Strategy Science

Enabling Technologies and the Role of Private Firms: A Machine Learning Matching Analysis

Jason M. Rathje, Riitta Katila

Please scroll down for article—it is on subsequent pages

Enabling Technologies and the Role of Private Firms: A Machine Learning Matching Analysis

Jason M. Rathje,[a,b] Riitta Katila[b]

[a] AF Ventures, Alexandria, Virginia 22310; [b] Department of Management Science and Engineering, Stanford University, Stanford, California 94305

Contact: jason.rathje@gmail.com (JMR); rkatila@stanford.edu (RK); https://orcid.org/0000-0001-8125-515X (RK)

Abstract. Investments in enabling technologies—including the fifth-generation technology standard for broadband cellular networks (5G), artificial intelligence (AI), and light detection and ranging (LIDAR) technology—are important strategic decisions for firms. This paper asks how inventions that private firms developed with (versus without) public-sector partners differ in their enabling technology trajectory. Using a novel method of machine learning matching, we compare patented technologies generated from more than 30,000 public–private relationships with comparable technologies invented by private firms alone during a 21-year period. To measure the enabling potential of a technology, we introduce a new enabling technology index. The findings show that private-firm relationships with the public sector—in particular cooperative agreements and grants with mission agencies (National Aeronautics and Space Administration and Department of Defense)—are likely starting points for enabling technology trajectories. We thus put a spotlight on organizational arrangements that combine the breadth of exploration (agreements, grants) with deep exploitation in a particular domain (mission agency). A key contribution is a better understanding of the types of private-firm efforts that are associated with enabling technologies. We also challenge the common assumption that enabling technologies have their origins only in public-sector projects and show how private firms are involved. Our significant contribution is to show how private firms can change evolution of ecosystems through technology development.

Introduction

Enabling technologies are strategically important. By definition, these technologies are novel, they enable complementary innovations often downstream in the value chain, and they have become widely used across industries and industry sectors (Teece 2018a, Gambardella et al. 2019). The list of much-discussed general-purpose technologies (GPTs) is short and includes only a select few that have game-changing impact across the economy as a whole (e.g., electricity, the internet; Bresnahan and Trajtenberg 1995, Lipsey et al. 2005). In contrast, the list of enabling technologies, that is, *junior GPTs*, includes many technologies that are disruptive and growth enabling in particular industries but not necessarily with measurable economy-wide impact (e.g., light detection and ranging (LIDAR) technology, the fifth-generation technology standard for broadband cellular networks (5G) (Teece 2018a)).[1] Recent case- and technology-specific evidence suggests that private (for-profit) firms are increasingly involved in the use of enabling technologies (Teece 2018b, Webb et al. 2018), but little research has examined the enabling technology trajectory of inventions developed by private firms (Rathje, 2019). Consequently, we ask in this study how enabling are the technologies developed by private firms.

To address this research question, we evaluate the technology outcomes of more than 30,000 public–private relationships and comparable private-firm-only efforts over a 21-year period from 1982 to 2002. We implement a hybrid machine learning–propensity score matching approach to identify treatment–control pairs and to enhance causal inference (Rathje and Katila 2019). In the data, private firms' relationships with the public sector—which are traced using the government interest statement in patent documents—differ primarily by public-sector partner (i.e., who the firm partners with) and by relationship type (i.e., how the relationship is structured). We investigate variation in mission agencies versus science agencies as public-sector partners[2] and include grants, contracts, and cooperative agreements as relationship types.[3] Altogether our data enable a comprehensive analysis of how private firms spawn enabling technologies.

There are several contributions. First, we unpack how enabling the inventions created by private firms—with public partners or not—become in the long run. We find that public–private relationships that combine the breadth of exploration (agreements, grants) with deep exploitation in a particular domain (mission agency) are likely to be conducive of enabling technologies. Interestingly, contracts—that agency theory designates as the most effective (Bruce et al. 2019) — have much weaker effects. For strategy, these results show that private firms can indeed create enabling technologies together with the government. Intriguingly, even when the government partner is a large end user (e.g., mission agency) and potentially narrowly focusing on a particular commercialization (Christensen et al. 2018), specific relationship types such as grants that offer greater latitude in the relationship can be a balancing force and serve as a starting point for an enabling technology trajectory.

Second, there are contributions to public policy. We add to the evidence that targeted government relations with private firms can support innovation (Bloom et al. 2019). We particularly find evidence of the positive role of public–private relationships in generating enabling technologies. If the goal of policy is to support innovation that generates knowledge spillovers and enables growth within and across sectors, we identify that agreements and grants with mission agencies in particular are aligned with this goal during the time period that we study.

Third, there are contributions to how ecosystems evolve (Adner and Kapoor 2016). We find that firms and their government collaborators can seed ecosystems with technologies that enable subsequent innovation by ecosystem participants and therefore shape how ecosystems evolve. More broadly, the results point in the direction that private firms, together with public partners, can be important in the evolution of ecosystems.

Finally, there are methodological contributions. We introduce a continuous enabling technology index (ETI) and its components. Prior work has often examined specific enabling technologies, such as artificial intelligence (AI), but has yet to provide a continuous operationalization of enabling technology. Because every technology does, to some degree, enable future innovation, we introduce a continuous, tractable index that enables us to quantify enabling as a fundamental attribute of a technology. Given that many enabling technologies are rapidly generating new opportunities and strategic challenges for firms and their decision makers, this index helps, for example, accurately track the potential of a firm's own efforts and those of the key organizations in the firm's ecosystem, including suppliers, competitors, collaborators, and complementors and their enabling technology potential.

Another methodological strength of our paper is a novel machine learning–propensity score matching (ML-PSM) method that enables enhanced causal inference for strategic management research (Stuart and Rubin 2008, Rathje and Katila 2019). Specifically, we build on the advantages of supervised machine learning to intelligently expand the set of observable covariates in the matching process. The result is more precise treatment–control pairs and a stronger push toward causality.

Private Firms, Public-Sector Relations, and Enabling Technologies

Public-sector governments have long worked in various capacities with private firms to enable research and development (R&D) that benefits society at large (Bloom et al. 2019). The public sector's support of private-firm R&D is based on two fundamental assumptions: (1) that the technical capabilities of private firms are essential to knowledge production and economic growth and (2) that the relatively high risk of funding high-impact, widely usable innovation that is often of the most benefit to society causes private firms to underinvest in them. Thus, policymakers hope to lower the costs of R&D for private firms in order to generate high-impact innovation—such as enabling technologies—that benefits not only the firm but also other firms and society at large.

Prior Work on Public-Sector Support of Private-Firm Research and Development

The public sector typically uses three main types of arrangements to support high-impact R&D in private firms: university funding, tax credits, and R&D relationships. In many nations, public funding agencies fund universities and national laboratories to conduct scientific discovery. The results are fully disclosed. The hope is that this scientific knowledge will spill over to private firms and inspire them to further invest in developing high-impact (rather than more incremental) innovation (Jaffe and Lerner 2001, Link et al. 2011). Although not directly in terms of enabling technologies, Azoulay et al. (2019b) illustrate this mechanism. They find that a $10 million increase in National Institutes of Health (NIH) funding to academics in a research area lead to 2.7 additional patents filed by private firms. Altogether the core premise is that university and national laboratory funding could set off technology projects (including enabling ones) in private firms, but this mechanism is indirect and does not provide a direct subsidy for private firms or for specific projects.

Second, R&D tax credits[4] enable firms to decrease the cost of R&D and fund more projects, but again, these public subsidies typically cannot be tied to specific projects. The risk, then, is that many credits are used for incremental R&D. For instance, Pless (2019)

shows that tax credits increase R&D effort, particularly by small UK firms, but it is unclear whether the efforts will have wide-ranging impact. Balsmeier et al. (2018) similarly show that although California firms' patents after the 1987 R&D tax credit are more highly cited by other firms—thus indicating greater public knowledge spillovers—tax credits also help firms engage in actions that only benefit the focal firm and not the community at large, such as strategic patenting. Thus, similar to university funding, the impact of R&D tax credits is likely to be indirect at best.

Third, and most significantly for our paper, the public sector engages in public–private R&D relationships.[5] In these public–private R&D relations, a public funding agency and a private firm jointly solve a complex, potentially high-impact technical problem (David et al. 2000, Zuniga-Vicente et al. 2014), and the partner private firm often retains exclusive rights to the results (Pahnke et al. 2015, Howell 2017, Bruce et al. 2019).

