Patenting and business outcomes for cleantech startups funded by ARPA-E

Anna Goldstein,[a,b] Claudia Doblinger,[c] Erin Baker,[a] Laura Diaz Anadon[b,d]

[a] University of Massachusetts Amherst, Department of Mechanical and Industrial Engineering, Amherst MA, 01003, USA

[b] Harvard University, Harvard Kennedy School, Belfer Center for Science and International Affairs, Cambridge MA, 02138, USA

[c] Technical University of Munich, Campus Straubing for Biotechnology and Sustainability, Innovation and Technology Management, 94315 Straubing, Germany

[d] University of Cambridge, Centre for Environment, Energy and Natural Resource Governance, Department of Land Economy, Cambridge CB3 9EP, United Kingdom

Abstract

Innovation to reduce the cost of clean technologies has large environmental and societal benefits, and governments can play an important role in helping cleantech startups innovate and overcome risks involved in technology development. Here we examine the impact of the U.S. Advanced Research Projects Agency – Energy (ARPA-E) on two outcomes for startup companies: innovation (measured by patenting activity) and business success (measured by venture capital funding raised, survival, and acquisition or initial public offering). We compare 25 startups funded by ARPA-E in 2010 to rejected ARPA-E applicants, startups funded by a related government program, and other comparable cleantech startups. We find that ARPA-E awardees have a strong innovation advantage over all the comparison groups. However, while we find that ARPA-E awardees performed better than rejected applicants in terms of post-award business success, we do not detect significant differences compared to other cleantech startups. These findings suggest that ARPA-E was not able to fully address the "valley of death" for cleantech startups within 10-15 years after founding.

ARPA-E was established in the U.S. Department of Energy (DOE) in 2009, in recognition of the urgent need for accelerating energy technology innovation. The agency was created to fund breakthrough energy research with greater flexibility to overcome the stovepiping that afflicts traditional DOE funding offices.[1–3] For over two decades, experts have called for increased government spending on energy research and demonstration (R&D)[4–7] and for stronger mission-oriented innovation support policies, including government action to create and shape markets.[8,9] It is well known that market failures due to knowledge spillovers lead to underinvestment in research by the private sector.[10,11] In addition, experimentation is fundamental to innovation and entrepreneurship, which are characterized by high uncertainty and low success rates,[12,13] and thus, public intervention can be essential for encouraging radical innovation when the cost of experimentation is high.[14] Particularly for cleantech, where societal benefits are generally high compared to private returns because of the environmental externalities of air pollution and climate change, governments can be seen as a risk-tolerant investor with the ability to de-risk or shape markets for new technology.[9]

However, concerns over the effectiveness of government subsidies for applied energy R&D in private firms have long been part of the political and academic debate. Very recently, for example, a proposal by the U.S. Office of Management and Budget to eliminate funding for ARPA-E claimed that "the private sector is better positioned to finance disruptive energy research and development and to commercialize innovative technologies."[15] Nevertheless, a vast amount of past research from various European countries has generally indicated that public R&D grants do not substitute private R&D spending.[16–19] Public direct investments can rather mobilize private investments in specific sectors such as renewable energy.[20] Studies have found that U.S. government partnerships in the form of joint development and licensing[21] and federal research subsidies for small businesses[22] can increase patenting and private financing for U.S. cleantech firms. In a recent study of UK firms across all sectors, small firms significantly increase their internal R&D spending in response to direct grants and tax credits for R&D, whereas the contrary holds for larger firms.[23]

Even though the positive effect of direct public R&D grants to startups has been highlighted in recent studies, we still know little about how specific government R&D programs impact cleantech startups—especially programs that provide more than just the financial resources, such as R&D grants and tax credits, offered by traditional subsidy programs. ARPA-E was designed in the model of the U.S. defense agency DARPA, with expert program staff who are hired on short-term rotation and empowered to craft solicitations in an area of technical need, select proposals, and actively manage projects.[24,25] Based on DARPA's long track record of accomplishments, this funding model was expected to produce breakthrough innovations with the potential to transform the energy market. In the case of ARPA-E, which has become the often-discussed posterchild of mission-oriented innovation in the U.S. and abroad,[26,27] it is important to understand whether and how ARPA-E has supported innovation, for the sake of learning and informing future initiatives.[28,29]

We are particularly interested in innovation outcomes for startup companies, which can be more flexible and agile than larger firms or public organizations. Startups have the ability to quickly respond to market opportunities and can provide solutions to urgent problems as long as they have access to adequate resources.[30,31] Public organizations and incumbent firms also play a key role for energy innovation, given their scale and breadth of capabilities. However, previous research suggests that larger organizations respond to innovation opportunities later than entrepreneurial firms, due to inertia and legacy systems; and when they do move into new products or markets, they often harvest innovation from startups.[32–35] Exploring the drivers for startup innovation and success is therefore a valuable pursuit, especially given these firms' growing potential for transforming the energy sector.[36] Moreover, startups typically operate under a short-term imperative for achieving success as a result of resource constraints. Thus, although innovation in the energy sector can take many years (and sometimes even decades), there are important outcomes for startups in cleantech that can be observed in the relatively short-term.[21]

In this paper, we investigate these short-term (5–10 years) outcomes—specifically innovation, as measured by patenting activity, and business success, as measured by acquisition or initial public offering (IPO), survival, and venture capital (VC) raised—for the earliest cohort of ARPA-E startups, funded in 2010. Using a dataset combining information from Cleantech Group's i3 platform, DOE financial assistance records, as well as substantive information on startups from public information sources, we ask whether and to what extent the performance of the cleantech startups funded by ARPA-E differs from that of comparable firms, controlling for key characteristics, including cleantech sub-sector, age, prior patenting, and prior VC investment. We find that ARPA-E's startup awardees produced significantly more patents post-2010 than similar companies did. However, while ARPA-E awardees performed better than the group of rejected ARPA-E applicants in terms of their ability to attract VC investment post-award, likelihood of surviving, being acquired, or going public, they had no advantage over the average similar cleantech company in these dimensions. These findings indicate that barriers remain for innovative cleantech companies in gaining market traction.

ARPA-E in the context of U.S. cleantech innovation

Before ARPA-E was established, there was a rise of VC funding for cleantech from 2006 to 2008 in the U.S. VC investors accustomed to funding commercialization of medical and software technology turned their attention to cleantech, expecting to earn positive returns from a few high-value companies in a portfolio. Returns on these cleantech investments, however, turned out to be comparatively low, due to longer development cycles and high capital intensity of energy innovation, as well as competition from incumbent technologies producing the same commodities (e.g. electricity and liquid fuels).[37] Subsequent VC funding for cleantech dropped dramatically.

