Public support for research in artificial intelligence: a descriptive study of U.S. Department of Defense SBIR Projects

By: Farhat Chowdhury, Albert N. Link, Martijn van Hasselt

Chowdhury, F., Link, A. N., & van Hasselt, M. (2022). Public support for research in artificial intelligence: a descriptive study of u.s. department of defense sbir projects. The Journal of Technology Transfer, 47(3), 762–774. https://doi.org/10.1007/s10961-022-09943-z

Abstract:

We describe public support for AI research in small firms using data from U.S. Department of Defense-funded SBIR projects. Ours is the first collection of firm-level project information on publicly funded R&D investments in AI. We find that the likelihood of an SBIR funded research project being focused on AI is greater the larger the amount of the SBIR award. AI-focused research projects are associated with a 7.6% increase in average award amounts. We also find suggestive evidence that the likelihood of an SBIR project being AI-focused is greater in smaller-sized firms. Finally, we find that SBIR-funded AI research is more likely to occur in states with complementary university research resources.

Keywords: artificial intelligence | machine learning | department of defense | small business innovation research program | agglomeration

Article:

> *Computers will overtake humans with AI within the next 100 years.*
> *When that happens, we need to make sure the computers have goals aligned with ours.*
>
> —Stephen Hawking.

1 Introduction

A recent Brookings Institution report (West and Allen, 2018, p. 1) suggested that:Footnote1

> AI [Artificial intelligence] is a technology that is transforming every walk of life. It is a wide-ranging tool that enables people to rethink how we integrate information, analyze data, and use the resulting insights to improve decisionmaking … AI already is altering the world and raising important questions for society, the economy, and governance.

There is no shortage of predictions about the impact that AI will have on economies and their economic agents, but there is little data to support such claims. More specifically, there is an absence of data on the level of investments in AI by private-sector firms as well as by public-sector agencies and organizations (including universities). Such data are a critical part of any policy motivated impact analysis that relates public expenditures to economic benefits to society from such social investments.

The OECD, through its AI Policy Observatory,Footnote2 adopted a global leadership role in collecting information about AI activities in some of the world's largest organizations—generally large private-sector organizations. OECD researchers are looking at job postings defining the AI-related skills required for hiring. The idea underlying the OECD's projects is that it is difficult to observe directly a firm's monetary investments in AI technology; therefore, identifying AI-related skills that firms demand is a first-order approximation for identifying firms that are AI focused—either conducting research on AI technologies or adopting AI technologies from other firms (e.g., Alekseeva et al., 2020, 2021; Babina et al., 2021; Baruffaldi et al., 2020; Nakazato and Squicciarini 2021; Samek et al. (2021); Squicciarini and Nachtigall, 2021). Relatedly, other researchers have creatively collected information on outputs from patents and trademarks related to AI, often at the country or industry level to identify AI-focused firms (e.g., Alderucci et al., 2020; Damioli et al., 2021; Dernis et al., 2021; Giczy et al., 2022; Lee et al., 2021).Footnote3 And, bibliometric studies are also revealing important information about the growth and international diffusion of AI research (Liu et al., 2021).Footnote4

In this paper, we begin the fill the void of data related to firm investments in AI research. We do this by looking at a slice of AI investment activity, namely public-sector R&D investments in private-sector firm projects related to AI technology. Our data are unique in that they pertain to public-sector investments in small private-sector firms. Specifically, our data relate to U.S. firms that received Small Business Innovation Research (SBIR) research awards from the U.S. Department of Defense (DOD), the largest SBIR funding agency.Footnote5

Using public information about SBIR Phase II awards, we are able to identify funded research projects, based on the summary abstracts of their funded projects, that are focused on AI; thereby, we are able to identify the level of public-sector research and development (R&D) investments in such projects. To the best of our knowledge, our dataset is the first such assemblage of information on U.S. public-sector investments in AI R&D at the firm level, and the first collection of firm-level project information on publicly funded AI R&D.Footnote6

The purpose of this paper is not only to describe our database of investments in AI project R&D but also to identify characteristics of Phase II SBIR-funded projects that are focused on AI.Footnote7 This paper emphasizes the relationship between a Phase II SBIR-funded project being AI-focused and the level of public-sector funding for the project, the size of the research firm, and the proximity of the firm to premier AI-focused research universities.

The remainder of the paper is organized as follows. In Section II, we offer context for our study through a brief overview of the SBIR program and the motivation to study the SBIR funding activity of the DOD. In Section III, we describe the construction of our DOD database of Phase II SBIR-funded AI and non-AI research projects, and we present descriptive statistics on Phase II projects that are AI focused over the fiscal years (FYs) FY2015 through FY2020. In Section IV, we characterize the firms that received DOD funding in terms of the amount of Phase II project funding received and their size, measured by number of employees in the firm at the time of the SBIR award. We estimate the relationship between the likelihood that a funded Phase II SBIR-funded research project will be focused on AI and these characteristics. Our paper concludes in

Section V with a brief summary of our descriptive findings. We offer words of caution about generalizing from our findings to other agencies' SBIR programs or to other public-sector AI investment activities; however, we do suggest a roadmap for possible future, policy-focused research on the nascent topic of public-sector investments in AI R&D.