Legislation typically defines the types of R&D relationships that the public sector has with private actors. In the United States, the Federal Grant and Co-operative Agreement Act of 1977 outlines three main types: grants, contracts, and cooperative agreements. Grants are direct financial subsidies given to a firm by a public funding agency as an investment toward an objective specified by the agency (U.S. Congress 1977). Grants are often used to advance a national objective, address a public problem, or stimulate an activity desired by the government. Grants allow considerable latitude, and the grant recipient (i.e., the firm) often defines the scope of work because there are no legally binding requirements to achieve results (Rathje 2019).[6]

By contrast, the second type of relationship, that is, contract, is a leader–follower relationship. A contract is a binding agreement between a buyer and a seller to provide goods or services in return for compensation (Rathje 2019). In government contracts with private firms, the contracting agency is looking to procure a good or a service that will be of direct benefit to the government. Contracts differ from grants because there is often a built-in customer, and payments are often based on deliverables and milestones. Although contracts can take many forms, the most common form is a cost-plus contract.[7] Public agencies use cost-plus contracts, as opposed to other contract forms, when they want to purchase an immature (i.e., potentially high-impact) end product. To derisk the firm's investment in a new technology with relatively uncertain development costs, the public agency agrees to pay both the entire cost of development and a standardized profit margin. The profit margin may vary significantly depending on the risk of the project, but it generally stays small. For example, in the United States, the maximum profit margin for a cost-plus contract is 15% (Arnold et al. 2008). Because they

are commonly used for more immature technologies that involve patents, cost-plus contracts serve as the basis for our understanding of contractual relations between private firms and the public sector.

Third, cooperative agreements (agreements for short) are public–private relationships in which firms and public agencies agree to work side by side toward a mutual objective (U.S. Congress 1977). They differ from grants and contracts by the degree to which the public and private entities are expected to cooperate after the award. In agreements, federal employees often are substantially involved in the execution of the work and participate closely in performing the work side by side with the private firm (Rathje 2019). In contrast, federal agencies usually take on purely monitoring and oversight roles once, for example, a grant is awarded. Another difference is that agreements cannot be used to acquire goods or services for the federal government. They differ from contracts in that regard and therefore often allow greater latitude in project scope. In cooperative agreements, the partnering firms and public agencies are free to determine their interaction pathways and schedules (Ham and Mowery 1998, Bruce et al. 2019).[8] In other words, interaction between public and private researchers is required, but the format for interaction, unlike in contracts, is not specified ex ante (Mowery 2009, Lerner 2012). Agreements can be attractive for firms because they allow collaborative, peer-to-peer working relationships between private firms and, for example, a wide variety of highly capable national laboratory technical talent at low or no cost to the firm.

In general, research is inconclusive on the effects of public–private R&D relationships. Some research has found positive associations with high-impact innovation (Azoulay et al. 2019a, Moretti et al. 2019), whereas other studies have not found significant effects (Pahnke et al. 2015). Some recent work points in the direction of suggesting that certain types of relationships could produce more high-impact innovation than others (Bruce et al. 2019). To sum up, whereas university funding and tax credits are indirect, general ways of government subsidy, R&D relationships can be targeted to specific purposes and thus possibly to support enabling technologies, but their effectiveness is debated. This third category of government R&D subsidies, that is, public–private R&D relationships, is the focus of this paper.

Gaps in Prior Work
To explain the occasionally conflicting findings of prior work, we focus on two major gaps in the study of private–public R&D relationships: incorporating the variety of partner and relationship types.

Partner Type. First, heterogeneity in public-partner agencies deserves more attention (Dasgupta and David 1994).

Prior work has taken a homogeneous view or focused on one agency at a time (Fuchs 2010, Pahnke et al. 2015). This work does not pay attention to differences in partner agencies or how each agency's unique goals may impact the relationship. This is particularly significant for our paper because an agency's goals are likely to influence the enabling technology trajectories that the relationship spawns. Distinguishing between different types of agencies is also significant because policy in many nations often treats agencies differently. Overall, our understanding of public–private relationships could benefit from more directly setting the agencies side by side.

The commonly used distinction between mission and science agencies is likely to be relevant (Ergas 1987). *Mission* agencies are defined as programmatic agencies with the goal of achieving practical goals; that is, they focus on an agency's mission. Ergas (1987) defines *mission-oriented agencies* as closely synonymous with agencies whose goal is national sovereignty and who use radical innovations to achieve national goals and are often centrally managed, including the Department of Defense (DOD) and the National Aeronautics and Space Administration (NASA) in the United States. In contrast, *science* agencies including the National Science Foundation (NSF), the NIH, and the National Institutes for Standards and Technology (NIST) are science focused (Ergas 1987), and their goal is to provide the scientific freedom for private firms to pursue their own research goals. Another difference is that mission agencies often act as lead markets, whereas science agencies do not.

Relationship Type. Second, incorporating a full range of public–private R&D relationship types deserves more attention (Feldman and Kelley 2006; Hiatt et al. 2017). By focusing on a single type of relationship in isolation—often grants—prior work is often agnostic to differences across relationship types. For example, our analysis of published work over the past 30 years in journals in the intersection of private-firm strategy and public policy shows that almost 70% of articles on public–private R&D relationships focus on grants, although less than 40% of the relationships are grants. We also found that grants were, on average, two to seven times more likely to be covered by each of the publications than contracts and collaborative agreements (i.e., 2.3, 7.4, 2.1, and 4.7 in *Strategic Management Journal*, *Research Policy*, *Management Science*,

and *Academy of Management Journal*, respectively). Altogether, grants are typically overrepresented in prior work.

Setting different relationship types side by side is useful because they often involve very different structures of interaction between public and private organizations (e.g., a grant is arm's length, but other relationships follow peer-to-peer or leader–follower structures; Rathje 2019). Altogether, as Bruce et al. (2019) note, considering the full range of types could be important for a comprehensive analysis of public–private relationships. Tables 1 and 2 summarize these differences that are likely to be meaningful for subsequent enabling technology trajectories.

Private–Public Relationships and Enabling Technologies

Core Components of Enabling Technologies. Building on Teece (2018a), we define enabling technologies as (1) pervasive, (2) novel and improvable, and (3) supportive to spawn complementary innovations. *Pervasive technologies* are those in wide use in an industry or with widespread application across multiple domains (Teece 2018a, Gambardella et al. 2019). Take, for example, LIDAR. LIDAR was created from a relatively simple concept—using light to measure distances between two objects (Neff 2018). Today, the technology is used in a broad spectrum of application sectors from autonomous driving to guided weapons and is therefore considered highly pervasive.

Second, enabling technologies are novel and capable of being continuously improved.[9] For instance, LIDAR has served as a foundational technology for many subsequent inventions, and these inventions have, in turn, spurred continual development in the core LIDAR technology. MonoLIDAR was the fundamental technology enabling the Defense Advanced Research Projects Agency's (DARPA's) 1985 autonomous land vehicle, thought to be the world's first fully autonomous vehicle (Burns 2018). Twenty years later, Ford engineers David Hall and Jim McBride introduced rotating LIDAR in DARPA's Grand Challenge, capturing LIDAR measurements in stereo. This advancement was widely acknowledged to be one of the most important catalysts for the creation of today's autonomous vehicle market. Because LIDAR is capable of ongoing improvement and is consistently introduced in new ways, it is considered a novel and improvable technology.

Table 1. Public Agencies

Classification	Definition	Examples	Contribution
Science	Goal is to increase scientific knowledge and understanding	National Science Foundation, National Institutes of Health	Freedom for scientific inquiry
Mission	Goal is to fulfill practical goals for the agency such as procure technologies	Department of Defense, NASA	Early monopsony buyer

Table 2. Relationship Types

Classification	Definition	Public-Sector contribution
Grant	No required direct interaction with public researchers	Financial resources
Contract	Required interaction with public researchers. Format for interaction is specified ex-ante	Financial resources, Structured knowledge flows
Agreement	Required interaction with public researchers. Format for interaction is not specified.	Financial resources, Diverse (i.e., public-private) R&D teams

Third, enabling technologies spawn uniquely complementary innovations; that is, they support future innovations that cannot be supported by other supplementary technologies. For example, since LIDAR first flew into space in the Apollo 15 mission, dozens of LIDAR-based atmospheric, mapping, and communications satellites have been created and launched, paving the way for the thriving commercial satellite business of today.[10] Furthermore, after significant technological codification and improvement, several consumer-oriented products such as the laser-based police speed gun were spawn from the original technology.

Given the attributes of enabling technologies, prior research has noted that designing a business model to capture value from enabling technologies is challenging (Arrow 1962).[11] Because enabling technologies are intermediate inputs in the value chain, they are often commercialized by downstream firms who own the required complementary assets (Teece 2000, 2018a). A prime example is the invention of float glass that drastically reduced the production cost of glass and facilitated downstream innovations such as flat-panel displays. However, as Teece (2000) documents, the inventor, Pilkington Glass, received only about 5% of the benefits, whereas the social rates of return of this new enabling technology were around 30%. In other words, pioneers in an enabling technology risk capturing only a fraction of the value that was created. This suggests that private firms may underinvest and that private firms' joint efforts with the government may be particularly relevant for spawning enabling technologies. Therefore, understanding how private firms are associated with enabling technologies, both together with and without the public sector, becomes important.

