $)	30500	2546	6141
Employment	30500	35.36	72.17
Award amount/employment$_{t=-1}$	30500	21880	33690
Average wage ('000 2010 $)	30500	64.15	38.55
Revenue ('000 2010 $)	13000	4834	11410
Log payroll growth (base is $t = -1$)	30500	-0.105	1.245
Log employment growth (base is $t = -1$)	30500	-0.082	1.008
Log wage growth (base is $t = -1$)	30500	-0.023	0.825
Log revenue growth (base is $t = -1$)	13000	-0.048	1.078
Key control variables			
	N	Mean	Std Dev
Firm age	30500	12.38	8.539
Subsequent patent citations (3 year window)	30500	2.071	10.81
Never previously won an award	30500	0.57	

Panel C: Additional firm-year variables

Probability in industry (most common 3 digit NAICS)

	N	Mean
Administrative and Support Services	30500	0.013
Chemical Manufacturing	30500	0.0167
Computer and Electronic Product Manufacturing	30500	0.079
Electrical Equipment, Appliance, and Component Manufacturing	30500	0.0324
Fabricated Metal Product Manufacturing	30500	0.0241
Machinery Manufacturing	30500	0.0495
Merchant Wholesalers, Durable Goods	30500	0.0257
Professional, Scientific, and Technical Services	30500	0.622

Worker-related variables

	N	Mean	Std Dev
Worker age	9600	43.1	8.398
Average worker tenure	9600	2.219	1.925
Share employees who are female	9600	0.223	
Share employees who are Asian	9600	0.173	
Share employees who are Black	9600	0.0273	
Share employees who are Hispanic	9600	0.0406	
Share employees who are White	9600	0.737	
Share employees with BA/advanced degree	9600	0.515	
Share employees with some college	9600	0.250	
Share employees with high school degree	9600	0.169	
Share employees with no high school degree	9600	0.0642	
Share employees who are US born	9600	0.714	

Panel D: Firm founder and worker-related firm-level variables

	N	Mean	Std Dev
Employment in application year	2000	33.31	25.64
Firm age in application year	2000	8.309	6.378
Average founder wage in application year	2000	136000	259300
Average founder age in application year	2000	51.28	12.14
Share founders female	2000	0.115	
Share founders Asian	2000	0.168	
Share founders Black or Hispanic	2000	0.0383	
Share founders white	2000	0.797	
Share founders US born	2000	0.718	
Share founders Eastern Europe/Russia born	2000	0.0363	
Share founders Western Europe born	2000	0.0457	
Share founders India born	2000	0.056	
Share founders China born	2000	0.0688	

Note: This table shows summary statistics about the SBIR data that were matched to U.S. Census data. Growth measures in Panel B use the year before the application year as the base year, denoted $t = -1$, where year t is the application year. The application year is first application year if the firm never won a grant, and first winning year if it ever won. Panel C contains the share of firms in the most common eight 3-digit NAICS codes are shown (there are a total of 99 3-digit NAICS). Firms may change NAICS codes across years. Worker-related variables in Panel C are from linked W-2-Individual Characteristics File data. "White" indicates non-Hispanic White. Panel D contains observations at the firm level and where worker information is available (i.e., linked to W-2-Individual Characteristics File data). The number of observations rounded to meet Census disclosure requirements.

For the matched SBIR-Census data, the average number of employees is 35. This is relatively similar to all firms in the U.S. In 2012, the average for all U.S. firms is 20 employees, and within establishments with 20-99 employees, the average is 39.[14] Average revenue is \$4.8 million; though the distribution is highly right skewed. The median (unreported due to disclosure limitations) is much lower. The average is also well-aligned with U.S. averages, which are \$779,000 for firms with less than 20 employees, and \$7.9 million for firms with 20-99 employees. Average payroll in the data is higher than the average for U.S. firms with 20-99 employees, at \$2.5 million relative to \$1.6 million.

The primary outcome variables are logged growth measures, defined as the log difference of an outcome in a given year relative to the year before application ($t = -1$): $Growth_{i,t} =$

[14] https://www.census.gov/data/tables/2012/econ/susb/2012-susb-annual.html

$\ln\left(\frac{Y_{i,t}}{Y_{i,t=-1}}\right)$. Logs are used to linearize the relationship between the independent (right-hand side) and dependent (left-hand side) variables in the regression. The underlying outcome data (dependent variables) are positively skewed, with a small number of observations that have large magnitudes. Taking the log smooths the distribution.

Table 2 Panel B shows that on average, these measures are near zero but negative. That is, size measures are larger in the year before application compared to other years. Note that years both before and after the application year are included, so the negative mean is consistent with a firms growing before the application and sometimes failing afterward. Note that the standard deviations are very large. On average, payroll in a given year is about 90 percent of its value in the year before application, employment 92 percent, wages 98 percent, and revenue 95 percent.

Additional firm and worker characteristics are in Table 2 Panels C and D. As might be expected for applicants to an R&D grant program, the most common NAICS 3-digit industry is Professional, Scientific, and Technical Services, at 62 percent of firms. The next most common is Computer and Electronic Product Manufacturing, at 7.9 percent. The table shows an additional seven industries. The average worker is 43 years old, while in the application year the average founder is 51 years old. Just 22 percent of employees are female, and only 11.5 percent of founders are female. There are also disparities relative to the population in ethnic makeup; only 2.7 percent of employees and 3.8 percent of founders are Black, for example. Seventy-one percent of employees and founders are US-born. Among founders, the next most common country of origin is China, with seven percent of founders.

2 Methods

This section presents the estimation strategy, which is called a regression discontinuity design (Section 2.1). It also describes specification tests (Section 2.2).

2.1 Regression Discontinuity Design

The estimation strategy relies on the ranks that DOE assigns to firms within competitions. It is assumed that firms ranked near the award cutoff are quite similar, and after validating this assumption with a litany of tests, comparing just-winners to just-non-winners is a local random experiment. The use of the ranks in this manner is an econometric estimation strategy called a sharp regression discontinuity (RD) design. As public agencies resist randomizing treatment to evaluate R&D subsidies (unlike new medicines), RD is the best approach to approximating an experiment (Jaffe 2002). The RD design estimates a local average treatment effect around the cutoff in a rating variable. Here, the rating variable is the applicant's rank. The critical assumption is that applicants cannot precisely manipulate their rank near the cutoff.[15]

Since the number of applicants and awards varies across competitions, applicant ranks are centered in each competition around zero at the cutoff. The lowest-ranked winner has centered rank $R_i = 1$, and the highest-ranked non-winner has $R_i = -1$. Each competition has at least this pair. As the bandwidth expands, higher ranked winners and lower ranked non-winners are included. Controls for rank eliminate ex-ante differences between winners and non-winners that may exist further away from the cutoff (for example, if higher ranked firms tend to be higher quality). In this context, however, rank turns out to be entirely uninformative about outcomes, so it is immaterial for the results what bandwidth is used.

2.1.1 Estimating Equation for Patenting

The patent analysis uses variants of Equation 1, where Y_i^{Post} is the outcome and dependent variable. The unit of observation is a firm i in a competition j.

$$Y_{i,j}^{\text{Post}} = \alpha + \beta Award_{i,j} + f\left(R_{i,j}\right) + \gamma_1 Y_{i,j}^{\text{Prev}} + \gamma_2 X_{i,j} + \lambda_j + \varepsilon_{i,j} \tag{1}$$

Following Aghion, Van Reenen, and Zingales (2013), the regression models focus on cite-weighted patents as an outcome ($Y_{i,j}^{\text{Post}}$), and use negative binomial and ordinary least squares (OLS) regression models. The coefficient of interest is β on an indicator for receiving a grant award (treatment). The pre-assignment outcome variable is Y_i^{Prev}. A level rather

[15]More specifically, a valid RD design must satisfy four conditions to be considered a local randomized experiment. First, it is necessary that treatment (a grant award) not lead to the rank assignment. This holds for the DOE SBIR program, as the award happens after ranking. To avoid contamination, applicants who previously won a grant within EERE/FE are excluded. Second, the cutoff must be exogenous to rank, which is true in my setting. Third, the functional form must be correctly specified, or else the estimator will be biased. A goodness-of-fit test shows that rank is uninformative. Finally, to meet the key assumption that applicants cannot precisely manipulate their rank in the region around the cutoff, all observable factors must be shown to be locally continuous. To establish the necessary weak smoothness (see Hahn et al. 2001), continuity of covariates is demonstrated below. For more on RD and the above tests, see Lee and Lemieux (2010) and Howell (2017).

than a change model (where the dependent variable would be $Y_{i,j}^{\text{Post}} - Y_i^{\text{Prev}}$) is preferred because it does not confound zero patenting with a zero difference between patents before and after the application. Also, it makes it easier to interpret the coefficients of interest and to compare treatment effects in linear and non-linear (here, binomial) models.

An SBIR winner is assigned the indicator $Award_{i,j} = 1$, and a non-winner is assigned $Award_{i,j} = 0$. $f(R_{i,j})$ is a polynomial controlling for the firm's rank within competition c. (As elsewhere, only first-time winners are included.) Rank is limited to various bandwidths around the cutoff. There is a full set of dummies for each competition $\ _j$, which controls for the date and a narrow sub-sector. X_i indicates other controls.[16] The estimations use OLS for binary dependent variables, negative binomial for count data, and two-part models for semi-continuous data.[17] Standard errors are robust and clustered by topic-year, to account for correlation in time and sector.

2.1.2 Estimating Equations for Firm Growth

For the firm growth measures, a panel data structure is used, where firms are observed over time. The panel approach exploits the richness of the U.S. Census data, permitting finer controls. The unit of observation is now a firm i in year t in competition j. The primary empirical specification for evaluating the effect of a grant award is shown in Equation 2.

$$\begin{aligned} G_{i,t} = {} & \beta PostAward_{i,j,t} + \quad Award_{i,j} + \delta Post_{i,j,t} \\ & + f\left(R_{i,j}\right) + \eta_3 Age_i + \eta_4 Age_i^2 \\ & + \quad _{j/i} + \tau_t + \varepsilon_{i,j,t} \end{aligned} \tag{2}$$

A firm that ever wins a grant is assigned the non-time varying indicator $Award_{i,j} = 1$. The variable $Post_{i,j,t}$ is an indicator for the year being after the year the firm applied, and $PostAward_{i,j,t}$ is the interaction between $Post_{i,j,t}$ and $Award_{i,j}$. As with patenting, winning

[16] The RD design does not require conditioning on baseline covariates but doing so can reduce sampling variability. Lee and Lemieux (2010) advise including the pre-assignment dependent variable as they are usually correlated. In tests reported in Howell (2017), rank is projected on observable covariates. Previous non-DOE SBIR awards are the strongest predictor of rank. A one standard deviation increase in previous SBIR wins (the mean is 11.4 and the standard deviation is 38) increases the rank by nearly one unit. Previous VC deals also have a small positive impact. These two variables are included in my primary specifications.

[17] OLS for binary outcomes is used because many of the groups defined by fixed effects (competitions) have no successes (e.g. no subsequent patents). Logit drops the groups without successes. Also, OLS with a binary variable is common in applied economics, following the arguments in Angrist (2001) that regression does as well as logit in estimating marginal effects and often better with binary treatment variables. My main results are intact with a logit specification (see Section 3.6).

firms are included only once. Similar results are observed in a non-panel setting, like that in Howell (2017), where each observation is an application rather than a firm-year.

In most cases, the dependent variable is a log growth measure, $\ln\left(\frac{Y_{i,t}}{Y_{i,t=-1}}\right)$, where $Y_{i,t}$ is, for example, employment or the average wage. Some specifications use levels ($Y_{i,t}$), in particular when comparing effects on new and incumbent employees ("change" is undefined for new employees). The growth specification increases power and ensures that unobserved time-invariant characteristics are controlled for, which is a conservative approach since specifications with narrow bandwidths around the cutoff are not reported due to disclosure limitations. However, all of the results are qualitatively robust to using narrow bandwidths around the cutoff, and as shown below, rank does not predict outcomes at all, as in Howell (2017).

The primary specification controls for rank within the competition quadratically, which means that rank and rank squared are included as covariates. This approach includes competition fixed effects ($_j$, as in Equation 1) and calendar year fixed effects τ_t. Other controls include the firm's age and its age squared. Beyond the primary model, two other specifications are shown for the main effects. One controls for rank separately among winners and non-winners. A second includes firm-application fixed effects ($_i$), which subsume rank and control more completely for pre-treatment differences, including all the characteristics of the application. Errors are clustered by competition, though the main effects are robust to a variety of error assumptions.

Graphed results are presented from two additional specifications that demonstrate robustness to alternative approaches. The first shows the effects by rank around the cutoff for the award using Equation 3.

$$\begin{aligned} Y_{i,t} = & \sum_{x=-6}^{x=3} \beta_x \left(PostAward_{i,j}\right)\left(Rank_{i,j} = x\right) \\ & + \eta_1 Age_i + \eta_2 Age_i^2 + \tau_t + \quad _j + \varepsilon_{i,j,t} \end{aligned} \tag{3}$$

Outcomes are in levels (e.g. log employment) to provide evidence that the findings from Equation 2 go through with levels. The effects are similar when growth outcomes are used in Equation 3 instead.