2 The SBIR Program

The U.S. Small Business Administration was established through the Small Business Act of 1953, Public Law 85–536. In response to the U.S. productivity slowdown in the early-1970s and then again in the late-1970s and early-1980s, Congress looked toward small businesses as a vehicle for stimulating the economy. The Small Business Innovation Development Act of 1982, Public Law 97–219, thus begins with a related motivation statement:

> [S]mall business is the principal source of significant innovations in the Nation … small businesses are among the most cost-effective performers of research and development and are particularly capable of developing research and development results into new products.

The purposes of the SBIR program, as stated in the 1982 Act are:Footnote9

> (1) to stimulate technological innovation,
>
> (2) to use small business to meet Federal research and development needs,
>
> (3) to foster and encourage participation by minority and disadvantaged persons in technological innovation, and.
>
> (4) to increase private sector commercialization innovations derived from Federal research and development.

The SBIR program is a set-aside program. Federal agencies with an extramural research budget that exceeds a threshold level ($100 million in 2021) are required to set aside a portion of their extramural research budget each year to support agency-specific research in small firms. The set aside percentage amount is 3.2%.

In 2021, 11 agencies participated in the SBIR program; the largest program is sponsored by the DOD.Footnote10 These programs fund Phase I and Phase II projects.Footnote11 The objective of a Phase I project is to establish proof of concept of the research as well as the commercial potential of the proposed R&D. These six-month projects are legislatively funded at no more than $150,000.Footnote12 The objective of a Phase II project is to continue the R&D funded in Phase I. Only Phase I awardees are invited to apply for a competitive Phase II award. It is anticipated that by the end of the two-year Phase II projects, legislatively funded in 2021 at no more than $1 million, the awardees will have attracted third-party funding to support the commercialization of their developed technologies. However, there are a number of exceptions to the upper-limit amounts on both Phase I and Phase II project awards.Footnote13

As mentioned, the DOD's SBIR program is the largest of the 11 SBIR programs. Table 1 shows that in FY 2020, the DOD accounted for almost 50% of the number of all Phase I and Phase

II awards. It also accounted for about that same percentage of the amount of all Phase II awards ($M).

Table 1. Descriptive Data on the DOD SBIR Program in Fiscal Year 2020

	Amount of Phase I Awards ($M)	Number of Phase II Awards	Amount of Phase II Awards ($M)	Number Phase I Awards
All Agency SBIR Programs	3,915	$669.52	2,061	$2,638.01
DOD SBIR Program	1,950	$229.13	1,027	$1,278.75
DOD SBIR Program as a Percent of All Agency Program	49.8%	34.2%	49.8%	48.5%

Source: https://www.sbir.gov/analytics-dashboard?program_tid%5B%5D=105791

3 DOD database of phase II funded projects

The six FYs considered for the construction of our DOD Phase II project database are FY2015 through FY2020.Footnote14 For each fiscal year, we identified abstracts from the SBIR awards database maintained by the Small Business Administration.Footnote15 A total of 5,484 Phase II abstracts were extracted over the six-year period.

Keyword searches were used to define whether the publicly funded SBIR Phase II project's research focused on AI. Following protocols established by OECD (e.g., Squicciarini and Nachtigall, 2021),Footnote16 the four relevant keywords applicable to U.S. firms doing research in the area of AI since the mid-2000s are:Footnote17AI and/or artificial intelligence, machine learning,Footnote18python,Footnote19 and data mining.Footnote20

Associated with each project abstract are the name and location of the funded firm, the number of employees in the firm at the time of the Phase II award, and the amount of the Phase II award. The number of employees was not available in 25 of the 5,484 abstracts; those projects were deleted from our sample, thereby reducing the database to 5,459 projects.Footnote21

Table 2. Number of DOD Phase II SBIR Projects and Percent of SBIR Funded Projects Being AI Focused, by Fiscal Year (n = 5,459)

Fiscal Year	Number of Phase II Funded Projects	Percent of Phase II Projects AI Focused*
2015	782	1.0%
2016	733	1.9%
2017	985	1.8%
2018	763	2.0%
2019	1,217	5.8%
2020	979	12.7%
Overall	5,459	4.6%

* For a project to be identified as conducting AI research, the project's abstract had to include at least 1 defining keyword

Table 2 shows the number of funded projects, by fiscal year, and the percent of Phase II projects identified as AI focused. Across the six-year sample period, the percent was 4.6. The

percent of AI projects, however, has increased over time. In FY2020, 12.7% of the 979 DOD-funded Phase II SBIR projects were focused on AI research.