Hypotheses
The hypotheses that follow focus on how private firms' relationships with public agencies versus the private firms' efforts alone can potentially lead the invention into a trajectory that becomes enabling. We also propose differential effects by public-partner type.

Enabling Technology and Public–Private Relationships
In Hypothesis 1, we propose that private-firm relationships with public partners have a positive impact

on enabling technology emergence. First, we expect public–private relationships to be more positively associated than private-firm-only efforts with pervasiveness. Technologies become pervasive through widespread application. Because the intent of public–private relationships is to ensure widespread public benefit, we expect that the relationships are likely to spawn technology that is more accessible and applicable to a broader set of segments (i.e., more pervasive) than private firms' efforts alone. One example of potential pervasiveness is the relationships between private firms and the public sector that are organized through grants. Given that grants are often based on public calls for proposals, that grant applications are sent to peer review, and that grant applications and final reports (at least the abstracts) are publicly available, it is likely that others can more readily build on the ideas that result from grants. Furthermore, because the data and results are often widely disseminated (e.g., data transparency rules by the NSF), we expect that inventions that result from grants can become widely used as stepping stones for new inventions and technologies (Leatherbee and Katila, 2020), thus fostering enabling technology trajectories. Cooperative agreements, in turn, can similarly increase pervasiveness because diverse R&D teams are involved (Helpman and Trajtenberg 1994). This is because the underlying expertise in using the technology is shared across a broader set of inventors and thus likely to be applied in a broader set of technology sectors. We therefore expect that tacit knowledge of any resulting technology would permeate both the partnering agency and the firm ecosystem. Finally, although contracts are designed for a specific procurement purpose, we still expect them to spawn more pervasive innovations than private firms' efforts alone because contracts necessitate sharing the technology with a public-sector partner.

Second, we propose that technologies developed together with public-sector partners have more potential to spawn complementary innovations than technologies developed without them. Because complementors must invest significant internal resources to learn the technology before they can take advantage of it (Teece 2018a), having the general public or even a particular government agency as a default customer,

as is typical in public–private relations, can be particularly helpful. For example, in contract relationships, the public procuring agency not only acts as a large lead market for enabling technologies but also provides translational help by codifying tacit aspects of the technology for potential builders of enabling technology (David and Hall 2000, Ahuja and Katila 2004). Similarly, in cooperative agreements, teams that work together often include a mix of scientists and lead-user practitioners that facilitates the understanding of any tacit knowledge related to the technology and the creation of complementary innovations. Although the role of the government partner is more limited in grants, grant recipients are typically asked to reflect on and outline broader impacts for research, which can facilitate spawning of complementary innovations relative to private firms' efforts alone.

Third, we expect that the diversity of the team when government is a partner—relative to private firms' efforts alone—is likely to produce novel and improvable technologies. By definition, in public–private R&D relations, the goal of the public agency and the private firm is to jointly solve a complex technical problem, thus increasing the diversity of perspectives (and resources) that are brought to bear on problem solving. Cooperative agreements are particularly likely to involve diverse interdisciplinary teams (e.g., a private firm's working relationship with a national laboratory) and thus likely to use more varied knowledge that often underlies novelty (Katila and Ahuja 2002). Grants and contracts also involve interactions with public funding partners and the focal firm that can be conducive of diversity in viewpoints and thus novelty. Thus, we expect public–private relationships to have a positive impact on the enabling trajectory of the technologies.

Hypothesis 1. *Technologies that are a result of a private firm's relationships with the public sector become more enabling than technologies that are a result of a private firm's efforts alone.*

Enabling Technology and Mission vs. Science Agency Partners
In Hypothesis 1, we treated all public–private R&D relationships as homogeneous and compared their effects to private-firm-only efforts. In Hypothesis 2, we propose that there may be differences across types of public partners—we focus on science versus mission agencies in particular—that potentially make some relationships even more likely to start enabling technology trajectories than others.

As noted previously, science agencies including the NSF and the NIH in the United States are science focused (Ergas 1987) and aim to accomplish scientific goals. Science agencies also typically have access to

substantial research funding internally and employ intramurally a wide variety of scientific personnel. They also own important national laboratory space, including sought-after capital-intensive research infrastructure. For example, in the United States, science agencies control the vast majority of national user facilities, that is, government-owned laboratory environments used by academic, government, for-profit, and not-for-profit organizations. Thus, science agencies are well equipped to support private firms' R&D activities with the intent of furthering scientific aims.

Mission agencies including the DOD and NASA are, conversely, generally built to support a national externality, for example, national security or space travel. Their mission is to fund and purchase technology to accomplish these practice-oriented (aka mission) goals. Mission agencies also differ from science agencies because they can typically provide access to lead users within government who can apply the novel technology to mission needs.

We argue that science agencies are more likely than mission agencies to serve as starting points for enabling technology trajectories for several reasons. Given the in-house research infrastructure, relationships with science agencies can likely support projects that cross geographic or technology boundaries and help firms engage diverse interdisciplinary teams, thus likely increasing the technology's pervasiveness. Because science agencies often support deep knowledge search and help overcome underinvestment in certain R&D areas, they are also likely associated with more novel technologies. We also expect relationships with science agencies to enhance complementarity because science agencies often encourage translational efforts across sectors, not only in a specific sector. Because these attributes are associated with more enabling technologies, as noted previously, it is likely, then, that private firms' relationships with science agencies can more positively influence the enabling trajectories of technologies than relationships with mission agencies. In contrast, the norms and goals of mission agencies are to fund demand side innovation by incentivizing the delivery of products (Edler and Georghiou 2007, Nemet 2009), which may limit pervasiveness, novelty, and wider complementarity beyond the mission agency's lead market. We propose the following hypothesis:

Hypothesis 2. *Technologies that are a result of a private firm's relationships with science agencies become more enabling than technologies from relationships with mission agencies.*

Method
Sample and Data Sources
The data set is the full population of U.S. patented technologies assigned to private firms between 1982

and 2002. We begin the sample in 1982. This allows sufficient time for the Federal Grant and Cooperative Agreement Act (U.S. Congress 1977) that established the three standard public–private R&D relationship types (grants, cooperative agreements, and contracts) to become institutionalized. We end the sample in 2002 to allow sufficient time for patent outcomes to be well understood. The final sample is 1,862,045 private-firm assigned patented technologies, 33,130 of which were the result of a public–private R&D relationship.

We chose to analyze patented technologies because they provide a useful, perhaps best, representation of private firms' technical problem-solving efforts over time (Jaffe and Trajtenberg 2002, Katila and Ahuja 2002). Patents by definition are both a detailed description of a complex technical problem and a solution to that problem (Walker 1995) and must be nontrivial, original, and useful, making them a particularly appropriate data source.

We focus on patents granted by the U.S. Patent and Trademark Office (USPTO), compiled by the USP-TO's PatentsView.org. We use the complete database. These data were further cross-referenced with Google's Patent Search. Google's Patent Search was used to specifically rectify the missing dates from less than 1% of the PatentsView data. By combining PatentsView and Google Patent Search data sets, the final data set includes the full population of patents and patent citations over the 30-year span, containing 4,218,252 unique patents. To focus on private firms, we further removed 1,100,000 patents that had no assignee (i.e., patented by an individual and not an organization) or were assigned to public organizations.[12]

We use the federal interest statement in patent documents to distinguish private-firm-only efforts from those with public partners. We supplement the USPTO data with federal contract information from USASpend.gov and cross-referenced with Google's Patent Search. Firms that file a patent in the United States are legally obligated to file a federal interest statement that outlines whether the patent was a result of a public–private relationship. "The contractor or grantee . . . in applying for a patent . . . must add a government interest statement that discloses the government's rights to the invention" (GAO/RCED-99-242 Federally Sponsored Inventions, 1999, p. 4).

The federal interest statement also contains other details about the relationship, including relationship type and partner agency. Rai and Sampat (2013) find that the federal interest statements in patents are a much more accurate data source than the records maintained by the government itself (e.g., iEdison). They further note that organizations are much more likely than individuals to comply with the requirement to note federal interest, therefore making the data source

appropriate for our data on firms. Corredoira et al. (2018) further review court cases that show that underreporting of government sponsorship undermines the inventor's legitimacy of ownership interest in patent infringement cases involving third parties, disincentivizing hiding government interest in a patent. Failure to properly report federal funding in patent documents may lead to a forfeit of the patent title to the government or withholding of additional grant funds (Corredoira et al. 2018), thus further corroborating the use of the federal interest statement as our data source.

Statistical Methods

We use a quasi-experiment to test the association of enabling technologies and public–private R&D relationships. A critical challenge in the comparison of private-only and public–private technologies is selection bias. An ideal approach would be to conduct an experiment by randomly assigning problems for different types of organizations and their partners to solve. As a result, selection bias would be eliminated. Unfortunately, a grand experiment of this kind is unavailable. We therefore rely on a three-step process to construct quasi-experimental conditions.

