Government-Backed Science and Venture Capital*

Roham Rezaei†
University of Amsterdam

Yufeng Yao‡
UNSW Business School

October 27, 2026

Abstract

We study whether government investment in basic science stimulates private investment in entrepreneurship. Our difference-in-differences setting examines the impact of the BRAIN Initiative (BI)—a coordinated federal investment program to map the human brain—on venture capital activity. We find that the BI increased both the probability and size of VC funding in neurotechnology ventures, with stronger effects for startups directly linked to BI. Neurotech startups also changed fundamentally: they were more likely to hire academics, file patents, and integrate complementary technologies like AI. Our findings suggest that government-backed science, as a public good, alleviates funding frictions in emerging technologies.

JEL Classification: G24, G18, O3, J24, L26.
Keywords: venture capital, innovation, academic entrepreneurship, government policy, machine learning.

*We would like to thank Ashish Arora, Tania Babina (discussant), Miyoung Chun, Silvia Dalla Fontana, Michael Ewens, Minmo Gahng (discussant), Suting Hong (discussant), Alexandre Jeanneret, Henny Jung (discussant), Mayank Kumar (discussant), Leo Liu, Song Ma, Ron Masulis, Ramana Nanda, Terry Pan (discussant), Peter Pham, Gordon Phillips, Kelly Posenau (discussant), David Robinson, Rik Sen, Janghoon Shon, Tereza Tykvova (discussant), Yizhou Xiao (discussant), Ting Xu, Jason Zein, seminar participants at HEC Montreal, Nanyang Technological University, National University of Singapore, University of Alberta, University of Amsterdam, and University of Georgia, conference participants at ABFER 2025, ACFOW, AFA 2025, CICF 2025, Columbia PE Conference, EUROFIDAI-ESSEC Paris December Finance Meeting 2024, FOM 2024, Global PhD Student Colloquium 2025, Monash Winter Finance Conference 2025, SFS Cavalcade North America 2025, and WEFI Fellows for helpful comments and suggestions. An earlier version of this paper was circulated under the title: *Venture Capital Response to Government-funded Basic Science.*
†Email: r.rezaei@uva.nl
‡Email: yufeng.yao1@unsw.edu.au

1 Introduction

Technical innovation relies fundamentally on investment in basic science. Seminal contributions from Nelson (1959) and Arrow (1962) argue that private markets underinvest in basic science due to its inherent uncertainty and the difficulty of fully appropriating its returns. For instance, while basic scientific research can generate positive externalities that benefit society as a whole, public disclosure of such research allows others to build upon it at low marginal cost. This prevents the initial funder from fully capturing the resulting value. Moreover, the cumulative and sequential nature of scientific progress, where each breakthrough builds on prior discoveries, leads to coordination frictions. Firms and researchers may delay investment, waiting for others to create enabling knowledge, datasets, or standards (Scotchmer, 1991; Bresnahan and Trajtenberg, 1995). As a result, many socially valuable projects may languish.

One view holds that non-market mechanisms, such as government funding of university research, offer an effective solution to this funding gap (Kerr, Nanda, and Rhodes-Kropf, 2014). By supplying public goods (e.g., new knowledge, open data and human capital), government investment in science alleviates the coordination and appropriability frictions that otherwise limit the financing of science-based innovation. In doing so, it expands the opportunity set of ideas that entrepreneurs can commercialize. A contrasting view questions whether this public good provision, which underpins the mission of U.S. science agencies (Bush, 1945; NSF, 2025), is necessary. The U.S. federal government has sharply cut science funding amid claims that markets alone can finance "disruptive" innovation (Garisto, 2025).

We contribute to this debate by examining the causal effect of government-backed science on venture capital (VC) activity. This is important because venture capital is a central engine of private-sector innovation (Kortum and Lerner, 2000). U.S. VCs, in particular, have successfully channeled resources toward entrepreneurial innovations and accelerated the diffusion of novel technologies (e.g., semiconductors, software and biotechnology) and are widely regarded as a model for the rest of the world. Yet it remains an open question how much this success rests on a foundation of publicly funded science. If VCs can independently identify and finance breakthrough opportunities, then government research spending may matter little for high-tech entrepreneurship. But if government-backed science lowers entry barriers and supplies the public goods for commercialization, then its enabling role for venture capital is more fundamental than often recognized.

To study the government's role in supporting entrepreneurial innovation, we turn to "big science" programs, which are large-scale, mission-driven research programs that pursue ambitious goals such as landing on the moon or developing nuclear weapons (Howell, 2024; Britannica, 2025).[1] These programs are often cited as settings where government can simultaneously address both the funding shortfall and the coordination frictions that hinder scientific progress (Mazzucato, 2021). Big science is an ideal setting for our study because it creates large and tractable spillovers: it mobilizes resources at scale, coordinates researchers across disciplines, and establishes shared infrastructures such as open datasets, standardized protocols, and training pipelines.

Yet, how these programs influence the financing of entrepreneurial projects is unclear. Knowledge, talent and data are essential inputs to R&D. By providing these as public goods, big science effectively subsidizes R&D, eases financing frictions, and expands the set of technologies investors can commercialize (Babina, He, Howell, Perlman, and Staudt, 2023b; Gross and Sampat, 2023). Together, these channels can "crowd in" VC. On the other hand, big science may "crowd out" VC if (i) reduced R&D costs lower the demand for dilutive private capital (Wallsten, 2000) or (ii) startups and university labs compete over scarce academic talent, IP and research facilities that are tied to the *big science* mission (Goolsbee, 1998; Lach and Schankerman, 2008; Hellmann, Mulla, and Qiam, 2025). Such programs may even have minimal impact if funds are coordinated inefficiently (Lerner, 2009). We therefore estimate the VC multiplier per dollar of government spending to understand the net effect of big science and identify the channels through which big science induces technological changes.

Our empirical investigation leverages the *Brain Research Through Advancing Innovative Neurotechnologies* (BRAIN) *Initiative*, or BI. Modeled after the Human Genome Project, BI's mission is to map the human brain. This mission, which was proposed by leading neuroscientists, was adopted by President Obama in 2013.[2] The program is funded through 2026 and had already invested more than $4.8 billion by 2022. The National Institutes

[1]The Human Genome Project, which sequenced the human genome and ignited genomics entrepreneurship, is a more recent example. Other examples include the internet, antibiotics, GPS, mRNA vaccines, and microwave radar (Arora, Belenzon, Patacconi, and Suh, 2020; Agarwal, Kim, and Moeen, 2021; Gross and Roche, 2023; Gross and Sampat, 2023).

[2]Through two influential papers, the scientists argued that advancing neuroscience required a large-scale, coordinated push to move beyond single-neuron studies and toward mapping whole-brain activity (Alivisatos, Chun, Church, Greenspan, Roukes, and Yuste, 2012; Alivisatos, Chun, Church, Deisseroth, Donoghue, Greenspan, McEuen, Roukes, Sejnowski, Weiss, and Yuste, 2013)

of Health (NIH) serves as the lead agency for the initiative, with an active working group coordinating its implementation. BI's coordinated, interdisciplinary approach distinguishes it from traditional research grants. This is reflected in our comparison of grant abstracts using two independent large language models; we find that BI projects draw far more heavily on computer science, engineering, and physics compared to traditional neuroscience funding, in line with BI's strategy of generating new tools and methods.

We exploit BI's launch in a difference-in-differences (DiD) design, comparing outcomes for a treated group of neurotechnology startups (those that benefit from the BI) to the control groups. Our sample consists of U.S. startups receiving their first VC funding round between 2000 and 2019, obtained from Pitchbook. We link our VC sample to Revelio Labs (LinkedIn) data for information on startup employees and their career histories. We also add information on the startup's patenting activity from USPTO and define a firm as a *Neuro startup* if it has at least one patent related to neuroscience. We also collect data on grants, output publications and their co-authors.

Venture capital is highly selective. VCs reject roughly 100 startups for every one they fund (Gompers, Gornall, Kaplan, and Strebulaev, 2020). Our first test, therefore, studies whether the likelihood of raising VC increased more for *Neuro startups* than for controls after the BI. This test requires a sample of all firms that may have sought external/VC funding, even if they did not ultimately receive it. To construct this sample, we supplement the Pitchbook data (which comprises mainly VC-funded firms) with the Ewens and Marx (2023) dataset, which covers the entire set of patenting firms in the economy. We find that VCs are significantly more likely to invest in *Neuro startups* post-BI, suggesting that the BRAIN Initiative led to a differential shift in VC allocation toward neurotechnology ventures.

Our DiD design employs a two-tiered control group to address knowledge spillovers. Because BI knowledge may diffuse beyond neuroscience, controls from adjacent fields (e.g., biotech) risk "silent" exposure, biasing estimates toward zero. We therefore use both a *pure* control group (from distant or green technologies) and an *adjacent* one (from healthcare or similar technologies) to bound the effect. Results are statistically and economically significant in both cases but attenuate with *adjacent* controls, consistent with BI impact beyond neuroscience.

The bulk of our analysis conditions on VC-backed firms and studies their financing outcomes. We find that relative to the *adjacent* and *pure* controls, *Neuro startups* receive 32%

to 53% larger VC investments post-BI, respectively. Turning to the efficiency of the program, our back-of-the-envelope calculations based on these estimates suggest that each BI dollar crowds in 1.7 to 3 dollars of VC investments. Such investments are also made at valuations that are 25% to 32% higher.

For a causal interpretation of these results, the parallel trends assumption is critical. While our dynamic estimations show no pre-trends, this assumption would be violated if neuroscience's inherent promise attracted VCs independent of the BI. In contrast, we find that *Neuro startups* with direct links to the BI raised larger investments at higher valuations. We measure BI link along three dimensions: labor: hiring BI-linked researchers; geographic: startup's location relative to BI-linked institutions; and IP: patents leveraging BI-generated knowledge. For the last measure, we fine-tune a SciBERT model to detect BI diffusion and estimate that at least 66% of startups' neurotech patents were influenced by the BI, consistent with its strong commercial impact.

Still, the government did not choose BI randomly. This raises an identification threat whereby unobserved scientific advances coincided with the BI or even prompted the government to join an already booming market. Contemporary reports, however, point in the opposite direction: leading up to the BI, pharmaceutical companies were cutting neuroscience R&D due to high failure rates.[3] We also examine this systematically by leveraging patent valuation data of listed firms developed by Kogan, Papanikolaou, Seru, and Stoffman (2017). The higher frequency of this data helps us isolate market reaction to the BI announcement. We find that *Neuro* patents were valued at a discount before the BI but revalued upward right after. This precise timing makes pre-existing technological breakthroughs an unlikely driver of our results. Lastly, we find no evidence of firms anticipating or lobbying for the BI.[4]

Our results show that the BI increased VC participation on the extensive margin and raised investment scale on the intensive margin. Investment outcomes demonstrate that this was beyond a "VC hype" and that the broader market also saw BI as a positive shock to the NPV of investment opportunities in neurotech: VCs achieved faster and more successful exits (via either IPOs or acquisitions). These patterns imply higher IRRs and shorter R&D

[3]See (Miller, 2010; Insel and Landis, 2013; Nuffield Council on Bioethics, 2013; Choi, Armitage, Brady, Coetzee, Fisher, Hyman, Pande, Paul, Potter, Roin, and Sherer, 2014)

[4]Two more observations support parallel trends. First, BI's scale (over $4.8 billion) dwarfs the investment capacity of typical VC funds. Also, the intrinsic coordination frictions in venture capital do not suit funding a non-excludable good like a comprehensive brain map (Nanda and Rhodes-Kropf, 2018).

timelines, consistent with better alignment to the finite horizons of VC funds.

But does this improved financing merely reflect a demand-side announcement effect, or a supply-side transformation driven by knowledge diffusion? While markets positively reacted to the BI announcement, our evidence supports technological transformation over time. We document three real effects of the BI that support the idea that BI shifted the composition of VC-funded *Neuro startups* toward frontier science: 1) a higher supply of skilled academic labor, 2) more and higher-quality patents, and 3) integration of neurotech with complementary BI-related technologies.

Funds from the BI were predominantly allocated to academic institutions. The labor flow from academia into entrepreneurship is thus a potent measure of its diffusion (Zucker, Darby, and Brewer, 1998). Contrary to the *crowd out* hypothesis, we find that post-BI, *Neuro startups* are 8-11 percentage points likelier to have STEM academics in senior positions three years after being founded – a 60-177% increase relative to the sample means. For VCs, who often lack the technical expertise to assess frontier science, the involvement of academics signals credibility and technological promise. This is consistent with Bernstein, Korteweg, and Laws (2017), who find that human capital is a central criterion for investors at early stages. Post-BI, these startups also hire twice as many academic inventors as controls, indicating an increased supply of skilled labor.

In terms of patenting, *Neuro startups* not only file for more patents, but their patents are more likely to be breakthrough inventions. This patenting profile makes them more attractive to VCs and increases the prospects for better exits (Farre-Mensa, Hedge, and Ljungqvist, 2020; Bowen, Frésard, and Hoberg, 2023). It also contrasts with a *crowd out* hypothesis where IP is entangled in university technology transfer offices.

We also find that BI shifted neurotechnology startups toward data science and AI. Using USPTO's AI Patent Dataset, we find that *Neuro startups* are twice as likely to integrate AI in their inventions, post-BI. Newly founded *Neuro startups* are also more likely to embed complementary technologies such as AI, big data, and smart wearables, consistent with an expanded investment opportunity set. BI's open data facilitated these transformations; to map the brain's 86 billion neurons and 100 trillion connections, the initiative coordinated labs nationwide to share live data and created a public data ecosystem. Like the Human Genome Project's open data policy that spurred genomics market (Williams, 2013), BI's datasets are public goods that lower entry barriers (Ngai, 2024). These datasets can be used

to train machine-learning models, integrate AI into neuroscience applications, and enable devices that interact with the brain directly. However, this raises a reverse causality concern whereby VCs invest in *Neuro startups* because of AI applications rather than the BI. While our sample ends before the current AI boom, we show our results persist even after excluding ventures that leverage AI or big data.

To provide external validity to our results, we examine a similar, but smaller, mission-oriented program: Cancer Moonshot. This program, launched in 2016 with a $1.8 billion budget, aimed to accelerate a decade's worth of progress in cancer research in five years. When we define our treated startups as those in the oncology verticals – the program's primary beneficiaries – we find a similar increase in the relative VC investment size post-2016. We also use this setting as a placebo test for our baseline identification. When we assign 2013 (the BI's launch) as the pseudo-start of a cancer mission, we find no effect on oncology startups. This further supports that our baseline results are driven by the BI rather than broader life-science momentum.

Our study speaks to the limits of venture capital as a "standalone" engine of innovation. Typical VC funds, which are time-limited and small in size, do not have the resources to independently support costly scientific experiments (Lerner and Nanda, 2020). As a result, VC disproportionately flows into areas like information technology, where innovation has become faster and cheaper (Ewens, Nanda, and Rhodes-Kropf, 2018), while allocating relatively little to science-based fields such as Alzheimer's therapeutics or green technologies. This pattern fits a broader literature arguing that innovation-intensive projects often suffer from financing frictions: they carry uncertain payoffs, lack collateral, and exhibit weak appropriability (Nanda and Rhodes-Kropf, 2017; Crouzet, Eberly, Eisfeldt, and Papanikolaou, 2022). These frictions discourage socially optimal levels of experimentation, but when constraints are relaxed—through hot markets, positive cash flow shocks, or public subsidies—innovation financing follows (Howell, 2017; Krieger, Li, and Papanikolaou, 2021; Babina, Bernstein, and Mezzanotti, 2023a). We contribute to this literature by showing that public science agencies such as NIH or NSF play a complementary role in relaxing these frictions.

In related work, Babina et al. (2023b) show that federal research funding spurs academic entrepreneurship. We extend their study in two ways. First, we shift the focus from lab-level outputs to startups and venture capital, an important mechanism for scaling scientific discoveries. Second, by highlighting BI's interdisciplinarity and open data, we demonstrate

that government can also serve as a "coordinator", not merely a funder of scientific R&D.

This connects us to the literature on the focus and design of innovation policies (Lerner, 1999; Bai, Bernstein, Dev, and Lerner, 2021; Bayar, Chemmanur, and Liu, 2024). Akcigit, Hanley, and Serrano-Velarde (2020) propose that the government's funding targeted at basic research is more efficient than applied research. Consistent with this distinction, our VC multiplier estimate of 1.7-3 for the BI suggests more success compared to the 0.4 local multiplier for NASA's "moonshot" (Kantor and Whalley, Forthcoming). The BI's basic science focus – with large spillovers – contrasts with the Apollo Mission – an applied program whose benefits were largely confined to its contractors. Similarly, Fieldhouse and Mertens (2024) find negligible gains from defense programs, whose classified nature limits knowledge diffusion, but substantial returns from non-defense R&D. In defense settings, where appropriability is strong, Howell, Rathje, Van Reenen, and Wong (Forthcoming) show that open, bottom-up competitions outperform centrally specified contracts. Basic science presents the opposite case. The successes of the BRAIN Initiative and the Human Genome Project suggest that when spillovers and coordination frictions are large, centrally directed programs can be especially effective.

2 Institutional Settings: BRAIN Initiative

History and agencies involved. The origin of the BRAIN Initiative traces back to a 2011 meeting called "*Opportunities at the Interface of Neuroscience and Nanoscience.*" This meeting was co-hosted by The Kavli Foundation, the Gatsby Charitable Foundation, and the Allen Institute for Brain Science, which are prominent nonprofit organizations dedicated to funding cutting-edge basic research. They had convened a group of 27 leading scientists to explore ways to spur interdisciplinary research. During a blue-sky discussion, the concept of creating a comprehensive map of brain-circuit activity emerged. This group subsequently published a paper in Neuron – the premier neuroscience journal – detailing their vision (Alivisatos et al., 2012). Up to that point, to understand neural activity, neuroscientists relied on electrodes that sparsely sampled brain activity, capturing signals from one to a few neurons in a specific region. The article argues that understanding brain functions such as thought or perception, which can involve millions of neurons, requires observation at a multi-neuron level. Single-neuron observations are insufficient, akin to understanding an HDTV

program by focusing on just one pixel on the screen. The article suggests that a large-scale effort to map neural circuits, which entails developing new tools and methods, could lead to major scientific breakthroughs:

> *"Emergent-level problems are not unique to neuroscience. Breakthroughs in understanding complex systems in other fields have come from shifting the focus to the emergent level. Examples include statistical mechanics, nonequilibrium thermodynamics, and many-body and quantum physics. Emergent-level analysis has led to rich branches of science describing novel states of matter involving correlated particles, such as magnetism, superconductivity, superfluidity, quantum Hall effects, and macroscopic quantum coherence. In biological sciences, the sequencing of genomes and the ability to simultaneously measure genome-wide expression patterns have enabled emergent models of gene regulation, developmental control, and disease states with enhanced predictive accuracy. We believe similar emergent-level richness is in store for circuit neuroscience. An emergent level of analysis appears to us crucial for understanding brain circuits. Likewise, the pathophysiology of mental illnesses like schizophrenia and autism, which have been resistant to traditional, single-cell level analyses, could potentially be transformed by their consideration as emergent-level pathologies."* (p.973)

These ideas were formally consolidated into an actionable proposal, known as the Brain Activity Map, published in Science by the same team (Alivisatos et al., 2013). Before publication, the scientists shared the proposal with the President's Office of Science and Technology. President Obama picked this initiative as his signature scientific project and announced the BRAIN Initiative in early April 2013, with vast bipartisan support in Congress later on.[5]

Interestingly, five months later, the European Union launched a brain research development program known as the Human Brain Project (HBP). Despite their similar focus, the two projects exhibit distinct characteristics. Theil (2015) and Modic and Feldman (2017) provide a detailed comparison of their backgrounds and differences. The overarching goal of the BI is to map the human brain, while HBP's goal was far more ambitious: simulating the human brain. Many neuroscientists found this unrealistic. BI emerged from a broad scientific consensus, was championed by leading philanthropic science institutions, and was coordinated and executed by the NIH. In contrast, HBP was mainly driven by a few individuals and lacked similar institutional support. Consequently, the BRAIN Initiative quickly

[5]For a detailed history and assessment of the BRAIN Initiative, see accounts by scientists who designed the program (Sejnowski, 2014; Chun, 2014; Newsome and Weiler, 2024). These sources emphasize that the President's announcement came largely as a surprise, not only to the broader neuroscience community but even to the program's architects, indicating that markets could not have anticipated the Initiative in advance.

gained popularity within the neuroscience research community, while HBP became controversial in the EU. In 2014, 750 European researchers signed an open letter to the European Commission criticizing the HBP's overly narrow focus and threatening to boycott the project (Guardian, 2014). Although the HBP continued until 2023, it appears to have had minimal impact on the European neuroscience community (Atlantic, 2019), whereas the BI, which is set to end in 2026, has already been applauded by the neuroscience community (Nature, 2021).

The U.S. government initially announced a BI funding commitment of $4.5 billion spread over 12 years (NIH, 2014). As detailed in Appendix A1, based on agency budget reports, grant data, and BI fact sheets, we estimate that the government has invested over $4.8 billion in basic neuroscience during our sample period. In the three years after our sample period (2023-2025), NIH has invested another $1.4 billion – peaking in 2023 and declining thereafter – in the BI, bringing the total to $6.2 billion. While multiple government agencies, including the National Science Foundation (NSF), the Department of Defence's DARPA, the intelligence agency IARPA, the Department of Energy, and the Food and Drug Administration (FDA) were involved in the initiative, not all were active funders.[6] NIH is the lead agency, accounting for 62% of total funding and overseeing a working group that coordinates and evaluates the program, followed by NSF with 25%.

Corporate Lobbying. As with other policy shocks, an endogeneity concern is the political economy of the policy adoption. If markets influenced policy adoption, the policy would become endogenous to the outcome we study. We consider this issue in Appendix A2 by examining lobbying data from Kim (2018) and do not find heightened corporate lobbying before BI's adoption (see Appendix Figure IA.2). Also, notes from scientists who established the program,[7] our interviews with them, and review of press coverage and public filings from 2011–2013 reach the same conclusion: the main promoters were academics and non-profits, not private markets. This suggests that the launch of the Initiative was orthogonal to private markets' lobbying pressures.

[6]For example, FDA provides non-financial support by enhancing the transparency and predictability of the regulatory landscape for neurological medical devices. Also, DoE has contributed $9m to the themes of neuroimaging and computational frameworks.

[7]See e.g., Sejnowski (2014); Chun (2014)

2.1 Scientific Impact of the BRAIN Initiative Science

Strategic Position in Neuroscience. We collect information on NIH and NSF BI grants from agency websites and through FOIA requests. We describe the data collection procedure in Appendix A3. Outside the BI, NIH also funds neuroscience through traditional, non-coordinated grants ("non-BI" hereafter). To understand BI's role against a similar benchmark, we compare NIH BI grants with non-BI grants. On average, BI grants received $1.10 million per project, 2.3 times larger than non-BI grants. Yet, our lower bound estimate is that as a whole, BI accounts for 8% of total NIH neuroscience funding.[8] Even so, directing $5 billion to one scientific bottleneck can be consequential. To put this in perspective, the Human Genome Project, which spawned the genomics market, cost about $3 billion. As with other big science initiatives, BI's importance is less in budgetary proportion than in catalytic impact. Two features of BI highlight this distinction: interdisciplinarity and public good provision.

Mapping large-scale brain activity required developing new tools and methods beyond what neuroscience could offer at the time, so the BI had to draw on other fields (Sejnowski, 2014; Mott, Gordon, and Koroshetz, 2018). To study whether BI-funded projects reflects this interdisciplinarity, we examine how the scientific focus of BI grants differs from non-BI grants. We classify grant abstracts into subject areas defined by Scopus, a widely used standardized taxonomy for bibliometric analysis, using two independent machine-learning models: GPT-4o mini (generative AI), and SciBERT (supervised multi-label classifier trained on scientific text). Each model has its advantages; generative AI mitigates performance issues related to SciBERT's limited training data, while SciBERT provides replicability and quantifiable performance metrics absent in generative AI. Appendix A4 offers a detailed discussion of these methodologies.

Tables 1 and Appendix Table IA.3 present Scopus subject areas identified through these methods. Both tables indicate that BI grants are more interdisciplinary compared to non-BI grants. Table 1 shows that neuroscience is a key focus for both. Nevertheless, BI projects incorporate Computer Science 4×, Engineering 3.4×, Mathematics 3.2× and Physics 4.5× more often than non-BI grants. In contrast, non-BI grants focus more on conventional medical research, such as Medicine, Immunology and Microbiology, and Health Professions.

[8] This is a lower bound because we account for NIH total neuroscience funding by capturing grants with even minor neuroscience elements, many of which are costly clinical trials.

The SciBERT classification, reported in Appendix Table IA.3 and discussed in Appendix A4, also support similar ratios.

Inspired by the Human Genome Project (Hood and Rowen, 2013), the BI has also spurred the creation of critical research infrastructure as public goods. Its foundational proposal in Science emphasized that mapping the brain would require new methods for storing, managing, and sharing large-scale imaging and physiology data, as well as tools for analyzing and modeling neuronal circuits, with open access for all investigators (Alivisatos et al., 2013). In line with this vision, the BI has developed new tools for capturing brain activity and platforms for disseminating these data openly.[9] Moreover, the BI's open-data policy requires awardees to deposit their outputs in designated archives, further promoting knowledge spillovers (NIH, 2019).