Table 3 shows the frequency with which the relevant defining keywords occurred in an abstract of an AI-focused Phase II SBIR project, by year. Generally, the most frequently found keyword was machine learning, followed by AI and/or artificial intelligence. The keywords python and data mining were found less frequently, especially since FY2018. Perhaps the frequency in later years is due to the more commonly used nomenclature machine learning and AI.Footnote22

Table 3. Frequency of Defining Keywords among AI-Focused DOD Phase II SBIR Projects, by Fiscal Year (n = 249)

Fiscal Year (number of Phase II projects)	***AI and/or Artificial Intelligence***	***Machine Learning***	***Python***	***Data Mining***
2015 (8)	12.5%	37.5%	12.5%	37.5%
2016 (14)	7.1%	64.3%	7.1%	21.4%
2017 (18)	22.2%	55.6%	0%	22.2%
2018 (15)	26.7%	73.3%	6.7%	6.7%
2019 (70)	37.1%	68.6%	1.4%	1.4%
2020 (124)	54.0%	78.2%	0.8%	0.8%

4 Determinants of AI focus in research projects

The data show that AI-focused projects, as defined by the presence of one or more defining keywords in an SBIR project abstract, were associated with a 7.6% increase in the average award amount, compared to projects without an AI focus.Footnote23 This "AI premium" possibly reflects funding priorities or the resource intensity (in terms of physical or human capital) of the AI-focused project. In the remainder of this section, we identify factors that are associated with the funding of an AI-focused research project.

To examine the relationship between project budget, firm size (employment size), university proximity, and the likelihood that a funded Phase II SBIR-funded research project will be focused on AI, we considered the model:

$$(1)\quad \text{Probability}\,(\textit{AnyAIproject}) = \text{f}\,(\textit{Award, Emp, AIUniv}, \mathbf{X})$$

where an AnyAIproject is measured by a dichotomous variable based on any of the four defining keywords being present in the project's abstract. The variable Award is measured as the Phase II SBIR award in 2020 dollars, Emp is the number of employees at the time of the Phase II SBIR award, and AIUniv is a binary variable indicating the states that are home to premier AI-focused research universities.

We hypothesize that Phase II projects with larger R&D budgets—that is, projects with more available research capital—are more likely to pursue research that is broader in scope than projects; AI-focused research is an area of inquiry that is broad in scope (Giczy et al., 2022; Squicciarini and Nachtigall, 2021). Thus, we hypothesize a positive relationship between the Phase II award amount and the probability that a funded Phase II project will be AI-focused.

We do not hypothesize the directional effect of employment (Emp) on the probability that a Phase II SBIR project will be focused on AI. On the one hand, firms with a larger level of employment have greater human capital research resources. Thus, they may be more likely to

pursue AI research because the likelihood of success is greater.Footnote24 On the other hand, AI research is risky because of the conceptual and changing nature of a nascent technology. Thus, smaller firms, that are generally more risk taking (i.e., less risk averse), are more likely to pursue AI research as a strategy to develop a first-entrant advantage into a market (Leyden and Link, 2004).

We account for an agglomeration effect on the likelihood that a Phase II SBIR-funded project will be focused on AI. Building on the research related to geographic proximity of research clusters and innovative activity (e.g., Audretsch 1998), we control for the proximity of AI-focused research clusters through the variable AIUniv. This variable equals 1 if the firm that received the Phase II SBIR award is in a state in the United States with a top-10 AI research university, according to the current ranking by U.S. News and World Report.Footnote25, Footnote26 We hypothesize that firms in such states will have a greater likelihood of receiving a Phase II SBIR award that is focused on AI because of the quality of the Phase II proposal as well as the promise of greater research success due to the potential knowledge spillovers from the juxtaposed AI-focused research university.

Finally, vector X includes an indicator for the gender of the owner of the firm that received the Phase II SBIR award, WomanOwned. Regarding gender, Link and Morrison's (2019) literature review reports selected evidence of research-related and commercialization-related differences between woman-owned and men-owned firms.