Step 1: Propensity score matching with machine learning. Matching to find comparable treated and control groups is particularly important for a study on public–private relationships for several reasons. First, public partners may target research areas with the most potential for enabling innovation, which could lead to a correlation between public relationships and enabling technology (patent) outcomes even if public relationships were unproductive. Similarly, the complex problems solved by public–private relationships could be inherently different from the problems solved by many private firms acting on their own and could again significantly bias the results. Matching methods attempt to address this selection bias by controlling for these preexisting differences across treated and control groups (Stuart and Rubin 2008) by identifying subsamples that are balanced with respect to observed covariates.

Coarsened exact matching (CEM) and propensity-score matching (PSM) help address endogeneity concerns by matching treated (public–private) and control (private-firm-only) groups on all potentially observable confounding covariates (Rosenbaum and Rubin 1983, Stuart and Rubin 2008). To implement PSM, researchers typically subjectively select a set of potentially confounding covariates. Next, they use the set of confounding covariates (technology class, application year, and grant year of patents constitute the standard set of covariates used in patent studies; Jaffe et al. 1993, Trajtenberg et al. 1997) to generate each patent's probability of being selected into treatment,

that is, propensity score, often using logistic regression, and match patents in the treatment and control groups with similar scores. Recent research has pointed out, however, that relying on a limited set of covariates leaves researchers unable to adequately control for selection bias (Hamilton and Nickerson 2003, Antonakis et al. 2010).

In response, we expand the confounding variable set used for matching significantly. Specifically, we use geographic location, number of inventors, and number of patent examiners in addition to the three standard covariates mentioned previously (technology class, application year, and grant year) and their interactions.[13] Given the expanded set of confounding covariates, however, traditional PSM methods break down.[14] Additionally, a large number covariates increases the likelihood of overfitting, which can increase bias (Caliendo and Kopeinig 2008). Overfitting can be particularly damaging because relying on overfit propensity scores can result in inflated standard errors, that is, lack of precision (Schuster et al. 2016), over- or underestimation of the second-stage effects (Cepeda et al. 2003), or "paradoxical associations" (i.e., significance in the wrong direction; Concato et al. 1996, p. 1373).

In response, we implement a hybrid machine learning (ML)–PSM approach. Specifically, our ML-PSM method incorporates three tenants of supervised machine learning to overcome limitations. First, advanced stochastic optimization overcomes intractability by exploiting online optimization, allowing us to include a much larger pool of potentially confounding dimensions. Second, regularization and cross-validation enable us to avoid overfitting. Regularization reweights unconfounding covariates (i.e., those that do not influence selection) to zero and therefore removes them from the regression. Cross-validation splits the sample between training, validation, and test sets, using the split sample to tune the regularization and model parameters such that the most accurate predictions can be generated. Combining these approaches, we substitute traditional logistic regression for a supervised machine learning approach to calculate patent propensity scores.

Next, we construct a subset of control patents with sufficiently high propensity scores (the matching sample) such that they resemble the patents that were the outcomes of public–private relationships, except for receiving the treatment. To determine which propensity scores were high enough for inclusion in the matching sample, we used a global optimum nearest-neighbor matching algorithm, matching one to one without replacement. Given that there were more than three million control patents to choose from, a fully matched sample can be generated (i.e., each treated patent is matched with a control). Robustness checks on various caliper[15] matching distances were run, per standard practice. The fully matched sample includes 66,260 patents, 33,130 of which are the result of a public–private relationship.

Step 2: Matched samples. We then analyzed the effectiveness of the matching. The effectiveness of the methodology to generate like distributions of technologies for treated (public–private relationship) versus control (private-firm-only) groups is illustrated using propensity score distribution overlap in logit form. Prematch, nearly two-thirds of the control group are not matched by the treatment group. Postmatch, the distributions are almost identical. Further evidence of balance across treatment and control groups is found using distributional difference in covariates pre- versus postmatch (figures available from the authors).

Step 3: Regression models. Using the matched sample, we then used ordinary least squares (OLS) and negative binomial regressions to analyze the impact of relationships and partner agencies on enabling technology trajectories. We included fixed effects for application year, location, and technology. Because U.S. patent technology categories were updated across all year groups at the same time, technology remains constant within time, and year-technology interaction effects are unnecessary.

Measures

Dependent Variable: *Enabling Technology.* We measured enabling technology by a continuous index (ETI), operationalizing three components of enabling technology (pervasiveness, novelty, and complementary innovations) using patent data as described below.

Pervasiveness, that is, how widespread the technology becomes across multiple domains, is measured by the breadth of technical fields in which the focal patent is subsequently cited. Three-digit U.S. patent classes are used to measure technical fields. Citations within 15 years of the patent's grant date are used. Breadth of technical fields is measured using a Herfindahl index that ranges from zero to one (Trajtenberg et al. 1997). The greater the breadth, the more widespread is the impact of the focal technology.

Novelty (i.e., the ability to be improved upon over time)[16] is measured by the breadth of technical fields of backward citations (backward citations are the patents that are cited in the focal patent). Although many measures of novelty exist (Verhoeven et al. 2016), we use the well-accepted approach of Trajtenberg et al. (1997) and measure novelty by the breadth of knowledge that is cited by the focal patent. Novelty is calculated using a Herfindahl index and ranges from zero to one, with a score of one indicating maximum novelty. Again, technical fields are operationalized using three-digit U.S. patent classes.

Complementarity is measured by the ratio of forward citations (citing patents) that uniquely cite the originating patent and not its predecessors (Funk and Owen-Smith 2017). By using the local network properties of the focal patent, the complementarity measure captures the change in citation patterns of future technologies given the introduction of the focal patent.[17] If the focal patent becomes the sole patent cited by future technologies, then the technology is completely unique in supporting complementary innovations and receives a score of 1, whereas, in contrast, if the citation patterns are reinforced, and all new technologies cite both the focal patent and its predecessors, then the technology is seen as consolidating and receives a score of −1. We calculate the index of unique complementarity at 15 years after the patent's grant date. Larger values indicate greater levels of unique complementarity.

The index, i.e. ETI is calculated by loading each factor (pervasiveness, novelty, and complementarity) onto a single factor using factor analysis (varimax with oblique rotation, eigenvalue > 1, Cronbach's alpha > 0.7). For each patent, the resulting ETI is mean centered at zero.

In order to construct the index, patents that did not receive any forward citations were excluded. To incorporate these patents in the analysis and to provide a comparison with a simple count-based measure of the general importance of the technology, we coded an alternate outcome variable, *Citation-Weighted Patents*—an often-used indicator of technology importance (Katila and Mang 2003). Patent classification and year fixed effects normalize this measure. We discuss the results using this alternative (*Importance of Technology*) as an outcome variable in robustness.

To take into account patents that receive no citations, we also ran a two-step Heckman analysis that attempts to take into account treatment effects (i.e., particular government relations could fly under the radar or, in contrast, be particularly widely disseminated, possibly influencing whether a particular invention is subsequently cited; Gross 2019). Results were consistent. Because OLS performs better than the Heckman method when it is challenging to find a valid exclusion restriction (Wolfolds and Siegel 2019), we report OLS as the main analysis.

Independent Variables
As noted previously, we use the mandatory federal interest statement in private firms' U.S. patents to code whether a public–private relationship underlies a technology (Selsky and Parker 2005, Flammer 2018, Bruce et al. 2019). We define a public–private relationship to exist for a patented invention if a federal interest statement is included in the patent document, indicating that a tie between a public agency and a for-profit firm existed to exchange resources (e.g., financial, infrastructure, or research support) and to create a

mutually desired R&D outcome (i.e. the patented technology).

We use the details provided in the federal interest statement to distinguish between three types of public–private relationships. *Grant* is a binary variable that takes a value of one if the patented technology was the result of a grant-based relationship. *Contract* is a binary variable that takes a value of one if the technology was the result of a procurement contract. *Agreement* is a binary variable that takes a value of one if the technology was the result of a cooperative agreement (U.S. Congress 1977).

We use the details provided in the federal interest statement to distinguish between two types of agency partners. *Mission Agency* is a binary variable that takes a value of one if the patented technology was a result of a relationship with the DOD (36% of public–private relationships in the original data) or NASA (8%). *Science Agency* is a binary variable that takes a value of one if the patented technology was a result of a relationship with the NIH (23%), NSF (10%), or NIST (0.5%). The remaining category is *Other*, that is, hybrid agencies that often share both science and mission goals (e.g., Department of Energy, 1.8%; Department of Agriculture, 0.4%).

Controls
We control for the *number of backward citations* in each patent because fewer citations may indicate less crowded (and less mature) fields with more room for enabling innovation (Scotchmer 2004). We control for the *number of inventors* measured by the count of the total number of inventors listed on the patent application. Larger teams working on a patentable technology likely have a larger network with which to share the innovation and more significant access to additional resources, such as knowledge and funding (Jaffe et al. 1993), thus suggesting a potential influence on the enabling nature of the technology.