Shortly after the exodus of cleantech VCs, there was a surge in U.S. funding for public energy R&D through DOE. The American Reinvestment and Recovery Act (ARRA), which was passed to address the financial crises of 2009, also authorized roughly USD 60 billion of stimulus spending on clean energy investments in the 2010 fiscal year (Oct. 1, 2009 – Sep. 30, 2010).[38] A portion of ARRA funding established ARPA-E to fund high-risk R&D on energy technologies. Other ARRA funds were distributed through applied technology offices such as the Office of Energy Efficiency and Renewable Energy (EERE); Supplementary Figure 1 illustrates the funding surge for U.S. DOE R&D and demonstration in 2010.

We construct a dataset containing 1,287 U.S. cleantech companies founded from 2005 to 2010, i.e. firms that were startups (five years old or younger) in 2010.[21] These firms are assigned to one of four groups: (i) startups funded by ARPA-E in 2010 (listed in Supplementary Table 2); (ii) startups whose applications for funding were denied by ARPA-E in 2010—importantly, our sample only includes rejected applicants who were "encouraged" by ARPA-E, suggesting that they were judged to be high quality relative to the larger pool of rejected applicants; (iii) startups funded by EERE in 2010; and (iv) startups that were not identified as ARPA-E applicants and also did not receive EERE funding in 2010, although they may have received other types of federal assistance (labeled as "other cleantech"). Supplementary Table 1 provides descriptive statistics of each group, and Supplementary Data 1 contains a de-identified dataset of all 1,287 firms and their post-2010 patenting and business outcomes.

We also collect data on companies that received awards from the DOE Small Business Innovation Research (SBIR) program. Previous research on this program found it had a positive impact on the likelihood of VC investment and patenting for small businesses.[22] However, in our data, we find that SBIR startups as a whole are fairly dissimilar to ARPA-E awardees and other cleantech startups. Of the 82 startups funded by DOE SBIR in 2010, we are only able to identify 60% of them as cleantech companies, compared to 93% of EERE startups and 100% of ARPA-E startups from the same year. When we perform coarsened exact matching on the ARPA-E 2010 awardees (see Methods for more detail), only five SBIR companies remain in the matched sample. For this reason, we exclude SBIR awardees from our main analysis; more information on these companies is provided in Supplementary Note 1.

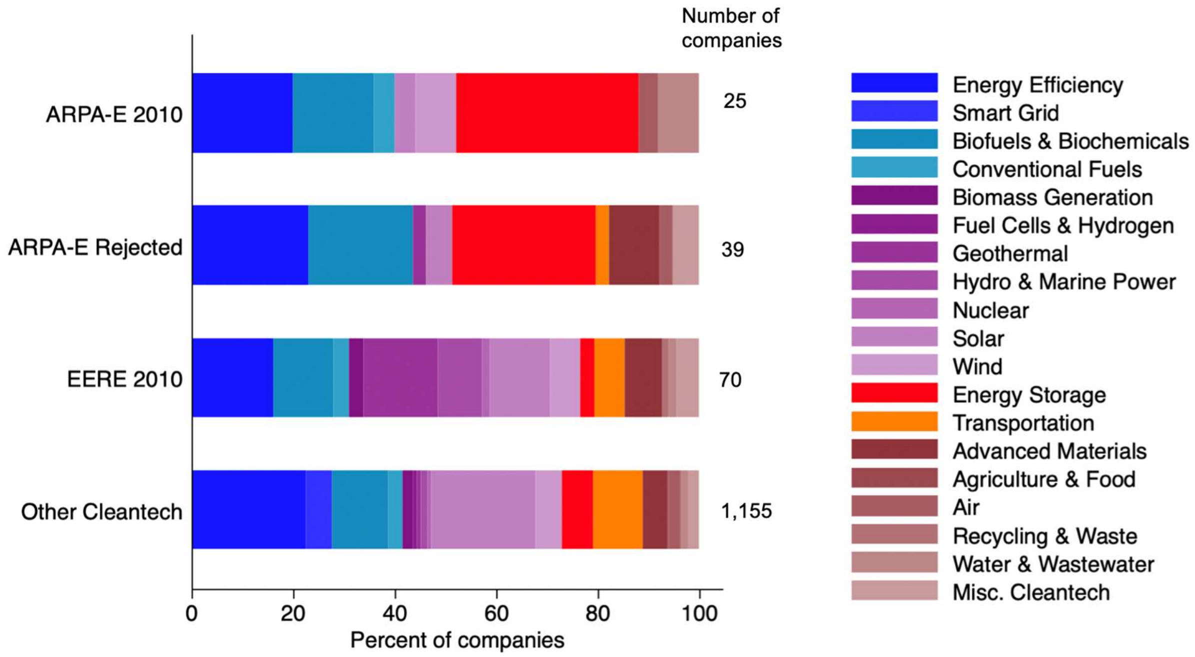

Figure 1. Sub-sectors represented in our dataset of U.S. cleantech startup firms in 2010. Each company was assigned to one of 19 cleantech sub-sectors used by Cleantech Group's i3 platform. See Methods for more detail on these categorizations.

Among the four groups of cleantech startups in our main dataset, we find differences in emphasis across various cleantech sub-sectors (Figure 1 and Supplementary Table 3). Both ARPA-E awardees and ARPA-E rejected applicants were concentrated in energy storage; ARPA-E funded nine storage startups in 2010—over one-third of its awards in that year, while such companies comprised only 6% of all cleantech startups. This emphasis arises largely from two targeted technical programs at ARPA-E for battery technology in 2010—one for electric vehicles and one for grid-scale storage.

One important reason for this emphasis is that technology priorities at ARPA-E are determined in a bottom-up fashion, as they aim to fund "white space", i.e. technical areas that are overlooked by other funding sources.[39] This is a complementary approach to that of other offices in DOE, like EERE, where funds are allocated based on road-mapping of existing technological trajectories—in other words, EERE funds are more likely to be subject to path dependencies in the programs when compared to ARPA-E. EERE's top-down approach is reflected in its 2010 startup portfolio, which was more similar to the overall makeup of U.S. cleantech companies. Solar photovoltaics, for example, comprised one-fifth of all cleantech startups in 2010; EERE funded eight solar startups, whereas ARPA-E funded only one.

We also find differences across groups in terms of the pre-2010 company profile: over half of ARPA-E awardees had already received private VC investment (56%), and over half had filed at least one patent (also 56%). Among ARPA-E rejected applicants, the proportion of companies with pre-2010 patenting and private funding were lower (38% and 18%, respectively). It appears that ARPA-E selections were preferential toward companies with pre-award patents or financing, either because these attributes were used directly by ARPA-E staff as a signal of quality, or because they happen to correlate with the actual selection criteria.

ARPA-E startups patented more than similar firms

Patents are an important part of the value chain for new ideas in energy technology, and protecting intellectual property allows startups to demonstrate their value to potential investors and/or companies that might acquire them.[40–42] If the ARPA-E funding mechanism is effective at supporting breakthrough R&D, then ARPA-E startups should show a higher rate of post-award patenting compared to rejected ARPA-E applicants as well as to similar cleantech startups.