Table 4. Definition of Variables in Eq. (1) and Descriptive Statistics (standard deviations in parenthesis)

Variable	Definition	Mean
AnyAIproject	= 1 if the firm received a Phase II SBIR award for an AI project as defined by the project's abstract having at least 1 defining keyword; 0 otherwise	0.046
Award	Amount of the Phase II SBIR award, deflated by the GDP deflator (2020 = 100) *	1,075,779 (775,877)
Emp	Number of employees in the firm at the time of the SBIR award	83.3 (144.7)
AIUniv	= 1 if the firm that received the Phase II SBIR award is located in a state with a top-10 university in AI research; 0 otherwise.	0.492
WomanOwned	= 1 if the firm that received the Phase II SBIR award was owned by a woman; 0 otherwise	0.121
D2016	= 1 if the Phase II SBIR award was in FY2016; 0 otherwise	0.134
D2017	= 1 if the Phase II SBIR award was in FY2017; 0 otherwise	0.181
D2018	= 1 if the Phase II SBIR award was in FY2018; 0 otherwise	0.140
D2019	= 1 if the Phase II SBIR award was in FY2019; 0 otherwise	0.223
D2020	= 1 if the Phase II SBIR award was in FU2020; 0 otherwise	0.179
Year	(FY of the Phase II SBIR award – 2015)	2.703
Post-2018	= 1 for FY2019 and FY2020; 0 otherwise	0.402
MLproject	= 1 if the firm received a Phase II SBIR award for an AI project as defined by the project's abstract having the phrase machine learning; 0 otherwise	0.033
AIproject	= 1 if the firm received a Phase II SBIR award for an AI project as defined by the project's abstract having the word or phrase AI/ artificial intelligence; 0 otherwise	0.019

* See Jankowski (1993)

Table 5. Probit Estimates of Alternative Specifications of Eq. (1) (n = 5,459)

Independent Variable	**Dependent Variable: AnyAIProject**			
	(1)	(2)	(3)	(4)
LnAward2020	0.1674***	0.1652***	0.1675***	0.1448**
	(0.0083)	(0.0091)	(0.0081)	(0.0198)
	[0.0146]	[0.0144]	[0.0148]	[0.0129]
LnEmp	-0.0304	-0.0307	-0.0351*	-0.0412**
	(0.1535)	(0.1502)	(0.0964)	(0.0499)
	[-0.0027]	[-0.0027]	[-0.0031]	[-0.0037]
AIUniv	0.1330**	0.1335**	0.1336**	0.1337**
	(0.0338)	(0.0332)	(0.0320)	(0.0311)
	[0.0116]	[0.0117]	[0.0118]	[0.0119]
WomanOwned	--	0.0720	--	--
		(0.4461)		
		[0.0066]		
D2016	0.2379	0.2376	--	--
	(0.1670)	(0.1677)		
	[0.0241]	[0.0241]		
D2017	0.2325	0.2321	--	--
	(0.1559)	(0.1565)		
	[0.0233]	[0.0233]		
D2018	0.2749	0.2759	--	--
	(0.1055)	(0.1042)		
	[0.0285]	[0.0286]		
D2019	0.7735***	0.7749***	--	--
	(0.0000)	(0.0000)		
	[0.0936]	[0.0938]		
D2020	1.2073***	1.2084***	--	--
	(0.0000)	(0.0000)		
	[0.1825]	[0.1827]		
Year	--	--	0.2667***	--
			(0.0000)	
			[0.0236]	
Post-2018	--	--	--	0.7944***
				(0.0000)
				[0.0734]
Log pseudo-likelihood	-911.52	-911.24	-921.68	-928.92
Wald χ 2	186.00*** (0.0000)	186.02*** (0.0000)	116.32*** (0.0000)	152.00*** (0.0000)
Pseudo-R 2	0.0993	0.0996	0.0893	0.0821

Notes: p-values, in parentheses, are calculated using robust standard errors; marginal effects are given in brackets

*** significant at 0.01-level, ** significant at 0.05-level, * significant at 0.10-level

In addition, the United Nations Educational, Scientific and Cultural Organization (UNESCO, 2020) pointed out the gender bias associated with the adoption of AI systems.Footnote27 The inclusion of WomanOwned as a control variable represents an effort to explore [our emphasis] gender differences in the receipt of Phase II SBIR project awards that focus on AI research.Footnote28

Also, in vector X is a variable to account for the rapidly accelerating growth in AI-focused research since FY2018 (see Table 2).

Each of the independent variables used to estimate Eq. (1) is defined in Table 4, and descriptive statistics are presented therein. Data on the independent variables, with the exception of AIUniv, come from Phase II SBIR award abstracts. The variables Award and Emp enter Eq. (1) as logarithms to account for possible non-linearity.

The Probit estimates from alternative specifications of Eq. (1) are presented in Table 5. The estimates in columns (1) and (2) control for FY fixed effects.

From Table 5, the estimates suggest that projects with larger Phase II research budgets are more likely to be associated with AI-focused Phase II projects. Estimates of this relationship are robust across the specifications in columns (1) and (2). An increase of 10% in the award amount is estimated to increase the probability of a funded Phase II projection being AI-focused by 0.14 to 0.15 (rounded) percentage points.

The results in columns (1) and (2) offer weak evidence of a statistical relationship between firm size and the likelihood that a Phase II SBIR project will be AI-focused ($p = .15$). There is no evidence that this likelihood differs by gender of ownership—see the estimated Probit coefficient on WomanOwned in column (2). There is, however, empirical support for the agglomeration effect hypothesis: proximity to a top-10 AI research university is associated with an increase of about 1 percentage point in the probability of a Phase II project being AI focused.