The second shows the effects by quarter around the award quarter using Equation 4, where

q denotes the quarter.

$$Y_{i,q} = \sum_{x=-13}^{x=13+} \left[\beta_x \left(Award_{i,j} = 1\right)\left(q = x\right) + \delta_x \left(q = x\right)\right] \tag{4}$$
$$+ \tau_q + \quad_i + \varepsilon_{i,j,q}$$

The equation includes indicators (δ_x) for 13 and more quarters before the award date, each quarter between 12 and two before the award date, the award quarter, each quarter after the award date through 12 quarters after, and 13 and more quarters after the award quarter. The coefficients of interest, β_x, are on these same indicators interacted with the award dummy, and these are shown in the graph. The equation includes firm-application fixed effects, which are the most stringent specification possible, as they control for all possible application and firm characteristics. Again, outcomes are in levels. There are similar effects using competition fixed effects or growth outcomes. In estimating both Equations 3 and 4, standard errors are clustered by competition.

2.2 Specification Tests

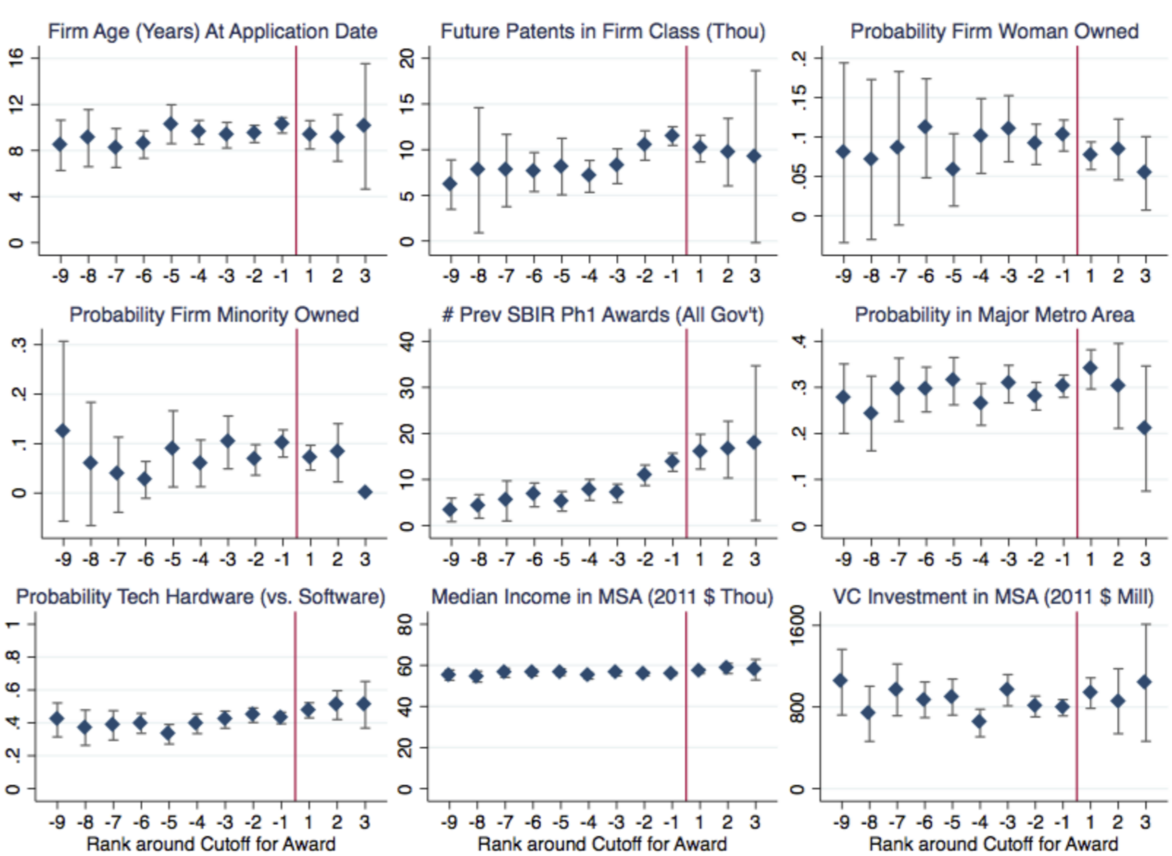
A rich array of specification and robustness tests are contained in Howell (2017). Crucial tests are discussed here.

A data limitation is that because competitions average only ten applicants, the rating variable is somewhat discrete. This discreteness means that we may worry about jumps in quality between ranks. If there were hundreds of applicants within each competition, we would expect the quality difference between adjacent ranks to be very small. Lee and Card (2008) note that discrete rating variables can require greater extrapolation of the outcome's conditional expectation at the cutoff. The fundamental econometrics are no different than with a continuous rating variable, however, as extrapolation is required in both cases. Howell (2017) demonstrates the robustness of my findings to this discreteness by, for example, separately considering competitions with low or high numbers of awards.

In order for the estimation strategy to be close to an experiment, it is important that firms seem to be equivalent immediately around the cutoff. One way to test for similarity around the cutoff is to examine whether observable characteristics are "smooth" (do not jump) around the cutoff. Howell (2017) demonstrates smoothness in observable baseline covariates in three ways: through an RD on baseline covariates, through differences in means, and most importantly visually. Figure 4 shows the visual evidence of the baseline covariate RD. Baseline covariates include pre-assignment outcome variables, average age as well as the probability a firm is located in a major metro area, is woman owned, and is minority owned. In none of the figures is there any discontinuity around the cutoff visible, nor is there any trend in rank. A ninth covariate, previous non-DOE SBIR wins, is the exception in terms of rank trends. When applicants have more previous wins, they tend to have a higher rank. Nonetheless, again there is no discontinuity around the cutoff.

Program officials observe more data than the econometrician, so it is impossible to fully test the assumption that firms are randomly assigned on either side of the cutoff. Nonetheless, this preponderance of evidence suggests the RD design is valid.

Figure 4: Continuity in Observable Covariates



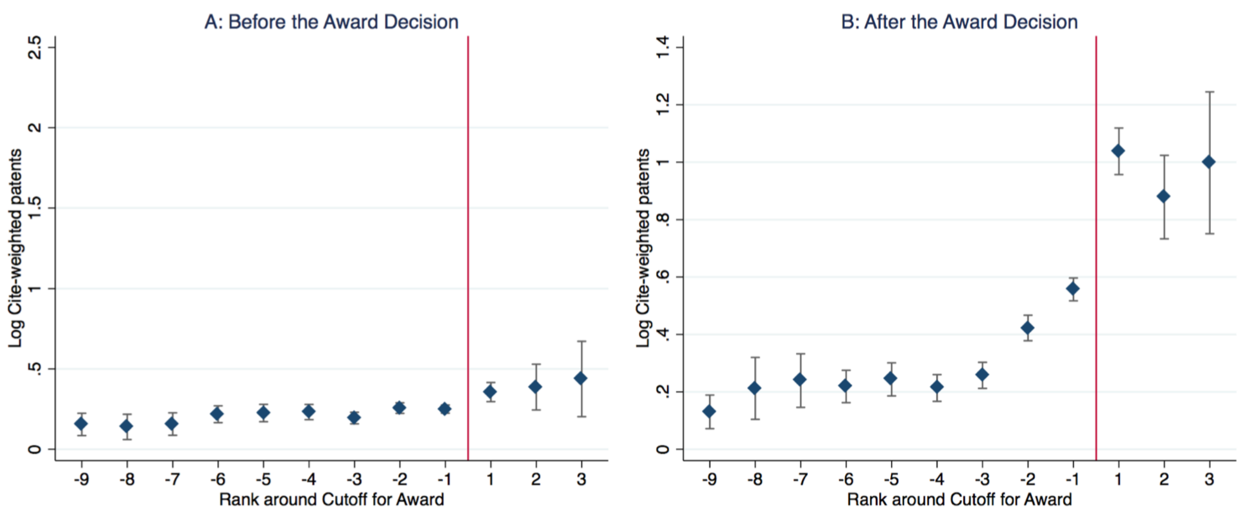
Notes: This figure shows covariates at the award date. 95% confidence intervals shown. The confidence interval is not included for the probability a firm is minority owned at the 3rd rank from the cutoff because it is very large.

3 The Effect of SBIR Grants on Patenting

3.1 Main Effect of the Phase 1 Grant

The best available measure for innovation is patenting. Though patents are not the only way that firms protect IP, they are positively associated with economic value creation and stock market returns (Hall, Jaffe, and Trajtenberg 2005). Figure 5 shows log cite-weighted patents before and after the Phase 1 grant award (all matching patents are included, so there is no time restriction around the award). The figure clearly indicates a strong effect of the award; we see a large jump around the cutoff for patents filed after the award. Comfortingly, there is no such jump around the cutoff before the award, indicating that the estimation strategy is valid. Ranks among high-ranking non-winners are predictive of future patenting. This relationship disappears for winners.

Figure 5: Phase 1 Effects on Cite-Weighted Patents by Rank Around the Award Cutoff



Notes: This figure shows $\ln\left(1 + Cites_i^{post}\right)$ before and after the Phase 1 grant award decision, using the patent application date. DOE's rank is centered so positive ranks indicate a firm won an award. 95% confidence intervals shown.

Results from estimating variants of Equation 1 are in Table 3. The bandwidth is one rank around the cutoff in each competition except in column 3 where all the data with quadratic rank controls are used. In the primary sample of no previous winners and using the negative binomial specification, an award increases cite-weighted patents by 2.5 times (panel A columns 1-4).[18] The OLS specification finds that the grant increases log cite-weighted

[18]Coefficients indicate, for a one unit increase in regressor, the difference in the logs of expected counts.

patents by about 30 percent (panel B columns 1-3). Note that the linearization reduces the effect.

All patents filed after the competition's award date are used, as competition fixed effects control for years since the grant. However, the time fixed effects do not necessarily control for when the firm patents. In unreported specifications (not shown because of limitations to the number of estimates that Census will permit to be disclosed), results are similar when citations are limited to three years after the patent. Also, the grant increases the probability of positive patenting by 9 percentage points (pp). Rank has no predictive power over patents in all models.

Centering ranks might obscure information in the raw rank. For example, firms with centered ranks of two might have different qualities in competitions with two and four awards. To address this, Panels A and B column 4 use dummies for the firm's rank quintile within the competition.[19] Conditional on award status, there is no information in rank visually or in regressions, regardless of bandwidth.

Column 5 permits previous winners to be included in the sample. As a result, the number of observations increases. The effect declines somewhat. The reason for this decline is highlighted by the two following columns. First, column 6 limits the sample to first time applicants, and yields a much larger effect than the main sample. Conversely, when only firms with more than two wins are included (column 7), the effect declines by more than half. Unreported regressions find that for each previous DOE SBIR award, the effect of an award on log cite-weighted patents declines by 20 percent. There is a similar pattern for other outcome variables examined in Howell (2017), but it is especially troubling for patenting. A small subset of applicants win many awards and may be dependent on grants. While such firms might naturally not be seeking VC or direct sales, if their R&D is productive it should yield patents. Instead, there is a a steeply declining patenting benefit of additional grants to the same firm. This supports DOE program officials' skepticism about permitting so-called "SBIR mills" to win many awards.

If is the Poisson rate (# patents), $\tau = \log\left(\frac{\lambda_{R_{ic}>0}}{\lambda_{R_{ic}<0}}\right)$. Exponentiating gives the incidence rate ratio (how many times more patents winners get than non-winners).

[19] There are similar results with the slope controlled for separately on each side of the cutoff (with a bandwidth of all specification) and using quartile ranks.