We count the successful U.S. patent applications that were filed by each company in a given year, and we find that 80% of ARPA-E-funded startups had filed a patent after their award in 2010—a greater proportion than any other group. However, it is not clear from these descriptive statistics whether ARPA-E's advantage in patenting is simply a result of selecting companies with more prior patenting activity. To test this possibility, we model post-2010 patenting activity using Poisson regression with a control variable for pre-2010 company patenting, as well as age, sub-sector, and pre-2010 VC activity (Table 1).

We find that the patenting advantage for ARPA-E startups holds in regression analysis controlling for various factors, including pre-award patenting behavior. A coefficient of 0.7 corresponds to an incidence rate ratio of 2, meaning that startups who received ARPA-E awards in 2010 went on to file successful patent applications at twice the rate of similar cleantech companies on average (Models 2 and 3 in Table 1), even accounting for their pre-2010 patenting activity. We repeat the regressions using a sub-sample of companies that has been balanced on observable co-variates between ARPA-E and non-ARPA-E companies, using coarsened exact matching.[43] The matched results confirm our finding (Models 4 and 5 in Table 1); after their award in 2010, ARPA-E companies filed patents at roughly double the rate that would be expected based on their age, sub-sector and pre-2010 company profile. Confidence intervals for *ARPA-E 2010* in Model 5 (not shown) indicate an incidence rate ratio between 1.5 and 3.2.

Table 1. ARPA-E patenting outcomes

Dependent Variable		Post-2010 Patents	Post-2010 Patents	Post-2010 Patents	Post-2010 Patents	Post-2010 Patents
Form		Poisson	Poisson	Poisson	Poisson	Poisson
Model		(1)	(2)	(3)	(4)	(5)
Group:	ARPA-E 2010	1.393***	0.716***	0.712***	0.805***	0.793***
		(0.391)	(0.223)	(0.208)	(0.193)	(0.196)
	ARPA-E Rejected	0.392		0.343		-0.525
		(0.365)		(0.312)		(0.377)
	EERE 2010	0.419		0.081		-0.171
		(0.370)		(0.175)		(0.184)
Company Age in 2010		0.160***	-0.140**	-0.141**	-0.128	-0.170
		(0.053)	(0.060)	(0.060)	(0.108)	(0.117)
Ln(Pre-2010 Patents + 1)			0.843***	0.842***	1.047***	1.073***
			(0.101)	(0.101)	(0.090)	(0.090)
Ln(Pre-2010 VC + $1)			0.035***	0.035***	0.007	0.006
			(0.012)	(0.012)	(0.020)	(0.021)
Sub-Sector Fixed Effects		Y	Y	Y	Y	Y
Coarsened Exact Matching					Y	Y
Observations		1287	1287	1287	238	238
Pseudo R^2		0.128	0.401	0.402	0.560	0.563

Note: Direct coefficients reported. Robust standard errors in parentheses. Incidence rate ratios for binary variables can be obtained by exponentiating the coefficient.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Non-matched sample is the full dataset of U.S. cleantech startups in 2010. Pre-2010 counts are cumulative from company founding through 2009. Post-2010 patent counts are cumulative 2011 through 2014. Base group is *Other Cleantech.* Coarsened exact matching for ARPA-E vs. non-ARPA-E companies was performed using the following criteria: sub-sector, company age (3 cohorts), any pre-2010 patent, and any pre-2010 VC.

Because companies that apply for government grants differ from other companies in several ways,[44,45] we compare ARPA-E awardees in 2010 to companies that were rejected by ARPA-E in the same year. These rejected applicants were publicly "encouraged" by ARPA-E, which indicates that they were close to receiving the award. We take this approach to be approximating the idea behind regression discontinuity design methods, although the agency does not assign differentiated scores within the rejected group or the awardees. We find that rejected ARPA-E applicants were no more likely to patent than other cleantech startups. This indicates that the increased post-award patenting activity of ARPA-E companies is not a product of self-selection of more innovative companies into the applicant pool. We also test whether increased patenting is observed in the set of companies funded by EERE in 2010; these companies showed no advantage over other companies in terms of post-award patenting activity.

Panel regressions of annual patents filed illustrate the consistently elevated patenting activity by ARPA-E-funded startups from 2011 to 2014 (Figure 2a; regression coefficients are in Supplementary Table 4). In Supplementary Table 5, we test for the likelihood of any patenting from the four groups of startups, and we compare citation-weighted patent counts as well. The increased patenting activity for ARPA-E startups compared to rejected applicants, EERE awardees, and the wider pool of cleantech startups is robust to these alternative measurements. Although our main analyses exclude SBIR awardees, results from an expanded dataset including those companies is shown in Supplementary Table 6; SBIR awardees produced patents at roughly half the rate of other cleantech startups, on average.

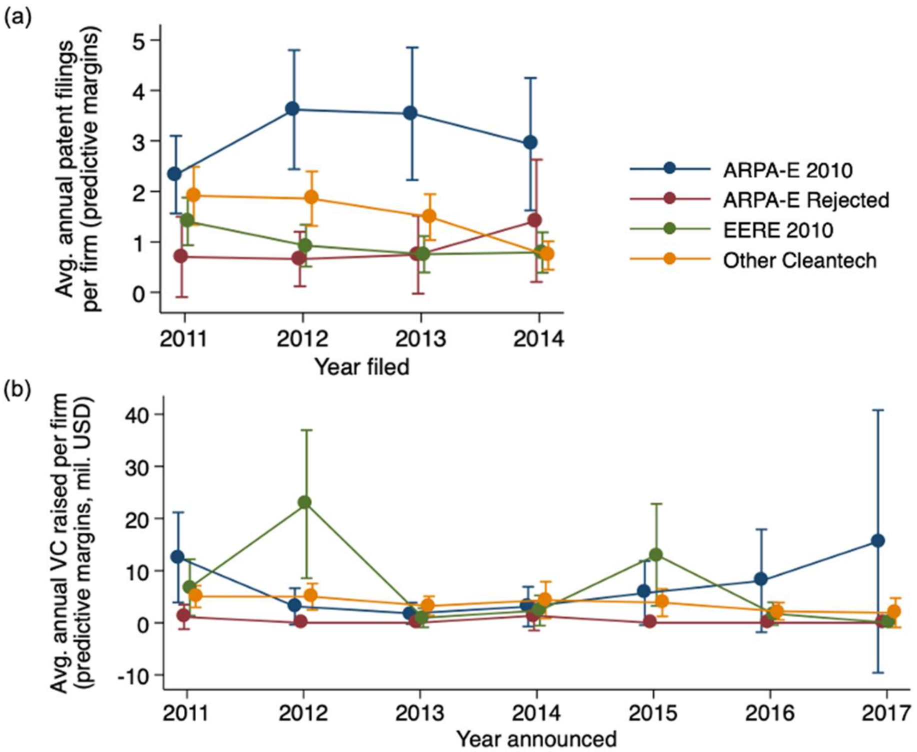

Figure 2: Annual outcomes for 2010 cleantech startups by company type. Plots of predictive margins for (a) patents filed and (b) VC funds raised are based on panel regressions controlling for sub-sector, company age, and pre-2010 levels of patenting and VC funding, within a matched sub-sample of cleantech companies; see Supplementary Table 4 for more detail. Outcome data are cut off in 2014 for patents and 2017 for funding due to a lag in observations. Predictive margins are the average predicted outcomes for companies in each group, using the observed values of all other variables. Error bars show 95% confidence intervals.