The specification for the estimates in column (3) characterizes the growth in AI-focused research by a trend variable, Year rather than by fiscal year fixed effects, to account for the apparent maturing of AI technology over time. The Probit estimates also show the significance of the award amount and the top-10 university research variables, and some evidence of a statistical relationship between firm size and the likelihood that a Phase II SBIR project will be AI-focused ($p = .09$).

5 Conclusions

In this paper, we have identified factors that predict an AI focus of a DOD-funded Phase II project. Our descriptive analysis is based on a unique database that, to the best of our knowledge, represents the first systematic effort to collect information on AI research projects among firms funded through public sector R&D. Our findings provide new, descriptive information about publicly funded AI projects, but also suggest additional policy-focused research, as noted below.

Applications of AI have become increasingly prevalent in recent years. Our analysis shows, however, that even after controlling for such a trend in multiple ways, some interesting patterns remain. Projects that received larger award amounts were more likely to have a focus on AI. We find suggestive evidence that firm size matters; smaller firms have a greater likelihood of pursuing AI-focused research. Also, we find evidence for an agglomeration or knowledge spillover effect; firms that are located in states with top-10 AI research universities are more likely to have an AI focus in their Phase II funded SBIR projects.

We recognize several limitations of this study, and these limitations underscore the exploratory nature of our analysis. First, our findings should be interpreted cautiously with respect

to generalizations about the funding patterns at other agencies, or about public-sector support of AI research more generally. While our motivation for studying DOD Phase II SBIR awards was that DOD's SBIR program is the largest such programs, other agencies' SBIR programs certainly merit study.

Second, our definition of an AI project, while based on OECD protocols, is not time invariant. The underlying technology and defining keywords change over time.

Third, our definition of an AI-focused project is based only on the published abstract for the Phase II SBIR project. As such, we are unable to control for the technology area of the AI research or the component of DOD that issued the solicitation. Award amounts might vary by the scope of research required across technology areas and solicitation areas. Also, firm size might vary across technology areas and solicitation areas and not within them.Footnote30 And, a project abstract does not reveal the role of direct collaborative research partners, such as universities.

Fourth, our agglomeration variable is exploratory in nature, and future research—perhaps case study research—should investigate knowledge spillovers from universities at a more disaggregated level.

Finally, there are likely between-firm differences in the extent of detail and the preciseness with which an abstract is prepared. Thus, an abstract is possibly an imperfect indicator of a project's research focus.

Our study represents, to the best of our knowledge, a first description of public support of AI research projects in small firms in the United States.Footnote31 It illustrates the feasibility of a search methodology applicable to other studies, and further investigation is clearly warranted. For example, the methodology herein can be mirrored to examine Phase II projects in the other 10 agencies that have SBIR programs. Additional research is also needed on public-sector and private-sector R&D investment patterns in small and large firms, keeping in mind that the large firms in our sample are in part constrained to have 500 or fewer employees. Given the nascent nature of AI research at the firm level, case studies might begin to overcome some of the limitations of our descriptive analysis that are noted above.Footnote32 Finally, additional research is needed on STTR (Small Business Technology Transfer) program awards, not only for comparative purposes but also for a more complete understanding of an agency's AI focus.

Lastly, we had no information about the economic outputs associated with AI-focused versus non-AI-focused Phase II SBIR projects. A policy-relevant next question to ask is: What are the social returns to AI-focused and non-AI-focused research supported by the public sector? This is an important question from both a public accountability perspective as well as from the perspective of anticipating the social implications of the development of new technology.Footnote33

Notes

1. The National Artificial Intelligence Initiative Act of 2020 defines the term artificial intelligence (AI) as follows: "The term 'artificial intelligence' means a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. Artificial intelligence systems use machine and human-based inputs to—(A) perceive real and virtual environments; (B) abstract such perceptions into models through analysis in an automated manner; and (C) use model inference to formulate options for information or action." See, https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210.

2. See https://www.oecd.ai/.

3. See also the research by the United States Patent and Trademark Office (USPTO) at: https://www.uspto.gov/ip-policy/economic-research/research-datasets/artificial-intelligence-patent-dataset.

4. Liu et al. (2021) provide an excellent review of AI studies in various other disciplines.

5. The history of the U.S. Department of Defense traces to the U.S. Constitution (1787) which empowered Congress to provide for the common defense, to declare war, to provide and maintain a navy, and to call forth the militia to suppress insurrections and repel invasions. Under the direction of President George Washington, Congress created the Department of War in 1789, the precursor to the Department of Defense. A detailed history of the development of the Department of Defense is at https://www.downsizinggovernment.org/defense/timeline.