Table 3: Phase 1 Grant Effect on Cite-Weighted Patents

Panel A: Negative binomial model; dependent variable is $Cites_i^{post}$

Sample:	No previous winners				All	No prev. apps	>2 prev. wins
Bandwidth:	1	1	All	All	1	1	1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Award	0.93***	0.91***	0.92***	0.94***	0.82***	2.1***	0.40***
	(0.21)	(0.19)	(0.33)	(0.26)	(0.13)	(0.34)	(0.14)
Normalized rank			0.052				
			(0.074)				
Normalized rank^2			-0.0072				
			(0.0045)				
Rank quintiles	N	N	N	Y	N	N	N
Controls†	N	Y	N	N	N	N	N
Sector-year f.e.††	Y	Y	Y	Y	Y	Y	Y
N	1871	1871	5021	5021	2714	972	1477
R^2	0.056	0.084	0.053	0.053	0.034	0.080	0.035

Panel B: OLS models; dependent variable is $\ln\left(1 + Cites_i^{post}\right)$

Sample:	No previous winners				All	No prev. apps	>2 prev. wins
Bandwidth:	1	1	All	All	1	1	1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Award	0.33**	0.27**	0.29***	0.22**	0.29**	0.49**	0.22*
	(0.15)	(0.13)	(0.11)	(0.087)	(0.087)	(0.21)	(0.13)
Normalized rank			-0.026				
			(0.02)				
Normalized rank^2			7.8e-4				
			(0.0011)				
Rank quintiles	N	N	N	Y	N	N	N
Controls†	N	Y	N	N	N	N	N
Competition f.e.	Y	Y	Y	Y	Y	Y	Y
N	1872	1872	5021	5021	2714	972	1477
Pseudo-R^2	0.46	0.60	0.53	0.351	0.63	0.71	0.75

Note: This table reports regression estimates of the Phase 1 award effect on cite-weighted patents using variants of Equation 1. The model is negative binomial in panel A and OLS in panel B. Columns 1-4 limit the sample to firms that have not previously won an award, but may have previously applied. Columns 5-7 respectively use the whole sample, only firms that have never previously applied, and only firms that have more than two previous wins. †Controls are $Cites_i^{prev}$ or $\ln(1 + Cites_i^{prev})$ and previous non-DOE SBIR awards. ††Competition fixed effects (f.e.), in the NB model do not permit convergence. Fixed effects mean that a separate control is included for each competition, so that the estimation result is within competition, Year 1995. Standard errors are clustered by sector-year. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

3.2 Variation in the Phase 1 Effect on Patents

The effect on innovation activity varies with firm characteristics. For this analysis, the number of patents is used, as this yields sharper results, though the effects are qualitatively similar using cite-weighted patents. First, it is expected that young firms have fewer internal resources and their R&D investment is likely more affected by capital market imperfections (Hall 2008). Columns 1-4 of Table 4 show that the grant effect on short-term patenting falls dramatically and loses all significance for older firms (at least 10 years old). The patent incidence rate ratio (IRR) is a staggering 12 for firms no more than two years old, statistically significant at the 1% level (column 1), whereas the IRR is only 1.75 for firms more than two years old and is not statistically significant. For firms less than 10 years old, the IRR is 4.5, whereas for firms older than 10, it is 0.62 - a negative effect - and statistically insignificant (columns 3 and 4). This suggests that young privately held firms may face greater R&D investment financing constraints than older private firms, supporting the findings on public firms in Brown, Fazzari and Petersen (2009).

As with age, there may be more information available about firms with patents. Hsu and Ziedonis (2008) and Conti, Thursby and Thursby (2013) show that patents improve entrepreneurs' access to finance by signaling potential investors about a firm's quality. Patents may also serve as collateral, as in Mann (2018) and Hochberg, Serrano and Ziedonis (2014). The latter paper finds that among VC-backed startups with patents, 36 percent used the patents to secure loans. Columns 5 and 6 of Table 4 shows that the treatment effect declines when firms have previous patents: with no patents, the grant leads a firm to produce 3.3 times more patents than it would otherwise. With at least one patent, the IRR is 2.7. This decline in the Phase 1 grant effect when a firm has more previous patents may indicate that more experienced, later stage firms with better access to debt finance benefit less from the grants.

There is wide variation in propensity to patent across technologies (Scherer 1983, Brouwer and Kleinknecht 1999). Table 4 columns 7 and 8 use an indicator for high propensity to patent from the USPTO (2012) patent intensity estimations.[20] In high propensity industries, the IRR is 8.1, statistically significant at the 1 percent level (Table 4 column 7). This means that a grantee produces 8.1 times as many patents as a non-winner. In contrast, in low propensity industries, the IRR is only 2.7, statistically significant at the 10 percent level

[20] These are based on patents per 1,000 jobs in an industry. The indicator takes a value of 1 if the firm is in one of the following sectors: Smart Grid, Sensors & Power Converters, Advanced Materials, Solar, or Batteries, and 0 otherwise.

(Table 4 column 8). For all three heterogeneity analyses in Table 4, difference (interaction) equations cannot be estimated because the maximum likelihood function does not converge.

Table 4: Variation in the Phase 1 Grant Effect on Patenting

Dependent Variable: $\#Patent_i^{3\text{ yrs Post}}$								
Sample:	Firm Age in Years				Firm # Previous Patents		Tech. Patent Propensity	
	≤ 2	> 2	≤ 9	> 9	0	1	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Award	2.5***	.56	1.5***	-.48	1.2***	1***	2.1***	.99***
	(.38)	(.42)	(.28)	(1.1)	(.39)	(.23)	(.46)	(.22)
Rank and rank^2 controls	Y	Y	Y	Y	Y	Y	Y	Y
Controls†	Y	Y	Y	Y	Y	Y	Y	Y
Topic f.e.	N	N	N	N	N	N	Y	Y
Year f.e.	Y	Y	Y	Y	Y	Y	N	N
N	576	2790	1410	1958	2308	1058	834	2532
Pseudo-R^2	.14	.092	.1	.1	.083	.067	.15	.2
Log likelihood	-383	-2221	-1220	-1367	-794	-1646	-719	-1640

Note: This table reports regression estimates of the effect of the Phase 1 grant award on patents using bandwidth (BW)=3 and a negative binomial model, focusing on variation by firm age, technology propensity to patent, and number of previous patents. Propensity to patent is calculated as described in the text. The specifications are variants of the model in Equation 1. The dependent variable $\#Patent_i^{3\text{ yrs Post}}$ is the number of successful patents that the firm applied for within three years of the grant award. The left panel divides the sample by an indicator for high propensity to patent, which is based on overall USPTO 2012 patent intensity by technology. It is 1 if the firm's technology sub-sector is Smart Grid, Sensors & Power Converters, Advanced Materials, Solar, or Batteries. The middle panel divides the sample by firm age, and the right panel by the firm's number of patents prior to applying for the grant. For all three, difference equations cannot be estimated due to non-convergence of the Poisson maximum likelihood. †Controls are $Patent_i^{prev}$ and previous non-DOE SBIR awards. Standard errors are robust. *** $p < .01$. Year 1995

Finally, Table 5 shows that there may be a precipitous drop in patents by previous non-DOE SBIR wins, but the coefficients are somewhat imprecise. A bandwidth of all firms is used to maximize the sample size, but similar results are found with narrower bandwidths. Relatedly, Howell (2017) finds a robust and sharp drop in the probability of receiving venture capital financing in the number of previous non-DOE SBIR awards.

Table 5: Variation in the Phase 1 Grant Effect by Number of Firm's Previous SBIR Awards

Dependent Variable: $\#Patent_i^{3 \text{ yrs Post}}$					
Sample:	Number of previous SBIR wins				
	< 2	< 5	> 0	2	5
	(1)	(2)	(3)	(4)	(5)
Award	0.896*	0.805*	-0.405	0.120	-0.257
	(0.531)	(0.481)	(0.369)	(0.444)	(0.576)
Rank and rank^2 controls	Y	Y	Y	Y	Y
Controls†	Y	Y	Y	Y	Y
Year f.e.	Y	Y	Y	Y	Y
N	4249	4651	1879	1444	1042
Pseudo-R^2	0.097	0.098	0.093	0.096	0.101

Note: This table shows estimates of the impact of the Phase 1 grant award on the firm's patent count within three years after grant award by number of previous non-DOE SBIR awards (from other government agencies, e.g. DOD, NSF), using BW=all and a negative binomial model. Each column includes only data from firms with the relevant number of wins. Unfortunately difference equations could not be estimated due to non-convergence of the Poisson maximum likelihood. †Controls are $Patent_i^{prev}$ and previous non-DOE SBIR awards. Standard errors robust and clustered at topic-year level. * $p < .1$. Year 1995

This supports the idea that efforts to avoid funding "SBIR mills" (firms with substantial recurring revenue from many SBIR awards) may be warranted.

3.3 The Phase 2 Grant Effect on Patenting

About nine months after receiving a Phase 1 award, a firm may apply for Phase 2. If successful, the firm receives Phase 2 in two increments of $500,000 roughly two and three years after the Phase 1 award. Any Phase 2 effect is local to the subset of Phase 1 winners. Many Phase 1 competitions have two or fewer Phase 2 applicants, so there is no control for rank, and only sector and year fixed effects are included. Phase 2 is therefore analyzed as though the Phase 2 grants were randomly assigned. If, contrary to this assumption, the DOE is selecting higher quality applicants to win Phase 2, the results should be biased upward. To further clarify, this means that the Phase 2 analysis is likely to produce positive results unless DOE is either randomly selecting applicants or selecting lower quality applicants as winners.

Using the negative binomial model and the standard sample of no previous winners, columns 1-3 of Table 6 show that Phase 2 doubles a firm's cite-weighted patents, relative to a mean of 20. Column 3 jointly estimates both Phases. The coefficient on Phase 2 drops to about 1.4 times as many cite-weighted patents. OLS results using $\ln\left(1 + Cites_i^{post}\right)$ in columns 7-9 find that Phase 2 increases cite-weighted patents by about 30 percent. Over half of Phase 2 applicants have won multiple DOE SBIR awards and so are excluded from the primary sample. Consistent with the Phase 1 findings, the positive Phase 2 effect on cite-weighted patents falls and loses significance when the sample includes previous winners.

The Phase 2 grant does generate new inventive activity in the form of cite-weighted patents. But its effects are much smaller than Phase 1 on a public dollar basis. Phase 1 involves only $150,000, but a grant yields about 1.4 cite-weighted patents. The Phase 2 grant is up to $1,000,000, and a high estimate for the Phase 2 impact is 2 cite-weighted patents. To illustrate how the per public dollar effect is so much larger for Phase 1, consider the following example. In 2012 the DOE spent $38 million on 257 Phase 1 grants and $112 million on 111 Phase 2 grants. If all the Phase 2 money were reallocated to Phase 1, the DOE could have provided 750 additional firms with Phase 1 grants, increasing by a factor of about 2.5 the program's impact on cite-weighted patents.

Note that this hypothetical is not a policy recommendation and comes with significant caveats. One is that discontinuing Phase 2 would likely affect the Phase 1 applicant pool.[21] In the absence of Phase 2 as a possibility in case the Phase 1 research did not quickly lead to external private finance, prospective applicants might not apply to the program at all. Another caveat is, as noted in Section 1.2, Phase 2 funds more extensive or later stage demonstrations than Phase 1. Phase 2 recipients may shift resources toward commercialization efforts and may not continue patenting at a high rate because they are busy working on commercialization. Finally, awarding many more Phase 1 grants would require awarding more lower ranked applicants. Although rank is not informative about outcomes (i.e. lower ranked applicants do not tend to do worse), this would affect the composition of awardees in unknown ways.

[21] According to the EERE SBIR Director, there is significant anecdotal evidence (e.g. Phase I grantees willingness to report on progress well beyond the minimum requirement) that the prospect of a Phase 2 grant is a strong motivation for applying for Phase 1.

Table 6: Phase 2 Grant Effect on Patenting

Model:	Negative binomial						OLS		
Dependent variable:	$Cites_i^{post}$						$\ln\left(1+Cites_i^{post}\right)$		
Sample:	No previous winners			All applicants		>1 prev. win	No previous winners		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Phase 2 Award	0.77***	0.69**	0.34*	0.28*	0.25	0.17	0.29**	0.32**	0.22*
	(0.29)	(0.30)	(0.17)	(0.15)	(0.17)	(0.15)	(0.13)	(0.15)	(0.12)
Phase 1 Award			0.33						0.17
			(0.10)						(0.061)
Controls†	Y	N	Y	Y	N	Y	Y	N	Y
Sector & year f.e.	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	408	408	5021	867	867	459	408	408	5021
[Pseudo-]R^2	0.13	0.091	0.21	0.092	0.042	0.10	0.51	0.35	0.42

Note: This table reports estimates of the Phase 2 grant effect on all outcomes, using variants of Equation 1. The model varies based on the dependent variable. There is no rank control due to the small number of applicants in each competition.
†Controls are $\ln\left(1+Cites_i^{prev}\right)$ and $SBIR_i^{prev}$ in both. Standard errors clustered by sector-year. Year 1995. Standard errors are clustered by sector-year. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

The Phase 2 sample is small because 37 percent of Phase 1 winners do not apply for Phase 2. There are four reasons for this, based on interviews with grantees and DOE officials. First, a firm is ineligible to apply if an outside investor owns more than 50 percent. Some firms that raise equity after Phase 1 may sell too much of the firm to be eligible to apply for Phase 2. Consistent with this possibility, among firms that receive VC within two years of the Phase 1 grant, 55 percent do not apply for Phase 2. Put another way, 19 percent of non-Phase 2 applicants received VC investment within two years of their initial Phase 1 award, but only 8 percent of Phase 2 applicants did (the statistics on VC can be found in Howell 2017).[22]

Second, firms might not apply if they changed business strategies. Phase 1 commercialization activities are known to DOE only if a firm applies for Phase 2 where such activities are required to be described. Third, the Phase 2 application and reporting processes are so onerous that once a firm receives external private finance, it may sometimes not be worthwhile to apply to Phase 2.[23] Relatedly, Gans and Stern (2003) suggest that private funding is preferred to SBIR funding. A firm's discount rate may increase once its initial R&D is successful, so that applying to Phase 1 is worthwhile but applying to Phase 2 is not, despite the larger sum at stake. This is consistent with the sharply decreasing risk premium in Berk, Green and Naik (2004) as an R&D project moves from initiation to completion. The adverse selection in the Phase 2 sample suggests that startups seeking to raise external finance whose Phase 1 R&D revealed positive information often secured private investment without needing further government support. Fourth, the Phase 1 "proof-of-concept research" may have failed to prove the concept and firms thus lack a rationale for continued Phase 2 funding.