To examine BI's emphasis on data infrastructure, we compare the prevalence of data-oriented terms in NIH BI and non-BI neuroscience funding. We classify a grant as data-oriented if its project terms include any of the following keywords: {*data science, machine learning, artificial intelligence, data set, data sharing, large datasets, large-scale data, deep learning, software, algorithm, open source, Python*}. By this measure, BI grants are three times more likely to be data-oriented compared to non-BI: 46.19% of grants in BI compared to 15.01% in non-BI.

Commercial Impact of the BI. BI was a basic science program. For it to be a relevant shock for venture investors, its knowledge diffusion should advance the commercialization of neurotech. We test this by comparing the patent influence of BI-linked papers with those funded under NIH's non-BI grants. We capture BI's commercial impact along three measures. The first one measures direct patent-paper citations, and its data is sourced from Marx and Fuegi (2020, 2022). We find that BI publications receive roughly four times as many patent citations. The second measure is the predicted *commercial potential* score developed by Masclans, Hasan, and Cohen (Forthcoming). BI publications score significantly higher on this ex-ante commercial potential, too. We elaborate on these results in Appendix B.

However, citation trails can miss spillovers (Myers and Lanahan, 2022). This is espe-

[9] A prominent example is the Cell Census Network, which maps diverse cell types in human, monkey, and mouse brains, enabling targeted research on neurological disorders. An editorial in Nature (2021) describes this project as a major advance in linking brain structure to function.

cially true for BI as a basic science program where appropriability is difficult. Patents tend to cite prior art immediately related to the invention, rather than the broader foundations. Another issue arises from the truncation issue in patent citations, which is particularly relevant because the BI began in the second half of our sample. A growing literature, including USPTO's economists, recommends using machine learning to overcome these limitations (Toole, Pairolero, Forman, and Giczy, 2020; Lerner and Seru, 2021; Masclans et al., Forthcoming). We follow this approach and in Appendix B fine-tune a SciBERT model to identify patents influenced by BI publications. The model estimates that BI-related knowledge has influenced at least 66% of all neuroscience-related patents of startups.

3 Data, Sample and Identification Strategy

3.1 Startups and Innovation

On the extensive margin we ask: *what is the likelihood that a newly founded neuro firm ever raises venture capital?* Answering this requires information on *both* firms that do and do not obtain VC. Pitchbook alone is insufficient for this purpose because it covers only the first set (VC recipients) but largely misses the universe of firms that do not raise institutional capital. To capture the broader universe, we use the patent-assignee dataset of Ewens and Marx (2023), which lists the founding year of 85% of U.S. patenting organizations. Patenting firms have demonstrated technological innovation, and are plausible candidates for VC funding (Farre-Mensa et al., 2020); leveraging this data thus lets us (i) observe startups that never raise external capital, and (ii) classify their technological focus. We focus on firms founded between 2000-2019 and merge them with Pitchbook to identify the year of first financing for those that do obtain VC. We drop firms that have received VC but their *first* patent is subsequent to the first VC round.[10] This leaves us with 56,817 firms, which we later partition into treated and control groups.

The bulk of our study is, however, focused on the intensive margin: *conditional on raising VC, how does neuro firms' financing differ from the rest?* For this purpose, we construct a new sample, which provides richer firm- and deal-level information from Pitchbook. We start with US-headquartered firms that received their *first* VC funding rounds between 2000

[10]This exclusion aligns the timing of innovative activity with potential fundraising, which suits this sample. However, our intensive-margin sample allows patents that are filed after VC investment but before its exit.

and 2019. We track the financing deals of these companies until 2022. We keep VC financing deals where investment size is non-missing and above $100k. We also drop deals where we do not know the number of investors in the round. After applying these filters, we are left with 81,359 VC investments corresponding to 44,413 firms.

Table 2, Panel A reports summary statistics for all VC investments, while Panel B focuses on first rounds only. Most rounds are either first or second round (29,294 are first and 20,452 are second), and less than 10 percent are fifth or larger. The first VC investment has an average round size of $5.4m at a pre-money valuation of $13.7m. When considering all financing rounds, the average round size goes up to $16.4m, at a pre-money valuation of $147.2m. This is unsurprising, as later rounds broadly reflect firm growth and life cycle. The distribution of both round size and pre-money valuations is highly right-skewed. The average number of investors per deal is 4.3 in all rounds and 3.6 in first rounds. 65.0% of startups in our sample are concentrated in just five of Pitchbook's industry groups. These industries are Software, Commercial Services, Pharmaceuticals and Biotechnology, Healthcare Devices and Supplies, and Media, accounting for 37.2%, 10.1%, 7.6%, 5.8%, and 4.2% of the total number of VC-backed startups, respectively.

We also track exit outcomes of VC investments through 2022 to evaluate their success and duration. We follow Ewens, Gorbenko, and Korteweg (2022) and define a *Successful Exit* as one where the startup has either IPOed or been acquired with a reported exit value greater than two times capital invested and zero for smaller.[11] *Successful Exit* is also zero for firms identified as bankrupt or out of business. Following Ewens, Nanda, and Stanton (2024), we classify "zombie firms" as startups that fail to raise capital within three years after their last financing round. For these firms, we impute a failure date between two and five years after the last round using a beta distribution. Our *Successful Exit* indicator is also set to zero for such cases. In Panel C, we report summary statistics at the firm level. Out of the 23,907 companies where we can determine the exit success, 15.7% have *Successful Exits*, which aligns with previous studies. To calculate *Time to Exit* (i.e., investment holding period), we count the number of days between the first VC round and an exit event, e.g., IPO, merger, buyout, bankruptcy or imputed failure. The average holding period for the 33,536 firms where we can calculate this information is 1953 days ($\approx$ 5.3 years).

[11] Like most VC studies, we cannot observe the exact amount returned to the VC to compare it to the amount invested.

Startup Patents. To construct the patent portfolios of startups, we link Pitchbook to USPTO's PatentsView. We employ a two-stage process, following Bernstein, Giroud, and Townsend (2016). First, we run a fuzzy name-matching algorithm developed by Tumarkin (2020) to identify the most likely assignee names corresponding to each startup's legal name. To avoid matching distinct entities with similar names, in the second step, we verify that the firm's state is the same across both datasets. A patent is linked to a startup if there is a match on both name and state. This link also allows us to map Ewens and Marx (2023) to Pitchbook. This is necessary because Ewens and Marx (2023) dataset does not contain the year a firm first received venture capital. We also use the Artificial Intelligence Patent Dataset (AIPD) constructed by Giczy, Pairolero, and Toole (2022) to identify patents containing AI-related components. In Panel D of Table 2, we report the patenting characteristics of VC-backed startups with at least one patent, based on a firm-year panel. These characteristics include the number of patents, citations, breakthrough patents, academic inventors and AI patents. We elaborate on these variables in Section 4.6.2, where we estimate them.

3.2 Treated Group

BI directly targets neuroscience, so we expect startups with neurotech exposure to receive the largest impact. To identify such startups, we rely on patent data and identify those with neuroscience components. This approach has several advantages. First, patents represent innovations that has been legally verified. They grant exclusive legal rights to the firm's intellectual property, and provide time-stamped records of a firm's innovation activities (Hall, Helmers, Rogers, and Sena, 2014). Second, patent data is available beyond VC-backed startups. This allows us to identify the treated group, across both VC-backed and non-VC backed startups, on the same basis. Third, there is no standard classification for neurotech in Pitchbook.[12]

Neuro Patents. We identify neuro-related patents based on a set of keywords, referred to as *Neuro keywords*. These keywords are selected following three steps. First, we collect all Pitchbook keywords that describe the business operations of startups and contain *neuro* or *brain* (or their variations). This results in a list of approximately 250 keywords.[13] Next, we

[12]More recent versions of Pitchbook, unavailable to us, list neurotech as an emerging category.

[13]These are not just one word and could be ngrams. For example: *Alzheimer testing, brainwave technology, neuromuscular disorder, vascularized tissue perfusion, or insurance automation*

ask ChatGPT to rank these keywords based on their relevance to neuroscience and consult with a neuroscientist for verification. Lastly, we manually check and filter out keywords that introduce noise.[14] Our final set of *Neuro keywords* are: {*neuro, nerve, brain, optogenetic, Parkinson, Alzheimer, dementia*}. We identify neuro-related CPC technology groups by searching for *Neuro keywords* in their titles. Technology group is the most granular level of CPC classification, allowing us to capture neurotech and reduce false positives from broad life science categories. In total, we identify 220 neuro-related CPC technology groups.[15] `G06F3/015` is an example with the following title:

> *"Input arrangements for transferring data to be processed into a form capable of being handled by the computer; Output arrangements for transferring data from processing unit to output unit, e.g. interface arrangements-Input arrangements or combined input and output arrangements for interaction between user and computer -Arrangements for interaction with the human body, e.g. for user immersion in virtual reality -Input arrangements based on nervous system activity detection, e.g. brain waves [EEG] detection, electromyograms [EMG] detection, electrodermal response detection"*

A patent can belong to multiple CPC groups. Our goal is to broadly identify technologies where BRAIN-related knowledge could diffuse. Therefore, we refer to patents with at least one neuro-related CPC technology group as *Neuro Patents*.

Neuro Startups. Our baseline definition of a *Neuro startup*, i.e., treated group, is one with at least one *Neuro Patent*. We believe this is a reasonable classification because patenting activity fairly reflects a startup's technological space; startups typically have a narrow technological focus and limited products due to resource constraints. In our main sample, i.e., conditional on raising VC, we exclude firms that file for their first *Neuro Patents* after a VC exit, as these firms are typically larger and tend to hold patents across diverse technologies. This sample contains 753 VC-backed *Neuro Startups*, which form the treated group and collectively hold over 4,000 *Neuro Patents*. Among these startups, 88% operate in the healthcare sector and 8% in IT. This sample includes well-known *Neuro Startups* such as Neuralink, Halo Neuroscience, Paradromics, Neurotrack, and Alector, all of which have received significant media or scientific attention.

[14] For example, the word *neural* could also pick up the AI related term *neural networks*. Therefore, we exclude the term neural.

[15] We use CPC classifications rather than raw patent text because CPC provides standardized terminology, ensuring consistent identification of neuro-related patents.

Our baseline results are robust to two alternative definitions of *Neuro startups*. Under the first definition, which we call *NeuroPreVC*, a *Neuro startup* must file for its first *Neuro patent* in or before the first year it receives VC. The limitation of this definition is that we lose startups with longer R&D timelines or those that reveal their IP through patents later in their life cycle. The number of treated startups under this definition falls to 288. This definition also maps to our extensive margin sample, where we study the probability of raising VC and require firms to have a *Neuro patent* before first VC investment.

The second alternative definition searches for *Neuro keywords* in Pitchbook's business descriptions and keywords instead of patents. Under this definition, we identify 1,088 startups, which we refer to as *NeuroDescript*. However, we prefer the patent-based approach for four reasons. First, it allows a uniform identification across VC-backed and non-VC firms, as the latter is not covered in Pitchbook. Second, patents capture technological innovations that are more likely to be science-based, whereas business descriptions reflect commercial focus.[16] Patent text allows us to measure the firm's technological exposure to the BI, which we cannot do with business descriptions. Third, Pitchbook's keyword tagging relies on proprietary methods and editorial discretion, while patent classifications are validated by examiners using standardized rules. Fourth, a startup's Pitchbook description can change over time, affecting the reproducibility of our results. Also, this may bias our classification if these changes are driven by the hotness of certain markets (e.g., neurotech becoming popular post-BI). In contrast, patents provide a stable and time-stamped record of a firm's technological activity, which offers a more objective and reproducible basis for classification.

3.3 Identification and Control Groups

3.3.1 Conceptual Motivation

We first conceptually motivate our choice of control groups. The DiD estimator captures the incremental effect of the treatment on the treated relative to a control group that would have followed the same path without the treatment. In our setting, neurotech is the treated group with the highest exposure to BI diffusion. The challenge is to choose a set of counterfactual firms that *truly* remain untreated. At first glance, ventures in neighboring technologies (e.g., biotech firms) appear natural. Yet this proximity generates two offsetting

[16] For instance, based on Pitchbook keywords, we pick startups that provide care facilities for people suffering from neurological disorders.

biases. The first one leads to an "underestimation": because biotech ventures are technologically *adjacent*, BI discoveries may diffuse into them in unobservable ways. Under this scenario, if even a subset of the control group quietly gains from BI knowledge, the "control" is partially treated, and the estimated effect is biased toward zero. The second bias works in the opposite direction, leading to "overestimation" for *adjacent* technologies: when capital is scarce, a capital reallocation because of BI also affects the pool of capital available to these *adjacent* ventures.

To present this logic, let $T \in \{0,1\}$ index time ($T=0$ for pre- and $T=1$ for post-BI) and $D \in \{0,1\}$ denote the group assignment in the data ($D=1$ for neurotech, $D=0$ for the chosen controls). With Y_{it} the realised outcome and $Y_{it}(1), Y_{it}(0)$ the potential outcomes with and without BI, the canonical DiD estimator is:

$$\begin{aligned}\widehat{\beta}_{\text{DiD}} = &\big\{\mathbb{E}[Y \mid D{=}1, T{=}1] - \mathbb{E}[Y \mid D{=}1, T{=}0]\big\} \\ &- \big\{\mathbb{E}[Y \mid D{=}0, T{=}1] - \mathbb{E}[Y \mid D{=}0, T{=}0]\big\} \end{aligned} \tag{1}$$

If parallel trends hold and no control firm is exposed – which we call *pure* controls – $\widehat{\beta}_{\text{DiD}}$ is unbiased for the average treatment effect on the treated, $\beta \equiv \mathbb{E}\big[Y(1) - Y(0) \mid D{=}1\big]$.

Underestimation scenario for adjacent controls. Now let $p \in (0,1)$ be the fraction of controls that actually receive BI knowledge (misclassified as $D=0$). Their expected post outcome is:

$$\mathbb{E}\big[Y \mid D{=}0, T{=}1\big] \;=\; (1-p)\,\mathbb{E}\big[Y(0) \mid D{=}0\big] + p\,\mathbb{E}\big[Y(1) \mid D{=}0\big].$$

In Appendix C, we substitute the above into Equation (1) and simplify under parallel trends to derive:

$$\widehat{\beta}_{\text{DiD}} \;=\; (1-p)\,\beta, \qquad \Longrightarrow \qquad \beta \;=\; \frac{\widehat{\beta}_{\text{DiD}}}{1-p}. \tag{2}$$

Because $p > 0$ for *adjacent* controls, the estimate in Equation (1) is mechanically biased toward zero; it equals the actual effect only when $p = 0$. Hence, results using the *adjacent* control group provide a lower bound, whereas estimates against the *pure* controls (where $p \approx 0$) recover an undiluted estimate.

Overestimation scenario for adjacent controls. Suppose instead that the BI crowds in capital for neurotech, which comes at the cost of crowding out capital for the control group. Let $q \in (0, 1)$ denote the fraction of controls that are affected, and let $\delta > 0$ capture the average drop in their outcome variable relative to the level they would have achieved absent any crowd out. Their post-period mean becomes: $\mathbb{E}[Y \mid D{=}0, T{=}1] = \mathbb{E}[Y(0) \mid D{=}0] - q\,\delta$, which implies:

$$\widehat{\beta}_{\text{DiD}} = \beta + q\,\delta, \qquad \beta = \widehat{\beta}_{\text{DiD}} - q\,\delta. \tag{3}$$

Because both q and δ are positive, the term $q\delta$ inflates the naive estimate.[17] Bounding the magnitude of this upward bias follows the same logic as that of the attenuation bias in Equation (2): when $q \approx 0$ (our *pure* controls), the bias disappears, whereas larger q or more severe substitution (δ) push the estimate above the true treatment effect.

SUTVA. The underestimation bias is a case of misclassification and does not breach the Stable Unit Treatment Value Assumption (SUTVA). SUTVA requires no *interference*: one unit's treatment altering another unit's potential outcomes. Under that scenario, the issue is simply that some controls are themselves treated and their own outcomes shift. However, under the overestimation scenario, SUTVA does not hold for our *adjacent* controls because the treatment transmits to the control firms through capital reallocation.

3.3.2 Empirical Construction of the Control Group

In our setting, overestimation is less likely because the post-BI period coincides with the boom years of VC, where capital was abundant. Still, to account for both possibilities, we build two tiers of control groups. A *pure* tier consists of technologies orthogonal to neuroscience, where BI spillovers and capital reallocation are implausible.[18] In this case, BI coefficients are unbiased. The second tier is *adjacent* controls, which can gain or lose from the BI due to the biases above. Our empirical tables report both tiers side-by-side to see which one exists. If the magnitude of the DiD coefficients for *pure* controls is greater, then the underestimation bias exists. However, in the opposite case ($\left|\hat{\beta}_{\text{pure}}\right| < \left|\hat{\beta}_{\text{adjacent}}\right|$) the

[17] Symmetrically, if BI crowds out capital for neurotech, some controls benefit. If a fraction q of control gains $\delta > 0$ then $\widehat{\beta}_{\text{DiD}} = \beta - q\,\delta$. This biases the estimate downward (more negative when $\beta < 0$).

[18] Capital reallocation is unlikely for the *pure group* because VCs typically invest in specific technologies. For example, while a neurotech startup can be a substitute for another healthcare startup, it is unlikely that neurotech draws capital away from green tech ventures (Popp and Newell, 2009).

overestimation is present.

Extensive Margin Control Groups. For the extensive margin sample (identified in Section 3.1) our only relevant information point is the firm's patenting profile. We therefore begin by taking all patenting firms as the unconditional control group. We then partition this group into two mutually exclusive sets: *pure* and *adjacent* controls. To implement this, we classify a patent as *Similar* if at least one of its CPC subclasses (one level above CPC group) also contains a *Neuro* CPC group. Building on the example in Section 3.2, any patent with subclass `G06F` ("Electric Digital Data Processing") sits in the same subclass as the neuro group `G06F3/015`. Such patents are close enough to neuroscience that BI knowledge could plausibly diffuse. This *adjacent* control group, called *SimilarTech*, comprises firms with at least one *Similar* patent; they represent 24,410 firms, or 43% of the sample. All remaining patents are called *Distant* and firms with exclusively *Distant* patents form the *DistantTech* group (the *pure* control). 30,556 startups belong to this group.

In Appendix D, we develop a text-based method to test the validity of these classifications by measuring the textual similarity of *Distant* and *Similar* patents to *Neuro* patents. We find that *Similar* patents (and firms) are significantly closer to *Neuro* patents (and firms) compared to the *Distant* patents (and firms).

We also construct a second *pure* control group: startups patenting in green technologies. Similar to neurotech, green innovation builds on basic science. VCs are also attracted to this area (Dalla Fontana and Nanda, 2023). Yet, as our topic classification in Tables 1 and Appendix Table IA.3 shows, BI knowledge is orthogonal to scientific subjects behind green technologies such as energy, materials science, chemistry and chemical engineering. As such, we follow the methodology of Cohen, Gurun, and Nguyen (2020), which tags green patents based on OECD environment-related technology classes. We find 5,566 firms with at least one green patent and classify them as *GreenTech* startups.

Intensive Margin Control Groups. Once we condition our analysis on firms that have raised VC, we can exploit the richer information set about the startup's business activities in Pitchbook (e.g., their industry or business description). We still follow the same intuition above, but now expand our control groups to non-patenting firms. As a baseline, we first report our results against all 43,660 untreated VC-backed companies in our sample as an unconditional control.

As a first *pure* control group, we exclude 6,978 companies that have a *Similar* patent (as defined above) any time before the VC exit and call the remaining 37,682 *Dissimilar* companies. The *Dissimilar* group is our preferred benchmark because it strikes the right balance between purity and representativeness. By excluding firms with *Similar* patents, we strip out companies that are technologically close to neuroscience or that could plausibly compete for the same specialized resources and capital. This ensures that the control group is not contaminated by potential spillovers of BI knowledge or by substitution effects in VC allocation. At the same time, the group remains large enough to provide meaningful statistical power and to capture broad VC allocation dynamics across sectors.

Our *adjacent* control group comprises the entire healthcare sector.[19] 88% of *Neuro startups* are in the healthcare sector. Also, 46% of the firms with a *Similar* patent are in this sector. Healthcare is the second-largest sector that VCs back after IT. A typical VC-backed company in this sector is R&D-intensive and life-science-oriented, like neurotech. This renders these companies as a reasonable counterfactual for neuotech. However, this sector is not orthogonal to BI's knowledge and may benefit from its diffusion. Our analysis in Section 2.1 support this possibility; BI knowledge can be classified under other areas in life sciences such as medicine, biochemistry, immunology and microbiology. For example, BI's brain imaging technologies may also benefit non-brain-focused or oncology startups. Because we cannot systematically rule out such spillovers, we call this sector an *adjacent* control.

To mirror the extensive margin control groups, our second *pure* control group are green/clean technologies. We label a company as *CleanTech* if its Pitchbook's business description or keywords contain any of: {*renewable, clean energy, climate, battery, environment*}. This yields 2,825 firms. Pitchbook also has two related classifications: the *Energy* sector and its own CleanTech vertical. The statistical and economic significance of our (unreported) results is similar if we use either of these as a control group instead of our text-based group. However, these classifications are subject to more measurement noise. Pitchbook's energy sector also contains startups focused on fossil fuels, which are not necessarily working on clean technologies. The problem with Pitchbook's CleanTech vertical is that a company can belong to multiple verticals, and therefore, we cannot identify the firm's primary focus, which is important in this setting. Note that for the intensive margin control groups, we use *CleanTech* instead of *GreenTech* and *Dissimilar* instead of *DistantTech* to differentiate this

[19] Pitchbook's industry sector classification includes: *Information Technology, Healthcare, B2B, B2C, Energy, Financial Services, Materials, and Resources.*

sample from the extensive margin sets, which are restricted to patenting firms.

3.4 Startup Employees

We collect data on employees of VC-backed startups[20] from Revelio Labs (augmented LinkedIn data), which provides résumés containing names, employment histories, and job titles. Revelio Labs provides a company-level link to Pitchbook, allowing us to identify 80% of all VC-backed startups and 86.48% of all *Neuro startups* in our sample. We link these employee records to two additional datasets: (i) patent inventors from PatentsView to identify academic inventors and (ii) authors of BI-linked publications from NIH and NSF sources.

Matching inventors to employees is nontrivial. We start with inventor names and assignee histories from PatentsView, and use assignee information as a proxy for employment since employees typically assign patents to their firms as the legal owners. We then match inventor names with individual names from Revelio Labs, following the name-matching procedure of Masulis and Yao (2025): first matching on surnames, then on closely similar first names. Lastly, we keep only those where the name match is accompanied by overlapping employment histories. For example, if inventor A has a patent assigned to company XYZ and employee A has worked at XYZ, we classify them as the same individual.

We identify BI-linked scientists by matching startup employees to authors of BI-funded publications. Using each publication's DOI or PMID, we retrieve author information from Scopus, which provide names, fields of expertise, affiliation histories, and publication records. Employees and authors are linked based on name similarity and overlapping employment histories. This procedure yields 60,371 startup employees identified as inventors and 2,983 as co-authors of BI publications.

3.4.1 Academic Startups and Inventors

We classify a startup as an *Academic Startup* if it has at least one senior employee with prior academic experience who joined the firm within three years of its founding. We focus on senior employees because they are more likely to shape a startup's innovation strategy

[20] We cannot link our extensive margin sample (any patenting firm) to Revelio Labs. Given the scale of both datasets and limited identifying information, fuzzy name matching would generate a high rate of false positives.

through their leadership roles. We use Revelio Lab's seven-point scale for seniority and classify those at level 5 or above as senior. An employee is considered *Academic* if they hold a PhD in a STEM field and have worked at an academic institution under job titles such as *professor, graduate, lecturer, academic, researcher, faculty, dean, instructor, scholar, scientist, postdoc, PhD, doctor, fellow, educator*. Academic institutions are identified based on whether the employer's name contains keywords such as *university, institute, or college,* specific abbreviations or names (e.g., *UCLA, MIT, Caltech*), or research-related terms such as *Lab, Research,* and names of landmark institutes (e.g., *Mayo Clinic*). We apply a similar procedure to identify *Academic Inventors*, with two modifications: (i) a PhD is not required, allowing us to include technicians and other non-PhD staff with experience at public research labs, and (ii) we do not impose the three-year post-founding limit, as the purpose of this measure is to capture any academic involvement in the firm's inventive activity, not just early managerial influence.

4 Main Results

4.1 Probability of Raising Venture Capital

Our first hypothesis is that the scientific spillovers from the BI crowds in venture capital. Raising VC is a key milestone for high-growth startups. According to Gompers et al. (2020), VCs reject roughly 100 firms for every one they decide to fund, implying that most startups seeking VC never receive it. This is especially true in emerging fields like neurotech, where VCs expect high costs for technological de-risking. If the BI's public goods have the promise of lowering these entry costs, VCs will be more willing to invest in neurotech. However, if BI diverts resources from commercialization, we expect lower VC participation. Thus, our analysis compares the probability of a *Neuro startup* raising VC before and after the BI relative to the control groups defined in Section 3.3. To test this, we estimate the following panel regression:

$$Y_{it} = \beta_1 \text{Neuro}_i \times \text{Post}_t \ + \beta_2 Ln(Age+1)_{it} + \lambda_i + \theta_t + \upsilon_{it} \tag{4}$$

where Y_{it} equals one if the startup i has ever raised VC up to year t and zero, otherwise. *Neuro* is an indicator for whether the startup is a *Neuro startup* as defined in Section 3.2.