6. Yamashita et al., (2021) reported public-sector investments in AI for a number of countries, and their report includes information on the percent of funded research projects by the U.S. National Institutes of Health and the National Science Foundation that are focused on AI.

7. To borrow from the title of Dernis et al., (2021), we are interested in who develops AI-related innovations.

8. Purpose statement (3) has been modified and it now (in 2021) reads: "to foster and encourage participation in innovation and entrepreneurship by women and socially or economically disadvantaged persons." See https://www.sbir.gov/about.

9. The agencies are the Environmental Protection Agency (EPA), National Aeronautics and Space Administration (NASA), National Science Foundation (NSF), and the Departments of Agriculture (USDA), Commerce (DOC), Defense (DOD), Education (ED), Energy (DOE), Health and Human Services (HHS), Transportation (DOT), and most recently the Department of Homeland Security (DHS).

10. To be eligible for an SBIR award, the small business must be: independently owned and operated; other than the dominant firm in the field in which it is proposing to carry out SBIR projects; organized and operated for profit; the employer of 500 or fewer employees, including employees of subsidiaries and affiliates; the primary source of employment for the project's principal investigator at the time of award and during the period when the research is conducted; and at least 51% owned by U.S. citizens or lawfully admitted permanent resident aliens. However, there are agency exceptions when the number of employees can be greater than 500. The National Research Council (2014) does not discuss this point; it simply reports the percent of awards to firms with 100 or more employees. See also Gallo (2021) on eligibility.

11. Several states give matching grants to firms that receive Phase I SBIR awards. See Hardin et al. (2020).

12. As referenced in Gallo (2021, p. 5), the 2019 Small Business Innovation Research (SBIR) Program Policy Directive "provides agencies with the authority to issue an award that exceeds this amount by as much as 50% … [A]gencies may request a waiver from the SBA [Small Business Administration] to exceed the Phase II award [and Phase I award] guideline by more than 50% for a specific topic. In general, the period of performance for Phase II awards is not to exceed two years, though agencies may allow for a longer performance period for a particular project. Agencies may make a sequential Phase II award to continue the work of an initial Phase II award. The amount of a sequential Phase II

award is subject to the same Phase II award guideline and agencies' authority to exceed the guideline by up to 50%. Thus, agencies may award up to $3 million, adjusted for inflation, in Phase II awards for a particular project to a single recipient at the agency's discretion, and potentially more if the agency requests and receives a waiver from the SBA." We thank Patrick Delehanty, Director of Economic Research at the U.S. Small Business Administration for confirming an additional source for these Phase II funding guidelines. He reported that the average award amount, across all agencies that funded Phase II projects, was about $1.68 million in FY 2019.

13. We thank Mariagrazia Squicciarini, former Senior Economist, Head of Unit at the OECD Directorate for Science Technology and Innovation (now at the UNESCO) for reminding us that more intense studies of the abilities associated with machine learning and AI began in the early 2010s. Prior to that time, computers had difficulty distinguishing among images, and a classic example of such is the visual distinction between a chocolate chip muffin and a Chihuahua: https://www.freecodecamp.org/news/chihuahua-or-muffin-my-search-for-the-best-computer-vision-api-cbda4d6b425d/. With this as background, and with some researcher judgement on our part, a decision was made to begin the database using information from FY 2015.

14. These abstracts are publicly available at https://www.sbir.gov/sbirsearch/award/all.

15. We thank Mariagrazia Squicciarini for personal correspondences on OECD's protocols. We realize that other researchers, who we noted in Section I, have used alternative protocols for defining a firm as being involved in AI research. At this point in the evolution of the study of AI activity alternative approaches are valuable, especially if the findings are similar.

16. As discussed in Squicciarini and Nachtigall (2021), relevant keyword search terms changed in the early 2000s.

17. The National Artificial Intelligence Initiative Act of 2020 defines the term machine learning (ML) as follows: "The term 'machine learning' means an application of artificial intelligence that is characterized by providing systems the ability to automatically learn and improve on the basis of data or experience, without being explicitly.

18. programmed." See, https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210.

19. Python is a programming language.

20. Data mining refers to the process of identifying patterns in large datasets.

21. With regard to footnotes 11 and 13 above, projects with more than 500 employees and projects funded at more than $1 M are retained in our database.

22. Relatedly, Abbas et al., (2019) document changes over time in the applications of AI in cyber security, and thus perhaps in the relevance of particular keywords to describe AI research over time.

23. We obtained this estimate from a linear regression of the logarithm of the award amount on an indicator for AI focus, the number of employees and a full set of fiscal year fixed effects.

24. See WIPO (2020, p. 207) in which it was argued that "human capital … is an indispensable condition for innovation to occur." Among SBIR projects that yield a new technology, success is defined in terms of whether the project was commercialized (Link and Ruhm, 2009; Link and Scott, 2010,

2012), and an innovation is commonly defined as a technology that is brought to market (Link and Cunningham, 2021).