The increased patenting by ARPA-E companies could be explained by either a treatment effect or a selection effect. If participation in the ARPA-E program causes companies to produce more patents, this treatment effect could be ascribed to ARPA-E's unique program design. ARPA-E and EERE both provide financial resources in the form of direct grants, but when ARPA-E was established in 2010, it was set apart by offering commercialization assistance and active project management by staff with technical expertise; these interventions can be considered additional resources for the firm. Furthermore, program directors are empowered to modify or terminate award agreements based on project performance.[46] This leverage may allow them to induce greater patenting simply by encouraging awardees to patent; patents are a celebrated output for ARPA-E as a whole.[47]

However, we are not able to rule out the explanation that ARPA-E chooses to fund companies with a higher propensity to patent. We address some obvious potential selection effects by controlling for pre-award company profiles and comparing to rejected applicants, but it is still possible that ARPA-E staff select more innovative companies based on attributes other than prior patenting activity or VC funds. A previous study of ARPA-E project selection practices revealed that the main driver of funding decisions at ARPA-E was the individual judgment of the program director, rather than observable characteristics or even external peer review scores.[48] This use of individual discretion by ARPA-E staff may allow them to judge research proposals on future patenting potential based on additional information beyond previous patents and private funding.

ARPA-E startups were no more successful than similar firms

ARPA-E has highlighted the ability of its awardees to leverage their ARPA-E award into follow-on funding.[47] One of the agency's goals is to fill a funding gap by de-risking technologies enough to make them suitable for private investment and commercialization. If this mission were fully accomplished, then we would expect to see an increased likelihood of success (measured in different ways) for ARPA-E startups compared to similar companies, perhaps commensurate with the increased patenting activity measured above.

We measure three post-2010 business outcomes associated with increased likelihood of market success for our sample of cleantech startups: acquisition/IPO, survival through 2019, and amount of VC funding raised through 2017 (Table 2). As with our patenting analysis above, we run regressions for these outcomes, with controls for firm age and cleantech sub-sector to account for age- and sector-specific effects, as well as prior patenting and VC funds raised. We find no statistically significant differences between ARPA-E companies and non-ARPA-E companies for the three market success indicators, in either the whole sample or the matched sub-sample. A greater proportion of ARPA-E startups (52%) raised VC funds post-2010 compared to other cleantech startups (40%), but this difference is within the error of measurement when accounting for the firm's pre-award profile. This non-significant difference could be masking an actual effect, due to the small sample size and large standard error;

confidence intervals for *ARPA-E 2010* in Model 2 of Table 2 (not shown) yield an incidence rate ratio between 0.7 and 5.1. Thus, this is a tentative result and we hope that future research with a larger sample of companies will be able to either identify an effect or better establish the lack thereof, for VC funding as well as acquisition/IPO and survival.

Table 2. ARPA-E business outcomes

Dependent Variable		Post-2010 VC		Acquired or IPO		Survival	
Form		Poisson	Poisson	Logit	Logit	Logit	Logit
Model		(1)	(2)	(3)	(4)	(5)	(6)
Group:	ARPA-E 2010	0.053	0.643	-0.663	-0.330	0.250	0.308
		(0.620)	(0.506)	(0.528)	(0.632)	(0.501)	(0.645)
	ARPA-E Rejected	-1.407**	-2.231**	-1.716***	-1.420*	-1.010***	-1.165**
		(0.641)	(1.068)	(0.615)	(0.784)	(0.347)	(0.589)
	EERE 2010	-0.665	0.569**	-0.057	-0.029	-0.321	-0.176
		(0.432)	(0.229)	(0.296)	(0.729)	(0.296)	(1.274)
Company Age in 2010		-0.620*	0.093	0.186***	0.217	0.065	-0.241
		(0.374)	(0.161)	(0.044)	(0.187)	(0.048)	(0.213)
Ln(Pre-2010 Patents + 1)		0.536	0.291**	0.109	0.315	0.234**	0.820**
		(0.339)	(0.129)	(0.092)	(0.253)	(0.114)	(0.327)
Ln(Pre-2010 VC + $1)		0.169***	0.044	0.010	0.018	-0.023**	-0.057*
		(0.046)	(0.032)	(0.010)	(0.028)	(0.011)	(0.035)
Sub-Sector Fixed Effects		Y	Y	Y	Y	Y	Y
Coarsened Exact Matching			Y		Y		Y
Observations		1287	238	1282	238	1287	228
Pseudo R^2		0.456	0.370	0.047	0.109	0.033	0.164

Note: Direct coefficients reported. Robust standard errors in parentheses. Incidence rate ratios for binary variables can be obtained by exponentiating the coefficient.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Non-matched sample is the full dataset of cleantech startups in 2010. Pre-2010 counts are cumulative from company founding through 2009. Post-2010 VC funding totals are cumulative U.S. dollars 2011 through 2017. Coarsened exact matching for ARPA-E vs. non-ARPA-E companies was performed using the following criteria: sub-sector, company age (3 cohorts), any pre-2010 patent, and any pre-2010 VC.

We show two sets of alternative tests in Supplementary Table 7: logit regressions for the binary outcome of any VC funding, and Poisson regressions for all private finance deal totals (including common stock and private equity as well as VC). The results are consistent—ARPA-E-funded companies were not measurably more successful at raising private funds than similar startups from *Other Cleantech*, despite their increased patenting activity post-award. Panel regressions also show no clear difference between ARPA-E companies and others in annual VC funding 2011-2017 (Figure 2b; regression coefficients are in Supplementary Table 4).

We find a positive and significant rate of VC funding for *EERE 2010* compared to *Other Cleantech*, but a Wald test fails to reject the hypothesis that the coefficients for *ARPA-E 2010* and *EERE 2010* are equal (p = 0.90). On the other hand, we find significantly less business success for firms whose application for ARPA-E was denied in 2010. The coefficient on *ARPA-E Rejected* is significantly negative even when accounting for their lower rates of prior VC funding; the rate of VC funding for this group is roughly one-tenth the funding rate of *Other Cleantech* (Model 2 of Table 2). A Wald test rejects the hypothesis that the coefficients for *ARPA-E 2010* and *ARPA-E Rejected* are equal (p = 0.014), meaning that ARPA-E rejected applicants performed significantly worse than awardees.