We include year (θ) and firm (λ) fixed effects, which absorbs the *Neuro* and *Post* coefficients. The main coefficient of interest (β_1) is the interaction between *Neuro* and *Post*. *Ln(Age+1)* is the log-transformed difference between year t and the startup's founding year plus one. A startup enters the panel in its founding year. In this panel, we have 17,010 firm-year observations for *Neuro startups*, corresponding to 1,455 firms.

Table 3 reports the results of this estimation. In the first column, the control group is unconditional: any patenting firm founded between 2000 and 2019, following the definitions in Section 3.1. The coefficient of 0.032 in Column (1) implies a 3.2 percentage points rise on a base probability of 4.8 percent–roughly a 67% increase. Columns (2-4) partition this sample, following the definitions in Section 3.3. Here, the controls are *DistantTech, GreenTech*, our *pure* tier, and *SimilarTech*, our *adjacent* tier. When the counterfactual is *DistantTech*, the coefficient grows to 4.1 percentage points, more than doubling the baseline probability: 108% increase. Against *GreenTech* firms, the gain is 2.8 percentage points, about a 49% increase. Finally, Column (4) turns to the adjacent group of *Similar* technologies. Here, the coefficient falls to 1.9 percentage points, which is still significant, but now a 30 percent increase relative to that group's mean probability of 6.4 percent. These results show that post-BI *Neuro* startups become more VC "investable". The systematic shrinkage of the estimates as we move from *pure* to *adjacent* controls supports the spillover (underestimation) story: once the comparison set itself can benefit from BRAIN knowledge, the measured treatment effect is mechanically attenuated, confirming both the presence and the reach of BI diffusion.

Figure 1 plots the dynamic DiD coefficients from the specification in Equation 4. All estimates hover around zero before 2013, indicating parallel pre-treatment trends. After the BRAIN announcement, the coefficients rise steadily, with the steepest climb when *Neuro* firms are compared with the *pure* control groups and a smaller, but still positive, response against the *adjacent* controls. Together with the table evidence, this pattern confirms the lower- and upper-bound logic: VC interest in neurotech strengthens progressively as BRAIN knowledge diffuses, and the gap between control sets bounds the extent of that diffusion.

4.2 Intensive Margin: VC Investments, Financing and Valuation

In the rest of the paper, we condition our analysis on startups that receive VC and study their financing and real outcomes. VCs typically deploy staged financing to manage downside risk. Each successive round is a real option that lets the investor abandon the

project if milestones are missed (Gompers, 1995; Cornelli and Yosha, 2003; Ewens et al., 2018). The value of this abandonment option is particularly high in science-heavy fields, with long R&D timelines and high failure rates. The "crowd in" story argues that BI would reduce information and coordination frictions, thereby lowering entry costs and R&D risk in neurotech. If so, rational VCs should respond to the increased NPV of such investments by writing larger checks earlier, because their abandonment option has become less valuable. A competing "crowd out" view predicts the opposite. BI dollars and public goods can reduce R&D costs and thereby substitute for dilutive VC dollars. They could also lock scientists, IP, and equipment in academia and bid up their acquisition costs. To preserve downside protection, VCs would then reduce check sizes. Lastly, if BI is merely a scientific project without commercial impact, we should expect no differential effect on VC investments in *Neuro startups.*

We test these hypotheses by examining the size and valuation of VC financing rounds for *Neuro startups*. We follow the estimation strategy in Ewens et al. (2018) and Ewens and Farre-Mensa (2020), who study cross-sectional variations in startups' financing events[21] and estimate the following equation at the financing round level:

$$Y_{i,j,t} = \beta_1 \text{Neuro}_i \times \text{Post}_t + \beta_2 \text{Neuro}_i + \beta_3 X_{i,j} + \tau_t + \omega_j + \upsilon_{ijt}, \tag{5}$$

In this regression, each observation is a financing round of startup i, with the sequence number of VC round denoted by j, and the year in which it occurs by t. $X_{i,j}$ is a vector of both round and firm-level characteristics. It includes fixed effects for the firm's headquarters location–because supply and demand of entrepreneurial finance is geographically clustered– and industry fixed effects, to control for 41 different industry groups in Pitchbook. It also controls for *Syndicate Size*, the logged-transformed number of investors in the round, because more investors in the round could capture the riskiness of the investments. ω_j denotes fixed effects for the sequence number of VC round to control for the financing life cycles. The model includes an indicator for whether the firm was a *Neuro startup* (i.e., treated) and a *Post* dummy for whether the financing occurred in or after 2013 (BI's announcement year). τ_t are year fixed effects corresponding to the year of the investment (which subsumes the noninteracted *Post* indicator). The main coefficient of interest (β_1) is the interaction between

[21]The average startup receives 2.5 rounds of financing with long gaps between rounds. Also, 36% of companies only receive one round. These reasons typically make it difficult to run a panel in this setting.

Neuro and *Post*.

In Table 4, we report the results of estimating Equation 5. The outcome variables, investment size (Panel A) and pre-money valuation (Panel B), are log-transformed to account for their skewness. Odd-numbered columns include all financing rounds, with the corresponding round number fixed effects, and the even-numbered columns are restricted to the first rounds. In the first two columns, the control group is all VC-backed startups as our unconditional control group. In the following columns, this group is partitioned based on the three control groups we specified in Section 3.3.

All coefficients are positive and statistically significant at 1% level. The DiD coefficient of 0.359 in Column (1) indicates VC investments in *Neuro startups* saw a $(e^{0.359} - 1) \approx 43\%$ increase relative to all other VC-backed startups, post-BI. In Column (3), where we examine our *Dissimilar* control group, this increase is 51%. This is our preferred specification as the main *pure* group. In Column (5), the increase is 53% relative to *CleanTech*, and relative to *Healthcare* (Column 7), it is 32%. Like our earlier extensive margin results, this decline when moving from *pure* to *adjacent* groups is consistent with BI diffusing into neighboring technologies and contrasts with BI diverting capital away from the control group.

We also conduct sub-sample analyses restricted to the first VC investment for two reasons. Uncertainty is greatest prior to the initial financing, when VCs must evaluate the startup without the benefit of insider information. Once the first investment occurs, the VC becomes actively involved and acquires private information about the firm's prospects, which can serve as a certification mechanism in subsequent rounds. Moreover, the initial investment is more likely to be based on the promise of the underlying technology rather than on demonstrated market traction. At this stage, uncertainty is tilted toward scientific and technological feasibility, precisely where the BI is most salient. The results for these first-round financings, reported in the even-numbered columns, are substantially larger. Investment sizes increase by 100% against all VC-backed startups, 109% against *Dissimilar*, 115% against *CleanTech*, and 75% against *Healthcare*. These magnitudes indicate that the BI's effect is strongest at entry, when capital is scarce but consequential, and highlight BI's role in easing financing frictions and enabling VCs to commit more resources upfront.

Nonetheless, these results do not directly translate into higher NPV. It could be that due to R&D or operational changes, *Neuro startups* have larger capital requirements. Panel B examines pre-money valuations to explore this possibility directly. The results are similar to

those for investment size: valuations for *Neuro* startups rise by 25-32% across all rounds and by 55-62% in first rounds post-2013. We use pre-money valuations because these valuations reflect the company's value right before the new cash is injected and therefore are not mechanically driven by the bigger checks documented above. The number of observations drops relative to Panel A, reflecting rounds with missing post-money valuations in Pitchbook.

To investigate the existence of a pre-trend in size and valuation, in Figure 2 we plot the dynamic estimations of Equation 5 for *pure* and *adjacent* control groups. None of the three series shows evidence of prior trends. Pre-BI, the coefficients are statistically insignificant at any conventional confidence level. This indicates that *Neuro startups* and their various control groups were on parallel trends for both investment size and valuation. The event-study coefficients also echo the patterns we reported in the table analysis. As the spillover (underestimation) story predicts, the coefficients are generally smallest when *adjacent* healthcare startups are the control group and are most pronounced for *pure groups*; when the counterfactual is technologically distant. Appendix Figure IA.3, which plots an event study plot for the first financing, also yields similar results.

Figure 2 also shows an earlier reaction in size and valuation post-2013, compared to Figure 1 (on the probability of raising VC). Together, this sequencing fits an initial announcement-driven revaluation strengthened over time with BI diffusion. First, the venture capital pool for "existing" neuro ventures deepens. Then, as BI knowledge, tools and talent flow begin seeding "new" ventures, the capital pool expands on the extensive margin. The results and figures in the following sections further support this narrative.

BI's Multiplier. We conduct a back-of-the-envelope calculation to gauge the "incremental VC dollars" invested in neurotech due to the BI. Dividing this increment by the government's total BI spending yields the policy multiplier for private VC investment. The calculation follows directly from the DiD estimates reported in Table 4. The DiD coefficient captures the causal change we expect for the treated group relative to its counterfactual path. Under the standard DiD identifying assumption (parallel trends), the counterfactual post-treatment mean for the treated group ($\bar{Y}^{\text{CF}}_{T,\text{post}}$) equals its pre-treatment mean ($\bar{Y}_{T,\text{pre}}$) plus the change observed in the control group ($\bar{Y}_{C,\text{post}} - \bar{Y}_{C,\text{pre}}$), where $\bar{Y}$ denotes the average VC investment size. Benchmarking against the *Dissimilar* group is our preferred estimation because it offers the broadest group of VC-backed firms, which are immune from BI spillovers. We therefore

focus on the sample in Column (3) for illustration. Below, we replace the average investment size values (all in $\$m$) for these groups and obtain a counterfactual average investment size of:

$$\begin{aligned}\bar{Y}^{\text{CF}}_{T,\text{post}} &= \bar{Y}_{T,\text{pre}} + (\bar{Y}_{C,\text{post}} - \bar{Y}_{C,\text{pre}}) \\ &= 16.20 + (15.67 - 8.54) = \$23.33\text{m}.\end{aligned}$$

The DiD's coefficient estimates that BI has led to a 51% *incremental* increase in this counterfactual. Thus, the DiD's mean effect per deal ($\Delta\bar{Y}_{\text{avg}}$) is $51\% \times \$23.33m = \$11.9m$. There are 1,219 post and treated deals ($N_{T,\text{post}}$), and thus the aggregate incremental VC dollars is: $\Delta VC_{\text{agg}} = 11.9 \times N_{T,\text{post}} = \$14,51$ *billion*. Since its announcement in 2013 and until 2022, we estimate that the U.S. government has spent $4.8 billion on the BI and therefore:

$$\text{Multiplier} = \frac{\Delta VC_{\text{agg}}}{\text{BI}} = \frac{14.51}{4.80} \approx 3.0.$$

We repeat this exercise for Column (7), where the control group is *Healthcare* and the coefficient is smallest to get a lower bound. The ΔVC_{agg} in that sample is $8.1b and therefore the multiplier becomes 1.7. As we argued before, this particular multiplier is an underestimation because BI likely also diffuses into control startups of the *adjacent* technologies.

Our multiplier estimates of $1.7-3 per BI dollar are significantly larger than the 0.40 multiplier on the Apollo Program estimated by (Kantor and Whalley, Forthcoming). The disparity can be explained by differences in (i) how the multipliers are measured and (ii) the programs' nature. Our multiplier assesses how basic science grants spur nationwide venture capital. The Apollo multiplier, however, measures how total mission spending, including R&D and hardware procurement, translates into local income for contractors. The two programs also have different natures. The BI is an open, basic science platform designed to generate non-rival knowledge and create broad spillovers. The Apollo program was an applied, defense-oriented mission, where the benefits were largely confined to its contractors. Recent macroeconomic evidence reinforces this conclusion. Fieldhouse and Mertens (2024) find social returns of 140-210 percent for non-defense federal R&D but negligible gains from defense programs, where secrecy inherently limits knowledge diffusion.

Robustness: Alternative Classification of Neuro startups. In Section 3.2, we discussed the advantages of our baseline definition of *Neuro startups*. Here, we test the sensitivity of our main result, investment size, against it. One possibility is that relying solely

on patents may cause us to overlook firms in neurotech that never file for a patent. Pitchbook's business descriptions and keywords give us an alternative perspective on the startup's business. We therefore search for the same set of *Neuro keywords* in Pitchbook's business descriptions or keywords and flag a startup as *NeuroDescript* if it has at least one of these keywords. A different concern with our baseline definition is look-ahead bias. We may classify a startup as *Neuro* too early because we require the firm to have a *Neuro patent* before the VC exit. Therefore, we introduce a more restrictive definition; *NeuroPreVC* is one if the startup files its first *Neuro* patent by the time of the "first" VC investment and zero otherwise.

Appendix Table IA.7 reports the results against both groups. To reduce misclassifications in the control groups, we exclude *Neuro startups* (original definition) that are untreated under these new definitions from the estimation. In Panel A, the treated group are *NeuroDescript* startups and in Panel B, these are *NeuroPreVC* startups. All coefficients are positive and significant in both panels, except for Column (7), the *adjacent* Healthcare sector, which is not significant at conventional levels. The point estimates under Panel A, however, are smaller than those in Table 4, consistent with measurement noise attenuating the signal in the treated set.

Robustness: AI Adoption. As we show later in Section 4.6.3, neurotech startups adopted AI and machine learning more frequently post-BI. This raises a reverse causality concern. While our sample period does not cover the current AI boom, advances in AI and ML have attracted VC attention in the last decade. Thus, an alternative explanation for our results is that improved VC financing of *Neuro startups* is because neuroscience is a fertile ground for AI applications. Under this scenario, our results would be primarily driven by *Neuro startups* in AI and big data verticals. To address this concern, we re-estimate our baseline results after excluding such startups. Appendix Table IA.8 shows that even without these startups, the relative increase in investment sizes for *Neuro startups* remains significantly larger than for the control groups (which themselves include many AI startups). The results also hold in Panel B, where we further exclude *Neuro startups* in other *ComplementaryTech* verticals (*DigitalHealth, HealthTech*, and *Wearables*) to align with the classification we later introduce in Section 4.6.3. These findings demonstrate that our results are not simply due to neuroscience's overlap with AI, but reflect a broader shift in VC financing induced by the

BI.

4.3 Heterogeneity in Exposure to the BRAIN Initiative

Our baseline results show that, after the BI's launch, *Neuro startups* attracted significantly more VC investment relative to control groups. To directly link this effect to the BI, we exploit the heterogeneity among *Neuro startups* in their exposure to the BI. We measure this exposure through three separate proxies: 1) labor exposure as observed through hiring BI-linked scientists, 2) IP exposure as observed through patents that use BI-generated knowledge and tools, 3) geographical exposure via proximity to BI-linked institutions. If the BI truly drives the surge in VC interest, then among *Neuro startups*, those with stronger BI links should exhibit superior financing outcomes. We cannot implement a complete DiD model, since proximity to the program is only meaningful after it starts. Instead, we examine cross-sectional variation in post-BI financing rounds and estimate the following regression model for *Neuro startups*:

$$Y_{i,j,t} = \beta_1 \times \text{BIExposure}_i + \beta_2 X_i + \tau_t + \omega_j + \upsilon_{ijt}, \tag{6}$$

where the outcome variable is either the log-transformed size or the pre-money valuation of the investment. *BIExposure* is defined using one of three proxies (*Emplyee Exposure*, *Patent Exposure* or *State Exposure*). The rest of the notation follows Equation 5.

Our first exposure proxy indicates if the startup has hired a BI scientist. For each financing round, we construct an indicator variable called *Employee Exposure*, which is equal to one if the startup has hired at least one BI scientist by that year. We identify these scientists following the process detailed in Section 3.4 and only include startups that can be matched to Revelio Labs. Among 435 such *Neuro startups* (1,087 financing events), 44 startups (86 financing events) have hired at least one BI scientist. Table 5 (Columns 1 and 4) shows that these BI-linked rounds were, on average, 71% larger ($e^{0.539} - 1$) and had 120% higher pre-money valuations.

Our second proxy, *Patent Exposure*, uses the large language model described in Section 2.1 to identify patents influenced by the BI. For every year, we count the number of startups' patents influenced by the BI up to that point and match this to the financing event in that

year. In this sample of 620 financing rounds,[22] the average round is associated with 1.22 BI-influenced patents. We then construct *Patent Exposure* as Ln(1+number of influenced patents). The coefficient of 0.285 in Column (2) of Table 5 thus indicates that having one BI-influenced patent (vs none) is associated with round sizes about 22% $\left(e^{0.285\times\ln(2)}-1\right)$ larger. The relationship with valuation, while positive, is not statistically distinguishable from zero.

The third proxy, *State Exposure*, is the count of BI grants awarded to any institution in the startup's home state in a given year, with the idea that greater geographical proximity intensifies BI spillovers. Similar to the previous measures, we log-transform the count of these grants in the state up to a given year and match it to the year of the financing event. The coefficients on *State Exposure* in Columns (3) and (6) imply that going from zero to one BI grant in a state is associated with approximately 26% larger financing rounds and 19% higher valuations. Because the explanatory variable is defined at the state level, we cannot include startup-state fixed effects. However, unreported estimates are even stronger when excluding California, a state with a strong academic and entrepreneurial ecosystem. These results indicate meaningful regional spillovers from BI activity, positively influencing financing outcomes for neurotech startups.

4.4 Identification Checks and External Validity

Public Market Reaction to the BI. A common concern in DiD analyses is that the shock captures omitted factors. In our setting, pre-existing breakthroughs could have independently raised the NPV of neurotech or even prompted the government to launch the BI in an already booming field. Three points address this concern. First, our evidence above supports direct links between BI and financing outcomes. Second, anecdotal evidence runs counter to the "booming field" view: leading up to the BI, large pharmaceutical companies, traditionally major investors in neuroscience R&D, were cutting their expenditures in the field due to high uncertainty and failure rates (Miller, 2010; Insel and Landis, 2013; Nuffield Council on Bioethics, 2013; Choi et al., 2014).

Third, we systematically examine this alternative story by studying the market's reaction in narrow windows around BI's announcement. If the field was already booming, then the BI

[22] To avoid look-ahead bias, we restrict the influence measure to publications available by year t on patents applied for up to year $t+1$, starting in 2015 (see Appendix Section B). This restriction reduces the number of rounds included in the regression.

announcement should have elicited no market response in a short window. Put differently, if such unrelated factors are driving our results, it is unlikely that their timing exactly matches that of the BI announcement. The higher frequency data of public firms allows testing this. We leverage patent valuation data developed by Kogan et al. (2017), which ties innovation announcements to their valuation.[23] We isolate the BI's impact by examining the cross-sectional variation in patent valuations. Brav, Jiang, Ma, and Tian (2018) estimate the effect of hedge fund activism on firm innovation based on the same intuition. At the VC level, this test is infeasible because (i) too few financing events occur in such short windows and (ii) patent values are unavailable for private firms.

We run a test that mimics our baseline test and estimate Equation 5 at the patent level, where the outcome is log-transformed patent value. We thus compare patents filed by the same firm, in the same filing year×technology class, where two key differences are (i) a plausibly exogenous variation in grant date (before or after the BI's announcement)[24] and (ii) being a *NeuroPatent*. This helps us isolate technological waves that might be correlated with the value of innovation. Our analysis includes patents granted within six-month, one-year and two-year windows around the BI announcement (April 2, 2013). We cannot narrow the window further than this because there are too few *NeuroPatnets* granted. We also control for patent citations as a measure of patent quality.

We report the results in Appendix Table IA.9. The coefficient on *NeuroPatent* is negative and significant in most columns. This indicates a lower valuation of neurotech innovation relative to the control group in the pre-period, consistent with anecdotal evidence and industry reports. However, the positive and significant coefficients on *NeuroPatent*×*Post* indicate that markets positively updated their expectations on neurotech investments only after the BI. The coefficients grow larger as the window around the shock expands. The relative increase in a six-month window is 4%, growing to 25% in a two-year window, consistent with an immediate announcement effect that is reinforced as BI-driven knowledge transforms the technology. Moreover, similar to the stronger effects we find in VC financing, the coefficients in Columns (4-6) – when the control group are pure from BI spillovers – are even larger than those in Columns (1-3). Together, these results show that BI's impact was beyond venture

[23]Note that we cannot simply examine the stock price reactions of pharmaceutical companies, because they often pursue R&D in multiple areas, with neuroscience representing only one.

[24]The average (median) lag between a patent's filing date and grant date in this sample is 1296 (1187) days. It is implausible that firms strategically file certain patents three to four years before the BI, in anticipation of its announcement.

capital and extended to public market valuations, too.

Cancer Moonshot. In 2016, the US government launched the *Cancer Moonshot*, a mission-oriented program aimed at achieving "*a decade's worth of progress in cancer prevention, diagnosis, and treatment in five years*" (National Cancer Institute, 2025). Its appropriation (approximately $1.8 billion) is smaller than the BI's funding and it is more focused on application and implementation than BI's tool-building and basic science focus. Still *Cancer Moonshot* shares key design features with the BI: an inter-agency working group to coordinate goals and an emphasis on open data to accelerate discovery. Like Alzheimer's, Parkinson's, and dementia, cancer represents a large market with unmet medical need, and cancer R&D is similarly rooted in academic science. These similarities make it a natural setting for two complementary empirical exercises.

First, as an external validity test, we ask whether venture capital responded to the Cancer Moonshot like it did to the BRAIN shock. We identify *Cancer startups* as those classified under Pitchbook's *Oncology* vertical. This vertical, typically a subset of healthcare, comprises startups focused on developing cancer treatment. We also introduce *Post16* to capture financing events in 2016 and later. Panel A of Table 6 shows that *Cancer* startups saw a statistically and economically significant increase in their VC financing. Across all VC rounds, the *Cancer*×*Post16* coefficient implies round sizes were 43-60% larger. These magnitudes are of the same order as the 32–53% increase we estimated for neurotech. Similarly, first rounds show larger increases of 71-89%. These results suggest that the crowd-in effect of mission-oriented science programs on VC is not unique to neurotech.

We also use the *Cancer Moonshot* to implement a timing placebo that assigns 2013 (BI's launch) as the pseudo-start of a cancer mission. The federal government chose cancer in 2016, right after the BI, so it is plausible that it could have selected cancer instead of the brain in 2013. If our neurotech estimates merely reflect a momentum in life-science investing, or the federal government opportunistically "joining the bandwagon" of already booming areas, oncology startups should exhibit an increase in VC funding post-2013, absent a government-led cancer-specific mission. To test this, Panel B restricts the sample to 2010-2015, re-labels 2013 as the pseudo-*Post* year, and re-runs the specification. To ensure that BI spillovers are not driving the results, we exclude *Cancer* startups that are also *SimilarTech* to *Neuro* startups. The *Cancer*×*Post* coefficients are small (e.g., 0.016 in Column 1) and statistically

indistinguishable from zero across every control group and round type. This means that oncology ventures did not experience the same increase in investment scale in 2013, despite being technologically and scientifically comparable to neurotech. This null result strengthens our identification strategy: the large VC response we document for neurotech reflects the BI, not the broader health-care trends or pre-existing momentum that the government opportunistically entered.

4.5 VC Exits

The surge in VC funding we document for *Neuro startups* may simply reflect a "VC hype". Thus, it is important to see if later-stage sources of capital also validate our baseline effect. A venture fund's main objective is to exit its investments successfully. Typically, these exits materialize through an IPO or sell-outs, or if things go poorly, a write-off. If the BI improves the NPV in neurotech, we should see a higher probability of successful exits and/or a faster realization of those exits. Since trade sales have dominated venture exits over the past two decades, we begin by comparing acquisition timing. Figure 3 plots the annual share of acquisitions for *Neuro startups* against *Dissimilar* startups in the healthcare control. Both series rise after the BI, but the increase for neurotech is markedly sharper, suggesting that strategic buyers also responded to the BI by stepping up neurotech acquisitions.

However, unsuccessful sell-outs (those with low returns) can mask underperformance (Puri and Zarutskie, 2012). We therefore follow the process detailed in Section 3.1 and define a *Successful Exit* as one where the startup has either IPOed or been acquired with a reported exit value greater than two times capital invested and zero for smaller. We then estimate the probability of a *Successful Exit* using our DiD specification (Equation 5) at the startup level, where *Post* equals one if the firm's first VC round occurred in 2013 or later. We control for the total capital the startup has raised prior to exit to compare startups of similar scale; otherwise, better-funded firms might exit more successfully (or take longer to exit) simply because they have had more resources or runway. We also include exit year fixed effects to account for time-varying market conditions (e.g. deal-making cycles, valuations) that affect exit probabilities and timing.

Table 7, Columns (1)–(4) report the results. Across specifications, the *Neuro*×*Post* coefficient is positive. In the broadest sample (Column 1), the point estimate of 0.155 implies that, relative to the pre-BI period, *Neuro startups* are 15.5 percentage points more likely

to deliver a successful outcome: about 100% increase relative to the 15.7 percent baseline success rate. The effect remains significant when we compare neurotech only with technologically *Dissimilar* firms (Column 2, 11.7 pp) and with other healthcare ventures (Column 4, 8.7 pp). Using *CleanTech* as the counterfactual, the coefficient is still positive, albeit not statistically significant.