4 The Effects of SBIR Grants on Firm Growth

4.1 Main Effect of the Phase 1 Grant

The effects of the Phase 1 grant award on firm growth (employees, payroll, revenue, and wages) are presented in Table 7. There are large effects on payroll, employment, and revenue. No effects were found on firm exit via acquisition or failure (unreported due to Census disclosure limitations).[24]

[22] A t-test of the difference of means strongly rejects the hypothesis that non-appliers and appliers have the same mean probability of VC investment within two years, with a t-statistic of 5.44.

[23] The time consuming and onerous process was given as a reason for not applying in interviews with grantees and investors.

[24] Exit via acquisition or merger is defined as an instance in which the last establishment year is later than the last firm year. This indicates that the establishment continues but the firm dies. Failure is defined as establishment and firm exit from the panel.

The effect of winning a grant on log employment after the application year relative to the base year is shown as the coefficient on $PostAward_{i,j,t}$ in columns 1-3. Put another way, the coefficient is the average effect of winning in years after the application year, controlling for whether the firm is a winning firm and whether the year is after the application year. The coefficients on quadratic rank (column 1) and on either side of the cutoff (column 2) are also shown. Firm-application fixed effects are included in column 3, which account for most of the controls used elsewhere. The coefficient of 0.27 on $PostAward_{i,j,t}$ indicates that a grant award increases employment relative to the base year by about 30 percent.[25] We can also calculate what the coefficient implies for employment growth among winners. Winners have about 19 percent more employees than non-winners, or on average 6.7 more employees, relative to the year before application.

Figure 6 Panel A demonstrates the effect on levels of log employment by rank around the cutoff. This figure provides visual evidence that there is a significant effect of the grant, as we see a jump at the cutoff for the award. Figure 7 Panel A demonstrates the effect on levels of log employment by quarter around the award quarter. This shows that the effect is quite immediate.

The effect on payroll in columns 4-6 (where the coefficients are 0.36-0.4) indicates a roughly 43 percent increase in payroll relative to the base year.[26] This translates to, at the mean, 29 percent more payroll than in the pre-application year. Figure 6 Panel B demonstrates the effect on levels of log payroll by rank around the cutoff, and Figure 7 Panel B demonstrates the effect on levels of log payroll by quarter around the award quarter. The effect on revenue in columns 7-9 indicates a 20 percent increase in revenue relative to the base year, or 15 percent more revenue than in the pre-application year. Figure 6 Panel C demonstrates the effect on levels of log revenue by rank around the cutoff. There is no quarterly graph because Census does not have quarterly revenue data.

[25]The coefficient gives the percentage change in $\frac{Y_{i,t}}{Y_{i,t=-1}}$ associated with being an award recipient relative to a non-winner. The exact effect is $100 * (e^{\beta} - 1)$. Note it is relative to the year before the application (that is, the effect is not an absolute increase).

[26]The coefficient gives the percentage change in $\frac{Y_{i,t}}{Y_{i,t=-1}}$ associated with being an award recipient relative to a non-winner. The exact effect is $100 * (e^{\beta} - 1)$. Note it is relative to the year before the application (that is, the effect is not an absolute increase).

Table 7: Phase 1 Grant Effect on Firm Growth Outcomes

Dependent variable:	Employment			Payroll			Revenue		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$PostAward_{i,j,t}$	.271***	.262***	.27***	.406***	.396***	.365***	.19***	.183***	.273***
	(0.0984)	(0.0968)	(0.0795)	(0.109)	(0.107)	(0.0898)	(0.0614)	(0.0613)	(0.0707)
$Award_{i,j}$	-0.073	0.0348	0.019						
	(0.0752)	(0.0889)	(0.0333)						
$Post_{i,j,t}$	-0.0333	-0.0321							
	(0.0374)	(0.0375)							
$Rank_{i,j}$	0.00352								
	(0.00773)								
$Rank^2_{i,j}$	0.000107								
	(0.000189)								
$Rank \mid win_{i,j}$		-0.0584							
		(0.036)							
$Rank \mid lose_{i,j}$		0.000996							
		(0.0024)							
Controls									
$Award_{i,j}$	-	-	-	Y	Y	N	Y	Y	N
$Post_{i,j,t}$	-	-	N	Y	Y	Y	Y	Y	Y
$Rank_{i,j}$, $Rank^2_{i,j}$	-	N	N	Y	N	N	Y	N	N
$Rank \mid win/lose_{i,j}$	N	-	N	N	Y	N	N	Y	N
$Age_{i,t}$, $Age^2_{i,t}$	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fixed Effects									
Year_t	Y	Y	Y	Y	Y	Y	Y	Y	Y
Competition_j	Y	Y	N	Y	Y	N	Y	Y	N
$\text{Firm-application}_{i,j}$	N	N	Y	N	N	Y	N	N	Y
N	30500	30500	30500	30500	30500	30500	13000	13000	13000
R^2	0.21	0.21	0.532	0.151	0.152	0.478	0.143	0.141	0.528

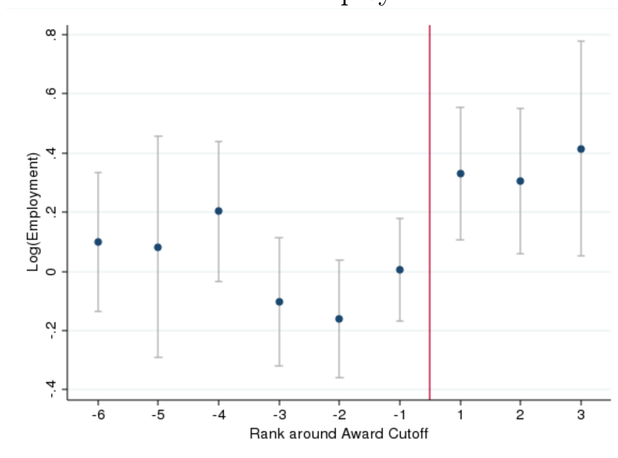
Note: This panel shows the effect of the grant on log growth outcomes, using Equation 2. The base year for the dependent variables is $t = -1$, the year before the application year. Control coefficients are shown only for employment and no other outcomes to minimize disclosure requirements. None of these variables have any statistical significance in subsequent columns. Data are observed at the firm-year level. Standard errors are clustered by competition. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

The effect is quite similar across the three specifications shown in Table 7. The results are also robust to a number of unreported approaches (unreported because Census disclosure requirements limit the number of estimates we may release). First, they are similar using narrow years around the application. Second, the results go through with a bandwidth of one firm around the cutoff. Third, when the sample is split roughly in half, around either 2005 or 2008, there are similar effects on either side. The magnitude of the effect is somewhat larger in the early period (i.e. before 2005 or 2008), but not statistically significantly so.

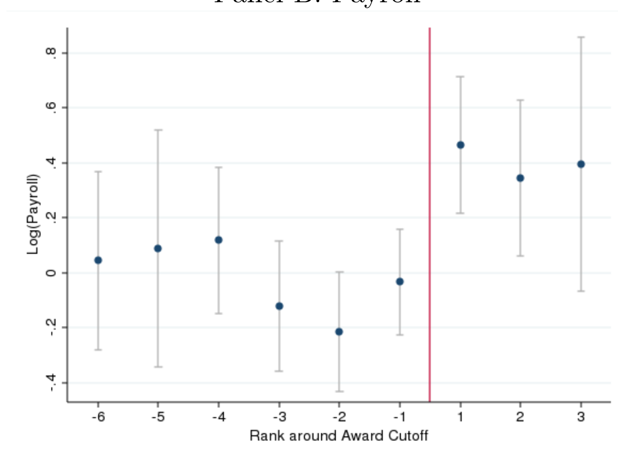
The effects occur quickly. Table 8 shows that about half the effect on employment occurs within two years of the grant application, and just more than half for payroll and revenue. The large magnitude of the effects relative to the size of the grant (just $150,000) implies that the grant is invested in productive activities.

Figure 6: Phase 1 Effects on Firm Growth by Rank Around the Award Cutoff

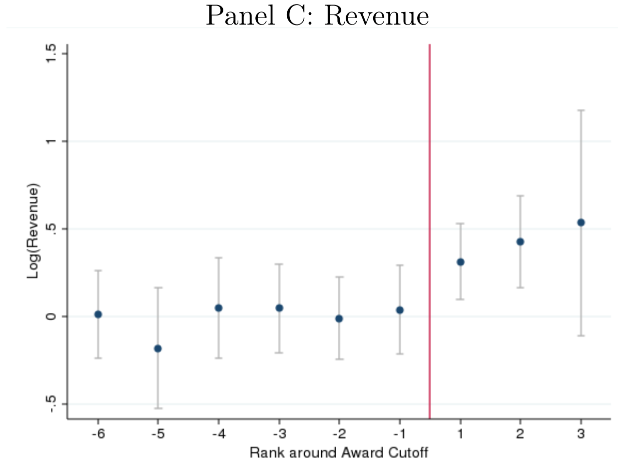
Panel A: Employment



Panel B: Payroll



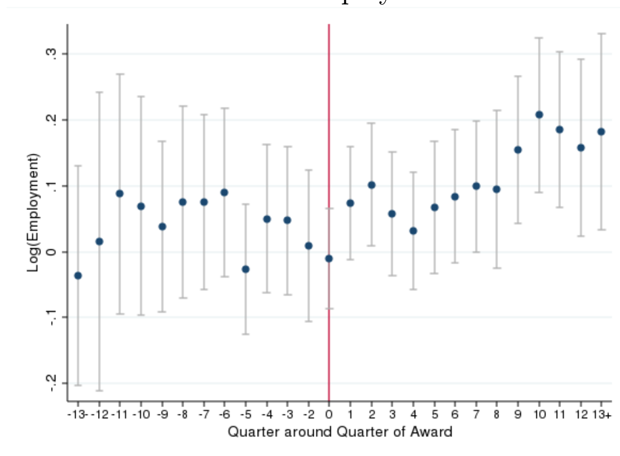
Panel C: Revenue



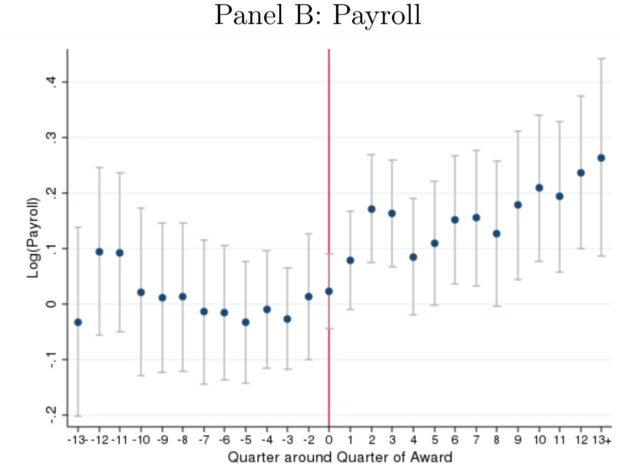
Note: These figures show the results from estimating Equation 3 on levels of log firm-year employment, payroll, and revenue. Each point is a coefficient on a specific DOE-assigned rank around the award cutoff, where positive ranks are winning applicant firms, and negative ranks are non-winning applicant firms. 95% confidence intervals are shown.

Figure 7: Phase 1 Quarterly Effects on Firm Growth

Panel A: Employment



Panel B: Payroll



Note: These figures show the results from estimating Equation 4 on quarterly levels of log firm-year employment and payroll. Each point is a coefficient on a quarter around the award quarter interacted with winning an award. The base quarter is -1 (immediately before the quarter of award). There are no quarterly revenue or employee-level wage (and thus inequality) data. 95% confidence intervals are shown.