We control for the geographic location because local infrastructure, cultural differences, and proximity to inventor networks can influence knowledge spillovers and propensity to cite (Jaffe et al. 1993). Location is defined by the categorical variables *State* and *Country* at the time of invention. We control for the three-digit U.S. patent *technology class* for each patented invention using dummy variables to take into account differences in technology fields, including the propensity to cite, build patent thickets, and protect intellectual property (Mansfield 1986). There are 496 technology classes in the data.

We include fixed effects for patent *application year* to control for technology-sector dynamism and tools and technologies available for subsequent innovators. We also control for *time to grant*, measured by the number of years from filing to grant of a patent,

Table 3. Descriptives and Correlations

Variable	Mean	Standard deviation	1	2	3	4	5	6	7	8	9	10
1 *Enabling Technology Index*	0.11	0.94										
2 *Public–Private Relationship*	0.50	0.50	0.02									
3 *Contract*	0.25	0.43	0.0003	0.58								
4 *Agreement*	0.01	0.11	0.01	0.11	−0.03							
5 *Grant*	0.13	0.34	0.04	0.38	−0.15	−0.01						
6 *Mission Agency*	0.13	0.34	0.01	0.39	0.52	0.04	−0.03					
7 *Science Agency*	0.09	0.29	0.03	0.31	−0.09	0.04	0.61	−0.12				
8 *Time to Grant*	2.61	1.54	0.10	−0.05	−0.07	0.02	0.12	−0.03	0.11			
9 *Patent Age*	19.81	8.21	−0.13	0.26	0.19	−0.03	−0.03	0.13	−0.05	−0.39		
10 *Number of Inventors*	2.45	1.65	0.05	0.01	0.01	0.05	0.08	0.00	0.06	0.15	−0.24	
11 *Number of Backward Citations* (logged)	1.77	1.09	0.04	−0.16	−0.06	0.04	−0.03	−0.04	−0.04	0.20	−0.45	0.14

to account for time-variant changes in the USPTO, such as changes in the examination processes or possibly even the underlying complexity of the patented technology not controlled for otherwise, which may affect enabling innovation. We also control for *patent age*, measured by the number of years since a patent was granted.

Results

Table 3 includes descriptive statistics and correlations. Table 4 presents the high-level associations of *Enabling Technology* and *Public–Private Relationships*. Private-only is the omitted category. Samples before matching (models 1 and 2) and after ML-PSM matching (models 3 and 4) are reported for comparison. Models 2 and 4 both indicate that public–private relationships are associated with technologies that start a more enabling trajectory in comparison with the private firms' efforts alone, supporting Hypothesis 1.

Table 5 includes the analyses by public agency (models 1–3) to test Hypothesis 2. Consistent with Table 4, the results show that public–private relationships have greater associations with enabling technology trajectories of technologies than private firms acting on their own, again supporting Hypothesis 1. However, we do not find support for Hypothesis 2, which predicted that science agencies as partners would result in more enabling technology trajectories than mission agencies. In fact, we find the opposite. In model 3, mission agency relationships relate to slightly more (not less) enabling technology than science agencies, thus contradicting Hypothesis 2.

Given the unexpected results regarding Hypothesis 2, Tables 6 and 7 further split the analysis by agency, respectively, to examine the possible differences across contracts, grants, and cooperative agreements within each agency relationship. Model 2 in both tables confirm that both science and mission agency relationships have a positive association with enabling technology

Table 4. Ordinary Least Squares Predicting Enabling Technology: Using Prematch vs. Postmatch Samples

	Dependent variable: *Enabling Technology Index*			
	Prematch		Postmatch	
Variable	(1)	(2)	(3)	(4)
Public–Private Relationship		0.113***		0.103***
		(0.102, 0.123)		(0.088, 0.117)
Time to Grant	0.037***	0.036***	0.015	0.013
	(0.032, 0.041)	(0.031, 0.041)	(−0.009, 0.039)	(−0.011, 0.037)
Patent Age	−0.012***	−0.012***	−0.019	−0.021
	(−0.016, −0.007)	(−0.017, −0.007)	(−0.043, 0.004)	(−0.044, 0.002)
Number of Inventors	0.003***	0.002***	0.003	0.002
	(0.002, 0.004)	(0.002, 0.003)	(−0.001, 0.008)	(−0.003, 0.013)
Intercept	−2.586***	−2.548**	0.245	0.310
	(−4.195, −0.978)	(−4.156, −0.940)	(−2.943, 3.432)	(−2.873, 3.493)
R^2	0.165	0.166	0.163	0.166
Adjusted R^2	0.165	0.165	0.156	0.159
Observations	1,862,045	1,862,045	66,260	66,260

Note. Fixed effects for application year, location, and technology class are included in all models.
　*p < 0.05; **p < 0.01; ***p < 0.001.

Table 5. Ordinary Least Squares Predicting Enabling Technology: Split by Agency

Variable	Dependent Variable: *Enabling Technology Index*		
	(1)	(2)	(4)
Intercept	3.457* (−0.397, 7.311)	3.565* (−0.283, 7.414)	3.581* (−0.269, 7.430)
Public–Private Relationship		0.102*** (0.081, 0.124)	
Mission Agency			0.104*** (0.079, 0.129)
Science Agency			0.101*** (0.070, 0.132)
Time to Grant (years)	−0.010 (−0.045, 0.026)	−0.013 (−0.049, 0.022)	−0.013 (−0.049, 0.022)
Patent Age (years)	−0.051*** (−0.086, −0.016)	−0.052** (−0.087, −0.017)	−0.052** (−0.087, −0.017)
Number of Inventors	0.002 (−0.004, 0.009)	0.0004 (−0.006, 0.007)	0.0004 (−0.006, 0.007)
Number of Backward Citations (logged)	−0.039*** (−0.050, −0.029)	−0.037*** (−0.047, −0.026)	−0.037*** (−0.047, −0.026)
R^2	0.171	0.173	0.173
Adjusted R^2	0.156	0.159	0.159
Observations	28,860	28,860	28,860

Note. Fixed effects for application year, location, and technology class are included in all models.
 *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

trajectories when private-firm-only efforts are the base-line. However, model 6 shows that cooperative agreements with mission agencies have the strongest positive association with enabling technology (Table 7). Grants in science agencies also have positive but relatively weaker effects (Table 6). Thus, particular types of relationships with mission agencies are likely starting points for enabling technology trajectory, which could explain why we find that mission agencies rather than science agencies are more likely to spawn more enabling technologies (i.e., the unexpected results regarding Hypothesis 2). We return to this intriguing finding in the discussion.

Robustness Tests
We ran several robustness tests to increase confidence in our findings. We first bolstered confidence in our dependent variable measure. Prior work argues that the more enabling the technology, the less likely are the private firms to capture value (Teece 2018a; Chesbrough and Appleyard 2007). We thus expected our dependent variable, that is, *Enabling Technology Index*, to have a negative relation with value capture. This expectation is strongly confirmed. Using knowledge spillover reabsorption (Belenzon 2012) to measure value capture, we find a negative relation that increases confidence in our dependent variable measure. These results are available from the authors.

We then used an alternate dependent variable, that is, *Importance of Technology*, measured by the sheer number of forward citations, to show that our original dependent variable, that is, *Enabling Technology Index*, is a distinctively different construct. As a dependent variable, *Importance of Technology* captures knowledge impact that is distinct from enabling technology impact. Using *Importance of Technology* as the dependent variable (i.e., citations made to the focal patent by subsequent patents, both with and without self-cites), the results differ from our original results, as expected. Grants and science agencies now have the largest positive coefficients, reflecting that they spawn important knowledge captured by the forward citation counts. Altogether, these alternative tests indicate that our dependent variable, that is, *Enabling Technology Index*, indeed captures a distinctively different trajectory of a technology than other measures.

Discussion
Despite the rise of interest in enabling technologies such as AI, our understanding of the role of private firms in developing enabling technologies is limited. We started this paper with the question of how enabling are the technologies developed by private firms with and without public-sector partners. Using machine learning matching, we compared technologies generated from more than 30,000 public–private relationships with comparable technologies invented by private firms alone during a 21-year period. Our results shed light on this important strategic option for firms.