The average rejected ARPA-E applicant was therefore less likely to succeed than other cleantech firms, while ARPA-E awardees performed on par with similar cleantech firms. As with our patenting results, we are not able to fully establish whether the difference between ARPA-E awardees and rejected applicants is due to treatment, selection, or both. However, also as with patenting, several factors would seem to contradict a strong selection effect. First, the rejected applicants in this study are relatively high quality, having been encouraged by ARPA-E. Furthermore, the comparison already accounts for prior patenting and VC funding, which eliminates some obvious mechanisms for selection. Finally, ARPA-E markets itself as a funder of high-risk ideas,[39] and previous research on ARPA-E selection practices indicates that they do in fact prefer to fund proposals with greater uncertainty.[48]

Discussion

Startups with an ARPA-E award were issued more patents in the years following their award compared to similar cleantech companies. Although we are cautious in over-interpreting the size of the effect, given the small sample of 25 ARPA-E companies, the trend toward greater patenting is significantly positive. We show that the higher rate of patenting is not shared by rejected ARPA-E applicants, indicating that the effect is not caused by a higher propensity to patent among firms that self-select into the ARPA-E applicant pool. Furthermore, the patenting advantage is not shared by similar startups with R&D awards from elsewhere in DOE, indicating that increased patenting is not caused simply by the infusion of public financial resources. However, two possible explanations remain: treatment (ARPA-E awards caused cleantech startups to produce more patents) and/or selection (ARPA-E provided a subsidy to more innovative startups who would have produced more patents anyway).

Without a randomized assignment of ARPA-E awards among a set of eligible firms,[49,50] we are not able to disentangle selection and treatment as possible explanations for the increased patenting post-2010 among ARPA-E awardees relative to other cleantech startups. Given our understanding that ARPA-E awardees are selected to be the best proposed ideas for solving a technical challenge, the treatment is far from random. Therefore, we face an unavoidable risk of bias from omitted variables, like project quality.

And yet, if ARPA-E program directors are skilled at predicting future patenting outcomes and choose to fund those companies with the greatest potential for future patenting, this could be an important result in itself given the environmental and societal importance of cleantech innovation. Predicting research outcomes and success of technical projects is notoriously difficult, even for industry experts or venture capitalists.[51] Although many ARPA-E startups had private funding prior to their award, those investors may not have been willing to support the particular R&D projects funded by ARPA-E—projects with transformative potential but greater risk. In a report from the U.S. Government Accountability Office, investment firms stated that they "generally do not fund projects that rely on unproven technologies."[52] One public company said that ARPA-E projects would not meet the criteria for internal investment because "the rate of return on investment required by its management was at least 20 percent per year."[52] ARPA-E funding allowed startups to pursue mission-aligned R&D projects without requiring short-term profitability, and we find that these startups were more inventive than others.

Despite the demonstrated advantage of ARPA-E awardees in producing inventions, we find no clear evidence that they perform differently from similar cleantech startups as a whole in terms of acquisition/IPO, survival, or VC funding post-award within 10-15 years of founding. We face the limitations of a small sample of ARPA-E companies, which partly explains the large standard error on these measurements, and a relatively short period of time following ARPA-E funding for cleantech innovations to pay off. Nonetheless, we do measure significantly better business outcomes for ARPA-E startups than our sample of high-quality rejected applicants. Although we cannot confirm the presence of a treatment effect, we note that such an effect could be consistent with the lack of business advantage for the kinds of firms that apply for ARPA-E support. If ARPA-E attracted applicants whose plans for technology development were riskier or more capital-intensive, then these firms may have been generally at a disadvantage in terms of business success. A positive treatment effect could have allowed ARPA-E awardees to overcome negative factors afflicting ARPA-E applicants and catch up with other firms, while not fully closing the gap between public and private funding sources for these innovative cleantech startups.

Hardware-based energy technology innovation requires large amounts of capital and long timescales to become fully commercialized. Private investors have shown some renewed interest in energy technologies, but their interest appears limited to digital innovations such as demand-management software.[53] More than half of the startups in our dataset are manufacturing firms, according to their North American Industry Classification System (NAICS) codes. It is not surprising if ARPA-E alone has not fully solved the "valley of death" problem for innovative cleantech companies, which has been shown to be especially acute in the demonstration phase.[54] Complementary innovation policies, such as increased funding for demonstration and commercialization, in-kind support from national laboratories, and targeted procurement programs, may be needed to allow scale-up beyond the R&D phase and to ensure that cleantech innovations can leverage additional private finance and transition to the market.

Ten years after ARPA-E first began issuing awards, our findings provide important contributions to the ongoing discussions in the U.S. about the role of governments as early investors in clean energy technologies. Our results will also inform other governments around the world that are interested in replicating or adapting the ARPA model as a mechanism for supporting mission-oriented innovation, as well as researchers interested in the design of public funding mechanisms for supporting R&D to advance energy-related and industrial policy goals.

Methods

Identification of cleantech companies

Cleantech is not a conventional market category; it cuts across many traditional industries and sectors, including agriculture, electricity, transportation, water and waste management. While "clean energy technology" does not fully encompass the range of technologies that can be considered cleantech, the two terms are sometimes used as interchangeably.[37] Cleantech companies are distinguished by offering reduced environmental impact compared to the status quo, although this reduced environmental impact of a cleantech company's product or service may or may not have market value. The label of "clean" has value to the extent that customers and investors recognize the environmental harm from existing technologies. For this reason, it can be difficult to identify a given company as cleantech, even knowing its product offerings, without knowing how the company signals its environmental value and impact.

In this work, cleantech firms were identified from three data sources: the Cleantech Group i3 dataset, recent DOE awardees, and applicants for ARPA-E funding in FY 2010. The commercial i3 dataset contains a list of cleantech companies, "focused on emerging and innovative companies." The Cleantech Group website states that they "typically track start-ups and SMEs from the seed stage to late-stage private equity – and everything in between."[55] This database was selected for its inclusion of detailed information on sub-sector and founding year and its focus on entrepreneurial firms.

To identify companies recently funded by DOE, we obtained transaction data on federal financial assistance from fiscal years 2009-2016 on USAspending.gov. Following the method of Goldstein and Narayanamurti,[3] we removed duplicate transactions and combined transactions to create an award-level dataset. This dataset includes grants, which entail relatively unfettered financial support, as well as cooperative agreements, which require substantial involvement of the funding agency; we refer to these collectively as "awards." We exclude contracts, which are used for government procurement. We limit our analysis to awards given to businesses by ARPA-E and EERE, with a start date in the range of FY 2010-2015. In our analysis, we pay special attention to awards that began in FY 2010 (Oct. 1, 2009 through Sep. 30, 2010), which was the year of the ARRA stimulus and the first year of operation for ARPA-E.