We also use the investment holding, i.e., time to exit, as an alternative measure of VC investment success. The holding period is important for two reasons. First, faster exits boost a VC fund's IRR, a key fund performance metric. Second, if VCs avoid science-heavy ventures due to their long R&D timelines, reducing this timeline should make such startups more attractive investments.

We test this hypothesis and report the results in Columns (5)-(8) of Table 7. The dependent variable is the natural log of the number of days from the first VC investment to the exit date, so negative coefficients indicate faster exits. We cannot simultaneously include fixed effects for both the first financing year and the exit year, as this would absorb all variation in the outcome. The estimates range from –0.177 to –0.294, corresponding to holding periods that are roughly 16%-26% shorter. As detailed in Section 3, we impute exit dates for some of the *failed* companies – those with no further financing three years after their last VC round. In unreported analysis, we drop these firms entirely, and our results remain robust. Overall, the evidence suggests VCs could exit these firms faster, which improves their attractiveness to VCs with finite fund lives.

4.6 Real Effects: Technological Shifts

Our results so far show that the BI turns *Neuro startups* into more attractive investment opportunities. We now turn to their real outcomes beyond financing to uncover whether BI diffusion drove a supply-side technological transformation, or whether the observed financing patterns merely reflect a demand-side announcement effect. We provide evidence consistent with the former by documenting three shifts that moved neurotech ventures closer to the scientific frontier: (i) a larger pool of specialized human capital, (ii) more and higher-quality innovation outputs, and (iii) deeper integration with complementary BI-linked technologies.

4.6.1 Academic Startups

The BI channels the bulk of its resources into university laboratories, so a natural commercial spillover is the movement of academic researchers into startups[25] (Babina et al., 2023b). In Section 4.3, we showed that *Neuro* startups hiring Bi-linked scientists raise significantly more capital. However, this diffusion channel need not be confined to the BI-linked academics. Once academic findings are published or circulated, they diffuse through the broader scientific community. As Zador, Escola, Richards, et al. (2023) and Arora, Belenzon, Cioaca, Sheer, and Zhang (2023) document, academic scientists are a crucial input for commercializing frontier research. From the investor perspective, human capital is also a key early-stage quality signal (Bernstein et al., 2017). Academics bring deep technical expertise that enhances a startup's credibility, and their salaried positions in research institutions imply a high opportunity cost of joining risky ventures (Hall and Woodward, 2010). When such individuals assume senior roles in startups, it signals a strong conviction in the underlying technology. As such, one channel for BI's knowledge diffusion can be an increase in the supply of academic human capital available to neurotech startups. Conversely, if the BI had no real technological impact and our results merely reflect an announcement effect, we would not expect greater academic involvement in *Neuro startups*.

We test this by estimating Equation 5, where the outcome variable is *Academic Startup*: an indicator variable with a value of one if the startup has hired an academic in a senior position within three years of its founding.[26] We focus on the founding year because we want to reflect the real effect of academics' presence in the startup's formative years, rather than the financial effect of VC money enabling startup to hire skilled labor. To reflect this timing, our *Post* variable is also constructed at the founding year level, as opposed to the year of the financing event.

Table 8 reports the results. Columns (1) show that, post-BI, *Neuro startups* are 10.6 percentage points more likely to be *Academic Startups* relative to all other VC-backed firms, which is a 141% increase relative to the mean of this sample. The 0.110 coefficient in Column (2), where the control groups are *Dissimilar* firms, shows a 177% increase relative to the mean. The 8.2 percentage points in Column (3) (against *CleanTech*) translates into

[25]For instance, 10% of NIH's BI funds were explicitly allocated to training skilled labor (e.g., postdoctoral researchers). Beyond these direct training grants, BI research grants can indirectly expand labor supply by enabling recipients to hire more researchers.

[26]See more on this classification in Section 3.4.1

a 60% increase relative to the mean. Column (4) yields a smaller, statistically insignificant 4 percentage points estimate versus *Healthcare* (a 23% increase), consistent with spillovers diffusing across *adjacent* biomedical sectors. As science-heavy ventures, *Neuro startups* were already more likely to involve academics prior to the BI. Yet the substantial post-BI increases indicate that the BI further amplified academic participation.

Figure 4, which plots dynamic DiD estimates for the Column (2)–(4) specifications, shows no pre-trend. More importantly, it exhibits a gradual increase in the likelihood that *Neuro Startups* are also *Academic Startups* following the BI. The trend becomes more pronounced in later years, suggesting that as the BI generated more knowledge, more academics chose to engage in entrepreneurship, facilitating the commercialization of neuroscience.

4.6.2 Innovation and Academic Inventors

As we showed in Section 2.1, BI-linked research has successfully advanced the frontiers of neuroscience. What remains unclear is how this scientific progress has translated into commercial innovations by *Neuro startups*. To examine this link, we analyze their patenting outcomes by studying the number of patents, citations, breakthrough patents, academic inventors, and AI-related patents. Table 2 reports summary statistics for these variables at the firm-year level. Unlike our earlier analyses of funding and exit outcomes, where most firms record only one or two events in total, patenting occurs often enough that we can observe firms repeatedly across years. This allows us to construct a firm-year panel with sufficient within-firm variation over time. We exploit this panel structure by estimating the same two-way fixed effects model as Equation 4. However, because patent counts are right-skewed and contain many zeros, we follow Cohn, Liu, and Wardlaw (2022) and use a Poisson specification.

We focus on startups with at least one patent as a control group because most VC-backed startups (77% in our sample) never file for patents. We do not further split this small pool into *pure* and *adjacent* controls because that would mostly reduce statistical power.[27]

Table 9 reports the results of this estimation. Odd-numbered columns use the entire firm-year panel from the founding year through VC exit. Even-numbered columns truncate the panel at the firm's first VC round (*Before VC*). This is for robustness because after the

[27]Because we estimate firm and year fixed effects, identification comes from within-firm changes relative to same-year controls and time-invariant, cross-firm differences are absorbed.

firm raises VC, it can access more resources, which may improve its patenting. Columns (1) and (2) show that *Neuro startups* produce more patents. The coefficient of the interaction in Column (1) is 0.576 and statistically significant at 1%, suggesting that *Neuro startups* produce 1.78 ($e^{0.576}$) times more patents relative to other *non-Neuro startups* after the BI.

A larger number of patents may not mean higher-quality innovation. One conventional measure of patent quality is patent citations, calculated as raw counts normalized by the average within a patent's technology class and grant year (Hall, Jaffe, and Trajtenberg, 2001). We aggregate this at the firm-year level by averaging adjusted citations per patent. Columns (3) and (4) show that the coefficients are statistically insignificant. This implies that the average patent of a *Neuro startup* is not of higher quality. Nonetheless, firms may build patent fences to protect one core innovation (Ziedonis, 2004). If the patent quality is right-skewed, the average patent may not reflect the firm's innovation quality. To capture this, we examine whether *Neuro startups* file for more *Breakthrough patents*, which are those receiving more citations than the 90th percentile within their technology class and grant year (Byun, Oh, and Xia, 2021). The coefficients in Columns (5) and (6) are positive and significant at the 1% level. This suggests that *Neuro startups* experienced a larger increase in the number of breakthrough patents compared to the control groups, consistent with BI shifting these firms closer to the scientific frontier (Krieger, Schnitzer, and Watzinger, 2024). The economic magnitude is also large; *Neuro startups* produce twice as many breakthrough patents even before VC funding.

To complement our analysis in the previous section on academics in *Neuro startups*, we also examine the background of inventors. Following the procedure detailed in Section 3.4, we link USPTO's inventor data to Revelio Labs to identify inventors with academic backgrounds. Columns (7) and (8) report the results of regressing the number of academic inventors on the interaction of *Neuro* and Post. The coefficients in both columns are positive and significant. Column (7), for example, indicates that post-BI *Neuro startups* hired two times more academic inventors relative to the control group.

The results above contrast with a *crowd out* hypothesis whereby BI's academic origins would trap talent or knowledge in universities.[28] Instead, we observe not only an increase in patent quantity and quality, but also a rise in the number of academic inventors contributing to startup innovation. These real effects also help explain the more favorable financing of

[28]Such frictions are commonplace in the commercialization of academic knowledge and are intertwined with university technology transfer processes (Lach and Schankerman, 2008; Hellmann et al., 2025)

neurotech firms. Patents are successful R&D outcomes and represent a richer portfolio of tangible IP-based assets. Patenting success has been associated with higher prospects for raising venture capital, being acquired or going public (Farre-Mensa et al., 2020; Bowen et al., 2023).

Lastly, we examine whether *Neuro startups* adopt AI more frequently in their inventions. As we elaborate in the next section, this hypothesis is motivated by BI's overarching goal of mapping the human brain, which requires significant support from AI and data science. To proxy for AI adoption, we draw on the AI Patent Dataset developed by USPTO economists; Giczy et al. (2022) have developed a machine learning method that determines whether a patent has used AI-related technologies. We aggregate the count of such patents at the firm-year level as the outcome variable in Columns (9) and (10). The coefficients suggest that *Neuro startups* are about 2.5 times more likely to adopt AI technology post-BI compared to the control group.

To examine the dynamics of these effects, in Figure 5, we plot the event study coefficients for the results above.[29] The trends follow a similar pattern to the prior results. While the difference between the treated and control group in the pre-period seems indistinguishable from zero, the coefficients slowly start to rise as BI knowledge diffuses and improves innovation in neurotech.

4.6.3 Integration with Complementary Technologies

In this section, we examine whether *Neuro startups* evolved in the direction of BI science by adopting BI-related technologies. In Section 2.1, we demonstrated that BI coordinated interdisciplinary projects, pooled neuroscientists with engineers and computer scientists and funded open data infrastructure. The human brain comprises 86 billion neurons, forming over 100 trillion connections. To map this network, BI supported the development of specialized multimodal machine learning algorithms to extract patterns from complex neural signals. Furthermore, BI coordinated multiple labs nationwide to share their live research data publicly and created a data ecosystem to disseminate it. Making such large-scale, complex datasets openly available and developing the analytic tools to use them can lower the barriers to AI adoption in neuroscience. This is similar to how the Human Genome Project's open-data policy spurred downstream innovation (Hood and Rowen, 2013; Williams, 2013).

[29] We do not plot citations. Reflecting Table 9, the yearly coefficients are small and insignificant.

NIH's BI director notes that before the Initiative, big data in neuroscience was the domain of a few well-resourced institutions (Ngai, 2024). But a decade after the program, modules such as the BRAIN Initiative Cell Census Network allow smaller players to ask targeted questions about the behavior of specific cell types in neurological disorders.

Ngai (2024) highlights several other technologies emerging from the BI; e.g., neural recording, neuromodulation therapies, brain-computer interfaces, neuro-inspired computer architectures, and precision circuit mapping. Understanding the neural system's interactions or what functions in the brain are responsible for thoughts and senses enables developing smart wearable devices that can interact with the brain.[30] Also, understanding how the human brain processes enormous data with little energy consumption inform low-power neuromorphic AI designs (Zador et al., 2023).

In Table 9, we showed that *Neuro startups* incorporate AI into their patents more frequently. We now extend this analysis to examine whether BI knowledge pushed the technological frontier more broadly, leading *Neuro startups* to adopt AI and other complementary technologies at higher rates. We use Pitchbook's vertical classifications to define complementary technologies. Verticals are taxonomies, often used in the entrepreneurial setting, that group startups by the technological niche or customer use-case they serve, and can span multiple industries. Pitchbook has 59 verticals, listed in Appendix E. Given the direction of BI's achievement described by the program's director (Ngai, 2024), we construct an indicator variable *ComplementaryTech* that captures startups with any of these verticals: *Artificial Intelligence and Machine Learning, Big Data, Wearables & Quantified Self, HealthTech, Digital Health*.[31]

We then estimate this outcome variable using Equation 5 at the startup level. *Post* is an indicator variable equal to one if the startup was founded after 2013 and zero otherwise. This choice is similar to the analysis of academic startups in Section 4.6.1, which aims to capture real technological change at formation rather than a financial effect.

Table 10 reports the results. In Columns (1)-(4) the outcome variable is *CompletementaryTech*, our broader definition. Columns (1)-(3) show that post-2013, *Neuro startups* are 6.6-8.1 percentage points more likely to adopt these technologies, representing significant

[30] Neuralink is a high-profile company in this regard, founded in 2016. It develops brain–computer interfaces with direct communication between the brain and external devices. Interestingly, Philip Sabes, one of the company's co-founders, has published under the BI.

[31] For example, Pitchbook classifies Neuralink under *HealthTech* and *Artificial Intelligence and Machine Learning* and Neurotrack under *Digital Health* and *HealthTech*.

magnitudes: 41% increase vs. the sample mean of all other VC-backed firms (Column 1), 46% vs. *Dissimilar* controls (Column 2), and 39% vs. *CleanTech* (Column 3). These estimates are all statistically significant. The coefficient becomes insignificant when the control group is the *adjacent* set of Healthcare ventures (Column 4). This may suggest that the increased adoption of broad verticals like *HealthTech* and *Digital Health* may have been a sector-wide trend.

To examine this possibility closely, in Columns (5)-(8), our outcome variable *ComplementaryTech2* is one only for the core technologies of *AI*, *Big Data*, and *Smart Wearables* and zero otherwise. Here, the coefficients slightly increase and, importantly, the effect against other Healthcare startups (Column 8) becomes statistically significant. This indicates that the BI enabled the integration of neurotech with AI and frontier hardware fields, which is distinct from broader trends within the healthcare sector. Lastly, note that as our robustness test in Section 4.2 shows, our baseline financing results are robust to dropping the startups with *ComplementaryTech*. This mitigates reverse causality concerns whereby *Neuro startups* financing can be explained by advances in AI and machine learning.

5 Conclusion

This study examines how strategic government investments in basic science can shape private investments in entrepreneurial innovation. Using a difference-in-differences design around the launch of the Brain Research Through Advancing Innovative Neurotechnologies (BRAIN) Initiative, we find that venture capitalists increased both the likelihood and scale of their investments in neurotechnology startups and placed financing and valuation premiums on the knowledge and labor supplied via the program. These ventures subsequently achieved faster and more profitable exits, indicating that the Initiative generated value beyond the venture capital market. The program also induced real technological changes that expanded the investment opportunity set: it drew more academics into entrepreneurship, spurred higher rates of breakthrough patenting, and, through its data infrastructure, lowered barriers to integrating frontier technologies such as AI.

Our findings suggest that it is no coincidence that many of the world's transformative technologies have emerged in the United States, where the federal government has long been the largest funder of basic science. Public science institutions fulfill a critical public goods

function for the U.S. markets, but their contribution is so foundational that it often goes unrecognized in downstream applications. By funding and coordinating early-stage academic discovery, public science institutions generate technological externalities that contribute to the global leadership of the U.S. high-tech entrepreneurship. Yet, these institutions currently face substantial budget cuts amid doubts about their efficiency. Our study indicates that these cuts may negatively impact the longstanding dynamism and innovation in the U.S. economy.

Our focus on the BRAIN Initiative leaves open the question of how similar initiatives perform in other emerging technological fields, especially those outside life sciences. Future research could compare the optimal design of such interventions across areas that vary in demand uncertainty (e.g., quantum computing or advanced materials). Such studies could inform policies aimed at bridging the gap between basic science and commercialization, ultimately maximizing the economic and societal impact of public investment in innovation.

References

Agarwal, Rajshree, Seojin Kim, and Mahka Moeen, 2021, Leveraging private enterprise: Incubation of new industries to address the public sector's mission-oriented grand challenges, *Strategy Science* 6, 385–411.

Akcigit, Ufuk, Douglas Hanley, and Nicolas Serrano-Velarde, 2020, Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth, *The Review of Economic Studies* 88, 1–43.

Alivisatos, A. Paul, Miyoung Chun, George M. Church, Karl Deisseroth, John P. Donoghue, Ralph J. Greenspan, Paul L. McEuen, Michael L. Roukes, Terrence J. Sejnowski, Paul S. Weiss, and Rafael Yuste, 2013, The brain activity map, *Science* 339, 1284–1285.

Alivisatos, Paul, Miyoung Chun, George M. Church, Ralph J. Greenspan, Michael L. Roukes, and Rafael Yuste, 2012, The brain activity map project and the challenge of functional connectomics, *Neuron* 74, 970–974.

Arora, Ashish, Sharon Belenzon, Larisa C Cioaca, Lia Sheer, and Hansen Zhang, 2023, The Effect of Public Science on Corporate R&D, Working Paper 31899, National Bureau of Economic Research.

Arora, Ashish, Sharon Belenzon, Andrea Patacconi, and Jungkyu Suh, 2020, The changing structure of American innovation: Some cautionary remarks for economic growth, *Innovation Policy and the Economy* 20, 39–93.

Arrow, K. J., 1962, *Economic Welfare and the Allocation of Resources for Invention*, 219–236 (Macmillan Education UK, London).

Atlantic, 2019, The human brain project hasn't lived up to its promise, The Atlantic, Published on 22 July 2019. Retrieved on 1 March 2024.

Babina, Tania, Asaf Bernstein, and Filippo Mezzanotti, 2023a, Financial disruptions and the organization of innovation: Evidence from the great depression, *The Review of Financial Studies* 36, 4271–4317.

Babina, Tania, Alex Xi He, Sabrina T Howell, Elisabeth Ruth Perlman, and Joseph Staudt, 2023b, Cutting the Innovation Engine: How Federal Funding Shocks Affect University Patenting, Entrepreneurship, and Publications, *The Quarterly Journal of Economics* 138, 895–954.

Bai, Jessica, Shai Bernstein, Abhishek Dev, and Josh Lerner, 2021, The dance between government and private investors: Public entrepreneurial finance around the globe, Working Paper 28744, National Bureau of Economic Research.

Bayar, Onur, Thomas J. Chemmanur, and Mark H. Liu, 2024, How to Motivate Fundamental Innovation: Optimal Interactions between Entrepreneurs, Venture Capitalists, and the Government, *Working paper* .

Bernstein, Shai, Xavier Giroud, and Richard R Townsend, 2016, The impact of venture capital monitoring, *The Journal of Finance* 71, 1591–1622.

Bernstein, Shai, Arthur Korteweg, and Kevin Laws, 2017, Attracting early-stage investors: Evidence from a randomized field experiment, *The Journal of Finance* 72, 509–538.

Bowen, Donald E., Laurent Frésard, and Gerard Hoberg, 2023, Rapidly evolving technologies and startup exits, *Management Science* 69, 940–967.

Brav, Alon, Wei Jiang, Song Ma, and Xuan Tian, 2018, How does hedge fund activism reshape corporate innovation?, *Journal of Financial Economics* 130, 237–264.

Bresnahan, Timothy F., and M. Trajtenberg, 1995, General purpose technologies 'engines of growth'?, *Journal of Econometrics* 65, 83–108.

Britannica, 2025, Big science, https://www.britannica.com/science/Big-Science-science, Retrieved 8 Aug 2025.

Bush, Vannevar, 1945, *Science, the Endless Frontier* (Princeton University Press).

Byun, Seong K., Jong-Min Oh, and Han Xia, 2021, Incremental vs. breakthrough innovation: The role of technology spillovers, *Management Science* 67, 1779–1802.

Choi, Dennis W., Robert Armitage, Linda S. Brady, Timothy Coetzee, William Fisher, Steven Hyman, Atul Pande, Steven Paul, William Potter, Benjamin Roin, and Todd Sherer, 2014, Medicines for the mind: Policy-based "pull" incentives for creating breakthrough cns drugs, *Neuron* 84, 554–563.

Chun, Miyoung, 2014, Igniting the brain initiative, *Association for Psychological Science, Presidential Column* .

Cohen, Lauren, Umit G Gurun, and Quoc H Nguyen, 2020, The esg-innovation disconnect: Evidence from green patenting, Working Paper 27990, National Bureau of Economic Research.

Cohn, Jonathan B., Zack Liu, and Malcolm I. Wardlaw, 2022, Count (and count-like) data in finance, *Journal of Financial Economics* 146, 529–551.

Cornelli, Francesca, and Oved Yosha, 2003, Stage financing and the role of convertible securities, *The Review of Economic Studies* 70, 1–32.

Crouzet, Nicolas, Janice C. Eberly, Andrea L. Eisfeldt, and Dimitris Papanikolaou, 2022, The economics of intangible capital, *Journal of Economic Perspectives* 36, 29–52.

Dalla Fontana, Silvia, and Ramana Nanda, 2023, Innovating to Net Zero: Can Venture Capital and Start-Ups Play a Meaningful Role?, *Entrepreneurship and Innovation Policy and the Economy* 2, 79–105.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, 2018, Bert: Pre-training of deep bidirectional transformers for language understanding, *arXiv preprint arXiv:1810.04805* .

Ewens, Michael, and Joan Farre-Mensa, 2020, The deregulation of the private equity markets and the decline in ipos, *The Review of Financial Studies* 33, 5463–5509.

Ewens, Michael, Alexander Gorbenko, and Arthur Korteweg, 2022, Venture capital contracts, *Journal of Financial Economics* 143, 131–158.

Ewens, Michael, and Matt Marx, 2023, Entrepreneurial patents, *Working Paper* .

Ewens, Michael, Ramana Nanda, and Matthew Rhodes-Kropf, 2018, Cost of experimentation and the evolution of venture capital, *Journal of Financial Economics* 128, 422–442.

Ewens, Michael, Ramana Nanda, and Christopher Stanton, 2024, Founder-ceo compensation and selection into venture capital-backed entrepreneurship, *The Journal of Finance* 79, 3361–3405.

Farre-Mensa, Joan, Deepak Hedge, and Alexander Ljungqvist, 2020, What is a patent worth? evidence from the u.s. patent "lottery", *The Journal of Finance* 75, 639–682.

Fieldhouse, Andrew J., and Karel Mertens, 2024, The Returns to Government R&D: Evidence from U.S. Appropriations Shocks, Federal Reserve Bank of Dallas Working Paper No. 2305, updated November 2024.

Garisto, Dan, 2025, Trump moves to slash NSF: why are the proposed budget cuts so big?, *Nature* 618, 12–13.

Giczy, Alexander V., Nicholas A. Pairolero, and Andrew A. Toole, 2022, Identifying artificial intelligence (AI) invention: a novel AI patent dataset, *The Journal of Technology Transfer* 47.

Gompers, Paul A, 1995, Optimal investment, monitoring, and the staging of venture capital, *Journal of Finance* 50, 1461–1489.

Gompers, Paul A., Will Gornall, Steven N. Kaplan, and Ilya A. Strebulaev, 2020, How do venture capitalists make decisions?, *Journal of Financial Economics* 135, 169–190.

Goolsbee, Austan, 1998, Does government r&d policy mainly benefit scientists and engineers?, *The American Economic Review* 88, 298–302.

Gross, Daniel P., and Maria P. Roche, 2023, Coordinated R&D Programs and the Creation of New Industries, *Working Paper* .

Gross, Daniel P, and Bhaven N Sampat, 2023, America, jump-started: World War II R&D and the takeoff of the US innovation system, *American Economic Review* 113, 3323–3356.

Guardian, 2014, Scientists threaten to boycott €1.2bn human brain project, https://www.theguardian.com/science/2014/jul/07/human-brain-project-boycott-mismanagement, The Guardian, Published on 7 July 2014. Retrieved on 1 March 2024.

Hall, Bronwyn, Christian Helmers, Mark Rogers, and Vania Sena, 2014, The choice between formal and informal intellectual property: A review, *Journal of Economic Literature* 52, 375–423.

Hall, Bronwyn H, Adam B Jaffe, and Manuel Trajtenberg, 2001, The NBER patent citation data file: Lessons, insights and methodological tools, Working Paper 8498, National Bureau of Economic Research.

Hall, Robert E, and Susan E Woodward, 2010, The burden of the nondiversifiable risk of entrepreneurship, *American Economic Review* 100, 1163–1194.

He, Haibo, and Edwardo A. Garcia, 2009, Learning from imbalanced data, *IEEE Transactions on Knowledge and Data Engineering* 21, 1263–1284.

Hellmann, T, Junia Mulla, and Matthias Qiam, 2025, The Burden of University Equity Stakes for Spin-outs: Evidence from the UK, *Working Paper* .

Hood, Leroy, and Lee Rowen, 2013, The human genome project: big science transforms biology and medicine, *Genome medicine* 5, 79.

Howell, Sabrina T., 2017, Financing Innovation: Evidence from R&D Grants, *American Economic Review* 107, 1136–64.

Howell, Sabrina T., 2024, Government intervention in innovation, *Annual Review of Financial Economics* 16, 367–390.

Howell, Sabrina T., Jason Rathje, John Van Reenen, and Jun Wong, Forthcoming, OPENing up Military Innovation: An Evaluation of Reforms to the U.S. Air Force SBIR Program, *Journal of Political Economy* Forthcoming.

Insel, T.R., and S.C. Landis, 2013, Twenty-five years of progress: the view from NIMH and NINDS, *Neuron* 80, 561–567.

Kantor, S., and S. Whalley, Forthcoming, Moonshot: Public RD and Growth, *American Economic Review* Forthcoming.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf, 2014, Entrepreneurship as experimentation, *Journal of Economic Perspectives* 28, 25–48.

Kim, In Song, 2018, Lobbyview: Firm-level lobbying & congressional bills database, Working paper available from http://web.mit.edu/insong/www/pdf/lobbyview.pdf.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *The Quarterly Journal of Economics* 132, 665–712.