Table 8: Grant Effect on Firm Growth Outcomes Within Two Years of Grant Application

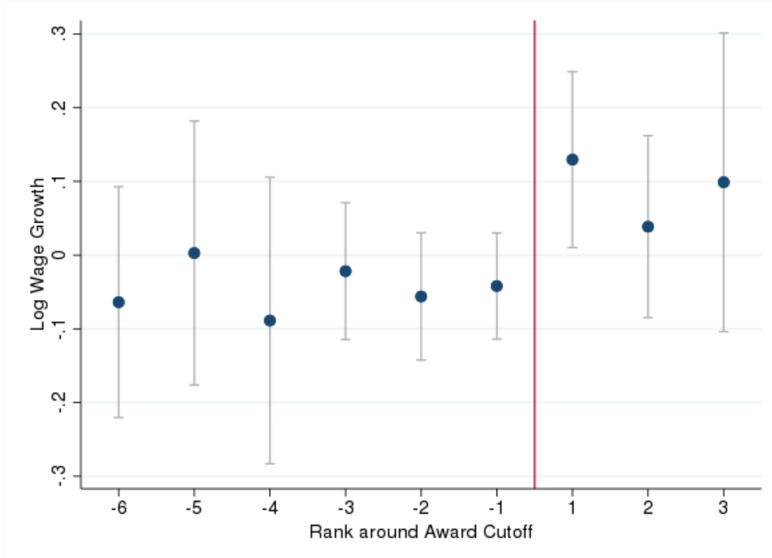
Dependent variable:	Employment	Payroll	Revenue
	(1)	(2)	(3)
$PostAward_{i,j,t}$	.142**	.268***	.159**
	(0.0573)	(0.0672)	(0.0624)
Controls			
$Award_{i,j}$	Y	Y	Y
$Post_{i,j,t}$	Y	Y	Y
$Rank_{i,j}$, $Rank^2_{i,j}$	Y	Y	Y
$Age_{i,t}$, $Age^2_{i,t}$	Y	Y	Y
Fixed Effects			
Year_t	Y	Y	Y
Competition_j	Y	Y	Y
N	20000	20000	9500
R^2	0.244	0.178	0.426

Note: This panel shows the effect of the grant on log growth outcomes in the two years after the grant application year (including the application year), using Equation 2. The base year for the dependent variables is $t = -1$, the year before the application year. Control coefficients are shown only for employment to minimize disclosure requirements. None of these variables have any significance in subsequent columns. Data are observed at the firm-year level. Standard errors are clustered by competition. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

4.2 The Effect of the Phase 1 Grant on Average Wages

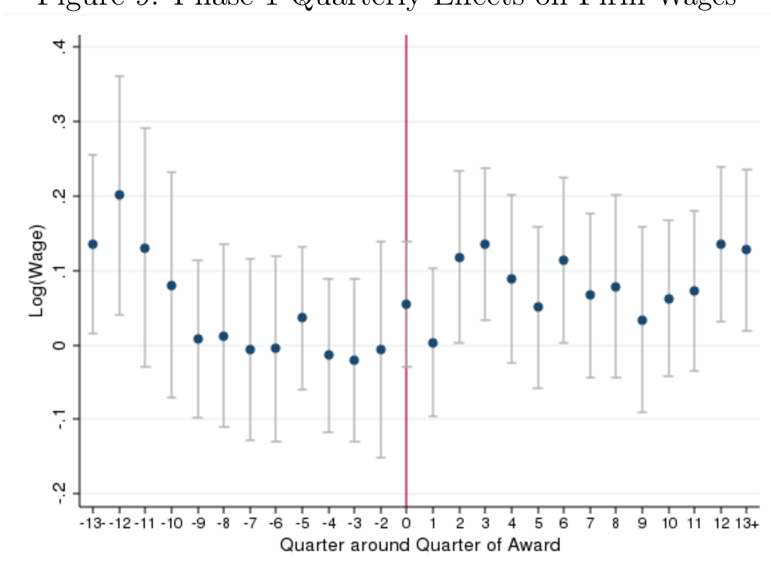
There may be interest in the effect of Phase 1 on the firm's average wage. Table 9 shows the grant effect on wage growth, with the three main specifications based on Equation 2 in columns 1-3. A grant increases wage growth in the years following the grant by about 10 percent in the most conservative result with firm-application fixed effects (column 3), and 14 percent in the main specification where rank is controlled for quadratically (column 1). (We control for rank quadratically to account for potential non-linearity in the effect of rank.) Using the larger result, the Phase 1 effect translates to about an 11 percent increase in the average wage relative to the pre-application year. Figure 8 demonstrates the effect on levels of log wages by rank around the cutoff, and Figure 9 shows the effect on levels of log wages by quarter around the award quarter.

Figure 8: Phase 1 Effects on Firm Wages by Rank Around the Award Cutoff



Note: This figure show the results from estimating Equation 3 on levels of log firm-year average wages. Each point is a coefficient on a specific DOE-assigned rank around the award cutoff, where positive ranks are winning applicant firms, and negative ranks are non-winning applicant firms. Only two positive ranks for inequality are reported, because the smaller sample led to a very large confidence interval for the firm three ranks away from the cutoff (which exists in competitions with at least three winners). The coefficient magnitude is in line with the previous two. 95% confidence intervals are shown.

Figure 9: Phase 1 Quarterly Effects on Firm Wages



Note: This figure shows the results from estimating Equation 4 on quarterly levels of log firm-year average wages. Each point is a coefficient on a quarter around the award quarter interacted with winning an award. The base quarter is -1 (immediately before the quarter of award). 95% confidence intervals are shown.

As with the Phase 1 effects on payroll, employment, and revenue, the wage increase occurs quickly. Almost the entire effect is observed within two years, as shown in column 4. Column 5 of Table 9 shows the wage growth of the firm founder. The Phase I grant has a much larger founder wage effect than average of about 47 percent. As the individual perhaps most responsible for the firm's past and future productivity growth, and for having won the grant, it is not surprising that the founder experiences a larger wage increase.

Table 9: Phase 1 Grant Effect on Wages

Dependent variable:	Wage growth					
				Within 2 years	Of firm founder	
	(1)	(2)	(3)	(4)	(5)	(6)
$PostAward_{i,j,t}$	.134***	.133***	.0946**	.126***	.39*	
	(0.048)	(0.0481)	(0.0391)	(0.0406)	(0.206)	
$PostAwardPerEmp_{i,j,t}$						.0128**
						(0.00519)
Controls						
$Award_{i,j}$	Y	Y	N	Y	Y	Y
$Post_{i,j,t}$	Y	Y	Y	Y	Y	Y
$Rank_{i,j},\ Rank^2_{i,j}$	Y	N	N	Y	Y	Y
$Rank \mid win/lose_{i,j}$	N	Y	N	N	N	N
$Age_{i,t},\ Age^2_{i,t}$	Y	Y	Y	Y	Y	Y
$AwardPerEmp_{i,j,t}$	N	N	N	N	N	Y
Fixed effects						
Year_t	Y	Y	Y	Y	Y	Y
Competition_j	Y	Y	N	Y	Y	Y
$\text{Firm-application}_{i,j}$	N	N	Y	N	N	N
N	30500	30500	30500	20000	1600	30500
R^2	0.0988	0.099	0.449	0.0738	0.288	0.0986

Note: This panel shows the effect of the grant on wage growth, using Equation 2. The base year is $t = -1$, the year before the application year. Columns 1-3 replicate the three main specifications from Table 7, with rank controlled for quadratically, on either side of the cutoff, or through firm-application fixed effects. Column 4 restricts the sample to the two years after the grant application year (including the application year). Column 5 examines the effect of the grant on the wage of the firm founder. Column 6 uses an alternative independent variable: $PostAwardPerEmp_{i,j,t}$ is $PostAward_{i,j,t}$ interacted with the logged grant amount per employee, where the number of employees is identified in year $t = -1$. Control coefficients are not reported to minimize disclosure requirements. Data are observed at the firm-year level. Except in column 5, wage is the computed as the average wage within the firm-year. Standard errors are clustered by competition. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

4.3 Variation in the Phase 1 Effect on Firm Growth

For all firm growth outcomes, potential variation in the effect was tested for a wide array of firm, firm founder, industry, and state characteristics. The only robust variation is shown in Table 10. The grant is more useful for smaller and younger firms, consistent with the findings for patenting shown above. A similar result was found for venture capital in Howell (2017). Columns 1 and 4 show the effect of winning a grant award interacted with log employment in the year before the grant application. The effect is large and negative, indicating that firms that are larger at the time of application experience a smaller effect.

Columns 2 and 5 similarly show that the effects of a grant on firm employment and payroll are decreasing in firm age at the time of application. Columns 3 and 6 interact winning with an indicator for a firm not having previous SBIR grants from other agencies. (Recall that only first-time winners are included in the panel data, so it is not possible to consider previous DOE SBIR wins.) The effect on the interaction is large and negative, albeit not statistically significant. The three characteristics in this table (size, age, and previous wins) operate in the same direction for wage growth but are statistically insignificant. There is no heterogeneity whatsoever on within-firm inequality.

It is worth mentioning key tests that yielded no statistically significant interaction effects. There is no effect of founder gender, age, tenure, or race/ethnicity, nor is there heterogeneity in the share of employees of a certain gender, age bin, tenure bin, or race/ethnicity. There is also no effect by worker tenure.

Table 10: Variation in the Phase 1 Grant Effect on Firm Growth Outcomes

Dependent variable:	Employment			Payroll		
	(1)	(2)	(3)	(4)	(5)	(6)
$PostAward_{i,j,t} \cdot Employment_{t=-1}$	-.167*			-.201**		
	(0.0942)			(0.0832)		
$PostAward_{i,j,t} \cdot Age_{t=-1}$		-.0315**			-.0339***	
		(0.0135)			(0.0131)	
$PostAward_{i,j,t} \cdot NoPrevWins_{t=-1}$			-0.226			-0.115
			(0.208)			(0.206)
$PostAward_{i,j,t}$	.859***	.733***	.364**	.861***	.679***	.255**
	(0.289)	(0.191)	(0.14)	(0.256)	(0.176)	(0.129)
$Employment_{t=-1}$	-.169***			-.19***		
	(0.0241)			(0.0183)		
$Age_{t=-1}$		.0167**			.0079**	
		(0.0076)			(0.0543)	
$NoPrevWins_{t=-1}$			.217***			.188***
			(0.062)			(0.0473)
Controls						
$Award_{i,j}$	Y	Y	Y	Y	Y	Y
$Post_{i,j,t}$	Y	Y	Y	Y	Y	Y
$Award_{i,j} \cdot Employment_{t=-1}$	Y	N	N	Y	N	N
$Post_{i,j,t} \cdot Employment_{t=-1}$	Y	N	N	Y	N	N
$Award_{i,j} \cdot Age_{t=-1}$	N	Y	N	N	Y	N
$Post_{i,j,t} \cdot Age_{t=-1}$	N	Y	N	N	Y	N
$Award_{i,j} \cdot NoPrevWins_{t=-1}$	N	N	Y	N	N	Y
$Post_{i,j,t} \cdot NoPrevWins_{t=-1}$	N	N	Y	N	N	Y
$Rank_{i,j},\ Rank^2_{i,j}$	Y	Y	Y	Y	Y	Y
$Age_{i,t},\ Age^2_{i,t}$	Y	Y	Y	Y	Y	Y
Fixed effects						
Year_t	Y	Y	Y	Y	Y	Y
Competition_j	Y	Y	Y	Y	Y	Y
N	30500	30500	30500	30500	30500	30500
R^2	0.189	0.175	0.175	0.244	0.214	0.214

Note: This table shows the effect of the grant interacted with firm characteristics on log growth outcomes, using Equation 2. The base year for the dependent variables is $t = -1$, the year before the application year. $Employment_{t=-1}$ is log employment in year $t = -1$, $Age_{t=-1}$ is firm age in years in year $t = -1$, and $NoPrevWins_{t=-1}$ is an indicator for the firm not having previously won an SBIR award from a non-DOE agency. Control coefficients are not reported to minimize disclosure requirements. Data are observed at the firm-year level. Standard errors are clustered by competition. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

4.4 The Phase 2 Grant Effect on Firm Growth

The Phase 2 grant was not found to have any statistically significant effects on firm growth. These null results are shown in Table 11. The coefficients are statistically insignificant and considerably smaller than the effects of the Phase 1 grant. As with patenting, because the sample is small, rank controls and competition fixed effects are omitted. The results are similar with these additional controls, or when Phase 1 and 2 are considered in the same regression. As explained in Section 3.3, the small sample and insignificant effects may reflect adverse selection in firms' decisions to apply to Phase 2.

Table 11: Phase 2 Grant Effect on Firm Growth Outcomes

Dependent variable:	Employment	Payroll	Revenue
	(1)	(2)	(3)
$PostAward_{i,j,t}$	0.0899	0.178	-0.0879
	(0.193)	(0.173)	(0.0666)
Controls			
$Award_{i,j}$	Y	Y	Y
$Post_{i,j,t}$	Y	Y	Y
Fixed Effects			
Year_t	Y	Y	Y
N	3100	3100	3100
R^2	0.384	0.464	0.272

Note: This panel shows the effect of the Phase 2 grant on log growth outcomes, using Equation 2. The base year for the dependent variables is $t = -1$, the year before the application year. Control coefficients are not reported to minimize disclosure requirements. Data are observed at the firm-year level. Standard errors are clustered by competition. $*$, $**$, and $***$ denote significance at the 10%, 5%, and 1% levels.

5 Conclusion

To the author's knowledge, this study offers the first large sample analysis of R&D subsidies to private firms that establishes a truly causal relationship between the grants and firm outcomes, in large part because ranked application data for a R&D subsidy program has never before been made available for research. This study may offer government agencies around the world a template for rigorous, external analysis of their R&D subsidy programs.