A final set of cleantech startups are identified as rejected ARPA-E applicants. Unfunded applicants are typically not publicly announced, but in 2011 and 2012, ARPA-E posted a list of "encouraged applicants" on its website. These applicants agreed to make a description of their proposed technology publicly available, in order to "help foster connections that will accelerate technology development." In total, 100 unfunded projects were listed prior to September 2011. Based on the dates when various ARPA-E awards were announced, we conclude that these applications were to one of ARPA-E's first seven technical programs, all of which were initiated in FY 2010: OPEN 2009, BEEST, Electrofuels, IMPACCT, ADEPT, BEETIT, and GRIDS. Based on internal data obtained from ARPA-E,[48] we find that the "encouraged applicants" sample is only 18% of the 565 rejected applications to the 2010 ARPA-E funding opportunities.

One startup in our dataset were funded by both ARPA-E and EERE in FY 2010, and one ARPA-E rejected applicant was funded by EERE in that same year. For the purpose of panel regressions and the bar plots in Figure 1, companies are assigned to only one group; these overlapping cases are assigned *ARPA-E 2010* and *ARPA-E Rejected* respectively.

Company characteristics

We combined several datasets to determine the founding year of each cleantech company identified by the methods above. The i3 dataset contains data on company founding year. USAspending.gov does not include information on the age of a business; DOE awardees and rejected ARPA-E applicants required manual research to determine founding year. When not available on the company's website, founding year was obtained from Orbis, Factset, public news sources, and/or state business registration documents. We limit our dataset to companies that were startups in 2010, i.e. with a founding year in the range 2005-2010.

We also used manual research to determine each company's sub-sector within cleantech. The i3 dataset includes a categorization for "primary sector" with 19 distinct labels; this label allows greater cohesion than the NAICS subsector, of which there are at least 50 unique categories represented within our dataset. The 19 cleantech sub-sectors include clean energy categories, which comprise the majority of the companies in our dataset (88%), as well as several categories that have broader application beyond clean energy: Advanced Materials, Agriculture & Food, Air, Recycling & Waste, Water & Wastewater, and Miscellaneous Cleantech.

We checked the categorization for each company, based on the company website and other publicly available information online. For companies that were not in the i3 dataset, the cleantech sub-sector was determined by a review of the company's website as of 2017. If the website was not active as of 2017, we used the most recent archive on The Wayback Machine (https://web.archive.org/). We focused on the company's self-description and degree of emphasis on various cleantech applications for their technology; our determination does not account for possible changes in technology area over time.

Some companies were found to operate primarily outside the 19 cleantech sub-sector classifications. We exclude these non-cleantech companies from our dataset.

Outcome variables – Patents

We searched the U.S. Patent and Trademark Office website for patents assigned to each cleantech startup, issued on or before April 20, 2017. Design patents were excluded from the dataset. Observations were manually reviewed if the name of the assignee was not an exact match for the name of the company, or if the year of the patent filing was prior to the recorded founding year of the company. Patents are counted by the year in which the application was filed. Observations were collected through Dec. 31, 2014, to account for the lag in observing a successful patent application. A citation-weighted count was also generated, following Trajtenberg's method of adding the annual patent count to a sum of all forward citations to those patents.[56] Citations to patents were observed through April 1, 2017.

Outcome variables – Private investments

We collected data on private funding from three sources: Thomson ONE, Crunchbase, and Preqin. To obtain the total funding received by a company at the end of a given calendar year, we use the maximum amount recorded from any of the three sources. We limited our observations to funding deals announced between Jan. 1, 2004 and Dec. 31, 2017. For VC funding, we include the following deal types: Venture Capital Equity Investment (Thomson ONE); Seed, Series A-I, Venture – Series Unknown (Crunchbase); Seed, Series A-I (Preqin). For private funding variables, we include all VC funding deals, plus the following deal types as well: Common Stock, Convertible Preferred Stock, Preferred Shares, and Warrants (Thomson ONE); Angel, Convertible Note, Corporate Round, Equity Crowdfunding, Funding Round, Post-IPO Equity, Pre-Seed, and Private Equity (Crunchbase); Add-on, Angel, Growth Capital/Expansion, Private Investment in Public Equity (PIPE), and Unspecified Round (Preqin).

Because funding decisions may be influenced by the signaling effect of government grants, we aim to observe fundraising activities for DOE awardees before their grants were announced publicly. We consider funding announced on or before Dec. 31, 2009 to be pre-award funding. This method introduces two countervailing sources of error. On one hand, it overestimates pre-award funding for companies whose awards were announced in 2009, by including investments that may have been influenced by that announcement; for example, ARPA-E grants in the OPEN 2009 program were announced on Oct. 26, 2009. On the other hand, our method underestimates pre-award funding for companies whose awards were announced in 2010, by excluding investments that were made before the announcement. For example, some ARPA-E programs announced awards in spring and summer of 2010 (April 29 for BEEST, Electrofuels, and IMPACCT; July 12 for BEETIT, GRIDS, and ADEPT.)

For consistency with the funding data, we limit pre-award patent observations to those filed on or before Dec. 31, 2009. For both patents and private funding, we ignore observations (patents filed or funding announced) in calendar year 2010. We test the main results for robustness with two alternative specifications in Supplementary Table 8: pre-award outcomes including 2010, and post-award outcomes including 2010.

Outcome variables – Exit and survival

We further collected data from public sources on two events for each company in the dataset: whether a company had been acquired or issued an IPO (indicating an "exit" for investors), and whether they remained in business as of December 2019. Data on acquisition and IPO were collected from i3, Crunchbase, Thomson ONE, Orbis, and complemented with a web search. The outcome was operationalized as a binary variable equaling 1 for either an acquisition or an IPO, and 0 otherwise.

Survival is an important indicator of performance in the startup context where many companies end in failure (bankruptcy, closures, or liquidations).[57,58] We compiled data on whether or not a company was still in business as of December 2019 in three steps. First, we searched for each company on the web using keywords for bankruptcy, closure, and liquidation. We also checked whether the firms still had webpages and when they were last updated, using The Wayback Machine and Google cached results. Finally, for companies with outdated or nonexistent online records, we searched their status in state-level business registration databases. Acquired companies may be absorbed into their new parent company and no longer observable as an independent entity, so we operationalized firm survival to also account for acquisition/IPO outcomes; survival is a binary variable that takes on the value 1 if a firm was still active or had undergone acquisition/IPO, and 0 otherwise.