Kortum, Samuel, and Josh Lerner, 2000, Assessing the contribution of venture capital to innovation, *The RAND Journal of Economics* 31, 674–692.

Krieger, Joshua, Danielle Li, and Dimitris Papanikolaou, 2021, Missing novelty in drug development*, *The Review of Financial Studies* 35, 636–679.

Krieger, Joshua L., Monika Schnitzer, and Martin Watzinger, 2024, Standing on the shoulders of science, *Strategic Management Journal* 45, 1670–1695.

Lach, Saul, and Mark Schankerman, 2008, Incentives and Invention in Universities, *The RAND Journal of Economics* 39, 403–433.

Lemaitre, Guillaume, Fernando Nogueira, and Christos K. Aridas, 2017, Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning, *Journal of Machine Learning Research* .

Lerner, Josh, 1999, The Government as Venture Capitalist: The Long-Run Impact of the SBIR Program, *The Journal of Business* 72, 285–318.

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

Lerner, Josh, Henry J Manley, Carolyn Stein, and Heidi L Williams, 2024, The Wandering Scholars: Understanding the Heterogeneity of University Commercialization, Working Paper 32069, National Bureau of Economic Research.

Lerner, Josh, and Ramana Nanda, 2020, Venture capital's role in financing innovation: What we know and how much we still need to learn, *Journal of Economic Perspectives* 34, 237–261.

Lerner, Josh, and Amit Seru, 2021, The Use and Misuse of Patent Data: Issues for Finance and Beyond, *The Review of Financial Studies* 35, 2667–2704.

Marx, Matt, and Aaron Fuegi, 2020, Reliance on science: Worldwide front-page patent citations to scientific articles, *Strategic Management Journal* 41, 1572–1594.

Marx, Matt, and Aaron Fuegi, 2022, Reliance on science by inventors: Hybrid extraction of in-text patent-to-article citations, *Journal of Economics & Management Strategy* 31, 369–392.

Masclans, Roger, Sharique Hasan, and Wesley M Cohen, Forthcoming, Measuring the Commercial Potential of Science, *Strategic Management Journal* Forthcoming.

Masulis, Ronald W., and Yufeng Yao, 2025, How scientists on the board enhance corporate innovation by bridging research and development, *UNSW Working Paper* .

Mazzucato, Mariana, 2021, *A Moonshot Guide to Changing Capitalism* (New York: Penguin).

Miller, G., 2010, Is pharma running out of brainy ideas?, *Science* 329, 502–504.

Modic, Dolores, and Maryann P. Feldman, 2017, Mapping the human brain: comparing the US and EU Grand Challenges, *Science and Public Policy* 44, 440–449.

Mott, Meghan C, Joshua A Gordon, and Walter J Koroshetz, 2018, The NIH BRAIN Initiative: Advancing neurotechnologies, integrating disciplines, *PLoS biology* 16, e3000066.

Myers, Kyle R., and Lauren Lanahan, 2022, Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy, *American Economic Review* 112, 2393–2423.

Nanda, Ramana, and Matthew Rhodes-Kropf, 2017, Financing risk and innovation, *Management Science* 63, 901–918.

Nanda, Ramana, and Matthew Rhodes-Kropf, 2018, Coordination frictions in venture capital syndicates, NBER Working Paper 24517.

National Cancer Institute, 2025, Cancer moonshot — progress on the national cancer moonshot initiative, https://www.cancer.gov/research/progress/moonshot-cancer-initiative, Updated April 29,2025.

Nature, 2021, Neuroscientists make strides towards deciphering the human brain, *Nature* 598, 7–7.

Nelson, Richard R, 1959, The simple economics of basic scientific research, *Journal of political economy* 67, 297–306.

Newsome, Bill, and Nicholas Weiler, 2024, The BRAIN Initiative: the national vision for the future of neuroscience is now in doubt, From Our Neurons to Yours: Podcast of Wu Tsai Neuroscience Institute, Stanford University.

Ngai, John, 2024, Brain @ 10: A decade of innovation, *Neuron* 112, 3003–3006.

NIH, 2014, NIH embraces bold, 12-year scientific vision for BRAIN Initiative, https://www.nih.gov/news-events/news-releases/nih-embraces-bold-12-year-scientific-vision-brain-initiative, Accessed: [28/Apr/2024].

NIH, 2019, Notice of data sharing policy for the brain initiative, https://grants.nih.gov/grants/guide/notice-files/NOT-MH-19-010.html, Notice Number: NOT-MH-19-010, Release Date: January 22, 2019.

NSF, 2025, The national science foundation: A brief history, https://www.nsf.gov/about/history/narrative, Accessed: 2025-07-05.

Nuffield Council on Bioethics, 2013, Novel neurotechnologies: Intervening in the brain, https://www.nuffieldbioethics.org/publications/novel-neurotechnologies.

Pham, Pham, Roham Rezaei, and Jason Zein, 2023, Venture capitalists vs. deep-pocketed incumbents: Startup financing strategies in the presence of competitive threats, *Working Paper* .

Popp, David, and Richard G Newell, 2009, Where Does Energy R&D Come From? Examining Crowding Out from Environmentally-Friendly R&D, Working Paper 15423, National Bureau of Economic Research.

Puri, Manju, and Rebecca Zarutskie, 2012, On the Life Cycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms, *The Journal of Finance* 67, 2247–2293.

Rose, Michael E., and John R. Kitchin, 2019, pybliometrics: Scriptable bibliometrics using a python interface to scopus, *SoftwareX* 10, 100263.

Scotchmer, Suzanne, 1991, Standing on the shoulders of giants: Cumulative research and the patent law, *Journal of Economic Perspectives* 5, 29–41.

Sejnowski, Terrence J., 2014, The brain initiative: Connecting the dots, Video Lecture: https://youtu.be/w3WSqp_Ck1s?si=xDb-l9LnsvxlHh0T, Joseph Leiter NLM/Medical Library Association (MLA) Lecture.

Theil, Stefan, 2015, Why the Human Brain Project Went Wrong—and How to Fix It, https://www.scientificamerican.com/article/why-the-human-brain-project-went-wrong-and-how-to-fix-it/.

Toole, Andrew A., Nicholas A. Pairolero, James L. Forman, and Alexander V. Giczy, 2020, The promise of machine learning for patent landscaping, *Institutions & Transition Economics: Microeconomic Issues eJournal* .

Tumarkin, Robert, 2020, Yenta, https://github.com/tumarkin/yenta.

Wallsten, Scott J, 2000, The effects of government-industry r&d programs on private r&d: the case of the small business innovation research program, *The Rand journal of economics* 82–100.

Williams, Heidi L, 2013, Intellectual property rights and innovation: Evidence from the human genome, *Journal of Political Economy* 121, 1–27.

Zador, A., S. Escola, B. Richards, et al., 2023, Catalyzing next-generation Artificial Intelligence through NeuroAI, *Nature Communications* 14, 1597.

Ziedonis, Rosemarie Ham, 2004, Don't fence me in: Fragmented markets for technology and the patent acquisition strategies of firms, *Management Science* 50, 804–820.

Zucker, Lynne G, Michael R Darby, and Marilynn B Brewer, 1998, Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises, *The American Economic Review* 88, 290–306.

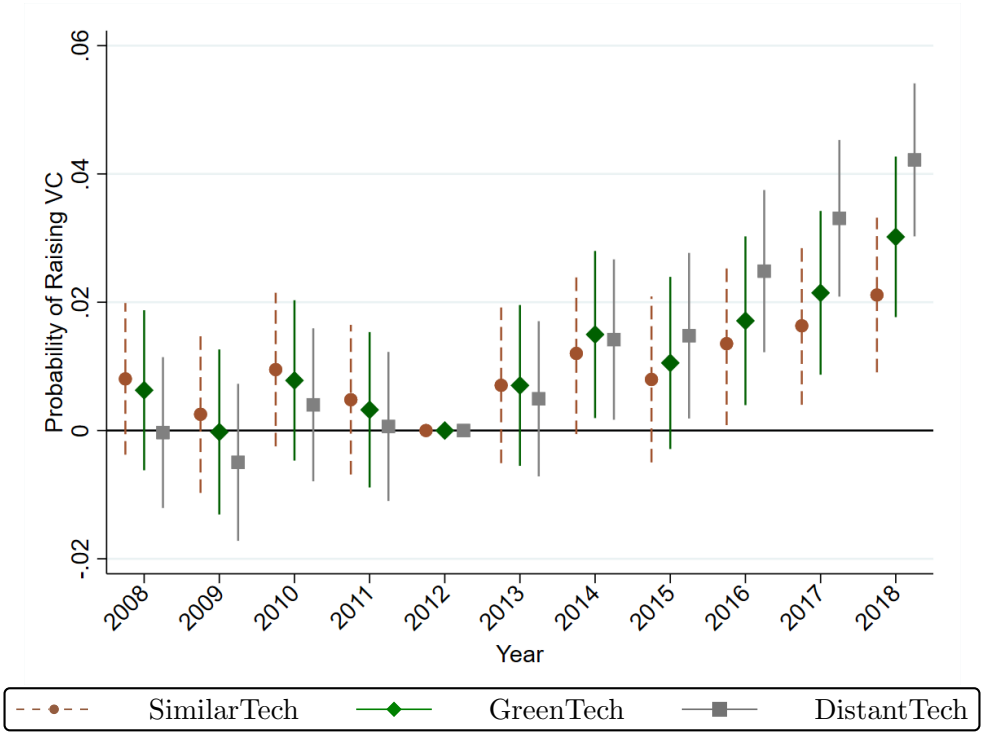


Figure 1: **Probability of Raising VC (Against Different Control Groups)**

The figure reports difference-in-differences coefficients from the dynamic (event study) specification of Equation 4, in three samples with different control groups. The unit of observation is a firm-year. For each calendar year, we interact the *Neuro startup* indicator with a year dummy and plot the resulting coefficient. The omitted (zero) year is 2012, the last pre-BRAIN year. Brown circles show coefficients when the control group is *SimilarTech*; green diamonds when it is *GreenTech*; gray squares when it is *DistantTech*. Vertical lines give 95 percent confidence intervals based on startup-clustered standard errors.

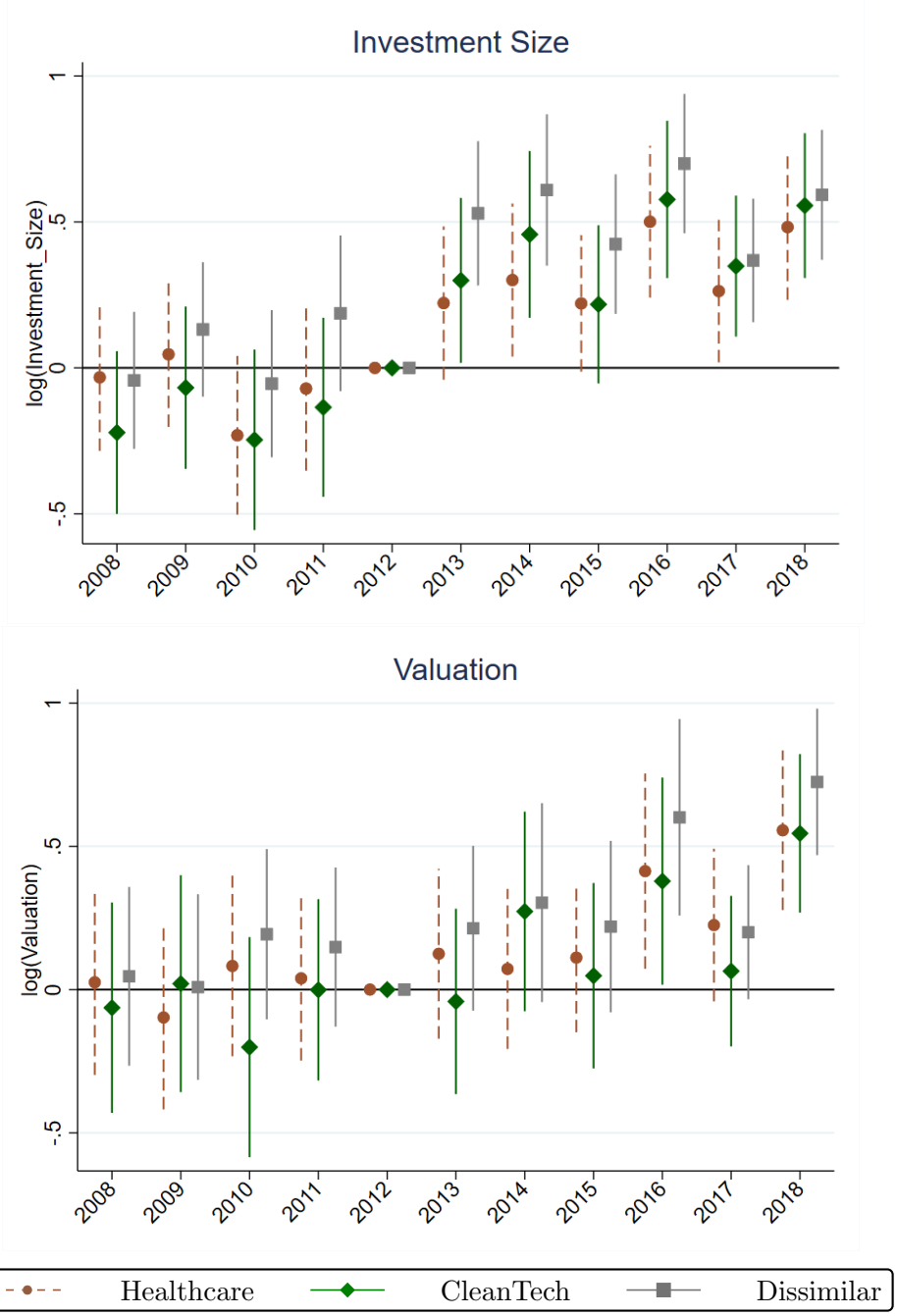


Figure 2: **Investment Size and Valuation (Against Different Control Groups)**

The figure reports difference-in-differences coefficients from the dynamic (event study) specification of Equation 5, in three samples with different control groups. The dependent variables are the log of the investment size (top) and the log of the pre-money valuation (bottom). The unit of observation is an entrepreneurial firm's financing event. For each calendar year, we interact the *Neuro startup* indicator with a year dummy and plot the resulting coefficient. The omitted (zero) year is 2012, the last pre-BRAIN year. Brown circles show coefficients when the control group is *Healthcare*; green diamonds when it is *CleanTech*; gray squares when it is *Dissimilar*. Vertical lines give 95 percent confidence intervals based on startup-clustered standard errors.

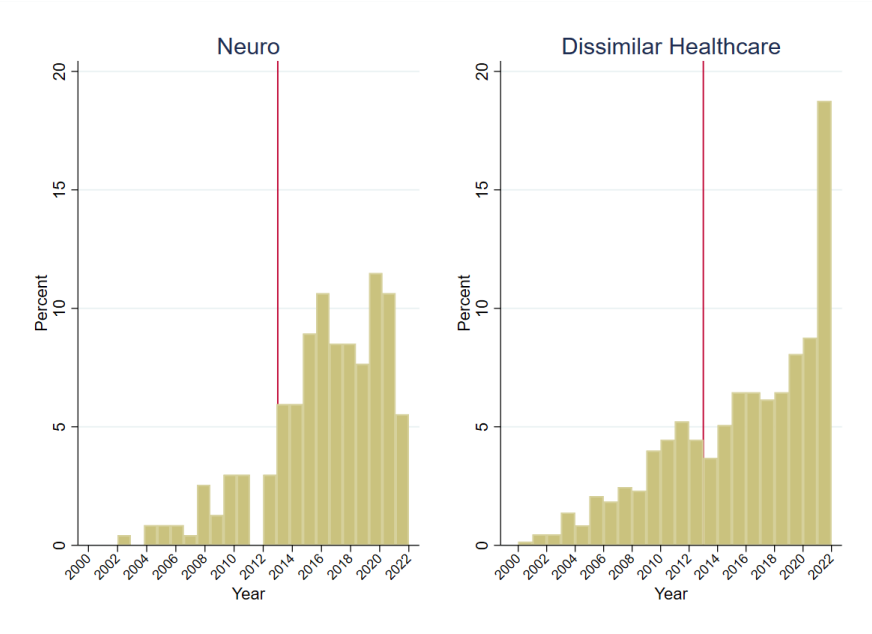


Figure 3: **Acquisitions of Neuro and Dissimilar Healthcare Startups**

The figure plots the distribution of acquisition years for *Neuro startups* (left panel) and for *Dissimilar* healthcare ventures (right panel). Each bar reports the *percentage* of all acquisitions in that group occurring in that calendar year. The red line marks 2013 (the launch of the BI.)

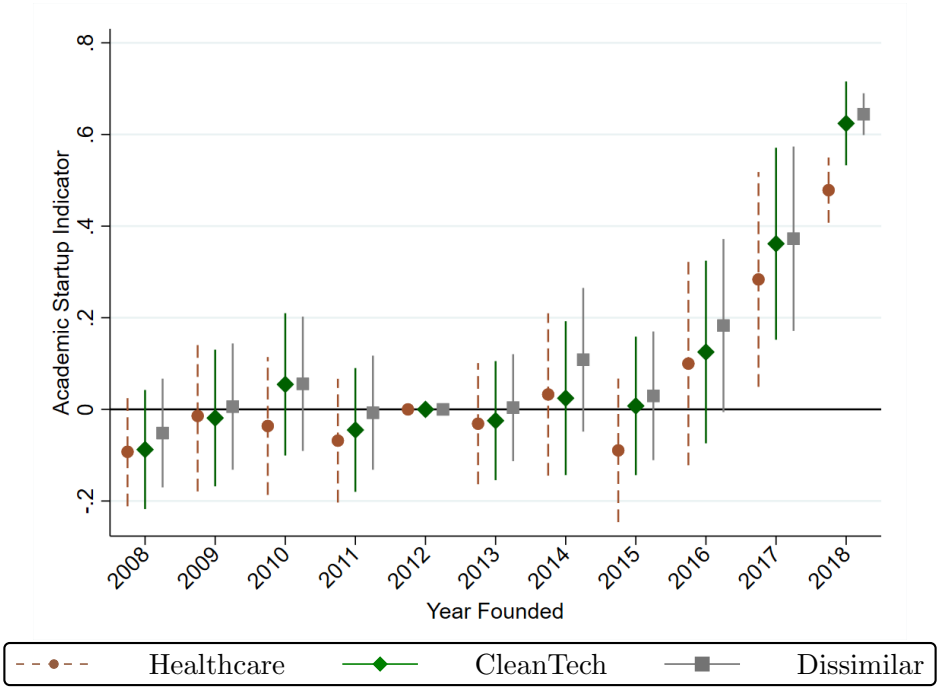


Figure 4: **Academic Startups (Against Different Control Groups)**

The figure plots the coefficients for the estimation of dynamic version of Equation 5, with interaction terms of each founding year and the *Neuro* dummy where the dependent variable is an indicator variable for *Academic Startup*: startups who have a STEM academic in senior positions in the first three years after being founded. The unit of observation is an entrepreneurial firm. For each calendar year, we interact the *Neuro startup* indicator with a year dummy and plot the resulting coefficient. The omitted (zero) year is 2012, the last pre-BRAIN year. Brown circles show coefficients when the control group is *Healthcare*; green diamonds when it is *CleanTech*; gray squares when it is *Dissimilar*. Vertical lines give 95 percent confidence intervals based on heteroskedasticity-robust standard errors.

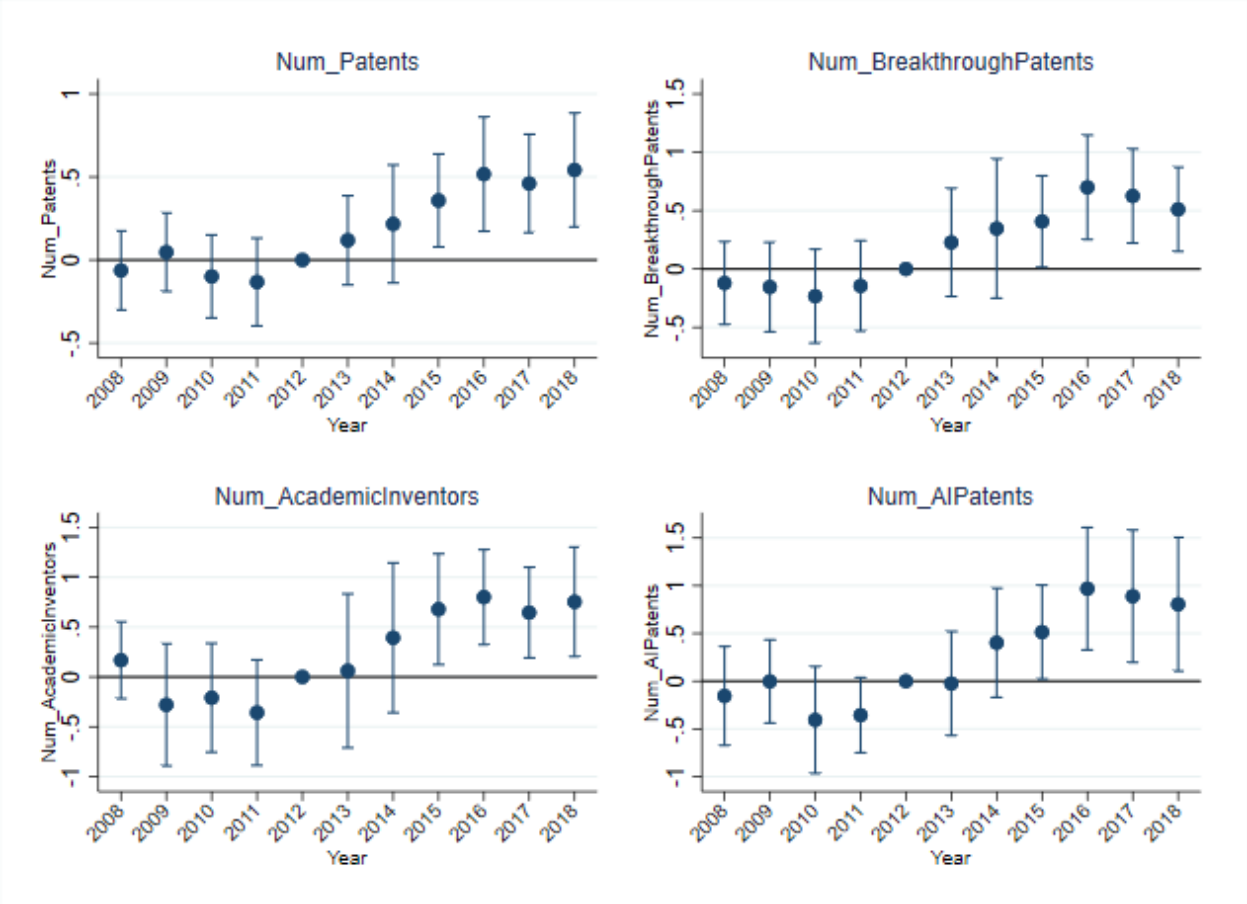


Figure 5: **Patenting of VC-Backed Startups**

The figure reports difference-in-differences coefficients from the dynamic (event study) specification of Equation 4. All outcome variables are count variables, and the models are estimated using Poisson regressions. The unit of observation is a firm-year. For each calendar year, we interact the *Neuro startup* indicator with a year dummy and plot the resulting coefficient. The omitted (zero) year is 2012, the last pre-BRAIN year. Vertical lines give 95 percent confidence intervals based on startup-clustered standard errors.

Table 1: Subject Areas of NIH Neuroscience Grants: BI vs. Non-BI This table compares the distribution of Scopus subject areas between BI grants and other non-BI NIH neuroscience grants. We utilize a prompt-based generative-AI model (OpenAI GPT-4o-mini) to classify each grant abstract into one or more relevant Scopus categories. For brevity, social-science subject areas are excluded. Columns (1) and (2) report the percentage of BI and non-BI grants that fall into that category for each subject area, respectively. The percentages in Columns (1) and (2) sum to more than 100%, because a single grant can belong to multiple subject areas. Column (3) reports the ratio of BI to non-BI shares for each subject area.