This study demonstrates that U.S. DOE Phase 1 SBIR grants have large, positive effects on firm innovation, growth, and average wages. In particular, receiving a Phase 1 grant award increases a firm's subsequent patents weighted by future citations by about 2.5 times, relative to an average of 12. The $1 million Phase 2 grant doubles a firm's cite-weighted patents, relative to a mean of 20. This Phase 2 effect is much smaller than the Phase 1 effect on a per-grant-dollar basis. Also, the effect of Phase 2 falls and loses significance when the sample includes previous winners.

The Phase 1 grant also leads firms to have about 19 percent more employees relative to the year of application than they would have had otherwise. The similar result for payroll is 29 percent, and the effect on revenue is 15 percent. The grant also increases firm wages by about 11 percent. The Phase 2 grant was not found to have any effects on firm employment, payroll, revenue, or wage growth.

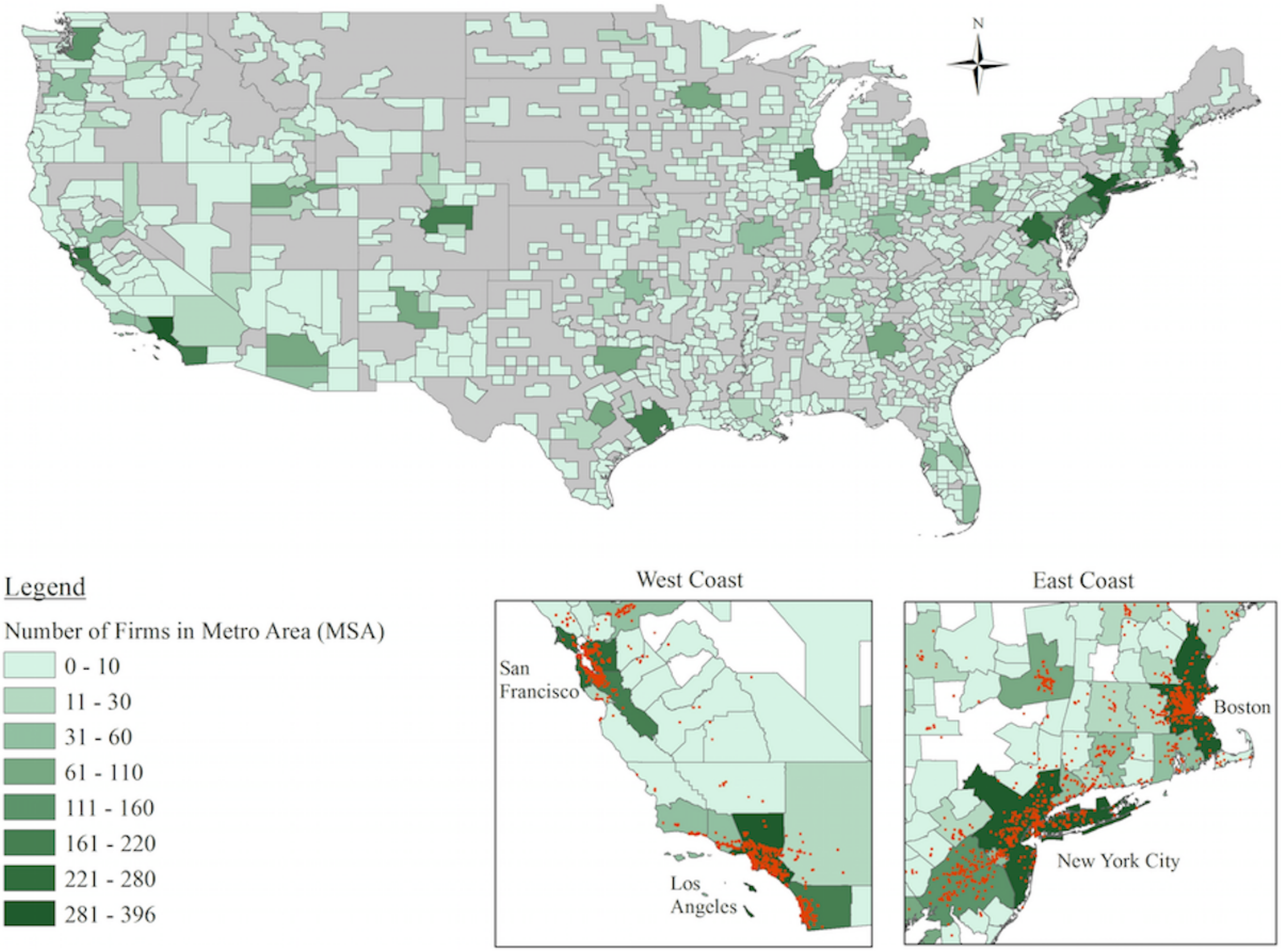
The Phase 1 grants are useful because they enable proof-of-concept or prototyping work. The firms that seem to benefit most from the SBIR Phase 1 are startups. With a successful proof-of-concept, a startup can show to investors that its technology concept is valid. A second avenue is certification; the government's decision might convey positive information to venture capitalists about the firm's technology. For additional discussion about the mechanism, see Howell (2017) provides additional discussion and evidence about the grants' mechanisms. However, future research could more affirmatively establish how grants are useful to firms, at what stage in the firm lifecycle, and for what types of firms.

One reason for the effectiveness of Phase 1 is that applicant firms may be financially constrained. For early-stage projects in general, it is difficult to access conventional small business bank loans, and prospective equity investors have substantial uncertainty about a technology's potential. Further, energy technology startups are more capital intensive, have longer lead times, and carry higher project finance and market risk than the startups VCs typically finance in IT and biotech (Nanda, Younge and Fleming 2013). Finally, commercializing clean energy technologies is challenging in the absence of a carbon price (Nordhaus 2013). All these reasons help explain why the grants do not appear to crowd out private investment. The importance of some of these factors could be tested by applying this type of analysis to other agencies' SBIR programs, such as those of the National Institutes of Health or the National Science Foundation.

6 Appendix: Geography and Firm Clustering

The geography of SBIR applicants corresponds to what might be expected of high-tech firms. Figure 10 shows the applicant concentrations by Metropolitan Statistical Area (MSA). Applicant locations were geocoded using address information. The largest clusters by far are Boston and San Francisco, and to a lesser extent New York, Los Angeles, Denver/Boulder, and the greater DC metro area.

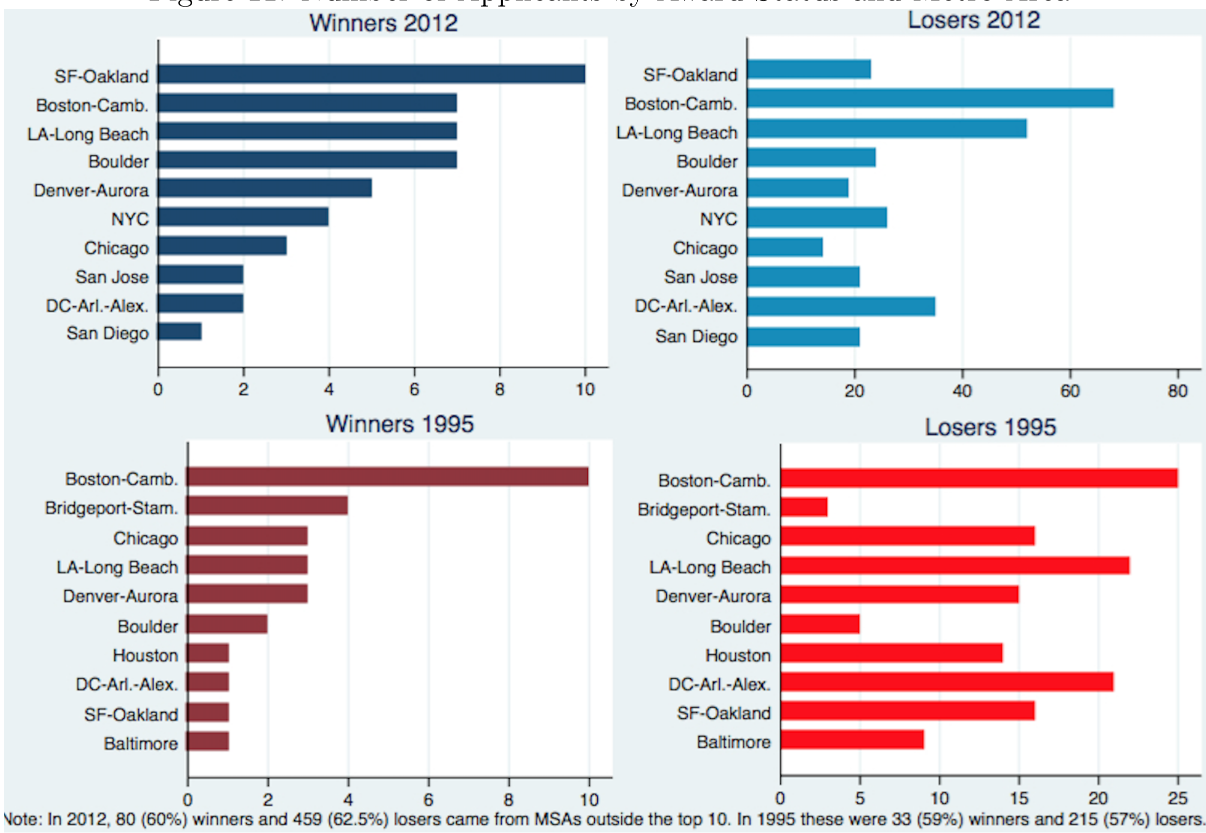
Figure 10: Applicant Firm Locations



Note: This figure shows the location of all applicant firms in the data. In the main figure, a darker color for a metropolitan statistical area (MSA) indicates higher firm density. In the insets, actual firm locations are overlaid as orange dots.

The number of applicants by award status from the top ten metro regions with the highest number of applicants in 1995 and in 2012 are in Figure 11. San Francisco moves from 9th place to 1st place, reflecting its growing role in the energy technology sector. LA and Boston are near the top of the list in both years. Bridgeport-Stamford and Baltimore fall completely off the list, while NYC and DC enter. This suggests that the changing agglomerations of SBIR winners over time may reflect cities' lifecycles. Graphs by year, not shown here, suggest that the concentration has not changed much over time except for the San Francisco area, which has increased in importance since the mid-1990s.

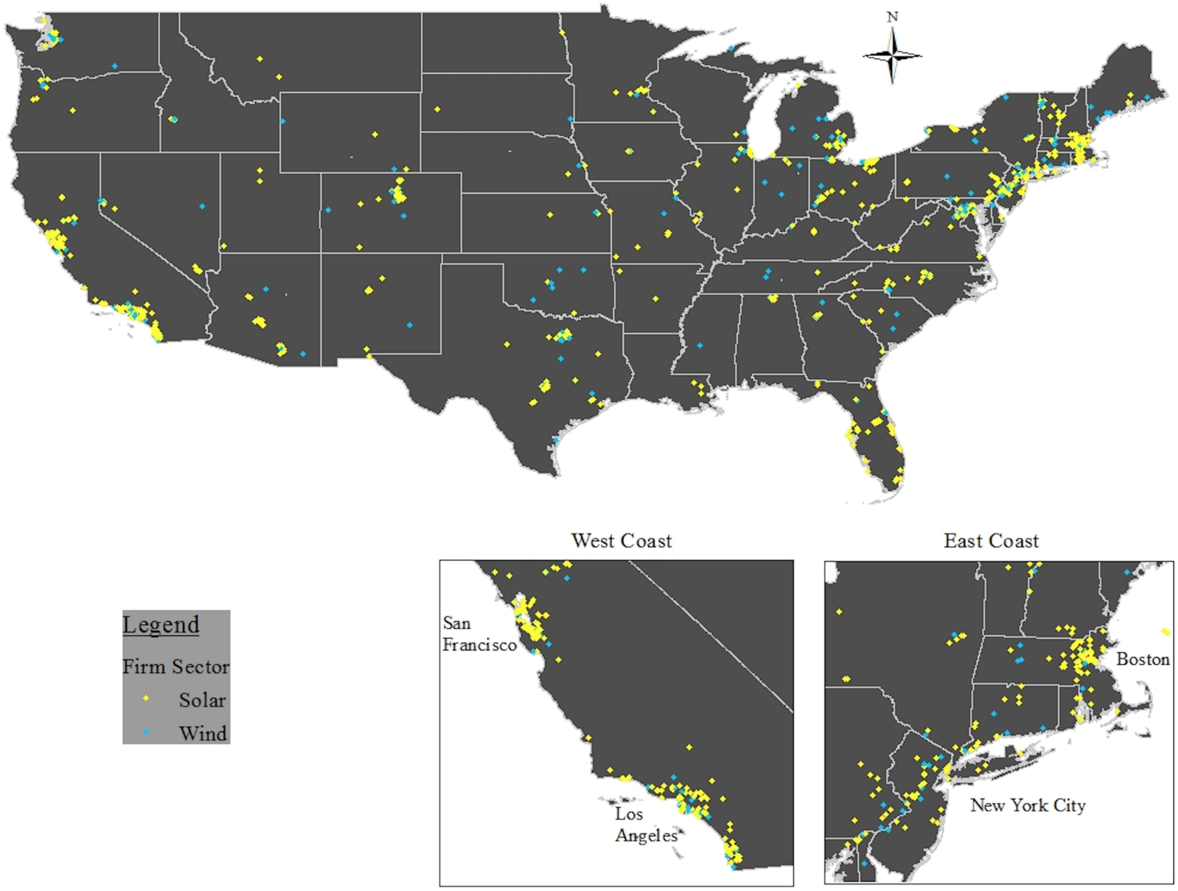
Figure 11: Number of Applicants by Award Status and Metro Area



Note: This figure shows the number of applicants by award status (whether they won or lost) from the top 10 EERE/FE SBIR applicant metropolitan statistical areas.