Regressions

To estimate differences between ARPA-E-funded firms and the other groups of cleantech startups, we conduct a set of regression analyses. We estimate post-2010 patent counts and funding amounts (Y_i) for firm i using the following Poisson model:

[1] $$ln(Y_{ij}) = \beta_0 + \beta_1 ARPA - E_i + \beta_2 ARPA - E\ Rej._i + \beta_3 EERE_i + \beta_4 Age + \beta_5 \ln(P_i + 1) + \beta_6 \ln(Z_i + 1) + \varphi_j$$

where $\beta_{1\text{-}3}$ are the coefficients of interest for three groups of cleantech startups, compared to *Other Cleantech*. P_i is the number of patents filed pre-2010, Z_i is the amount of pre-2010 VC funding, and φ_j is a fixed effect for the cleantech sub-sector of the firm (j). Sub-sector fixed effects are not reported for brevity, but outcomes differ dramatically across sub-sectors; descriptive statistics are shown in Supplementary Table 9.

All regression tables report the coefficients directly. For Poisson regressions, coefficients can be converted to incidence rate ratios between firms in different groups—for example, ARPA-E and non-ARPA-E firms—by calculating exp(β). We estimate likelihood of acquisition/IPO and firm survival using a logit model, with the same set of explanatory and control variables in Equation 1. Here, the exponentiated coefficient for binary independent variables (e.g. *ARPA-E 2010*) indicates the odds ratio of an event for startups in a given group compared to *Other Cleantech*.

Coarsened exact matching is performed to reduce imbalance between ARPA-E and non-ARPA-E cleantech startups. The *cem* package in Stata is used to determine importance weights.[59] The criteria for company matching are: 19 individual cleantech sub-sectors, 3 age cohorts (founded 2005-2006, 2007-2008, and 2009-2010), whether they had any pre-2010 patents, and whether they had any pre-2010 VC funding. The matched dataset contains 238 companies (including 23 of 25 ARPA-E companies) across 17 matched strata. The multivariate L_1 distance is reduced to 1.25e-16 from 0.731 pre-matching.[43]

Alternative regression models are used to check for robustness. We convert our cross-sectional data on startups to a panel dataset, and we repeat the Poisson regressions above with company-years as the unit of analysis, for years starting in 2011 (Supplementary Table 4). We estimate logit regressions for whether or not a firm patented or received any VC funding (Supplementary Tables 5 and 7). We conduct Tobit analysis to account for left-censored data on the latent variable of propensity to patent or raise VC funds (Supplementary Table 10), with a lower limit of 0 patents filed or $0 raised. The coefficient in this case is an estimate of linear increase in the uncensored latent variable. We also estimate difference-in-differences to account for pre-2010 and post-2010 time trends (Supplementary Table 11), where the coefficient is an estimate of linear increase in the outcome variable.

Data Availability

The data on cleantech firm identification used in this study were made available to us by Cleantech Group under a restricted license and are therefore not publicly available. Supplementary Data 1 contains an anonymized version of our company-level analytical sample with the names of the firms removed. Supplementary Data 2 contains the company-year panel dataset. Supplementary Code 1 contains the Stata code that enables the reproduction of our main analysis. The full dataset is available from the authors upon reasonable request and with permission of Cleantech Group.

References

1. 110th Congress. *America COMPETES Act*. (2007).
2. Anadón, L. D., Chan, G. A., Bin-nun, A. Y. & Narayanamurti, V. The pressing energy innovation challenge of the US National Laboratories. *Nat. Energy* 16117 (2016). doi:10.1038/nenergy.2016.117
3. Goldstein, A. P. & Narayanamurti, V. Simultaneous Pursuit of Discovery and Invention in the US Department of Energy. *Res. Policy* (2018). doi:10.1016/j.respol.2018.05.005
4. Anadón, L. D., Bunn, M. & Narayanamurti, V. *Transforming U.S. Energy Innovation*. (Cambridge University Press, 2014).
5. President’s Council of Advisors on Science and Technology (PCAST). *Accelerating the Pace of Change in Energy Technologies Through an Integrated Federal Energy Policy*. (2010).
6. American Energy Innovation Council. *A Business Plan for America’s Energy Future*. (2010).
7. Jaffe, A. B., Newell, R. G. & Stavins, R. N. A tale of two market failures: Technology and environmental policy. *Ecol. Econ.* **54**, 164–174 (2005).
8. Kattel, R. & Mazzucato, M. Mission-oriented innovation policy and dynamic capabilities in the public sector. *Ind. Corp. Chang.* **27**, 787–801 (2018).
9. Mazzucato, M. *The Entrepreneurial State: Debunking Public vs. Private Sector Myths*.

(2013).

10. Arrow, K. Economic welfare and the allocation of resources for invention. in *The Rate and Direction of Inventive Activity: Economic and Social Factors* **I**, 609–626 (1962).
11. Nelson, R. R. The Simple Economics of Basic Scientific Research. *J. Polit. Econ.* **67**, 297–306 (1959).
12. Kerr, W. R., Nanda, R. & Rhodes-Kropf, M. Entrepreneurship as Experimentation. *J. Econ. Perspect.* **28**, 25–48 (2014).
13. Hall, R. E. & Woodward, S. E. The Burden of the Nondiversifiable Risk of Entrepreneurship. *Am. Econ. Rev.* **100**, 1163–1194 (2010).
14. Nanda, R. & Rhodes-Kropf, M. Financing Risk and Innovation. *Manage. Sci.* **63**, 901–918 (2017).
15. Office of Management and Budget. *America First: A Budget Blueprint to Make America Great Again*. (2017).
16. Bronzini, R. & Piselli, P. The impact of R&D subsidies on firm innovation. *Res. Policy* **45**, 442–457 (2016).
17. Aerts, K. & Schmidt, T. S. Two for the price of one? Additionality effects of R&D subsidies: A comparison between Flanders and Germany. *Res. Policy* **37**, 806–822 (2008).
18. González, X. & Pazó, C. Do public subsidies stimulate private R&D spending? *Res. Policy* **37**, 371–389 (2008).
19. Hussinger, K. R&D and Subsidies at the Firm Level: An Application of Parametric and Semiparametric Two-Step Selection Models. *J. Appl. Econom.* **23**, 729–747 (2008).
20. Deleidi, M., Mazzucato, M. & Semieniuk, G. Neither crowding in nor out : Public direct investment mobilising private investment into renewable electricity projects. *Energy Policy* 111195 (2019). doi:10.1016/j.enpol.2019.111195
21. Doblinger, C., Surana, K. & Anadón, L. D. Governments as partners: The role of alliances in U.S. cleantech startup innovation. *Res. Policy* **48**, 1458–1475 (2019).
22. Howell, S. T. Financing Innovation: Evidence from R&D Grants. *Am. Econ. Rev.* **107**, 1136–1164 (2017).
23. Pless, J. Are “Complementary Policies” Substitutes? Evidence from R&D Subsidies in the UK. (2018). Available at: http://jacquelynpless.com/research/. (Accessed: 12th December 2018)