Scopus Subject Areas	%BI Grants	%Non-BI Grants	$\frac{(BI)}{(Non-BI)}$ $\frac{(Column1)}{(Column2)}$
	(1)	(2)	(3)
Neuroscience	99.80	79.47	1.26
Medicine	87.82	96.80	0.91
Biochemistry, Genetics and Molecular Biology	62.28	57.27	1.09
Computer Science	22.20	5.54	4.01
Engineering	8.06	2.34	3.44
Psychology	5.70	12.71	0.45
Immunology and Microbiology	5.11	27.70	0.18
Health Professions	3.14	18.04	0.17
Mathematics	2.95	0.92	3.21
Physics and Astronomy	1.18	0.26	4.54
Pharmacology, Toxicology and Pharmaceutics	1.18	6.38	0.18
Veterinary Science	0.20	0.30	0.67
Nursing	0.00	0.47	0.00
Agricultural and Biological Sciences	0.00	0.17	0.00
Materials Science	0.00	0.07	0.00
Energy	0.00	0.01	0.00
Earth and Planetary Sciences	0.00	0.00	NA
Dentistry	0.00	0.20	0.00
Chemistry	0.00	0.41	0.00
Chemical Engineering	0.00	0.00	NA
Environmental Science	0.00	1.15	0.00

Table 2: Summary Statistics of Startups. This table shows summary statistics for 44,413 unique startups receiving their first venture capital financing between 2000 and 2019, whose financing outcomes have been followed until 2022. Panel A presents financing information for all rounds where round size is not missing, while Panel B focuses on the first rounds. Panel C presents data at the startup level. Panel D offers summary statistics of patenting characteristics of firms with at least one patent at the firm-year level. All variables are defined in Appendix Table IA.1

	N	Mean	Std Dev	p10	Median	p90
Panel A: VC Financing Round Level						
Investment Size	81,359	16.38	77.73	0.53	4.80	33.00
Pre-Money Valuation	50,288	147.22	1401.38	3.50	17.00	175.00
Deal Year	81,359	2014.08	5.46	2006	2015	2020
# Investors	81,359	4.29	4.22	1.00	3.00	9.00
VC Round Number	81,359	2.51	1.74	1.00	2.00	5.00
Neuro	81,359	0.03	0.16	0.00	0.00	0.00
Employee Exposure	1,087	0.08	0.27	0.00	0.00	0.00
Patent Exposure	620	1.22	4.71	0.00	0.00	2.00
State Exposure	1,217	106.65	111.18	1.00	63.00	283.00
Panel B: First VC Financing Round Level						
Investment Size	29,294	5.44	20.42	0.30	2.00	10.60
Pre-Money Valuation	15,707	13.72	139.62	2.00	6.15	20.96
Neuro	29,294	0.02	0.13	0.00	0.00	0.00
Deal Year	29,294	2012.45	5.30	2004	2014	2018
# Investors	29,294	3.59	3.99	1.00	2.00	8.00
Panel C: Startup Characteristics (Firm Level)						
Neuro Startup	44,413	0.02	0.13	0.00	0.00	0.00
Successful Exit	23,097	0.16	0.36	0.00	0.00	1.00
Time to Exit (days)	33,565	1952.75	1237.27	685	1656	3657
Exit Year	33,565	2016.88	4.74	2010	2018	2022
ComplementaryTech	43,339	0.19	0.39	0.00	0.00	1.00
AcademicStartup	41,037	0.07	0.26	0.00	0.00	0.00
Panel D: Patenting Characteristics of VC-backed Startups (Firm-Year level)						
# Patents	106,418	0.93	3.55	0.00	0.00	2.00
# Average Citations	106,418	0.62	2.61	0.00	0.00	1.54
# Breakthrough Patents	106,418	0.28	1.54	0.00	0.00	1.00
# Academic Inventors	106,418	0.23	2.02	0.00	0.00	0.00
# AI Patents	106,418	0.21	1.70	0.00	0.00	0.00

Table 3: Probability of Raising Venture Capital. This table reports results from OLS panel regressions estimating Equation 4. The unit of observation is a firm-year. The dependent variable equals one if the startup has raised any VC by year t. *Neuro* is an indicator variable equal to one if the startup is a *Neuro Startup* and zero otherwise; *Post* equals one in 2013 (the BRAIN Initiative announcement year) and thereafter, and zero otherwise. Column (1) uses *All Patenting* firms as the control group. In Columns (2)–(4), this is restricted to: *DistantTech*, *GreenTech* and *SimilarTech* respectively. Section 3.3 defines these groups. Each specification includes startup and calendar-year fixed effects plus log-transformed age. Robust t-statistics (clustered by startup) are in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Raised VC			
Control Group:	All Patenting	DistantTech	GreenTech	SimilarTech
	(1)	(2)	(3)	(4)
Neuro×Post	0.032***	0.041***	0.028***	0.019***
	(5.12)	(6.64)	(4.34)	(3.10)
Ln(Age+1)	0.041***	0.029***	0.051***	0.059***
	(40.95)	(25.37)	(16.16)	(33.59)
Observations	656282	373698	81977	299584
Mean Outcome	0.048	0.038	0.057	0.064
Adj R-squared	0.724	0.739	0.706	0.712
Firm FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y

Table 4: Funding Size and Valuation. This table presents the results of OLS regressions estimating Equation 5. The dependent variable is ln(Investment Size) in Panel A and ln(Pre-money Valuation) in Panel B. The unit of observation is a VC financing event. Columns (1), (3), (5), and (7) include all financing rounds and other columns are restricted to the first VC rounds. *Neuro* is an indicator variable equal to one for *Neuro startups* and zero otherwise. *Syndicate Size* is ln(#investors). *Post* equals one for financing events on or after 2013: announcement year of the BRAIN Initiative. Control groups for Columns (1-2) are all untreated VC-backed startups, which in Columns (3-4), (5-6) and (7-8) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. These groups are defined in Section 3.3. Fixed effects for both panels are presented at the bottom of the table and include dummies for the year of the financing event, Pitchbook's 41 industry groups, the state of the startup's headquarters and the sequence of financing event (odd columns). The t-statistics (in parentheses) are clustered by firm in Columns (1), (3), (5), and (7) and are based on heteroskedasticity-robust standard errors in Columns (2), (4), (6), and (8). ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Panel A:	Ln(Investment Size)							
Control Group:	All		Dissimilar		CleanTech		Healthcare	
VC Round #:	Any	1st	Any	1st	Any	1st	Any	1st
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neuro×Post	0.359***	0.691***	0.412***	0.739***	0.428***	0.764***	0.278***	0.562***
	(4.78)	(5.92)	(5.42)	(6.28)	(5.10)	(5.65)	(3.67)	(4.20)
Neuro	0.057	0.011	0.128**	0.064	0.171*	0.006	0.058	0.046
	(1.03)	(0.16)	(2.21)	(0.87)	(1.76)	(0.05)	(1.04)	(0.60)
Syndicate Size	0.535***	0.341***	0.516***	0.337***	0.614***	0.392***	0.697***	0.558***
	(83.07)	(38.72)	(72.84)	(36.22)	(30.51)	(11.19)	(47.13)	(22.00)
Observations	81359	29294	65663	25196	7967	2344	16855	5613
Mean Outcome	1.502	0.692	1.373	0.626	1.855	0.910	1.759	0.978
Adj R-squared	0.386	0.194	0.380	0.198	0.424	0.241	0.396	0.235
Panel B:	Ln(Valuation)							
Neuro×Post	0.241**	0.437***	0.275***	0.464***	0.279***	0.484***	0.227**	0.469***
	(2.48)	(3.82)	(2.85)	(4.05)	(2.70)	(3.47)	(2.43)	(3.49)
Neuro	0.160**	-0.013	0.207***	0.012	0.483***	0.117	0.066	-0.055
	(2.35)	(-0.17)	(2.90)	(0.16)	(4.01)	(0.92)	(0.98)	(-0.68)
Syndicate Size	0.301***	0.157***	0.283***	0.164***	0.327***	0.195***	0.303***	0.241***
	(37.36)	(16.08)	(32.78)	(15.72)	(13.56)	(5.09)	(18.12)	(10.14)
Observations	50493	15869	39210	13189	5153	1359	10760	3263
Mean Outcome	3.028	1.835	2.906	1.798	3.274	1.941	3.049	1.907
Adj R-squared	0.504	0.081	0.503	0.087	0.518	0.149	0.474	0.121
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y
VC Round # FE	Y	N	Y	N	Y	N	Y	N

Table 5: Exposure to the BRAIN Initiative and Funding Variation within Neuro startups. This table presents the results of estimating Equation 6, which examines the relation between *Neuro startups*' exposure to the BI and their financing characteristics. The dependent variable is ln(Investment Size) in Columns (1–3) and ln(Pre-money Valuation) in Columns (4–6). *Employee Exposure* is an indicator variable equal to one if the firm has hired at least one BI scientist up to the financing year and zero otherwise. *Patent Exposure* is the number of patents influenced by the BI up to the financing year, as identified by our large language model (Section B). *State Exposure* is the number of BI grants awarded to the institutions in the startup's state. *Syndicate Size* is ln(#investors). *Year FE* are dummies for financing year, *Industry FE* for Pitchbook's 41 industry groups, *State FE* for the startup's headquarters state and *VC Round #* for the sequence number of VC financing rounds. The t-statistics (in parentheses) are based on standard errors clustered at the startup level, with ***, ** and * representing significance at the 1%, 5%, and 10% levels, respectively.

	Ln(Investment Size)			Ln(Valuation)		
	(1)	(2)	(3)	(4)	(5)	(6)
Employee Exposure	0.539**			0.789***		
	(2.37)			(2.85)		
Ln(Patent Exposure+1)		0.285***			0.102	
		(2.84)			(0.98)	
Ln(State Exposure+1)			0.330***			0.251***
			(5.44)			(3.77)
Syndicate Size	0.703***	0.686***	0.735***	0.330***	0.344***	0.348***
	(15.33)	(12.11)	(16.08)	(5.87)	(4.41)	(6.45)
Observations	1087	620	1217	809	457	891
Mean Outcome	2.664	2.762	2.611	4.044	4.152	4.003
Adj. R2	0.461	0.495	0.401	0.559	0.562	0.510
Industry FE	Y	Y	Y	Y	Y	Y
VC Round # FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
State FE	Y	Y	N	Y	Y	N

Table 6: Cancer Moonshot and VC Investments. This table presents estimates using the 2016 Cancer Moonshot as a policy shock (Panel A) and cancer-related startups as a placebo (Panel B). The dependent variable is ln(Investment Size) in both panels. The unit of observation is a VC financing event. Columns (1), (3), (5), and (7) include all financing rounds and other columns are restricted to the first VC rounds. *Cancer* is an indicator variable equal to one for startups classified in Pitchbook's *Oncology* vertical. *Syndicate Size* is ln(#investors). In Panel A, *Post16* equals one for financing events occurring in or after 2016, the year of the Cancer Moonshot announcement. Panel B serves as a placebo test by using 2013, the year of the BRAIN Initiative, as a pseudo-shock for *Cancer startups*; this sample is limited to the 2010–2015 period and checks whether *Cancer startups* experienced an increase in VC investment post-2013. Control groups for Columns (1-2) are all untreated VC-backed startups, which in Columns (3-4), (5-6) and (7-8) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. These groups are defined in Section 3.3. Fixed effects for both panels are presented at the bottom of the table and include dummies for the year of the financing event, Pitchbook's 41 industry groups, the state of the startup's headquarters and the sequence of financing event (odd columns). The t-statistics (in parentheses) are clustered by firm in Columns (1), (3), (5), and (7) and are based on heteroskedasticity-robust standard errors in Columns (2), (4), (6), and (8). ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Cancer Moonshot	Ln(Investment Size)							
Control Group:	All		Dissimilar		CleanTech		Healthcare	
VC Round #:	Any	1st	Any	1st	Any	1st	Any	1st
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cancer×Post16	0.360***	0.534***	0.362***	0.544***	0.471***	0.638***	0.386***	0.551***
	(5.66)	(5.51)	(5.62)	(5.56)	(6.34)	(5.57)	(5.89)	(5.42)
Cancer	0.045	0.090	0.137***	0.172***	0.099	0.162	0.040	0.091
	(0.93)	(1.46)	(2.63)	(2.62)	(0.84)	(1.19)	(0.85)	(1.48)
Syndicate Size	0.536***	0.342***	0.514***	0.338***	0.645***	0.463***	0.698***	0.562***
	(83.12)	(38.85)	(72.22)	(36.14)	(31.73)	(13.39)	(47.37)	(22.24)
Observations	81359	29294	64991	25069	8663	2732	16855	5613
Mean Outcome	1.502	0.692	1.362	0.625	1.825	0.982	1.759	0.978
Adj R-squared	0.386	0.194	0.377	0.199	0.411	0.262	0.399	0.240
Panel B: Placebo Test	Ln(Investment Size)							
Cancer×Post	0.016	-0.006	0.055	0.002	0.021	0.064	-0.034	-0.061
	(0.11)	(-0.03)	(0.37)	(0.01)	(0.13)	(0.25)	(-0.23)	(-0.25)
Cancer	-0.016	-0.046	0.113	0.090	0.326	0.302	-0.004	0.009
	(-0.13)	(-0.24)	(0.89)	(0.46)	(1.62)	(0.96)	(-0.04)	(0.04)
Syndicate Size	0.385***	0.202***	0.345***	0.205***	0.463***	0.264***	0.607***	0.368***
	(36.92)	(15.36)	(30.06)	(14.62)	(12.46)	(4.16)	(21.72)	(8.16)
Observations	25711	11085	19920	9456	2100	792	4445	1707
Mean Outcome	1.065	0.327	0.846	0.238	1.354	0.538	1.359	0.659
Adj R-squared	0.341	0.098	0.300	0.087	0.401	0.211	0.366	0.168
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y
VC Round # FE	Y	N	Y	N	Y	N	Y	N

Table 7: Success of the Exits. This table reports results from OLS regressions estimating Equation 5, where the dependent variable is an indicator variable for a *Successful Exit* in Columns (1-4) and log-transformed *Time to Exit* in Columns (5-8). *Successful Exit* is defined as an IPO or a M&A at a reported value at least twice the total capital invested. *Time to Exit* is the difference between the number of days between the date of the first VC investment and the exit event. A unit of observation is an entrepreneurial firm. *Neuro* is an indicator variable equal to one for *Neuro startups* and zero otherwise. *Post* equals one for startups receiving the first VC financing event on or after 2013 (BI's announcement year). Control groups for Columns (1) and (5) are all untreated VC-backed startups, which in Columns (2) and (6), (3) and (7), (4) and (8) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. These groups are defined in Section 3.3. *Ln(Raised before exit)* is the log-transformed total amount of financing that the startup has raised before the exit. *First VC Financing Year FE* (*Exit Year*) indicate dummies for financing (exit) year, *Industry FE* are dummies for Pitchbook's 41 industry groups. *State FE* are dummies for entrepreneurial firm headquarters state. The *t*-statistics (in parentheses) are based on heteroskedasticity-robust standard errors ***, ** and * representing significance at the 1%, 5% and 10% levels, respectively.

	Successful Exit				Ln(Time to Exit)			
	All	Dissimilar	CleanTech	Healthcare	All	Dissimilar	CleanTech	Healthcare
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neuro×Post	0.155***	0.117***	0.058	0.087**	-0.256***	-0.294***	-0.177**	-0.182***
	(3.74)	(2.89)	(1.31)	(2.20)	(-4.19)	(-4.82)	(-2.42)	(-2.84)
Neuro	0.049*	0.114***	0.103**	0.054**	0.216***	0.330***	0.322**	0.203***
	(1.82)	(4.29)	(2.52)	(2.12)	(6.17)	(9.11)	(2.39)	(5.39)
Ln(Raised before exit)	0.060***	0.057***	0.076***	0.100***	0.078***	0.073***	0.040***	0.046***
	(53.13)	(46.77)	(18.24)	(38.48)	(42.83)	(36.58)	(4.43)	(11.12)
Observations	23097	20312	1741	4598	33609	29411	1536	6157
Mean Outcome	0.157	0.136	0.235	0.263	7.359	7.306	7.573	7.469
Adjusted R2	0.231	0.242	0.391	0.381	0.175	0.143	0.207	0.209
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
First VC Year FE	Y	Y	Y	Y	Y	Y	Y	Y
Exit Year FE	Y	Y	Y	Y	N	N	N	N
State FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 8: Academic Startup. This table reports results from OLS regressions estimating Equation 5, where the dependent variable is an indicator variable for *Academic Startups*. An *Academic Startup* is one with at least one STEM academic holding a senior position within the startup up to three years after the startup's founding year; see Section 3.4.1 for more details. A unit of observation is an entrepreneurial firm. *Neuro* is an indicator variable equal to one for *Neuro startups* and zero otherwise. *Post* equals one for startups founded in or after 2013, the announcement year of the BRAIN Initiative. Control group for Column (1) is all untreated VC-backed startups, which in Column (2), (3) and (4) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. *Industry FE* are dummies for Pitchbook's 41 industry groups. *State FE* are dummies for entrepreneurial firm headquarters state. *Founding Year FE* indicate dummies for founding year. The t-statistics (in parentheses) are based on heteroskedasticity-robust standard errors ***, ** and * representing significance at the 1%, 5% and 10% levels, respectively.

Acadedmic Startup				
Control Group:	All	Dissimilar	CleanTech	Healthcare
	(1)	(2)	(3)	(4)
Neuro×Post	0.106***	0.110***	0.082**	0.040
	(2.74)	(2.86)	(1.97)	(0.93)
Neuro	0.038**	0.053***	0.068**	0.046**
	(2.09)	(2.91)	(2.16)	(2.49)
Observations	41037	35400	3287	8226
Mean Outcome	0.075	0.062	0.136	0.173
Adj R-squared	0.080	0.078	0.099	0.080
Industry FE	Y	Y	Y	Y
Founding Year FE	Y	Y	Y	Y
State FE	Y	Y	Y	Y

Table 9: Patents, Academic Inventors and AI. This table reports Poisson regression results estimating Equation 4, based on a firm-year panel of VC-backed patenting startups. The unit of observation is a firm-year. Columns (1), (3), (5), (7), and (9) use the entire firm-year panel from the founding year through VC exit. *Before VC* in Columns (2), (4), (6), (8), and (10) indicates that the panel is truncated at the first year the firm receives venture capital. The outcome variable in Columns (1) and (2) is the number of patents filed in year t; in Columns (3) and (4) it is the average number of adjusted citations received by these patents; in Columns (5) and (6) it is the number of *breakthrough patents*, defined as those receiving more citations than the 90th percentile within their technology class and year; in Columns (7) and (8) it is the number of *academic inventors* hired in year t; in Columns (9) and (10) it is the number of *AI-related patents* filed (and later granted) in year t. *Neuro* is an indicator variable equal to one for *Neuro startups* and zero otherwise. *Post* equals one for 2013 (BI announcement year) and later years and zero otherwise. Each specification includes startup and calendar-year fixed effects plus log-transformed age. t-statistics (in parentheses) are clustered at the startup level. Significance levels are denoted by ***, **, and * for 1%, 5%, and 10%, respectively.

	# Patents		Avg. Patent Citations		# Breakthrough patents		# Academic Inventors		# AI Patents	
		Before VC		Before VC		Before VC		Before VC		Before VC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Neuro × Post	0.576***	0.586***	-0.014	-0.014	0.747***	0.673***	0.739**	0.536*	0.892**	0.953***
	(3.33)	(2.98)	(-0.11)	(-0.12)	(3.60)	(2.93)	(2.32)	(1.73)	(2.35)	(2.72)
Ln(Age+1)	1.052***	1.076***	0.704***	0.702***	1.157***	1.178***	-0.163**	-0.184***	1.016***	1.095***
	(20.46)	(21.07)	(12.90)	(11.65)	(16.94)	(15.93)	(-2.40)	(-2.88)	(10.44)	(12.10)
Observations	102973	75942	90393	65911	50585	34112	42042	28379	40700	27591
Psuedo-R2	0.488	0.446	0.413	0.409	0.435	0.390	0.445	0.347	0.471	0.405
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Startup FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 10: Integration with Complementary Technologies. This table reports results from OLS regressions estimating Equation 5. In Columns (1-4), the dependent variable is an indicator variable for whether the startup uses at least one *ComplementaryTech* according to Pitchbook's vertical classifications. These technologies include: *Artificial Intelligence and Machine Learning, Big Data and Wearables & Quantified Self, Digital Health and HealthTech*. In Columns (5-8), the outcome variable is *ComplementaryTech2*, which includes startups in *Artificial Intelligence and Machine Learning, Big Data and Wearables & Quantified Self*. A unit of observation is an entrepreneurial firm. *Neuro* is an indicator variable equal to one for *Neuro startups* and zero otherwise. *Post* equals one for startups founded in or after 2013, the announcement year of the BRAIN Initiative. Control group for Column (1) and (5) is all untreated VC-backed startups, which in Column (2) and (6), (3) and (7), (4) and (8) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. *Industry FE* are dummies for Pitchbook's 41 industry groups. *State FE* are dummies for entrepreneurial firm headquarters state. *Founding Year FE* indicate dummies for the year the company was founded. The t-statistics (in parentheses) are based on heteroskedasticity-robust standard errors ***, ** and * representing significance at the 1%, 5% and 10% levels, respectively.

	ComplementaryTech				ComplementaryTech2			
Control Group:	All	Dissimilar	CleanTech	Healthcare	All	Dissimilar	CleanTech	Healthcare
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neuro×Post	0.077**	0.081**	0.066*	-0.022	0.089***	0.092***	0.076**	0.053*
	(2.13)	(2.23)	(1.72)	(-0.57)	(2.83)	(2.94)	(2.31)	(1.73)
Neuro	0.052***	0.062***	0.108***	0.052***	0.037***	0.043***	0.103***	-0.001
	(3.08)	(3.53)	(3.43)	(2.92)	(3.97)	(4.37)	(4.13)	(-0.17)
Observations	43339	37435	3525	8658	43339	37435	3525	8658
Mean Outcome	0.190	0.177	0.171	0.391	0.124	0.117	0.111	0.094
Adj R-squared	0.208	0.219	0.213	0.292	0.093	0.091	0.140	0.098
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Founding Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y

Appendix: Government-Backed Science and Venture Capital

A Institutional Details of the BRAIN Initiative

A1 BRAIN Initiative Funding Sources

This section provides more details on the BI's funding levels and organizational structure. The funding level for BI was initially announced at $4.5 billion over a period of 12 years (NIH, 2014). However, the exact funding levels and budget were updated annually. In addition to NIH, five other federal agencies were involved in the Initiative. Ordered by their contribution, these agencies are: NSF, DARPA, IARPA, FDA, and DoE. Given the variety of agencies funding the program, there is no single source aggregating the overall budget appropriation. Therefore, we collect this information from three sources: 1) BI fact sheets, 2) agency budget reports, and 3) the sum of individual grants publicly available. The information from the last source is only available on the NIH and NSF websites; other agencies do not publicly report their funded projects and amounts.[32] In cases of conflicting information from these three sources, we report the highest amount.

Figure IA.1 presents the funding levels for NIH, NSF, DARPA, and other organizations grouped together. This *Other* category includes IARPA, DoE, and other non-profit organizations such as universities and private research institutes. From 2014 to 2022, BI's budget across agencies adds up to $4.8 billion. NIH provides the largest amount of funding, with $3.1 billion in appropriations up to 2022 (and an additional $1.4 billion in 2023-2025). In the first four years of the program, DARPA is the second-largest funding agency. In 2018, five years after its announcement, the program underwent a review, leading to BRAIN 2.0, which included a revised version and updated scientific priorities. After 2018, there are no reports of DARPA and IARPA's involvement in the initiative, while NSF's funding level increased.

[32] We could obtain information on DARPA total funding amounts only from 2014 to 2017.

A2 Lobbying Activity

Corporate lobbying is a principal channel through which firms attempt to shape the regulatory environment. The Lobbying Disclosure Act of 1995 and the Honest Leadership and Open Government Act of 2007 mandate detailed reporting of such activity. Kim (2018) consolidate these disclosures a dataset called LobbyView. We exploit this database to study lobbying behaviour in the run-up to the BRAIN Initiative. Each filing reports a topic for the lobbying issue, a brief description of the issue, the dollar amount, and, when available, the registrant's Compustat GVKEY. We focus on corporate lobbying reports (those with a GVKEY) classified either under SCI (Science/Technology) or MED (Medical/Disease Research/Clinical Labs) codes. We label a filing as *neuro-related* if its free-text description contains any of the *Neuro* keywords we use across the paper.[33] We aggregate spending to the calendar-quarter level. Figure IA.2 presents the resulting series. It shows that listed firms spent roughly $40 million per quarter on science- and medicine-related lobbying during the five years preceding the BI (vertical line at 2013 Q1), yet devoted virtually none of those resources to neuroscience specifically. This evidence supports the view that the BI's policy momentum emerged outside the private sector.

A3 BI vs non-BI Grants in Neuroscience

The primary funding institutes of BI are NIH, NSF, and DARPA. NIH has a dedicated website that discloses information, including funded projects, on the BRAIN Initiative.[34] Each of the NIH's BI grants belongs to one of the 13 sub-programs, with the overarching goal of mapping the brain.[35] Because NSF does not have a dedicated website or a flag for BI projects, we obtain information on the BI grants through a FOIA request. Other than the funding amount, each grant provides data such as the project abstract and the publications that emerge from it. The majority of these research grants produced scientific publications. More specifically, 82% of NIH-funded projects resulted in publications, yielding 7,448 unique outputs. Meanwhile, the NSF's 694 BI grants led to 6,138 publications. We

[33] These keywords include: *neuro*, *brain*, *Alzheimer*, *Parkinson*, *dementia*, or *optogenetics*

[34] The NIH BI website is https://braininitiative.nih.gov/

[35] These sub-programs are: Armamentarium, Brain Behavior Quantification and Synchronization, BICAN, BRAIN CONNECTS, Cell Census Network (BICCN), Dissemination, Informatics, Investigative Human Neuroscience, Neuroethics, Public-Private Partnerships, Therapeutic Human Neuroscience, Training, Understanding Circuits.

link these publications to Scopus (Rose and Kitchin, 2019) to collect information.