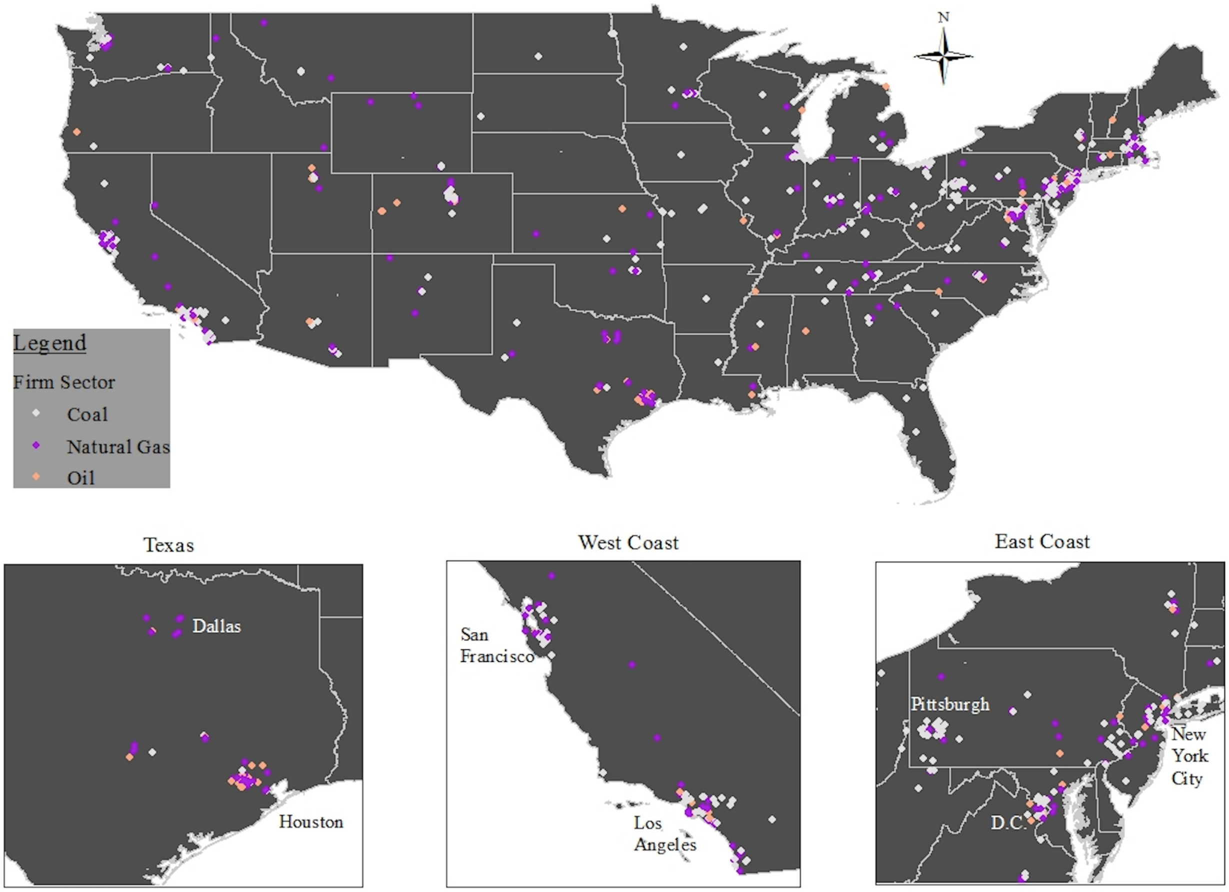
Wind and solar applicants (categorized by the competition topic area) are in Figure 12, and oil, gas and coal applicants in Figure 13. It is clear that although both renewable and conventional fuel companies locate in major cities, some clustering relates to the area's resource base. Clusters of coal companies in Pittsburgh, PA and oil and gas companies in Houston, TX contrast with clusters of solar firms in Tampa, FL and Orlando, FL.

Figure 12: Solar and Wind SBIR Firms



Note: This figure shows the location of unique EERE/FE SBIR applicant firms in the solar and wind industries.

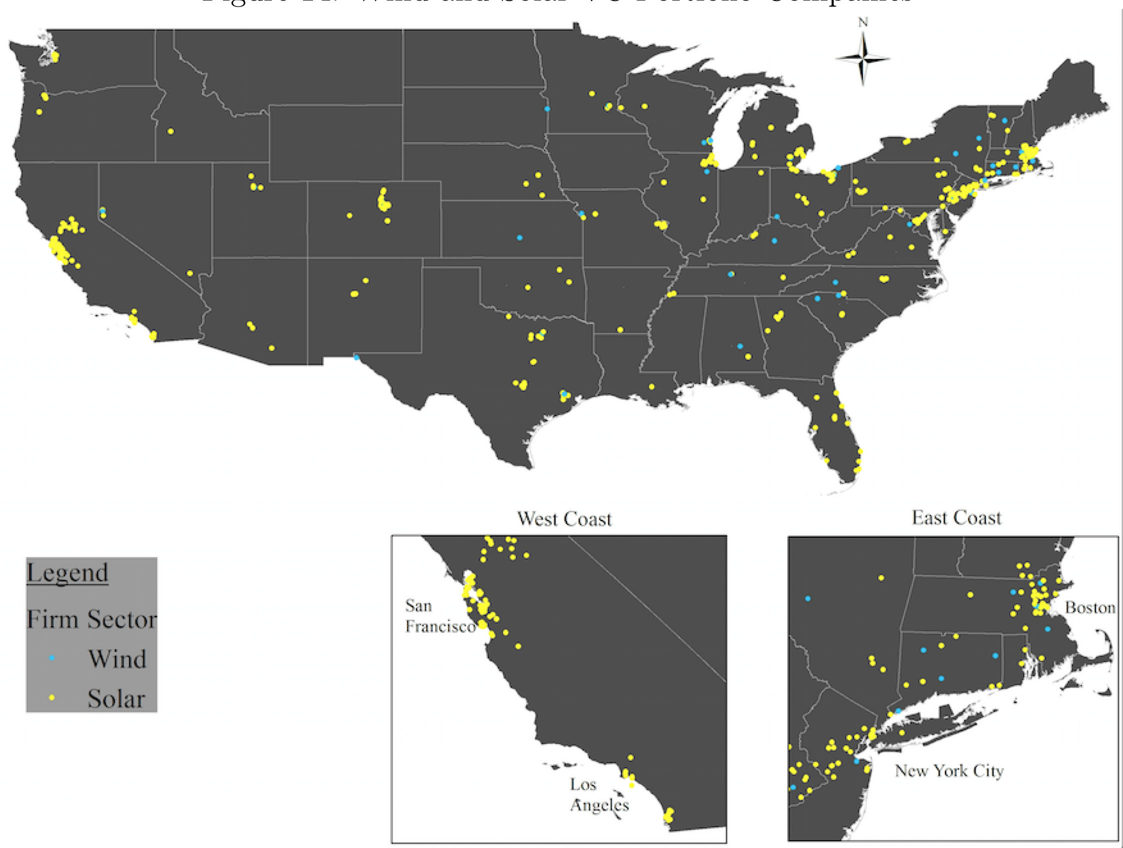
Figure 13: Oil, Natural Gas, and Coal SBIR Firms



Note: This figure shows the location of unique EERE/FE SBIR applicant firms in the oil, natural gas, and coal industries.

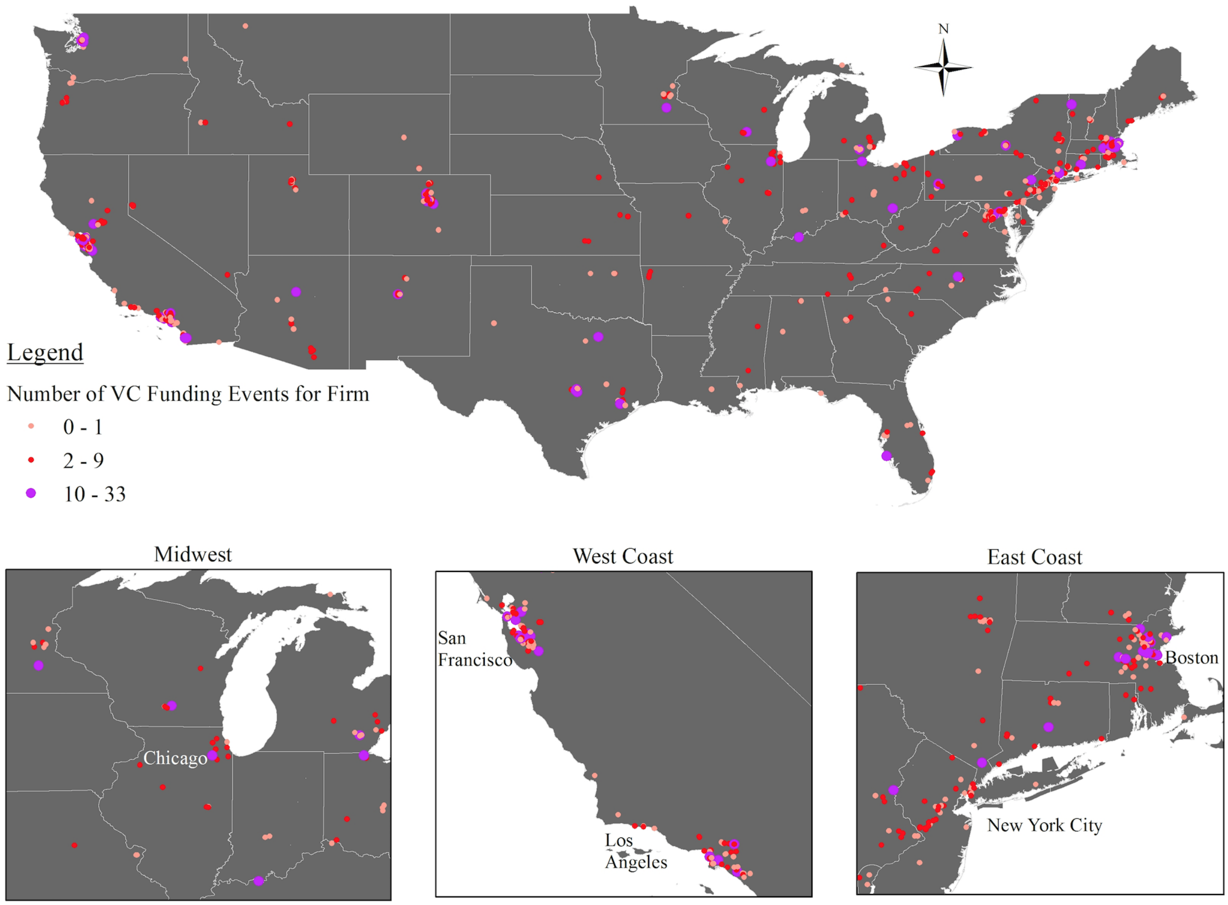
Figure 12 shows the location of all wind and solar companies that venture capital (VC) firms have invested in, based on the ThompsonOne database. Agglomeration in Boston and San Francisco, and to a lesser degree in New York, Denver and Chicago are common to both the SBIR applicants (Figure 16) and the VC portfolio companies (Figure 14) in these clean energy sectors. However, the portfolio companies are more concentrated in the major cities where VC firms are also located. We can see this by examining the location of applicants with at least one VC deal (Figure 15), and comparing to the concentration by MSA of all active VC firms in the Preqin database that, according to Preqin, invest in clean technology (Figure 16).