Table 6. Ordinary Least Squares Predicting Enabling Technology: Relationships with Science Agencies

Variable	Dependent variable: *Enabling Technology Index*					
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	3.622**	3.556**	3.592**	3.619**	3.635**	3.576**
	(0.504, 6.740)	(0.442, 6.670)	(0.475, 6.710)	(0.500, 6.738)	(0.520, 6.750)	(0.461, 6.691)
Science Agency Relationship		0.102***				
		(0.064, 0.140)				
Contract			0.060*			0.091***
			(−0.007, 0.127)			(0.023, 0.159)
Agreement				0.008		0.040
				(−0.169, 0.185)		(−0.137, 0.217)
Grant					0.090***	0.099***
					(0.051, 0.128)	(0.060, 0.138)
Time to Grant (years)	−0.030	−0.034	−0.030	−0.030	−0.034	−0.034
	(−0.089, 0.030)	(−0.093, 0.025)	(−0.089, 0.030)	(−0.089, 0.030)	(−0.094, 0.025)	(−0.094, 0.025)
Patent Age (years)	−0.076**	−0.077***	−0.076**	−0.076**	−0.078***	−0.077***
	(−0.134, −0.018)	(−0.135, −0.019)	(−0.134, −0.018)	(−0.134, −0.018)	(−0.136, −0.020)	(−0.136, −0.019)
Number of Inventors	−0.004	−0.005	−0.004	−0.004	−0.005	−0.005
	(−0.014, 0.005)	(−0.015, 0.005)	(−0.014, 0.005)	(−0.014, 0.005)	(−0.014, 0.005)	(−0.015, 0.005)
Number of backward citations (logged)	−0.039***	−0.035***	−0.038***	−0.039***	−0.036***	−0.035***
	(−0.054, −0.023)	(−0.051, −0.019)	(−0.054, −0.023)	(−0.054, −0.023)	(−0.051, −0.020)	(−0.051, −0.019)
R^2	0.186	0.188	0.186	0.186	0.188	0.188
Adjusted R^2	0.154	0.156	0.154	0.154	0.155	0.15
Observations	10,544	10,544	10,544	10,544	10,544	10,544

Note. Fixed effects for application year, location, and technology class are included in all models.
*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

Table 7. Ordinary Least Squares Predicting Enabling Technology: Relationships with Mission Agencies

Variable	Dependent variable: *Enabling Technology Index*					
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	2.330	2.355	2.309	2.431	2.398	2.511
	(−2.452, 7.113)	(−2.420, 7.131)	(−2.470, 7.088)	(−2.353, 7.214)	(−2.383, 7.179)	(−2.266, 7.288)
Mission Agency Relationship		0.101***				
		(0.072, 0.130)				
Contract			0.071***			0.084***
			(0.042, 0.101)			(0.054, 0.114)
Agreement				0.213*		0.253**
				(−0.011, 0.437)		(0.029, 0.477)
Grant					0.109***	0.139***
					(0.048, 0.170)	(0.077, 0.201)
Time to Grant (years)	−0.008	−0.010	−0.009	−0.009	−0.009	−0.012
	(−0.057, 0.040)	(−0.059, 0.038)	(−0.057, 0.039)	(−0.058, 0.039)	(−0.058, 0.039)	(−0.060, 0.036)
Patent Age (years)	−0.041*	−0.043*	−0.042*	−0.043*	−0.042*	−0.044*
	(−0.089, 0.006)	(−0.090, 0.005)	(−0.090, 0.005)	(−0.090, 0.005)	(−0.089, 0.006)	(−0.091, 0.003)
Number of Inventors	0.006	0.003	0.004	0.006	0.005	0.003
	(−0.003, 0.015)	(−0.006, 0.011)	(−0.005, 0.013)	(−0.003, 0.015)	(−0.003, 0.014)	(−0.006, 0.012)
Number of backward citations (logged)	−0.035***	−0.034***	−0.034***	−0.035***	−0.035***	−0.034***
	(−0.050, −0.020)	(−0.049, −0.019)	(−0.049, −0.020)	(−0.050, −0.020)	(−0.049, −0.020)	(−0.049, −0.019)
R^2	0.186	0.189	0.188	0.187	0.187	0.189
Adjusted R^2	0.163	0.165	0.164	0.163	0.163	0.165
Observations	16,046	16,046	16,046	16,046	16,046	16,046

Note. Fixed effects for application year, location, and technology class are included in all models.
  $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

First, our results help unpack the enabling technology trajectories of inventions created by private firms. The findings show that technologies that are a result of private firms' relationships with the public sector—in particular cooperative agreements with mission agencies (e.g., NASA, DOD)—have a stronger association with enabling technology trajectories than private firms' efforts alone. Interestingly, contracts—which agency theory traditionally considers the most effective—have weaker effects.

The new ETI that we introduce to measure enabling technology helps us quantify a firm's strategic decisions about enabling technology. Overall, firms are involved in enabling technologies in unexpectedly many ways, warranting a study on this increasingly important option for strategists.

Our findings also challenge two assumptions widely held about enabling technologies. The first assumption is that enabling technologies have their origins only in the public sector. The second assumption is that enabling technologies originate with science agencies. Our findings challenge both assumptions.

First, as noted previously, we find that private firms have an important role in enabling technology creation. Second, we contradict the belief that mission rather than science agencies may not adequately support the development of enabling technologies in private firms but rather support narrow exploitation (Christensen et al. 2018). Our results indicate a more nuanced interpretation. Mission agencies have the strongest overall effect when collaborative agreements are used. It seems that when the relationship type (agreement) with mission agencies affords more latitude and joint work, private firms' technology is sent into an enabling technology trajectory. Joint peer relationships with mission agency partners may be particularly important because they likely give a firm a better understanding of the customer's unique and specialized demands—demands that may not easily be summarized in public communications from the government, such as the statement of objectives (for contracts). This reduced information asymmetry inherent in agreements thus may enable customer relations with the mission agency instead of falling into a "valley of death." Overall, these results suggest that a balance of exploratory and exploitative components in a relationship may be most beneficial.

A third unexpected result was that we did not find contracts to be particularly helpful for enabling technology trajectories. One reason may be that relationships between private firms and the public sector that are organized through contracts have a prespecified format that often involves formal channels of communication between public and private researchers in order to cement access and insight into the firm's R&D process. Contracts thus may be effective for generating prespecified types of knowledge tailored to the relationship but less helpful for enabling technologies that gain widespread use. The norms of interaction are instantiated by public agencies primarily to reduce the risk (agency costs) in the investment. This structured interaction would explain the limited positive impact of contracts on more high-impact enabling technology. For example, in contracts, public and private researchers are assigned roles that force them to collaborate throughout the R&D process toward a specific procurement need. In other words, the public partner controls the R&D process for a specific procurement purpose, and the technology developed is possibly then less likely to be targeted for widespread use. The constrained R&D process inherent in contracts will also most likely result in less novel technologies than if less restrictive types of relationships were used. This may be true even if the firm had proposed to develop a more novel technology in its response to solicitation, given the restricted freedom inherent in the structure of the contract relationship, again explaining the more limited impact of contracts on enabling technologies.

The findings are significant for public policy because they add to the evidence (Bloom et al. 2019) that targeted government relations with private firms can support enabling technologies. We particularly find the positive role of public–private relationships that generate technologies that become enabling for subsequent inventions in many sectors. If the goal of policy is to support private-firm innovation that generates spillovers that enable growth across sectors, we have identified that particular government research relations with private firms (e.g., grants, cooperative agreements, particularly with mission agencies) are aligned with this goal.

Our results also add to the evidence about how to build technology ecosystems. We find that firms and their government collaborators can seed ecosystems with technologies that enable subsequent innovation by ecosystem participants and so shape how ecosystems evolve. More broadly, the results point in the direction that private firms, together with public partners, can be important in the creation of technology ecosystems.

As with all research, these findings should be interpreted with the usual caution and, as such, open up several intriguing opportunities for future work. Although we focused on patented technologies, one possibility is, as Murray (2010) noted, that projects with a science focus would have a lower propensity to patent in the first place. If this propensity would be particularly likely, for example, in grant and science agency relationships, it could potentially underestimate the grants' impact on enabling technologies if the projects that are missing would be of high enabling technology potential. Because we find a positive relationship between grants and science agencies

with our data, this type of missing data would be unlikely to qualitatively change our interpretation of the findings. However, if we are missing the least promising projects, that is, the low end of the distribution, because they are not patented and therefore overestimate, our findings could be affected. There is, however, no clear theoretical reasoning why grantees would be likely to not patent projects with low enabling technology value (one would actually expect the opposite because enabling technologies are difficult to capture value from). Nevertheless, these issues at the intersection of public policy and private-firm R&D present several interesting opportunities for follow-up work.

Finally, we also provide a methodological contribution by introducing a novel ML-PSM method. Extending the work that applies machine learning in strategy and economics (Illari et al. 2011, Athey and Imbens 2015), we used a new inference strategy that applies machine learning to matching. The machine learning techniques applied in this paper allow scholars to simultaneously increase efficiency and avoid overfitting, minimizing the potential for selection bias.

Acknowledgments

Endnotes

[1] General-purpose technologies (Bresnahan and Trajtenberg 1995) are defined as a high-end subset of enabling technologies, that is, technologies that rise to have substantive cumulative economy-wide impact.

[2] We use the definition of mission agencies of Ergas (1987) as agencies with the goal of achieving programmatic, practical goals. Science agencies, in contrast, focus on scientific goals.