25. These states are (alphabetically) California, Georgia, Illinois, Massachusetts Michigan, New York, Pennsylvania, Texas, and Washington. See, https://www.usnews.com/best-graduate-schools/top-science-schools/artificial-intelligence-rankings.

26. For a recent overview of the economic arguments supporting the clustering of knowledge bases, see Rosenthal and Strange (2020) and Moretti (2021). Relatedly, see the recent collection of papers on proximity relations in Torre and Gallaud (2022).

27. UNESCO (2020, p. 4) points out that AI "risks having a negative impact on women's economic empowerment and labour market opportunities by leading to job automation." UNESCO also emphasized the need to (p. 39): "Strengthen gender diversity in AI research both in academia and the private sector."

28. As noted in footnote 9 above, a purpose of the SBIR program is to support women-owned firms, but one should not infer that the purpose statement necessarily applies to woman-owned firms pursuing AI R&D.

29. To examine the robustness of the findings from this specification of Eq. (1), we considered two alternative dependent variables: MLProject and AIProject as defined in Table 4. The estimated Probit coefficient on the award amount remains statistically significant at conventional levels; firm size is at best weakly related to the likelihood that a Phase II project is AI-focused when AIProject is the dependent variable, and the agglomeration effect is also significant only when AIProject is the dependent variable. These results are available on request from the authors.

30. Even if we were able to control for such fixed effects, the statistical consequences would show up in the intercept term in our regressions and not the slope term on LnAward2020.

31. The existing research on AI that focuses on patents as a way of identifying AI-active businesses is important, but the patent approach does not necessarily capture the AI activities of small, often entrepreneurial AI firms. We believe that our methodology is relevant to future studies of AI activities in such firms.

32. And, reflecting on the scholarship of Zhang et al., (2019), case studies might represent an excellent opportunity to introduce to application of altmetric indices to a longitudinal analysis of U.S. firm data.

33. As Chowdhury et al., (2022, p. 6), pointed out: "In FY 2020, the National Science Foundation established seven National AI Research Institutes … Seven initial institutes were established in four of the states with a top 10 AI research university: California, Illinois, Massachusetts, and Texas. In FY 2021, 11 more new institutes were established and now each state with a top 10 AI research university has a National Science Foundation (NSF) institute. This NSF policy response to the anticipated social benefits associated with the development of AI technology and its diffusion brings about the need for future research related to the estimation of the net social benefits associated with the use of supporting public research resources. Certainly, the net social benefits associated with the implementation of AI technology are varied not only over time but also over affected networks. Perhaps future NSF and/or Congressional initiatives to support AI research will include the establishment of centers to estimate such net social benefits through multiple case studies and to codify associated data."

References

Abbas, N. N., Ahmed, T., Shah, S. H. U., Omar, M., & Park, H.W. (2019). Investigating the Applications of Artificial Intelligence in Cyber Security. Scientometrics, 121, 1189–1211

Alderucci, D., Branstetter, L., Hovy, E., Runge, A., & Zolas, N. (2020). "Quantifying the Impact of AI on Productivity and Labor Demand: Evidence from U.S. Census Microdata," Carnegie Mellon University working paper, Pittsburgh, PA

Alekseeva, L., Azar, J., Gine, M., Samila, S., & Taska, B. (2021). The Demand for AI Skills in the Labor Market. Labour Economics. DOI: https://doi.org/10.1016/j.labeco.2021.102002

Alekseeva, L., Gine, M., Samila, S., & Taska, B. (2020). AI Adoption and Firm Performance: Management versus IT. SSRN Electronic Journal. DOI: https://doi.org/10.2139/ssrn.3677237

Audretsch, D. B. (1998). Agglomeration and the Location of Innovative Activity. Oxford Review of Economic Policy, 14, 18–29

Babina, T., Fedyk, A., He, A. X., & Hodson, J. (2021). Artificial Intelligence, Firm Growth, and Industry Concentration. SSRN Electronic Journal. DOI: https://doi.org/10.2139/ssrn.3651052

Baruffaldi, S., van Beuzekom, B., Dernis, H., Harhoff, D., Rao, N., Rosenfeld, D., & Squicciarini, M. (2020). "Identifying and Measuring Developments in Artificial Intelligence: Making the Impossible Possible," OECD Science, Technology and Industry Working Paper 2020/05, Paris

Chowdhury, F., Link, A. N., & van Hasselt, M. (2022). The Spatial Distribution of Public Support for AI Research. Science and Public Policy. DOI: https://doi.org/10.1093/scipol/scac008

Damioli, G., Van Roy, V., & Vertesy, D. (2021). The Impact of Artificial Intelligence on Labor Productivity. Eurasian Business Review, 11, 1–25