We also identify traditional NIH neuroscience grants (non-BI NIH) via the NIH RePORT API based on three criteria: (1) the grants must contain *Neuro* keywords in their project terms, (2) we exclude SBIR and STTR grants, and (3) the grants must be managed by the same NIH institutes and Centers that also are managing BI grants.[36] From 2014 to 2022, NIH allocated $64 billion to non-BI projects, with an average of $0.47 million per project. In Column (1) of Table IA.2, we report total annual funding levels for both BI and NIH-funded neuroscience projects outside the BI ("non-BI NIH Grants"). BI grants, on average, received $1.10 million per project, 2.3 times larger than NIH non-BI grants. NSF's BI projects are smaller on average than NIH BI projects but still significantly larger than comparable NIH non-BI grants in several years. Across every year, NIH BI grants exceed the size of NIH non-BI grants, and the differences are statistically significant at the 1% level in most years.

A4 Classifying the Subject Areas of Grants Using AI Tools

To classify the subject areas of neuroscience grants, we use two large language models (AI tools): prompt-based generative AI and SciBERT. Because each grant can fall into multiple scientific domains, this is a multi-label classification task: a grant may have zero, one, or several subject area labels. One advantage of Generative AI is that they can perform the classification task with high accuracy due to the very large data they have been trained on. As our generative AI model, we use GPT-4o-mini, and set the temperature to 0 to maximize reproducibility. We then input every grant abstract into its API with the following prompt and report the output in Table 1.

> "Classify the NIH grant abstract below into one or more of the Scopus 2-digit subject areas. Return a semicolon-separated string of Scopus 2-digit abbreviations, e.g. 'neur;medi;bioc'. Do not add keys, comments, or extra text. Areas: "agri": "Agricultural & Biological Sciences","arts": "Arts & Humanities","bioc": "Biochemistry, Genetics & Molecular Biology","busi": "Business, Management & Accounting","ceng": "Chemical Engineering","chem": "Chemistry","comp": "Computer Science","decs": "Decision Sciences","eart": "Earth & Planetary Sciences","econ": "Economics, Econometrics & Finance","ener": "Energy","engi": "Engineering","envi": "Environmental Science","heal": "Health Professions","immu": "Immunology & Microbiology","mate": "Materials Science","math": "Mathematics","medi": "Medicine","neur": "Neuroscience","nurs": "Nursing","phar": "Pharmacology, Toxicology & Pharmaceutics","phys": "Physics & Astronomy","psyc": "Psychology","soci": "Social Sciences","vete": "Veterinary","dent": "Dentistry" ABSTRACT:NIH Grant Abstract"

[36] These NIH institutes and Centers are NCCIH, NEI, NIA, NIAAA, NIBIB, NICHD, NIDA, NIDCD, NIMH, and NINDS.

Prompt-based generative AI faces two key limitations in our context: (i) inconsistent replicability; classification results can vary even with the temperature set to 0; (ii) evaluation challenges, as we lack standard metrics to assess performance when expert knowledge is required for proper subject area labeling. To address these issues, we trained a supervised multi-label classifier using SciBERT. Because Scopus is a query-based database (requiring DOI, PMID, or Scopus EID), our publication data is limited. To broaden coverage, we used publications authored by scientists on corporate boards from Masulis and Yao (2025) as training data. This dataset contains NIH neuroscience-funded research but also spans multiple scientific areas, which is helpful given BI's interdisciplinary nature.

A challenge with this dataset is class imbalance: most publications cluster in only eight fields. Masulis and Yao (2025) addressed this by excluding fields representing less than 5% of publications,[37] but we could not do so because BI covers a wide range of disciplines. Instead, we report performance for all subject areas in Table IA.3, distinguishing between qualified models (F1-score ≥ 0.55) and unqualified models. Eleven subject areas meet this threshold, including all major BI-relevant fields: Neuroscience; Biochemistry, Genetics & Molecular Biology; Engineering; Computer Science; Medicine; Physics & Astronomy; Chemistry; Immunology & Microbiology; Pharmacology, Toxicology & Pharmaceutics; Earth & Planetary Sciences; and Dentistry.

Comparing the outputs, the generative AI method assigns an average of 3.13 subject areas per grant, while SciBERT assigns only 1.4. This difference reflects model priorities: generative AI tends to over-include labels (not penalized for false positives), whereas supervised models focus on precision and are penalized for incorrect labels. Despite these technical differences, both methods point to the same conclusion. BI grants are more interdisciplinary – frequently combining neuroscience with computer science, engineering, and physics – while non-BI grants concentrate in traditional biomedical domains such as medicine, immunology, and pharmaceutics. Interestingly, the SciBERT model (Table IA.3) also predicts BI grants draw on engineering and computer science 3 times more than non-BI grants, similar to the ChatGPT model (Tables 1).

[37] Their rationale was that since their objective was to link patents to the core fields in which directors actively publish, they treated sparsely represented areas as relatively unimportant to their specific research question and omitted them from their model.

B Commercial and Scientific Impact of BI's Output

In this section, we focus on measuring the commercial impact of BI-funded publications through three distinct measures. Two of these measures build on the principle that academic publications demonstrate greater commercial impact when they influence follow-on patents. Specifically, we examine (1) the number of patents citing a publication and (2) the commercial potential metric developed by Masclans et al. (Forthcoming). As a third measure, we employ a machine learning (ML) model to predict BI influence on startups' Neuro patents. We first discuss the two measures based on patent citations, then examine the ML model.

Commercial Value of BI Based on Patent Citations Our first measure uses Reliance on Science data by Marx and Fuegi (2020, 2022) to count patents citing each academic article. The second measure uses Masclans et al. (Forthcoming) data, which predicts an academic article's commercial potential based on its probability of being cited by a renewed patent. We compare these impact measures with those of NIH non-BI-funded publications.[38] This group serves as a close benchmark that controls for both the unique features of the field and research funded by the main public agency behind the BI. For example, medical research publications generally receive more citations than those in other disciplines, and funding availability across different fields can affect their commercialization.

BI publications receive about 4 times more patent citations: an average of 0.44 patent citations for BI compared to 0.12 for non-BI publications, with this difference being statistically significant. We test this more formally in Panel A of Table IA.4 via a Poisson regression model.[39] In Column (1), we regress the number of citations across all publications on a dummy variable for BI publications and include year fixed effects to capture time trends in citations. In Column (2), we repeat this exercise for a subsample including only BI publications and non-BI NIH-backed neuroscience papers.

Furthermore, we analyze the commercial potential of BI and non-BI publications using the measurement from Masclans et al. (Forthcoming), who argue that patent citations reflect the ex post commercialization of an academic article but not its ex ante potential. They therefore develop a large language model trained on a dataset of renewed patents—such patents are presumably more commercialized. This trained model generates two scores for

[38] See Appendix A3 for more detail.

[39] Since most publications never receive a patent citation, we use a Poisson model to accommodate an outcome variable with many zero values (Cohn et al., 2022).

each publication: *Commercial Potential* and *Scientific Potential*. We link these scores to the academic articles in our sample. In Column (3) of Panel A, we repeat the exercise in the last two columns, this time controlling for the two variables. Our significant positive coefficients persist. Overall, these results indicate that BI publications receive substantially higher patent citations.

Panel B directly reports the summary statistics on *Commercial Potential* and *Scientific Potential*. The average commercial potential of non-BI grants after 2014 is 0.69, which is lower than the commercial potential of BI publications at 0.78. The difference between the commercial potential of BI and non-BI grants is statistically significant at the 1% level across both mean and median, suggesting that research under BI grants is more commercializable than similar research grants in the sample period.[40]

Influence Based on Machine Learning Direct patent citations are unlikely to fully capture BI's influence for two reasons. First, the BI occurred in the recent half of our sample, causing truncation issues with patent citations (Lerner and Seru, 2021). For example, Lerner, Manley, Stein, and Williams (2024) find that it takes 3.7 years for a publication to receive its first patent citation. A more important issue is that relying on citation paper trails overlooks a significant amount of spillovers (Myers and Lanahan, 2022). This is exacerbated in our setting due to the imperfect appropriability of basic science. Patents tend to cite prior art immediately related to the invention, meaning the broader scientific foundation, open data, tools, and methods used in the invention process are less frequently cited. For example, while the BI-funded Cell Census Network (BICCN) helps identify cells that stop functioning in Parkinson's disease, the statistical models underlying BICCN may be too abstract for citation in Parkinson-related patents.

USPTO's chief economist argues that Machine Learning (ML) is a powerful tool to overcome these limitations (Toole et al., 2020). We therefore fine-tune a large language model to identify *Neuro patents* that are influenced by the BI publications. We use SciBERT – a BERT model trained on 1.14M scientific articles (Devlin, Chang, Lee, and Toutanova, 2018) – as a well-equipped model to understand scientific text. We fine-tune the model to calculate a score between 0 and 1 to measure BI's influence on a given patent. To avoid a look-ahead bias, we develop a separate model for every year from 2015 to 2020, such that the patents

[40] We cannot perform a DiD analysis here because BI grants did not exist before 2014.

are only influenced by the knowledge generated up to that point. For example, a patent in 2017 cannot be influenced by the knowledge generated in 2018.

Supervised ML models typically require a balanced labeled dataset for effective training (Lemaitre, Nogueira, and Aridas, 2017; He and Garcia, 2009). In our training data, the positive cases are BI-funded publications, constituting a relatively small fraction of the overall publication pool. We face the sample imbalance issue if we include all other publications as negative cases. To address this, we follow the standard practice of randomly selecting papers funded by non-BI grants for this pool. Publications from non-BI grants are a strong baseline for distinguishing between the influence of BI and other grant-funded research (Masclans et al., Forthcoming; Giczy et al., 2022). All the negative cases are papers published between 2009 and 2013, which proxy for pre-existing neuroscience. Our results are robust when using alternative randomly selected negative cases.

We divide our labeled sample into three sets: 80% for training, 10% for testing, and 10% for validation. Table IA.5 reports model performance. All models achieve weighted average F1-scores above 0.9, demonstrating high prediction accuracy. Beyond conventional ML performance metrics, we validate our model using patent citations. We test whether our model can predict patents that directly cite BI publications as patents influenced by BI research. Our model correctly predicts that 87.54% of these patents are influenced by BI research. Using our trained model, we find that 66% of *Neuro patents* from VC-backed startups in our sample are influenced by BI research.

C Bias from Exposed Controls

This appendix derives Equation (2). Let Δ captures the parallel evolution of outcomes across groups absent treatment, under the parallel trends assumption:

$$\mathbb{E}[Y(0) \mid D{=}1, T{=}1] - \mathbb{E}[Y(0) \mid D{=}1, T{=}0] = \mathbb{E}[Y(0) \mid D{=}0, T{=}1] - \mathbb{E}[Y(0) \mid D{=}0, T{=}0] \equiv \Delta$$

Let $\beta \equiv \mathbb{E}[Y(1) - Y(0) \mid D{=}1]$ denote the average treatment effect on the treated. We assume a homogeneous treatment effect across exposed units:[41] $\mathbb{E}[Y(1) - Y(0) \mid \text{exposed}] = \beta$. If a fraction $p \in (0, 1)$ of control firms are exposed in the post period (but still labeled

[41] While heterogeneous treatment effects for the exposed group may alter the magnitude of the bias term, the same downward attenuation of the DiD estimate whenever some controls are exposed will hold.

$D = 0$), we can write the four terms in Equation (1) as:

$$\begin{aligned}
\mathbb{E}[Y \mid D{=}1, T{=}0] &= \mathbb{E}[Y(0) \mid D{=}1, T{=}0], \\
\mathbb{E}[Y \mid D{=}1, T{=}1] &= \mathbb{E}[Y(0) \mid D{=}1, T{=}1] + \beta, \\
\mathbb{E}[Y \mid D{=}0, T{=}0] &= \mathbb{E}[Y(0) \mid D{=}0, T{=}0], \\
\mathbb{E}[Y \mid D{=}0, T{=}1] &= (1-p)\ \mathbb{E}[Y(0) \mid D{=}0, T{=}1] + p\ \mathbb{E}[Y(1) \mid D{=}0, T{=}1] \\
&= \mathbb{E}[Y(0) \mid D{=}0, T{=}1] + p\,\beta.
\end{aligned}$$

Now we can write the within-group differences as:

$$\begin{aligned}
\text{Treated diff} &= \mathbb{E}[Y \mid D{=}1, T{=}1] - \mathbb{E}[Y \mid D{=}1, T{=}0] \\
&= \big(\mathbb{E}[Y(0) \mid D{=}1, T{=}1] + \beta\big) - \mathbb{E}[Y(0) \mid D{=}1, T{=}0] = \Delta + \beta, \\
\text{Control diff} &= \mathbb{E}[Y \mid D{=}0, T{=}1] - \mathbb{E}[Y \mid D{=}0, T{=}0] \\
&= \big(\mathbb{E}[Y(0) \mid D{=}0, T{=}1] + p\beta\big) - \mathbb{E}[Y(0) \mid D{=}0, T{=}0] = \Delta + p\beta.
\end{aligned}$$

which allows deriving the DiD estimator as

$$\widehat{\beta}_{\text{DiD}} = \underbrace{(\Delta + \beta)}_{\text{treated diff}} - \underbrace{(\Delta + p\beta)}_{\text{control diff}} = (1-p)\,\beta.$$

Hence, the DiD estimate is attenuated by the fraction of contaminated controls, and the true effect is

$$\beta = \frac{\widehat{\beta}_{\text{DiD}}}{1-p}.$$

D Similarity to NeuroTech: SimilarTech vs. DistantTech

In Section 3.3, we classify patents into *SimilarTech* or *DistantTech* relative to *Neuro* patents based on CPC codes. This classification assumes that *SimilarTech* patents are technologically closer to *Neuro patents* than *DistantTech* patents. Here, we test this assumption by measuring the textual similarity between *Neuro* and these patents in the same spirit as (Pham, Rezaei, and Zein, 2023).

We start with the patents of firms in our extensive margin sample (both VC-backed and non-VC-backed startups). For this sample, we combine the title and abstract to capture the patent's technological content. We then remove common English stop words, as well as extremely rare or overly frequent terms, to reduce noise. The cleaned texts are vectorized using Term Frequency–Inverse Document Frequency (TF–IDF), and cosine similarity is computed between each Neuro startup patent and all patents from other startups filed in the same year, to ensure contemporaneous comparisons. This yields pairwise similarity scores for each *Neuro patent* against all other patents. Based on these scores, we conduct two tests, one at the patent and one at the firm level.

In our patent-level test, we compare each *Neuro patent*'s closest technological proximity to patents in the *Similar* and *Distant* groups. For each *Neuro patent*, we identify two values: (i) the maximum textual similarity to any *Similar* patent, and (ii) the maximum textual similarity to any *Distant* patent. Using these two observations per *Neuro patent*, we regress the maximum similarity score on an indicator variable, *Similar*, which equals 1 if the observation reflects the highest similarity with the *Similar* group and 0 if it reflects the highest similarity with the *Distant* group. The specification includes fixed effects for the focal *Neuro patent* and the filing year. This enables us to compare the two similarity measures for the same *Neuro patent* and capture time trends.

The results, reported in Column (1) of Table IA.6, show that the coefficient on *Similar* is 0.084 and significant at the 1% level. This implies that, on average, the closest textual match between a *Neuro patent* and *SimilarTech* patents is 0.084 points higher than the closest match with *Distant* patents. Relative to the sample median similarity score of 0.121, this represents roughly 69% of the median value, indicating a substantial economic magnitude.

Nonetheless, our *SimilarTech* and *DistantTech* control groups are defined at the startup (not patent) level (Section 3.3.2). A *Neuro startup* may own patents outside neuroscience. The similarity between such patents may also bias the classification in either way. To check this, we move the analysis to the startup level. For every *Neuro startup* and every other startup, we compute the *Max Similarity of Firm Pair* as the maximum cosine similarity between any patent from the Neuro startup and any patent from the other startup filed in the same year (so the comparison is contemporaneous). This gives us the closest technological

proximity between each startup pair.[42]

We then regress this measure on an indicator, *SimilarTech Startup*, which equals 1 when the comparison startup is in the *SimilarTech* group and 0 when it is in the *DistantTech* group. The specification includes *Neuro startup* fixed effects and filing-year fixed effects for the year of the closest patent pair. Column (2) of Table IA.6 shows a coefficient of 0.004 on *SimilarTech Startup* significant at 1%. This equals the sample median of the dependent variable (0.004), i.e., one full median unit of the max-similarity measure. This supports that, even at the startup level, *DistantTech* firms are in fact technologically distant too.

E Pitchbook Verticals

The Pitchbook verticals are the following: Manufacturing, FoodTech, TMT, Mobile, Oil & Gas, Industrials, AudioTech, Virtual Reality, FinTech, InsurTech, LOHAS & Wellness, Internet of Things, SaaS, Gaming, Space Technology, CleanTech, Real Estate Technology, Big Data, Robotics and Drones, CloudTech & DevOps, Mobility Tech, Life Sciences, Nanotechnology, Digital Health, HealthTech, Wearables & Quantified Self, 3D Printing, Artificial Intelligence & Machine Learning, AdTech, Marketing Tech, HR Tech, Climate Tech, Beauty, E-Commerce, Infrastructure, Supply Chain Tech, Oncology, EdTech, Cybersecurity, AgTech, Legal Tech, Impact Investing, Cryptocurrency/Blockchain, Esports, Mobile Commerce, Advanced Manufacturing, Cannabis Construction Technology, Augmented Reality, Restaurant Technology, B2B Payments, Mortgage Tech, Micro-Mobility, Ridesharing, FemTech, Autonomous cars, Car-Sharing Pet Technology, Ephemeral Content.

[42] With 1,455 Neuro startups (extensive margin) and 55,363 non-Neuro startups overall, there are in principle 80,553,165 pairs, but because our pairs are restricted to firms that share at least one filing year, we have 18,786,887 valid pairs.

Appendix Figures and Tables

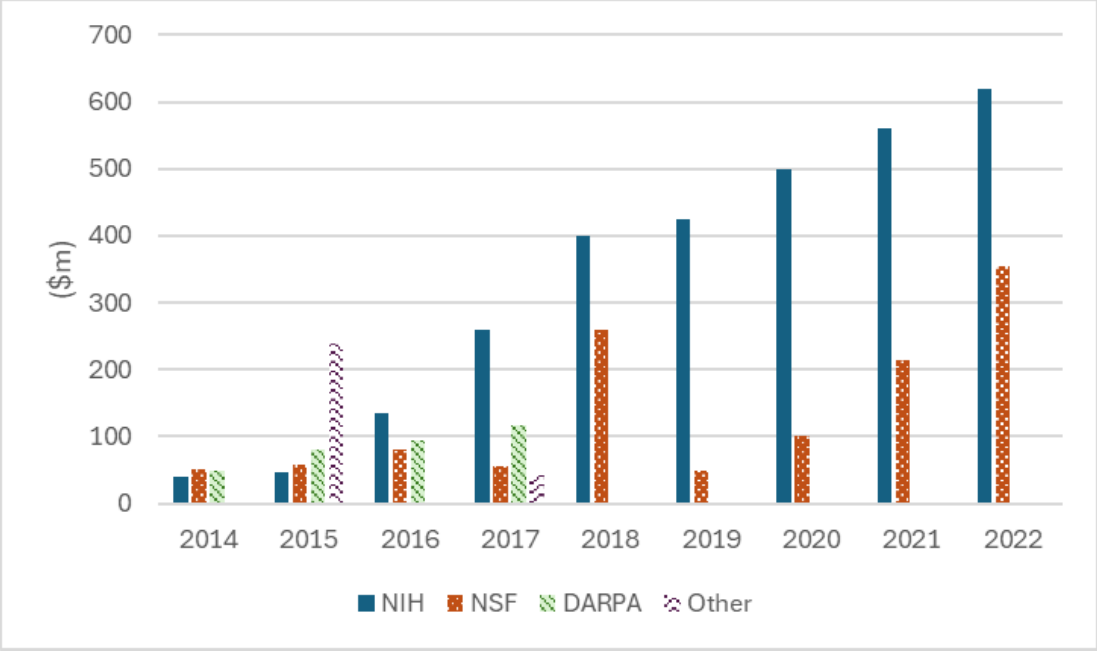


Figure IA.1: **Total BRAIN Initiative Funding per Agency**

This figure shows the total funding of the BRAIN Initiative (BI) by the funding organization.

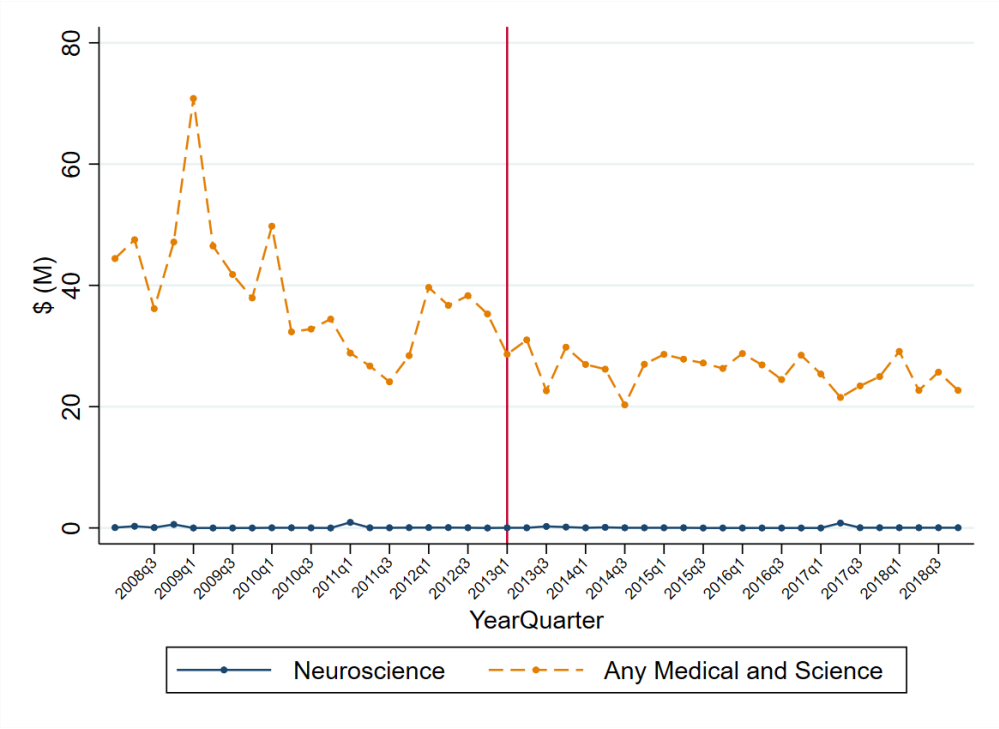


Figure IA.2: **Science and Neuroscience Lobbying Activity**

This figure shows quarterly lobbying spending by Compustat-listed companies. The orange dashed line shows total spending on issues coded SCI (Science/Technology) or MED (Medical/Disease Research). The navy solid line isolates the subset of that spending that explicitly references neuroscience keywords. The vertical red line at 2013 Q1 marks President Obama's public launch of the BRAIN Initiative. Data is sourced from LobbyView (Kim, 2018).

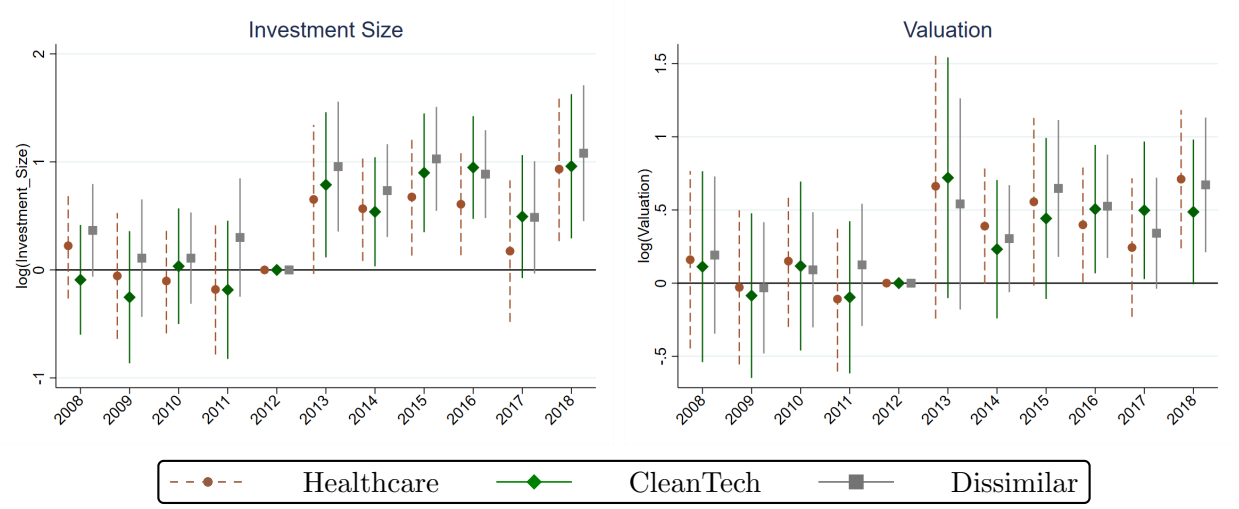


Figure IA.3: **Investment Size and Valuation (1st Rounds)**

The figure reports difference-in-differences coefficients from the dynamic (event study) specification of Equation (5), in three samples with different control groups. The dependent variables are the log of the financing amount (left) and the log of the pre-money valuation (right). The unit of observation is an entrepreneurial firm's financing event (only first VC rounds). For each calendar year, we interact the *Neuro startup* indicator with a year dummy and plot the resulting coefficient. The omitted (zero) year is 2012, the last pre-BRAIN year. Brown circles show coefficients when the control group is *Healthcare*; green diamonds when it is *CleanTech*; gray squares when it is *Dissimilar*. Vertical lines give 95 percent confidence intervals based on firm-clustered standard errors.