[3] Prior research often ties partner agency to relationship type, but this is not necessarily accurate. For example, in our data, mission agencies are involved in contracts and agreements in approximately 80% and in grants in approximately 20% of cases, whereas in science agencies the numbers are reversed.

[4] R&D tax credits are typically broad spectrum, whereas loans are more targeted. For example, the U.S. Federal R&D Tax Credit Program on R&D expenditures covers up to 14% of any corporate expenditures on R&D. In contrast, federal loan programs are instituted by specific government sponsors to support specific areas of interest. For example, the U.S. Clean Energy Loan Program guarantees 2% loans for B-rated firms, with loans as low as 0.375% for AAA-rated firms. The loans are granted only to firms that apply the loaned capital in the development of innovative technologies that meet the requirements as set by the public agency.

[5] The National Science Board (NSB) reported that more than 60% of U.S. R&D spending is allocated to public–private R&D relationships (UNESCO 2016, NSB 2018).

[6] For example, in the interviews of Pahnke et al. (2015, p. 608), an NIH informant stated, "We are spending billions of dollars on projects—what are they producing? They produce knowledge. We fund inquiry. Following up after funding would question and undermine the whole system."

[7] Government agencies often switch between cost-plus, incentive-based, and fixed-price contracts depending on the level of technical maturity presumed in the procured end item (Ng et al. 2009). Cost-plus contracts are associated with the least mature technologies, whereas fixed-price contracts are tied with the most mature products.

[8] Other relationships that are similar to cooperative agreements but are used more sparingly include research consortia [Sematech in the United States, the Consortium R&D Project for Regional Revitalization in Japan], innovation grand challenges, and NASA's crowdsourcing platforms (Lifshitz-Assaf 2017).

[9] The concept of capable of ongoing improvement (Bresnahan and Trajtenberg 1995) is also frequently used. We use novelty, given its widespread use and conceptualization in the R&D literature.

[10] LIDAR was used primarily in scientific experiments until the late 1980s, with the exception of military operations. The first recorded use of LIDAR in the military was in 1972 (Neff 2018).

[11] Macroanalysis shows that social returns from enabling technologies are two to seven times greater than private returns (Lichtenberg 1992, Hall et al. 2010). Similarly, research finds that licensing models almost always undervalue enabling technology, thus leaving much of the value with other firms (e.g., suppliers, manufacturers, competitors; Arrow 2012). One way in which knowledge leaves firms is negative knowledge spillover, in which the same collaborative R&D that leads to enabling technology development may transfer tacit knowledge external to the organization, bleeding the firm of future rents.

[12] The authors thank an anonymous reviewer who noted that the focus of the data naturally limits the generalizability of the findings to private-firm inventions. Future work could expand to compare with inventions that are a result of intramural efforts of the federal government. Although patented inventions are important to examine in their own right, future work could also expand the analysis to inventions that are not patented.

[13] Patent examiners in particular are an important covariate to include in the matching process. Patent examiners are coded as a discrete variable at each patent examiner name (first_last). Patent examiners are responsible for reviewing patent applications and adding patent citations when necessary. Examiners are considered potentially confounding variables because (a) they are assigned to evaluate specific technologies for which they have unique expertise (Righi and Simcoe 2019) and (b) they vary greatly in their propensity to add patent citations. Because patent examiners therefore may be seen as highly correlated with both treatment and outcome, they are included. There are 26,698 patent examiners and 496 technology classes in the data.

[14] The absorb command in Stata is an example of a computational trick that makes regressions computationally manageable with a large number of dummy variables (Athey and Stern 2002). However, the advance of machine learning methods is particularly useful to identify potentially significant interactions that may be confounding (e.g., potential interdependency of patent examiners, technology class, and team size). Thank you to an anonymous reviewer for pushing us to clarify this.

[15] Caliper distances are defined as the maximum radius under which two variables are considered a good match. For our strictest test, we use 0.02 multiplied by one standard deviation of the distribution of all

matching distances (Austin 2011) because matching with additional neighbors may increase the bias given that additional neighbors will necessarily be worse matches.

[16] Trajtenberg et al. (1997) use the labels of generality and originality, respectively, for pervasiveness and novelty.

[17] We define $UniqueComplementarity = \frac{1}{n_t} \sum_{i=1}^{n} \frac{-2f_{it}b_{it}+f_{it}}{w_{it}}, w_{it}>0$ where n_t is the number of forward citations of both the focal patent and its predecessors, f_{it} is one if a new patent cites the focal patent, b_{it} is one if a new patent cites the focal patent's predecessors, and w_{it} is an optional weighting parameter that indexes a matrix W of weights for the focal patent i at time t. For simplicity, in our analysis, $w_{it} = 1$. See Funk and Owen-Smith (2017).

References

Adner R, Kapoor R (2016) Innovation ecosystems and the pace of substitution: Re-examining technology S-curves. *Strategic Management J.* 37(4):625–648.

Ahuja G, Katila R (2004) Where do resources come from? *Strategic Management J.* 25(8–9):887–907.

Antonakis J, Brendahan S, Jacquart P, Lalive R (2010) On making causal claims: A review and recommendatins. *Leadership Quart.* 21:1086–1120.

Arnold SA, Harmon BR, Tyson KW, Fasana KG, Wait CS (2008) *Defense Department Profit and Contract Finance Policies and Their Effects on Contract and Contractor Performance* (Institute for Defense Analysis, Alexandria, VA).

Arrow K (1962) Economic welfare and the allocation of resources for invention. *The Rate and Direction of Inventive Activity: Economic and Social Factors* (National Bureau of Economic Research, Cambridge, MA), 609–626.

Arrow K (2012) The economics of inventive activity over fifty years. Lerner J, Stern S, eds. *The Rate and Direction of Inventive Activity Revisited* (University of Chicago Press, Chicago), 43–48.

Athey S, Imbens GW (2015) Machine learning methods in economics and econometrics: A measure of robustness to misspecification. *Amer. Econom. Rev.* 105(5):476–480.

Athey S, Stern S (2002) Impact of information technology on emergency healthcare outcomes. *RAND J. Econom.* 33:399–432.

Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav. Res.* 46(3):399–424.

Azoulay P, Fuchs E, Goldstein A, Kearney M (2019a) Funding breakthrough research: Promises and challenges of the "ARPA model." *Innovation Policy Econom.* 19(1):69–96.

Azoulay P, Zivin G, Li D, Sampat B (2019b) Public R&D investments and private-sector patenting: Evidence from NIH funding rules. *Rev. Econom. Stud.* 86(1):117–152.

Balsmeier B, Kurakina M, Fleming L (2018) R&D tax credits: Mechanisms of private and public value. Working paper, University of Luxembourg, Luxembourg.

Belenzon S (2012) Cumulative innovation and market value: Evidence from patent citations. *Econom. J. (London)* 122(559):265–285.

Bloom N, Van Reenen J, Williams W (2019) A toolkit of policies to promote innovation. *J. Econom. Perspect.* 33(3):163–184.

Bresnahan TF, Trajtenberg M (1995) General purpose technologies: "Engines of growth"? *J. Econometrics* 65:83–108.

Bruce JR, de Figueiredo JM, Silverman BS (2019) Public contracting for private innovation: Government capabilities, decision rights, and performance outcomes. *Strategic Management J.* 40(4): 533–555.

Burns LD (2018) *Autonomy: The Quest to Buld the Driverless Car—And How It Will Reshape Our World* (HarperCollins, New York).

Caliendo M, Kopeinig S (2008) Some practical guidance for the implementation of propensity score matching. *J. Econom. Survey* 22(1):31–72.

Cepeda MS, Boston R, Farrar JT, Strom BL (2003) Comparison of logistic regression vs. propensity score when the number of events is low and there are multiple confounders. *Amer. J. Epidemiology* 158(3):280–287.

Chesbrough HW, Appleyard MM (2007) Open innovation and strategy. *California Management Rev.* 50(1):57–76.

Christensen C, McDonald R, Altman E, Palmer J (2018) Disruptive innovation: An intellectual history and directions for future research. *J. Management Stud.* 55:1043–1078.

Concato J, Feinstein AR, Peduzzi P, Kemper E, Holford TR (1996) A simulation study of the number of events per variable in logistic regression analysis. *J. Clinical Epidemiology* 49(12): 1373–1379.

Corredoira RA, Goldfarb BD, Shi Y (2018) Federal funding and the rate and direction of inventive activity. *Res. Policy* 47(9): 1777–1800.

Dasgupta P, David PA (1994) Toward a new economics of science. *Res. Policy* 23(5):487–521.

David PA, Hall BH (2000) Heart of darkness: Modeling public–private funding interactions inside the R&D black box. *Res. Policy* 29(9):1165–1183.

David PA, Hall BH, Toole AA (2000) Is public R&D a complement or substitute for private R&D? A review of the econometric evidence. *Res. Policy* 29(4):497–529.

Edler J, Georghiou L (2007) Public procurement and innovation: Resurrecting the demand side. *Res. Policy* 36(7):949–963.