Figure 14: Wind and Solar VC Portfolio Companies



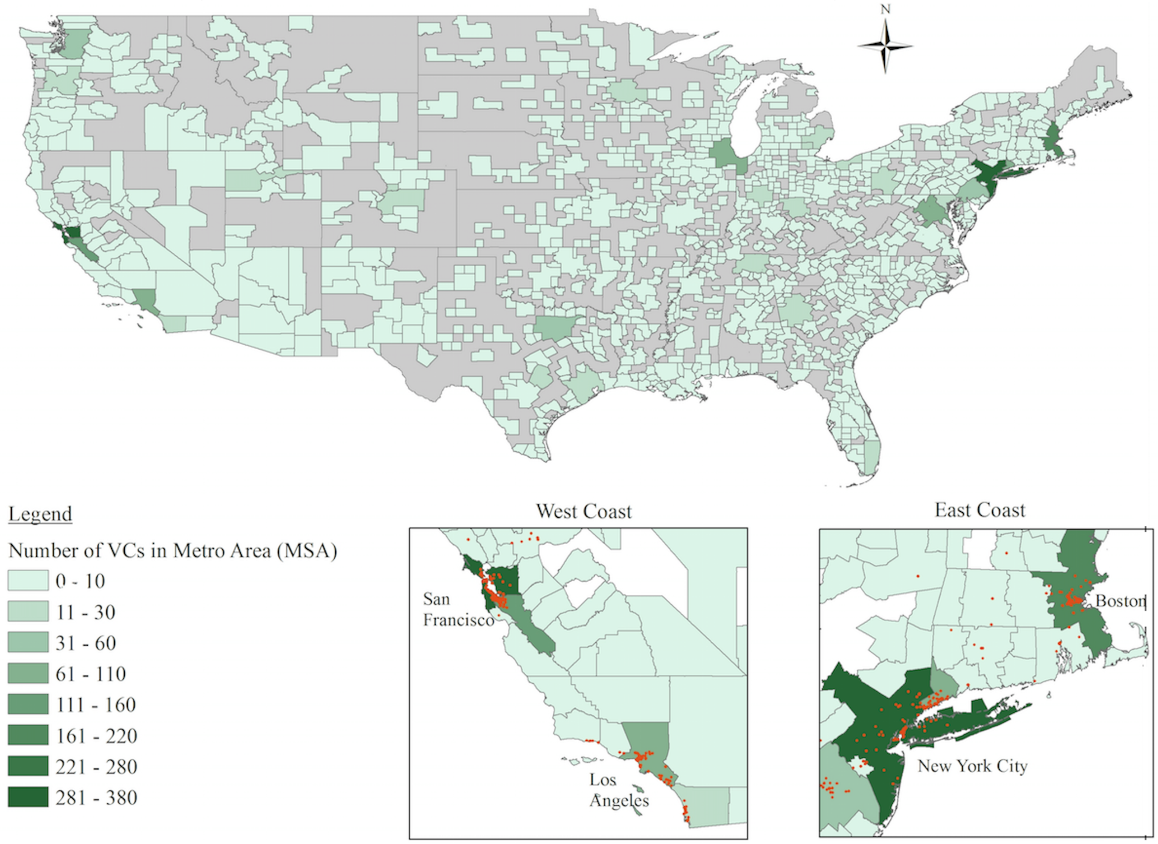
Note: This figure shows the location of unique VC-backed firms in the solar and wind industries from the ThompsonOne database.

Figure 15: SBIR Applicant Firms with at Least 1 VC Deal



Note: This figure shows the location of unique EERE/FE SBIR applicant firms that have at least one VC deal.

Figure 16: Concentration of Clean Tech-Investing VC Firms



Note: This figure shows the location by metropolitan statistical area of VC firms that invest in clean technology using the whole Preqin database.

In general, the clustering of both VC firms and VC-funded DOE SBIR applicants aligns fairly closely with the clustering of the overall applicant pool. However, there is considerably more clustering of both VC firms and VC-funded companies in Boston and San Francisco, especially VC-funded companies that have received many VC investment rounds. Meanwhile, there are far more SBIR applicants from Los Angeles and from the greater DC metro area than portfolio companies. For the subset of firms that focus their resources on government grants and procurement contracts, the DC concentration makes sense. Los Angeles also seems be a long-term hub of government-oriented tech companies. For example, Physical Optics - the largest SBIR winner in my data - is located in Torrance, CA, within the Los Angeles MSA. The Los Angeles government provides supporting activities, such as regular workshops on applying for SBIR grants and other informational and convening resources through its PortTech Los Angeles program, Los Angeles Regional Small Business Development Center, and others.

References

Aghion, P., Van Reenen, J., & Zingales, L. (2013). Innovation and institutional ownership. American Economic Review, 103(1), 277-304.

Akcigit, U., & Kerr, W. R. (2018). Growth through heterogeneous innovations. Journal of Political Economy, 126(4), 1374-1443.

Almus, M., & Czarnitzki, D. (2003). The effects of public R&D subsidies on firms' innovation activities: The case of eastern Germany. Journal of Business & Economic Statistics, 21(2), 226-236.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. Journal of Business & Economic Statistics, 19(1), 2-16.

Arora, A., Ceccagnoli, M., & Cohen, W. M. (2008). R&D and the patent premium. International Journal of Industrial Organization, 26(5), 1153-1179.

Audretsch, D. B., Keilbach, M. C., & Lehmann, E. E. (2006). Entrepreneurship and economic growth. Oxford University Press.

Azoulay, P., Jones, B., Kim, J. D. & Miranda, J. (2018). Age and high-growth entrepreneurship. Working paper, available at https://siepr.stanford.edu/system/files/Age%20and%20High%20Growth%20Ent

Babina, T., & Howell, S. T. (2018). Entrepreneurial spillovers from corporate R&D (No. w25360). National Bureau of Economic Research, available at https://www.nber.org/papers/w25360.

Benavente, J. M., Crespi, G., Figal Garone, L., & Maffioli, A. (2012). The impact of national research funds: A regression discontinuity approach to the Chilean FONDECYT. Research Policy, 41(8), 1461-1475.

Berk, J. B., Green, R. C., & Naik, V. (2004). Valuation and return dynamics of new ventures. Review of Financial Studies, 17(1), 1-35.

Blasio, G., Fantino, D., & Pellegrini, G. (2011). Evaluating the impact of innovation incentives: Evidence from an unexpected shortage of funds. Industrial and Corporate Change, 23(5), 1-30.

Bloom, N., Griffith, R., & Van Reenen, J. (2002). Do R&D tax credits work? Evidence from a panel of countries 1979–1997. Journal of Public Economics, 85(1), 1-31.

Bloom, N., Schankerman, M., & Van Reenen, J. (2013). Identifying technology spill-overs and product market rivalry. Econometrica, 81(4), 1347-1393.

Busom, I. (2000). An empirical evaluation of the effects of R&D subsidies. Economics of Innovation and New Technology, 9(2), 111-148.

Bronzini, R., & Iachini, E. (2014). Are incentives for R&D effective? Evidence from a regression discontinuity approach. American Economic Journal: Economic Policy, 6(4), 100-134.

Brouwer, E., & Kleinknecht, A. (1999). Innovative output, and a firm's propensity to patent: An exploration of CIS micro data. Research Policy, 28(6), 615-624.

Brown, J. R., Fazzari, S. M., & Petersen, B. C. (2009). Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. Journal of Finance, 64(1), 151-185.

Clausen, T. H. (2009). Do subsidies have positive impacts on R&D and innovation activities at the firm level?. Structural Change and Economic Dynamics, 20(4), 239-253.

Conti, A., Thursby, J., & Thursby, M. (2013). Patents as signals for startup financing. Journal of Industrial Economics, 61(3), 592-622.

Czarnitzki, D., & Bento, C. L. (2012). Evaluation of public R&D policies: A cross–country comparison. World Review of Science, Technology and Sustainable Development, 9(2), 254-282.

Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K. T., & Van Reenen, J. (2016). Do tax incentives for research increase firm innovation? An RD design for R&D (No. w22405). National Bureau of Economic Research.

Duguet, E. (2004). Are R&D subsidies a substitute or a complement to privately funded R&D?. Revue d'économie politique, 114(2), 245-274.

Eaton, J., & Kortum, S. (1999). International technology diffusion: Theory and measurement. International Economic Review, 40(3), 537-570.

Gans, J. S., & Stern, S. (2003). The product market and the market for "ideas": Commercialization strategies for technology entrepreneurs. Research Policy, 32(2), 333-350.

González, X., Jaumandreu, J., & Pazó, C. (2005). Barriers to innovation and subsidy effectiveness. RAND Journal of Economics, 930-950.

González, X., & Pazó, C. (2008). Do public subsidies stimulate private R&D spending? Research Policy, 37(3), 371-389.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica, 69(1), 201-209.

Hall, B. H. (1993). R&D tax policy during the 1980s: Success or failure?. Tax Policy and the Economy, 7, 1-35.

Hall, B. H. (2008). The financing of innovation. In S. Shane (Ed.), Handbook of Technology and Innovation Management (pp. 409-430). Oxford: Blackwell Publishers, Ltd.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. RAND Journal of Economics, 16-38.

Hall, B., & Van Reenen, J. (2000). How effective are fiscal incentives for R&D? A review of the evidence. Research Policy, 29(4-5), 449-469.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who creates jobs? Small versus large versus young. Review of Economics and Statistics, 95(2), 347-361.

Henningsen, M. S., Hægeland, T., & Møen, J. (2015). Estimating the additionality of R&D subsidies using proposal evaluation data to control for research intentions. Journal of Technology Transfer, 40(2), 227-251.

Hochberg, Y. V., Serrano, C. J., & Ziedonis, R. H. (2018). Patent collateral, investor commitment, and the market for venture lending. Journal of Financial Economics, 130(1), 74-94.

Howell, S. T. (2017). Financing innovation: Evidence from R&D grants. American Economic Review, 107(4), 1136-64.

Hsu, D. H., & Ziedonis, R. H. (2008). Patents as quality signals for entrepreneurial ventures. In Academy of Management Proceedings (Vol. 2008(1), pp. 1-6). Briarcliff Manor, NY: Academy of Management.

Jacob, B. A., & Lefgren, L. (2011). The impact of research grant funding on scientific productivity. Journal of Public Economics, 95(9-10), 1168-1177.

Jaffe, A. B. (2002). Building programme evaluation into the design of public research-support programmes. Oxford Review of Economic Policy, 18(1), 22-34.

Kerr, S. P., & Kerr, W. R. (2016). Immigrant entrepreneurship. In J. Haltiwanger, E. Hurst, J. Miranda, & A. Schoar (Eds.), Measuring Entrepreneurial Businesses: Current Knowledge and Challenges (pp. 187-249). University of Chicago Press.

Lach, S. (2002). Do R&D subsidies stimulate or displace private R&D? Evidence from Israel. Journal of Industrial Economics, 50(4), 369-390.

Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. Journal of Econometrics, 142(2), 655-674.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. Journal of Economic Literature, 48(2), 281-355.

Lerner, J. (2000). The government as venture capitalist: The long-run impact of the SBIR program. Journal of Private Equity, 3(2), 55-78.

Lerner, J. (2009). Boulevard of broken dreams: Why public efforts to boost entrepreneurship and venture capital have failed–and what to do about it. Princeton University Press.

Link, A. N., & Scott, J. T. (2010). Government as entrepreneur: Evaluating the commercialization success of SBIR projects. Research Policy, 39(5), 589-601.

Mamuneas, T. P., & Nadiri, M. I. (1996). Public R&D policies and cost behavior of the US manufacturing industries. Journal of Public Economics, 63(1), 57-81.

Mann, W. (2018). Creditor rights and innovation: Evidence from patent collateral. Journal of Financial Economics, 130(1), 25-47.

Nanda, R., Younge, K., & Fleming, L. (2013). Innovation and entrepreneurship in renewable energy. In A. Jaffe & B. Jones (Eds.), The changing frontier: Rethinking science and innovation policy. University of Chicago Press.

Nemet, G. F., & Kammen, D. M. (2007). US energy research and development: Declining investment, increasing need, and the feasibility of expansion. Energy Policy, 35(1), 746-755.

Nordhaus, W. D. (2013). The climate casino: Risk, uncertainty, and economics for a warming world. Yale University Press.

National Academies of Sciences, Engineering, and Medicine (NAS). (2016). SBIR/STTR at the Department of Energy. Washington, DC: The National Academies Press.

Oliver, M. (2012). Overview of the DOE's Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programs. DOE Webinar, November 30, 2012.

Popp, D., & Newell, R. (2012). Where does energy R&D come from? Examining crowding out from energy R&D. Energy Economics, 34(4), 980-991.

Rao, N. (2016). Do tax credits stimulate R&D spending? The effect of the R&D tax credit in its first decade. Journal of Public Economics, 140, 1-12.

Scherer, F. M. (1983). The propensity to patent. International Journal of Industrial Organization, 1(1), 107-128.

Serrano-Velarde, N. (2008). The financing structure of corporate R&D: Evidence from regression discontinuity design. Working paper, available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.

Takalo, T., Tanayama, T., & Toivanen, O. (2013). Estimating the benefits of targeted R&D subsidies. Review of Economics and Statistics, 95(1), 255-272.

U.S. Department of Commerce. (2012). Intellectual Property and the U.S. Economy: Industries in Focus. Retrieved from http://www.uspto.gov/news/publications/IP_Report_March_2012.pdf

Wallsten, S. J. (2000). The effects of government-industry R&D programs on private R&D: The case of the Small Business Innovation Research program. The RAND Journal of Economics, 82-100.

Wilson, D. J. (2009). Beggar thy neighbor? The in-state, out-of-state, and aggregate effects of R&D tax credits. Review of Economics and Statistics, 91(2), 431-436.

Wu, Y. (2008). State R&D tax credits and high-technology establishments. Economic Development Quarterly, 22(2), 136-148.

Zhao, B., & Ziedonis, R. H. (2012). State governments as financiers of technology startups: Implications for firm performance. Working paper, available at SSRN https://ssrn.com/abstract=2060739.