Dernis, H., Moussiegt, L., Nawa, D., & Squicciarini, M. (2021). "Who Develops AI-related Innovations, Goods and Services? A Firm Level Analysis," OECD Science, Technology and Industry Policy Papers Number 121, Paris

Gallo, M. E. (2021). "Small Business Research Programs: SBIR and STTR," Congressional Research Service report R43695, Washington, DC

Giczy, A. V., Pairolero, N. A., & Toole, A. A. (2022). Identifying Artificial Intelligence (AI) Invention: A Novel AI Patent Dataset. Journal of Technology Transfer, 47, 476–505

Hardin, J. W., Kaiser, D. J., & Link, A. N. (2020). Public Support of Private Innovation: An Initial Assessment of the North Carolina SBIR/STTR Phase I Matching Funds Program. Annals of Science and Technology Policy, 4, 1–81

Jankowski, J. E. Jr. (1993). Do We Need a Price Index for Industrial R&D? Research Policy, 22, 195–205

Lee, S., Hwang, J., & Cho, E. (2021). Comparing Technology Convergence of Artificial Intelligence on the Industrial Sectors: Two–way Approaches on Network Analysis and Clustering Analysis. Scientometrics. DOI: https://doi.org/10.1007/s11192-021-04170-z

Leyden, D. P., & Link, A. N. (2004). Transmission of Risk-Averse Behavior in Small Firms. Small Business Economics, 23, 255–259

Leyden, D. P., & Link, A. N. (2015). Public Sector Entrepreneurship: U.S. Technology and Innovation Policy. New York: Oxford University Press

Link, A. N., & Cunningham, J. A. (2021). Advanced Introduction: Technology Policy. Northampton, MA: Edward Elgar

Link, A. N., & Morrison, L. T. R. (2019). Innovative Activity in Minority-Owned and Women-Owned Business: Evidence from the U.S. Small Business Innovation Research Program. New York: Springer

Link, A. N., & Ruhm, C. J. (2009). Bringing Science to Market: Commercializing from NIH SBIR Awards. Economics of Innovation and New Technology, 18, 381–402

Link, A. N., & Scott, J. T. (2010). Government as Entrepreneur: Evaluating the Commercialization Success of SBIR Projects. Research Policy, 39, 589–601

Link, A. N., & Scott, J. T. (2012). Employment Growth from Public Support of Innovation in Small Firms. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research

Liu, N., Shapira, P., & Yue, X. (2021). Tracking Developments in Artificial Intelligence Research: Constructing and Applying a New Search Strategy. Scientometrics, 126, 3153–3192

Moretti, E. (2021). The Effect of High-Tech Clusters on the Productivity of Top Inventors. American Economic Review, 111, 3328–3375

National Research Council. (2014). SBIR at the Department of Defense. Washington, DC: The National Academies Press

Nakazato, S., & Squicciarini, M. (2021). "Artificial Intelligence Companies, Goods and Services: A Trademark-based Analysis," OECD Science, Technology and Industry Working Paper 2021/06, Paris

Rosenthal, S. S., & Strange, W. C. (2020). How Close Is Close? The Spatial Reach of Agglomeration Economies. Journal of Economic Perspectives, 34, 27–49

Samek, L., Squicciarini, M., & Cammeraat, E. (2021). "The Human Capital Behind AI: Jobs and Skills Demand from Online Job Postings," OECD Science, Technology and Industry Policy Papers Number 120, Paris

Squicciarini, M., & Nachtigall, H. (2021). "Demand for AI Skills in Jobs: Evidence from Online Job Postings," OECD Science, Technology and Industry Working Paper 2021/03, Paris

Torre, A., & Gallaud, D. (2022). Handbook of Proximity Relations. Northampton, MA: Edward Elgar

United Nations Educational, Scientific and Cultural Organization. (UNESCO, 2020). Artificial Intelligence and Gender Equality: Key Findings of UNESCO's Global Dialogue. UNESCO: Paris

West, D. M., & Allen, J. R. (2018). "How Artificial Intelligence is Transforming the World," Brookings Institution Report (https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/)

WIPO (World Intellectual Property Organization). (2020). Global Innovation Index 2020: Who Will Finance Innovation?. Geneva: World Intellectual Property Organization

Yamashita, I., Murakami, A., Cairns, S., & Galindo-Rueda, F. (2021). “Measuring the AI Content of Government-Funded R&D Projects: A Proof of Concept for the OECD Fundstat Initiative,” OECD Science, Technology and Industry Working Paper 2021/09, Paris

Zhang, X., Wang, X., Zhao, H., de Pablos, D. O., Sun, Y. & Xiong, H. (2019). “An Effectiveness Analysis of Altmetrics Indices for Different Levels of Artificial Intelligence Publications,” Scientometrics, 119: 1311–1344