Ergas H (1987) Does technology policy matter? *Technology and Global Industry: Companies and Nations in the World Economy* (The National Academies Press, Washington, DC).

Feldman MP, Kelley MR (2006) The ex ante assessment of knowledge spillovers: Government R&D policy, economic incentives and private firm behavior. *Res. Policy* 35(10):1509–1521.

Flammer C (2018) Competing for government procurement contracts: The role of corporate social responsibility. *Strategic Management J.* 39(5):1299–1324.

Fuchs E (2010) Rethinking the role of the state in technology development: DARPA and the case for embedded network governance. *Res. Policy* 39(9):1133–1147.

Funk RJ, Owen-Smith J (2017) A dynamic network measure of technological change. *Management Sci.* 63(3):791–817.

Gambardella A, Novelli E, Heaton S, Teece D (2019) Profiting from enabling technologies? Working paper, University of California, Berkeley, Berkeley.

GAO/RCED-99-242 Reporting Requirements for Federally Sponsored Inventions, U.S. General Accounting Office, Washington, DC, 1999.

Gross D (2019) The consequences of invention secrecy: Evidence from the USPTO patent secrecy program in World War II. NBER Working Paper No. 25545, National Bureau of Economic Research, Cambridge, MA.

Hall BH, Mairesse J, Mohnen P (2010) Measuring the returns to R&D. Rosenberg N, Hall B, eds. *Handbook of the Economics of Innovation* (North Holland, Amsterdam), 679–730.

Ham RM, Mowery DC (1998) Improving the effectiveness of public–private R&D collaboration: Case studies at a U.S. weapons laboratory. *Res. Policy* 26(6):661–675.

Hamilton BH, Nickerson JA (2003) Correcting for endogeneity in strategic management research. *Strategic Organ.* 1(1):51–78.

Helpman E, Trajtenberg M (1994) A time to sow and a time to reap: Growth based on general purpose technologies. Working Paper No. 4854, National Bureau of Economic Research, Cambridge, MA.

Hiatt SR, Carlos CW, Sine WD (2017) Manu militari: The institutional contingencies of stakeholder relationships on entrepreneurial performance. *Organ. Sci.* 29(4):633–652.

Howell ST (2017) Financing innovation: Evidence from R&D grants. *Amer. Econom. Rev.* 107(4):1136–1164.

Illari P, Russo F, Williamson J (2011) *Causality in the Sciences* (Oxford University Press, Oxford, UK).

Jaffe AB, Lerner J (2001) Reinventing public R&D: Patent policy and the commercialization of national laboratory technologies. *RAND J. Econom.* 32(1):167–198.

Jaffe AB, Trajtenberg M (2002) *Patents, Citations and Innovations: A Window on the Knowledge Economy*, 1st ed. (MIT Press, Cambridge, MA).

Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localization of knowledge spillovers as evidenced by patent citations. *Quart. J. Econom.* 108(3):577–598.

Katila R, Ahuja G (2002) Something old, something new. *Acad. Management J.* 45(6):1183–1194.

Katila R, Mang P (2003) Exploiting technological opportunities: the timing of collaborations. *Res. Policy* 32(2):317–332.

Leatherbee M, Katila R (2020) The lean startup method: Early-stage teams and hypothesis-based probing of business ideas. *Strategic Entrepreneurship Journal* in press.

Lerner J (2012) *The Architecture of Innovation* (Harvard Business School Publishing, Boston).

Lichtenberg FR (1992) R&D investment and international productivity differences. Working Paper No. 4161, National Bureau of Economic Research, Cambridge, MA.

Lifshitz-Assaf H (2018). Dismantling knowledge boundaries at NASA: The critical role of professional identity in open innovation. *Administrative Science Quarterly* 63(4):746–782.

Link A, Siegel DS, Van Fleet DD (2011) Public science and public innovation: Assessing the relationship between patenting at U.S. National Laboratories and the Bayh-Dole Act. *Res. Policy* 40(January):1094–1099.

Lipsey R, Carlaw K, Bekar C (2005) *Economic Transformations* (Oxford University Press, New York).

Mansfield E (1986) Patents and innovation: An empirical study. *Management Sci.* 32(2):173–181.

Moretti E, Steinwender C, Van Reenen J (2019) *The Intellectual Spoils of War?* (National Bureau of Economic Research, Cambridge, MA).

Mowery DC (2009) *Plus ca change*: Industrial R&D in the "third industrial revolution." *Industrial Corporation Change* 18(1):1–50.

Murray F (2010) The oncomouse that roared. *Am. J. Sociol.* 116: 341–388.

Neff T (2018) *The Laser That's Changing the World* (Prometheus Books, Amherst, MA).

Nemet GF (2009) Demand-pull, technology-push, and government-led incentives for non-incremental technical change. *Res. Policy* 38(5):700–709.

Ng ICL, Maull R, Yip N (2009) Outcome-based contracts as a driver for systems thinking and service-dominant logic in service science: Evidence from the defense industry. *Eur. Management J.* 27(6):377–387.

NSB (2018) Science & Engineering Indicators 2018. Accessed October 29, 2020, https://www.nsf.gov/statistics/2018/nsb20181/.

Pahnke EC, Katila R, Eisenhardt KM (2015) Who takes you to the dance? How partners' institutional logics influence innovation in young firms. *Admin. Sci. Quart.* 60(4):596–633.

Pless J (2019) Are complementary policies substitutes? Evidence from R&D subsidies in the UK. Working paper.

Rai A, Sampat B (2013) Accountability in patenting of federally funded research. *Natural Biotech.* 30(10):953–956.

Rathje J (2019) Essays on private funding for public innovation. PhD dissertation, Stanford University, Stanford, CA.

Rathje J, Katila R (2019) Supervised machine learning and quasi-experimental designs: Machine-learning—Matching methods for causal inference in strategy research. Working paper, Stanford Technology Ventures Program, Stanford, CA.

Righi C, Simcoe T (2019) Patent examiner specialization. *Res. Policy* 48:137–148.

Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55.

Schuster T, Lowe WK, Platt RW (2016) Propensity score model overfitting led to inflated variance of estimated odds ratios. *J. Clinical Epidemiology* 80:97–106.

Scotchmer S (2004) *Innovation and Incentives* (MIT Press, Cambridge, MA).

Selsky JW, Parker B (2005) Cross-sector partnerships to address social issues: Challenges to theory and practice. *J. Management* 31(6): 849–873.

Stuart E, Rubin D (2008) Best practices in quasi-experimental designs: Matching methods for causal inference. Osborne J, ed. *Best Practices in Quantitative Methods* (Sage, Thousand Oaks, CA), 155–176.

Teece DJ (2000) *Managing Intellectual Capital: Organizational, Strategic, and Policy Dimensions* (Oxford University Press, Oxford).

Teece DJ (2018a) Profiting from innovation in the digital economy: Enabling technologies, standards, and licensing models in the wireless world. *Res. Policy* 47(8):1367–1387.

Teece DJ (2018b) Reply to Nelson, Helfat and Raubitschek. *Res. Policy.* 47(8):1400–1402.

Trajtenberg M, Henderson R, Jaffe A (1997) University vs. corporate patents: Window on the basicness of invention. *Econom. Innovative New Techn.* 5:19–50.

UNESCO (2016) UNESCO Science Report. Accessed October 29, 2020, https://en.unesco.org/wssr2016.

U.S. Congress (1977) Federal Grant and Cooperative Agreement Act, Public Law 95-224.

Verhoeven D, Bakker J, Veugelers R (2016) Measuring technological novelty with patent-based indicators. *Res. Policy* 45(3):707–723.

Walker RD (1995) *Patents as Scientific and Technical Literature* (Scarecrow Press, Ann Arbor, MI).

Wallsten SJ (2000) The effects of government-industry R&D programs on private R&D: The case of the small business innovation research program. *RAND J. Econom.* 31(1):82–100.

Webb M, Short N, Bloom N, Lerner J (2018) Some facts of high tech patenting. NBER Working Paper No. 24793, National Bureau of Economic Research, Cambridge, MA.

Wolfolds SE, Siegel J (2019) Misaccounting for endogeneity: The peril of relying on the Heckman two-step method without a valid instrument. *Strategic Management J.* 40(3):432–462.

Zuniga-Vicente J, Alonso-Borrego C, Forcadell FJ, Galan JI (2014) Assessing the effect of public subsidies on firm R&D investment: A survey. *J. Econom. Survey* 28(1):36–67.

Jason M. Rathje is the director of AFVentures, the commercial investment arm of the U.S. Air Force. His research focuses on the overlap of government funding, entrepreneurship, and technical innovation, with special attention on how firms form research and development relationships with government agencies.

Riitta Katila is professor of management science and engineering at Stanford University and research director of the Stanford Technology Ventures Program. Her research focuses on technology strategy, innovation, and entrepreneurship in large organizations.