24. Azoulay, P., Fuchs, E. R. H., Goldstein, A. P. & Kearney, M. Funding Breakthrough Research: Promises and Challenges of the 'ARPA Model'. *Innov. Policy Econ.* **19**, 69–96 (2019).
25. Fuchs, E. R. H. Rethinking the role of the state in technology development: DARPA and the case for embedded network governance. *Res. Policy* **39**, 1133–1147 (2010).
26. Stokstad, E. UK cues up big funding increases for R&D. *Science* (2020). doi:10.1126/science.abb6928
27. Bonvillian, W. B. DARPA and its ARPA-E and IARPA clones : a unique innovation organization model. *Ind. Corp. Chang.* **27**, 897–914 (2018).
28. Chan, G. A., Goldstein, A. P., Bin-Nun, A., Anadón, L. D. & Narayanamurti, V. Six Principles for Energy Innovation. *Nature* **552**, 25–27 (2017).
29. Pless, J., Hepburn, C. & Farrell, N. Bringing rigour to energy innovation policy evaluation. *Nat. Energy* (2020). doi:10.1038/s41560-020-0557-1
30. Audretsch, D. B., Keilbach, M. C. & Lehmann, E. E. *Entrepreneurship and economic growth*. (Oxford University Press, 2006).
31. Bettis, R. A. & Hitt, M. A. The New Competitive Landscape. *Strateg. Manag. J.* **16**, 7–19 (1995).
32. Hockerts, K. & Wüstenhagen, R. Greening Goliaths versus emerging Davids — Theorizing about the role of incumbents and new entrants in sustainable entrepreneurship. *J. Bus. Ventur.* **25**, 481–492 (2010).
33. Dushnitsky, G. & Lenox, M. J. When do incumbents learn from entrepreneurial ventures?: Corporate venture capital and investing firm innovation rates. *Res. Policy* **34**, 615–639 (2005).
34. Markard, J. & Truffer, B. Innovation processes in large technical systems: Market liberalization as a driver for radical change? *Res. Policy* **35**, 609–625 (2006).
35. Doblinger, C. & Soppe, B. Change-actors in the U.S. electric energy system: The role of environmental groups in utility adoption and diffusion of wind power. *Energy Policy* **61**, 274–284 (2013).
36. Popp, D., Pless, J., Hascic, I. & Johnstone, N. Innovation and Entrepreneurship in the Energy Sector. in *The Role of Innovation and Entrepreneurship in Economic Growth* forthcoming (2020). doi:10.4324/9780203117170-1

37. Gaddy, B. E., Sivaram, V., Jones, T. B. & Wayman, L. Venture Capital and Cleantech: The wrong model for energy innovation. *Energy Policy* **102**, 385–395 (2017).
38. Council of Economic Advisors. *A Retrospective Assessment of Clean Energy Investments in the Recovery Act*. (2016).
39. National Academies of Sciences Engineering and Medicine. *An Assessment of ARPA-E*. (2017). doi:10.17226/24778
40. Islam, M., Fremeth, A. & Marcus, A. Signaling by early stage startups: US government research grants and venture capital funding. *J. Bus. Ventur.* **33**, 35–51 (2018).
41. Conti, A., Thursby, J. & Thursby, M. Patents as Signals for Startup Financing. *J. Ind. Econ.* **61**, 592–622 (2013).
42. Guzman, J. & Stern, S. Where Is Silicon Valley? *Science* **347**, 606–609 (2015).
43. Iacus, S. M., King, G. & Porro, G. Multivariate Matching Methods That Are Monotonic Imbalance Bounding. *J. Am. Stat. Assoc.* **106**, 345–361 (2011).
44. Feldman, M. P. & Kelley, M. R. The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Res. Policy* **35**, 1509–1521 (2006).
45. Conti, A. Entrepreneurial finance and the effects of restrictions on government R & D subsidies. *Organ. Sci.* **29**, 134–153 (2018).
46. Goldstein, A. P. & Kearney, M. Know When to Fold 'Em: An Empirical Description of Risk Management in Public Research Funding. *Res. Policy* **49**, 103873 (2020).
47. ARPA-E. *ARPA-E Impacts: A Sampling of Project Outcomes, Volume III*. (2018). doi:10.2172/1464210
48. Goldstein, A. P. & Kearney, M. Uncertainty and Individual Discretion in Allocating Research Funds. (2017). doi:10.2139/ssrn.3012169
49. Azoulay, P. Research efficiency: Turn the scientific method on ourselves. *Nature* **484**, 31–32 (2012).
50. Jaffe, A. B. Building Programme Evaluation into the Design of Public Research-Support Programmes. *Oxford Rev. Econ. Policy* **18**, 22–34 (2002).
51. Scott, E., Shu, P. & Lubynsky, R. Entrepreneurial Uncertainty and Expert Evaluation: An Empirical Analysis. *Manage. Sci.* **66**, (2018).
52. Government Accountability Office. Advanced Research Projects Agency- Energy Could

Benefit from Information on Applicants' Prior Funding. (2012). Available at: https://www.gao.gov/products/GAO-12-112. (Accessed: 27th December 2017)

53. Bumpus, A. & Comello, S. Emerging clean energy technology investment trends. *Nat. Clim. Chang.* **7**, 382–385 (2017).
54. Nemet, G. F., Zipperer, V. & Kraus, M. The valley of death, the technology pork barrel, and public support for large demonstration projects. *Energy Policy* **119**, 154–167 (2018).
55. Cleantech Group. Our coverage. (2017). Available at: https://www.telstra.com.au/coverage-networks/our-coverage. (Accessed: 13th July 2017)
56. Trajtenberg, M. Patents, Citations and Innovations: Tracing the Links. (1987). doi:10.3386/w2457
57. Clarysse, B., Wright, M., Bruneel, J. & Mahajan, A. Creating value in ecosystems: Crossing the chasm between knowledge and business ecosystems. *Res. Policy* **43**, 1164–1176 (2014).
58. Hannan, M. T. & Freeman, J. *Organizational ecology*. (Harvard University Press, 1989).
59. Blackwell, M., Iacus, S., King, G. & Porro, G. Cem: Coarsened exact matching in Stata. *Stata J.* **9**, 524–546 (2009).

Correspondence should be addressed to A.P.G. at goldstein@umass.edu.

Acknowledgements

The authors are grateful to Cleantech Group for providing access to their i3 database. We also acknowledge excellent research assistance by Alen Amini, Benedikt Dirscherl, Pepita Gigglinger, Anna Hammerstingl, Susanne Kurowski, Claudio Lex and Nico Schlosser. We acknowledge funding from the EU Framework Programme for Research and Innovation H2020 under the Grant Agreement No. 730403 (INNOPATHS). We also acknowledge a grant from John and Elizabeth Armstrong.

Author Contributions

A.P.G., C.D., and L.D.A. designed the study. A.P.G. and C.D. collected data. A.P.G. analysed data and ran statistical tests. A.P.G. and C.D. wrote the paper. L.D.A and E.B. guided the study and edited the paper.

Competing Interests

The authors declare no competing interests.