Table IA.1: Variables Definitions

Variable name	Definitions	Tables
Independent variables		
Neuro	The indicator variable equals one if the startup is a *Nuero Startup*; zero otherwise. *Neuro Startup* is identified as startups granted at least one patent within neuroscience-related technology groups, as detailed in Section 3.2.	Table 3, 4, 7, 8, 9, 10, IA.8
Syndicate Size	The natural logarithm of the number of VCs in the financing round.	Table 4, 5, 6, IA.7, IA.8
Employee Exposure	The indicator variable equals one if the firm has hired at least one BI scientist up to the financing year and zero otherwise. BI scientists are authors of BI publications.	Table 5
Ln(Patent Exposure+1)	The number of patents influenced by the BI up to the financing year, as identified by our large language model (Section B)	Table 5
Ln(State Exposure+1)	The number of BI grants awarded to the institutions in the startup's state.	Table 5
Ln(Raised before exit)	The natural logarithm of the total amount of financing that the startup has raised before the exit of VC.	Table 7
Ln(Total Raised)	The natural logarithm of the total amount of financing that the startup has raised up to the year.	Table 9
Ln(Age+1)	The natural logarithm of one plus the number of years since the founding year.	Table 3, 9
Cancer	The indicator variable equals to one for startups classified in Pitchbook's Oncology vertical.	Table 6
Post16	The indicator variable equals to one for financing events occurring in or after 2016, the year of the Cancer Moonshot announcement.	Table 6
BI	The indicator variable equals one for publications from BI grants and zero otherwise.	Table IA.4
Non-BI Neuro	The indicator variable equals one for publications from NIH-funded non-BI neuroscience grants and zero otherwise.	Table IA.4
Commercial Potential	The commercial potential of scientific publications from Masclans et al. (Forthcoming)	Table IA.4
Scientific Potential	The scientific potential of scientific publications from Masclans et al. (Forthcoming)	Table IA.4
Similar	The indicator variable equals 1 if the patent shares a technology class with *Neuro patents*, and 0 if it does not.	Table IA.6
SimilarTech	The indicator equals 1 if the startup has at least one *Similar* patent	Table IA.6
NeuroDescript	The indicator variable for startups whose Pitchbook's business description or keywords contain any of the Neuro keywords defined in Section 3.2	Table IA.7
NeuroPreVC	The indicator variable for startups who have filed for their first Neuro patent at or prior to the first year they raise VC.	Table IA.7
NeuroPat	The indicator variable equals one for a neuro-related patent and zero otherwise.	Table IA.9
Ln(Citations+1)	The natural logarithm of the number of patent citations plus one.	Table IA.9
Dependent Variables		
Raised VC	The indicator variable equals one if the startup has raised any VC funding by year t, and zero otherwise.	Table 3
Ln(Investment Size)	The natural logarithm of VC financing amount, excluding below $0.1m financing amounts.	Table 4, 5, 6, IA.7, IA.8
Ln(Valuation)	The natural logarithm of VC Pre-Money Valuation, excluding below $0.1m valuations.	Table 4, 5, 6, IA.7, IA.8
Successful Exit	The indicator variable equals one for startups' successful exit. A successful exit is an IPO or an M&A at a reported value at least twice the total capital invested.	Table 7
Ln(Time to Exit)	The natural logarithm of the difference between the number of days between the first VC investment and the VC exit date.	Table 7
Academic Startup	The indicator variable equals one if the startup is an Academic Startup, defined as a startup with at least one senior employee with an academic background within the three years after being founded, and zero otherwise.	Table 8
#Patents	Startup i's the total number of patents filed (and eventually granted) in year t	Table 9
#Breakthrough Patents	Startup i's the number of breakthrough patents filed (and eventually granted) for the next n years. The breakthrough patents at the 90 percentile are patents that received more citations than the citations at the 90 percentile within the same technology class and year.	Table 9
Avg. Patent Citations	Startup i's the average adjusted cites of patents filed (and eventually granted) in year t. The adjusted cites are the number of cites over the average cites of patents in the same technology field and granted year.	Table 9
#Academic inventors hired	The number of Academic inventors hired by the startup at year t. Academic inventors are inventors who begin working in startups following their academic roles or upon finishing their doctoral degrees.	Table 9
#AI Patent	The number of AI-related patents filed (and later granted) in year t	Table 9
ComplementaryTech	The indicator variable that equals one if the startup operates in at least one complementary technology sector (Artificial Intelligence and Machine Learning, Big Data, Wearables & Quantified Self, Digital Health, or HealthTech) according to PitchBook's vertical classifications, and zero otherwise.	Table 10
ComplementaryTech2	The indicator variable that equals one if the startup operates in at least one of the core complementary technology sectors (Artificial Intelligence and Machine Learning, Big Data, or Wearables & Quantified Self) according to PitchBook's vertical classifications, and zero otherwise.	Table 10
Commercial Potential	The commercial potential of scientific publications from Masclans et al. (Forthcoming)	Table IA.4
#Patent Citations of Publications	The number of patent citations received per publication from Marx and Fuegi (2022, 2020)	Table IA.4
Max Similarity of Patent Pair	The maximum pairwise similarity between neuro patents and other patents.	Table IA.6
Max Similarity of Firm Pair	The maximum pairwise similarity between neuro startups and other startups.	Table IA.6
Ln(Patent Value)	The natural logarithm of value of a patent that obtained from Kogan et al. (2017).	Table IA.9

Table IA.2: BI grants vs. Non-BI grants This table compares neuroscience funding under the BRAIN Initiative with comparable NIH non-BI neuroscience grants. All amounts are in millions of U.S. dollars. Column (1) reports the total BI budget by fiscal year; Column (2) reports the total for non-BI grants. *Average Amount per Project* is calculated as the total awarded amount divided by the number of projects funded by each agency. Columns (6) and (7) report differences in average amounts: (3)–(5) and (4)–(5), respectively. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively, based on two-sided t-tests.

	Total Funding ($m)		Average Amount per Project ($m)			Diffs	
FY	BI	NIH non-BI	NIH BI	NSF BI	NIH Non-BI	(3)-(5)	(4)-(5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2000		1,251					
2001		1,469					
2002		1,711					
2003		1,902					
2004		2,076					
2005		2,136					
2006		2,124					
2007		3,529					
2008		3,706					
2009		4,552					
2010		4,586					
2011		4,129					
2012		4,351					
2013		4,205					
2014	142	4,513	0.8	0.37	0.38	0.42***	0
2015	425	4,618	0.57	0.44	0.39	0.18***	0.05
2016	312	5,921	0.65	0.65	0.42	0.23***	0.23***
2017	477	6,489	1.47	0.49	0.45	1.03***	0.04
2018	659	7,239	0.89	2.4	0.47	0.42***	1.93***
2019	472	8,249	1.07	0.63	0.5	0.57***	0.13
2020	600	8,658	1.22	0.75	0.53	0.69***	0.22***
2021	775	8,996	1.27	1.73	0.56	0.71***	1.17***
2022	974	9,450	1.66	2.79	0.57	1.09***	2.22***
Total Amount	4,837	64,132					

Table IA.3: Subject Areas of NIH Neuroscience Grants from SciBERT This table presents the SciBERT supervised multi-label model classification of Scopus subject areas for the abstract of BI grants and NIH neuroscience grants outside BI (non-BI). The construction and output is discussed in A4. We split all Scopus subject areas into those associated with qualified models, defined as models with an F1-score greater than 0.55, and those associated with unqualified models, which have an F1-score below 0.55. Columns (1) and (2) report the percentage of BI and non-BI grants that fall into that category for each subject area, respectively. The percentages in Columns (1) and (2) sum to more than 100%, because a single grant can belong to multiple subject areas. Column (3) reports the ratio of BI to non-BI shares for each subject area. Columns (4), (5), and (6) show the model performance, including precision, recall, and F1-score.

Panel A: Qualified Models (F1-Score $\geq$ 0.55)

	BI vs. Non-BI Grants			Model Performance		
	(1)	(2)	(3)	(4)	(5)	(6)
Scopus Subject Areas	%BI Grants	%Non-BI Grants	BI/Non-BI	Precision	Recall	F1-score
Neuroscience	62.08	29.23	2.12	0.63	0.54	0.58
Biochemistry, Genetics and Molecular Biology	35.36	31.10	1.14	0.83	0.73	0.78
Engineering	15.72	5.00	3.14	0.76	0.71	0.73
Computer Science	11.98	3.71	3.23	0.82	0.83	0.83
Medicine	5.11	32.65	0.16	0.85	0.85	0.85
Physics and Astronomy	3.54	0.38	9.32	0.73	0.62	0.67
Chemistry	1.18	0.57	2.07	0.71	0.52	0.60
Immunology and Microbiology	0.20	1.05	0.19	0.70	0.57	0.62
Pharmacology, Toxicology and Pharmaceutics	0.00	1.00	0.00	0.71	0.45	0.55
Earth and Planetary Sciences	0.00	0.00	NA	0.82	0.52	0.64
Dentistry	0.00	0.04	0.00	0.72	0.54	0.62

Panel B: Unqualified Models (F1-Score < 0.55)

	BI vs. Non-BI Grants			Model Performance		
	(1)	(2)	(3)	(4)	(5)	(6)
Scopus Subject Areas	%BI Grants	%Non-BI Grants	BI/Non-BI	Precision	Recall	F1-score
Agricultural and Biological Sciences	1.97	0.40	4.93	0.70	0.44	0.54
Mathematics	1.57	0.65	2.42	0.61	0.07	0.13
Materials Science	1.57	0.36	4.36	0.66	0.44	0.53
Psychology	0.00	1.52	0.00	0.54	0.17	0.25
Nursing	0.00	0.30	0.00	0.70	0.36	0.48
Environmental Science	0.00	0.43	0.00	0.60	0.23	0.34
Health Professions	0.00	0.08	0.00	0.54	0.08	0.14
Energy	0.00	0.00	NA	0.58	0.42	0.49
Chemical Engineering	0.00	0.17	0.00	0.62	0.24	0.34
Veterinary Science	0.00	0.00	NA	0.77	0.27	0.40

Table IA.4: Commercial Potential of BI research This table compares the commercial potential of BI research against research from non-BI neuroscience grants. Panel A presents the results of the Poisson regression of the *#Patent Citations* received by publications on the BI grant indicator. A unit of observation is a publication. The key dependent variable, *#Patent Citations*, represents the number of patent citations each publication receives. The BI indicator variable equals one for publications resulting from BI grants and zero otherwise. Control variables include *Non-BI Neuro*, an indicator for publications from NIH-funded non-BI neuroscience grants, as well as measures of commercial and scientific potential. All regressions include year-fixed effects. Columns (1) and (3) report regression results for the full sample, while Column (2) focuses specifically on publications from NIH-funded grants in neuroscience. Panel B utilizes the predicted ex-ante commercial potential of publications from Masclans et al. (Forthcoming). Column (1) in Panel B comprises publications from BI grants, whereas Column (2) includes all publications from non-BI neuroscience grants after 2014. The last column in Panel B presents the difference between Groups A and B and reports the statistical significance of the t-test for mean differences and the Wilcoxon rank-sum test for median differences.

Panel A: Patent Citations of Publications			
	(1)	(2)	(3)
	#Patent Citations	#Patent Citations	#Patent Citations
BI	2.022	1.192	1.485
	(0.301)***	(0.307)***	(0.301)***
Non-BI Neuro			0.428
			(0.056)***
Commercial Potential			4.361
			(0.065)***
Scientific Potential			0.383
			(0.052)***
Observations	2,274,602	83,838	2,274,602
Year FE	Y	Y	Y

Panel B: Predicted Commercial Potential of Publications			
	(1)	(2)	(3)
	BI	Non-BI	(1)-(2)
Mean	0.78	0.69	0.09***
Median	0.84	0.78	0.06***
SD	0.17	0.26	

Table IA.5: Performance of Machine Learning Models for BI Influence on Patents This table presents the performance metrics for our fine-tuned SciBERT models of the influence of BI knowledge on patents from 2015 to 2020. The metrics are based on a testing dataset that was not used for training. For each class (positive and negative), we report precision, recall, and F1-score. Precision measures the proportion of predicted positive cases that are truly positive. Recall reflects the proportion of actual positive cases correctly identified. The F1-score, the harmonic mean of precision and recall, provides a balanced measure of model performance. We also report macro and weighted averages across classes: macro average calculates each metric independently for each class and then takes the unweighted average, while the weighted average accounts for class size when averaging the metrics. Further details on how we fine-tuned the SciBERT model are provided in Section 2.1.

2015

	Precision	Recall	F1-Score	Support
0	0.90	0.96	0.93	28
1	0.97	0.90	0.93	31
Maro avg	0.93	0.93	0.93	59
Weighted avg	0.93	0.93	0.93	59

2016

	Precision	Recall	F1-Score	Support
0	0.92	0.90	0.91	81
1	0.89	0.92	0.91	74
Maro avg	0.91	0.91	0.91	155
Weighted avg	0.91	0.91	0.91	155

2017

	Precision	Recall	F1-Score	Support
0	0.92	0.91	0.92	166
1	0.89	0.90	0.89	130
Maro avg	0.90	0.90	0.90	296
Weighted avg	0.91	0.91	0.91	296

2018

	Precision	Recall	F1-Score	Support
0	0.96	0.90	0.93	268
1	0.89	0.95	0.92	241
Maro avg	0.93	0.93	0.93	509
Weighted avg	0.93	0.93	0.93	509

2019

	Precision	Recall	F1-Score	Support
0	0.90	0.92	0.91	373
1	0.93	0.91	0.92	409
Maro avg	0.92	0.92	0.92	782
Weighted avg	0.92	0.92	0.92	782

2020

	Precision	Recall	F1-Score	Support
0	0.92	0.96	0.94	559
1	0.95	0.92	0.94	551
Maro avg	0.94	0.94	0.94	1110
Weighted avg	0.94	0.94	0.94	1110

Table IA.6: Textual Proximity of NeuroTech to SimilarTech vs. DistantTech. This table tests whether patents and startups in the *SimilarTech* group are technologically closer to *NeuroTech* than those in the *DistantTech* group. Column (1) uses a patent-level design: for each *Neuro patent*, we compute its maximum cosine similarity (based on TF–IDF of title and abstract) to any *Similar* patent and to any *Distant* patent filed in the same year. The dependent variable is this maximum similarity score across each group, and the regressor *Similar* equals 1 for the *SimilarTech* match and 0 for the *DistantTech* match; *Neuro patent* and filing-year fixed effects control for patent-specific and time variation. Column (2) shifts the analysis to the startup level: for each *Neuro startup* and other startup pair with overlapping filing years, we record their highest patent-to-patent similarity of their patents filed in the same year. The dependent variable is this firm-pair maximum similarity; *SimilarTech Startup* equals 1 if the comparison firm is in *SimilarTech*, and 0 if in *DistantTech*. The fixed effects and clustering levels are listed at the bottom of the table. t-statistics are in parentheses, with standard errors clustered as reported. Significance levels: *** 1%, ** 5%, * 10%.

Outcome Variable:	Max Similarity of Patent Pair	Max Similarity of Firm Pair
	(1)	(2)
Similar	0.084***	
	(29.04)	
SimilarTech Startup		0.004***
		(31.65)
Observations	8,960	18,786,887
Mean Outcome	0.166	0.010
Median Outcome	0.121	0.004
Adj R-squared	0.22	0.22
Neuro Patent FE	Yes	No
Neuro Startups FE	No	Yes
Year FE	Yes	Yes
Cluster	Neuro Patent	Neuro & comparison Startups

Table IA.7: Funding Size and Valuation (Alternative Definitions of Neuro Startups).

This table presents the results of OLS regressions estimating Equation 5. The dependent variable is ln(Investment Size) in both panels. The unit of observation is a VC financing event. Columns (1), (3), (5), and (7) include all financing rounds and other columns are restricted to the first VC rounds. In Panel A, *NeuroDescript* is an indicator variable for startups whose Pitchbook's business description or keywords contain any of the *Neuro keywords* defined in Section 3.2. In Panel B, *NeuroPreVC* is an indicator variable for startups that have filed for their first *Neuro patent* at or prior to the first year they raise VC. *Syndicate Size* is ln(#investors). *Post* equals one for financing events on or after 2013: announcement year of the BRAIN Initiative. Control groups for Columns (1-2) are all untreated VC-backed startups, which in Columns (3-4), (5-6) and (7-8) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. These groups are defined in Section 3.3. Fixed effects for both panels are presented at the bottom of the table and include dummies for the year of the financing event, Pitchbook's 41 industry groups, the state of the startup's headquarters and the sequence of financing event (odd columns). The t-statistics (in parentheses) are clustered by firm in Columns (1), (3), (5), and (7) and are based on heteroskedasticity-robust standard errors in Columns (2), (4), (6), and (8). ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Keywords	Ln(Investment Size)							
Control Group:	All		Dissimilar		CleanTech		Healthcare	
VC Round #:	Any	1st	Any	1st	Any	1st	Any	1st
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NeuroDescript×Post	0.125*	0.355***	0.165**	0.384***	0.264***	0.521***	0.109	0.267**
	(1.68)	(3.29)	(2.18)	(3.55)	(3.05)	(4.16)	(1.39)	(2.24)
NeuroDescript	-0.043	-0.213***	0.025	-0.158*	-0.170*	-0.406***	-0.022	-0.134
	(-0.68)	(-2.59)	(0.37)	(-1.87)	(-1.69)	(-3.18)	(-0.33)	(-1.54)
Syndicate Size	0.531***	0.339***	0.511***	0.336***	0.598***	0.428***	0.692***	0.547***
	(81.65)	(38.25)	(71.70)	(35.91)	(28.55)	(12.18)	(44.59)	(20.99)
Observations	79888	28953	64623	24978	7793	2487	15690	5333
Mean Outcome	1.483	0.681	1.350	0.614	1.699	0.816	1.706	0.936
Adj R-squared	0.382	0.189	0.374	0.192	0.404	0.226	0.391	0.225
Panel B: Patent before VC	Ln(Investment Size)							
NeuroPreVC×Post	0.236*	0.619***	0.275**	0.667***	0.353**	0.652***	0.206	0.511**
	(1.75)	(3.22)	(2.01)	(3.45)	(2.51)	(3.14)	(1.51)	(2.46)
NeuroPreVC	0.180*	0.160	0.246**	0.207	0.235	0.306*	0.140	0.155
	(1.68)	(1.17)	(2.21)	(1.49)	(1.59)	(1.68)	(1.30)	(1.13)
Syndicate Size	0.531***	0.339***	0.510***	0.335***	0.586***	0.358***	0.692***	0.550***
	(81.57)	(38.32)	(71.21)	(35.80)	(25.06)	(9.46)	(44.21)	(21.15)
Observations	79804	28961	64108	24863	6412	2011	15529	5326
Mean Outcome	1.484	0.684	1.347	0.616	1.716	0.829	1.707	0.952
Adj R-squared	0.382	0.192	0.374	0.194	0.405	0.229	0.389	0.231
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y
VC Round # FE	Y	N	Y	N	Y	N	Y	N

Table IA.8: Funding Size and Valuation (Excluding AI-related Neuro Startups). This table presents the results of OLS regressions estimating Equation 5. The dependent variable is ln(Investment Size) in both panels. The unit of observation is a VC financing event. Columns (1), (3), (5), and (7) include all financing rounds and other columns are restricted to the first VC rounds. In Panel A, *Neuro startups* in *Artificial Intelligence and Machine Learning* and *Big Data* verticals are excluded. In Panel B, those in *CompletementaryTech*, including *Artificial Intelligence and Machine Learning, Big Data and Wearables & Quantified Self, Digital Health and HealthTech*, are excluded. *Syndicate Size* is ln(#investors). *Neuro* is an indicator variable equal to one for *Neuro startups* and zero otherwise. *Post* equals one for financing events on or after 2013: announcement year of the BRAIN Initiative. Control groups for Columns (1-2) are all untreated VC-backed startups, which in Columns (3-4), (5-6) and (7-8) are restricted to *Dissimilar*, *CleanTech* and *Healthcare* startups, respectively. These groups are defined in Section 3.3. Fixed effects for both panels are presented at the bottom of the table and include dummies for the year of the financing event, Pitchbook's 41 industry groups, the state of the startup's headquarters and the sequence of financing event (odd columns). The t-statistics (in parentheses) are clustered by firm in Columns (1), (3), (5), and (7) and are based on heteroskedasticity-robust standard errors in Columns (2), (4), (6), and (8). ***, ** and * represent significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Dropping AI and ML	**Ln(Investment Size)**							
Control Group:	All		Dissimilar		CleanTech		Healthcare	
VC Round #:	Any	1st	Any	1st	Any	1st	Any	1st
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neuro×Post	0.359***	0.691***	0.412***	0.739***	0.428***	0.764***	0.278***	0.562***
	(4.78)	(5.92)	(5.42)	(6.28)	(5.10)	(5.65)	(3.67)	(4.20)
Neuro	0.057	0.011	0.128**	0.064	0.171*	0.006	0.058	0.046
	(1.03)	(0.16)	(2.21)	(0.87)	(1.76)	(0.05)	(1.04)	(0.60)
Syndicate Size	0.535***	0.341***	0.516***	0.337***	0.614***	0.392***	0.697***	0.558***
	(83.07)	(38.72)	(72.84)	(36.22)	(30.51)	(11.19)	(47.13)	(22.00)
Observations	81359	29294	65663	25196	7967	2344	16855	5613
Mean Outcome	1.502	0.692	1.373	0.626	1.855	0.910	1.759	0.978
Adj R-squared	0.386	0.194	0.380	0.198	0.424	0.241	0.396	0.235
Panel B: Dropping CompletemntaryTech								
Control Group:	All		Dissimilar		CleanTech		Healthcare	
VC Round #:	Any	1st	Any	1st	Any	1st	Any	1st
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Neuro×Post	0.326***	0.934***	0.382***	0.988***	0.444***	0.997***	0.290***	0.792***
	(3.64)	(6.38)	(4.20)	(6.71)	(4.63)	(6.17)	(3.37)	(5.06)
Neuro	0.047	0.056	0.104	0.105	0.126	0.144	0.077	0.117
	(0.77)	(0.71)	(1.64)	(1.29)	(1.18)	(1.08)	(1.30)	(1.42)
Syndicate Size	0.533***	0.341***	0.512***	0.338***	0.594***	0.403***	0.696***	0.564***
	(82.36)	(38.68)	(72.07)	(36.18)	(27.85)	(11.07)	(46.11)	(22.00)
Observations	80711	29160	65018	25062	7321	2210	16384	5509
Mean Outcome	1.495	0.691	1.363	0.625	1.809	0.913	1.749	0.982
Adj R-squared	0.383	0.195	0.377	0.198	0.409	0.255	0.394	0.240
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
State FE	Y	Y	Y	Y	Y	Y	Y	Y
VC Round # FE	Y	N	Y	N	Y	N	Y	N

Table IA.9: **Valuation of Listed Firms' Patents Around the BI Announcement**. This table reports results from OLS regressions where the dependent variable is the log valuation of patents obtained from Kogan et al. (2017). A unit of observation is a listed firm's patent. *NeuroPat* is an indicator variable with a value of one for a neuro-related patent and zero for the control group. *Post* equals one for grant dates on or after April 2, 2013, (the announcement date of the BRAIN Initiative) and zero otherwise. Columns (1) and (4) examine a six-months window around the BI's announcement date, Columns (2) and (5) a one-year window, and Columns (3) and (6) examine a two-year window. The analysis in Columns (1–3) use any patent and (4-6) use *Distant Patents* as the control group. All analysis include firm and patent technology class×filing year fixed effects. The *t*-statistics (in parentheses) are clustered at patent technology class ***, ** and * represent significance at the 1%, 5% and 10% levels, respectively.

Ln(Patent Value)						
Control Group	Any Patent			Distant Patens		
Grant Window (relative to BI)	[-6m – +6m]	[-1y – +1y]	[-2y – +2y]	[-6m – +6m]	[-1y – +1y]	[-2y – +2y]
	(1)	(2)	(3)	(4)	(5)	(6)
NeuroPatent×Post	0.042*	0.143***	0.248***	0.084*	0.180***	0.315***
	(1.68)	(4.03)	(11.88)	(1.91)	(4.37)	(12.54)
NeuroPatent	-0.034	-0.117***	-0.149***	-0.004	-0.086**	-0.181***
	(-1.12)	(-9.03)	(-13.88)	(-0.07)	(-2.48)	(-6.17)
Post	0.034***	0.007	-0.096***	0.014	-0.008	-0.131***
	(4.78)	(0.75)	(-4.49)	(1.44)	(-0.70)	(-5.57)
Ln(Citations+1)	0.000	0.003***	0.008***	-0.001	0.004*	0.011***
	(0.04)	(3.15)	(4.70)	(-0.25)	(1.88)	(5.64)
Observations	29716	64278	126720	17438	37749	73818
Mean Outcome	0.749	0.770	0.765	0.560	0.572	0.573
Adjusted R2	0.953	0.951	0.939	0.959	0.957	0.947
Firm FE	Y	Y	Y	Y	Y	Y
Tech Class × Filing Year FE	Y	Y	Y	Y	Y	Y