Public and Private Biomedical R&D: Complements or Substitutes?

Luís Cabral
New York University, Stern School of Business and *CEPR*

Chiara Gardenghi
University of Rochester, Simon Business School

Deepak Hegde
New York University, Stern School of Business and *NBER*

Junrui Lin
University of Texas at Austin

March 2026

Abstract. We develop a two-player multi-period game in which a public innovator and a private innovator allocate effort across research and development over time. We show that, when decisions are simultaneous, innovation efforts are strategic substitutes, as each player's effort reduces the marginal value of the other's. When decisions are sequential, however, innovation efforts are strategic complements, as knowledge accumulated through spillovers reduces future innovation cost.

We test these predictions in the context of U.S. biomedical research over the 2000-2019 period. Leveraging the granting of FDA Accelerated Approval status as a positive shock to private clinical trial activity, we find that the public sector reduces research funding allocated to affected disease areas in the short run—consistent with substitutability—but increases it in the medium-to-long run—consistent with complementarity. A symmetric pattern emerges when exploiting federal initiatives and executive actions as positive shocks to public funding: private clinical trial activity declines in the short run before expanding in the medium-to-long run.

Taken together, our findings suggest that substitutability and complementarity are not competing characterizations but rather two faces of strategic interaction.

We are grateful to seminar participants at the University of Maryland, University of Rochester, University of Toronto Rotman, NYU, and the University of California Irvine for their helpful comments. We also gratefully acknowledge Jiayi Chen for excellent research assistance. All remaining errors are our own.

1. Introduction

In many areas, from defense to healthcare, economists have long debated how public and private innovation efforts interact (Li et al., 2017, Chakravarthy et al., 2016, Toole, 2007, Cohen et al., 2002). The question has direct implications for public policy (should the public sector compensate for the absence or excess of private sector entry?) and for firm strategy (to what extent should a firm account for public investment when allocating its own innovation effort?)

Prior research in healthcare and beyond has typically examined only one direction of this interaction (either the private sector's response to public decisions, or viceversa) and has concluded in favor of either complementarity or substitutability, implicitly treating one as the definitive characterization. For example, Azoulay et al. (2019) estimate that a $10 million boost in public funding for biomedical research leads to a net increase of 2.7 private-sector patents, evidence of complementarity. By contrast, Wallsten (2000) finds that government grants crowd out firm-financed R&D spending, evidence of substitutability.[1]

Which one is it: are private and public innovation efforts complements or substitutes? Our first contribution is to propose a unifying framework that encompasses both simultaneous decisions (where strategic substitutability is the norm) and sequential decisions (where strategic complementarity is the norm). Our second contribution is to empirically test these theoretical predictions in the context of biomedical research.

To begin, we develop a theoretical framework in which a private innovator and a public innovator each decide how much to invest in research and development in every period of a multi-period game. We thus have both within-period simultaneous choices and across-period sequential choices. We assume different objective functions: the public innovator cares about successful innovation by any player, whereas the private innovator receives a positive payoff only if it is the sole innovator.

We show that the nature of the strategic interaction depends not on the type of strategic variable—research versus development—but on the temporal relationship between decisions. When choices are simultaneous, innovation effort levels are strategic substitutes: best-response curves are downward sloping. The intuition for the public innovator is that private innovation reduces the marginal value of public innovation effort through duplication; for the private innovator, public innovation reduces the profitability of private activity through competition. When, instead, one player's effort decision follows the other's, an increase in the leader's effort induces an increase in the follower's. The intuition here is that knowledge accumulated through innovation spillovers reduces the follower's innovation cost, generating complementarity.

We test these theoretical predictions in the context of biomedical innovation. Using supervised machine learning (ML) techniques, we assemble a longitudinal dataset that maps private and public sector activities to major disease-specific areas over twenty pre-COVID years (from 2000 to 2019). We measure U.S. public sector activity through research funding allocated annually by the National Institute of Health (NIH) by disease area, as well as through publicly-funded clinical trials and the resulting research publications. Since private corporations do not disclose their R&D investments by disease area, we focus on privately-sponsored clinical trials as a measure of their drug development effort. For each disease and

1. In a different context, Kingma (1989) shows that a $10,000 increase in government funds for public radio results in a total crowd-out of $1,350, also evidence of substitutability.

year, we additionally collect data on mortality rates, a proxy for disease burden; research publication intensity, a proxy for the fertility of research efforts in each area (e.g., Lichtenberg, 2001; Hegde and Sampat, 2015); and total annual medical expenditure by patients and insurers, a proxy for disease revenue potential.

Testing empirically our theoretical predictions presents several challenges. A first challenge is mapping game-theoretic concepts—such as sequential versus simultaneous decisions—to time periods that reflect the actual technological and regulatory steps that players must undertake to implement their innovation effort (e.g., the time elapsed between receiving funding and running a clinical trial). A second challenge is distinguishing exogenous from endogenous changes in innovation effort. To address these challenges, we employ a staggered event study approach that exploits plausibly exogenous shocks to innovation decisions (Sun and Abraham, 2021). A further advantage of this design is the ability to separately identify short-run and medium-to-long-run effects of each shock (Clemens and Gottlieb, 2014), by estimating the impact on innovation effort at each point in time.

We first examine the *public* sector's best response to increases in private innovation effort. As a source of plausibly exogenous variation in private clinical trial activity, we exploit the granting of FDA Accelerated Approval status to drugs targeting specific disease areas. This designation allows faster approval of drugs treating serious conditions with limited existing therapies, on the basis of early clinical evidence. Diseases in which one or more drugs receive this status are therefore expected to attract increased late-stage private trial activity as firms move to finalize approval. Importantly, the timing of these designations is driven by pre-clinical and early-phase scientific evidence and the regulatory process rather than by broader economic or strategic motives, lending credibility to the identification strategy.

We find that, as expected, receiving the FDA Accelerated Approval designation induces a substantial increase in private clinical trial activity for treated diseases relative to untreated ones in the year following the shock. Moreover, public NIH funding allocated to those same diseases declines beginning approximately two years after the shock—particularly for new research projects and early-stage explorative research. The number of publicly-funded early-stage clinical trials and associated publications also falls around three years after the shock. Taken together, these results indicate that the NIH acts as a substitute in the short run following a positive shock to private innovation. In the medium-to-long run, however, the picture reverses: public funding to treated diseases rises relative to untreated ones around five years following the shock, consistent with the public sector acting as a complement by building on the knowledge accumulated through prior private trial activity.

We then turn to the *private* sector's best response to increases in public funding. As plausibly exogenous shocks to NIH's disease-specific resource allocation, we exploit a sequence of high-profile federal initiatives and executive actions that directed public resources toward particular scientific domains. These include Executive Order 13505, signed by President Obama in 2009, which lifted prior restrictions on NIH funding for human embryonic stem cell research, and the BRAIN Initiative, announced in 2013 to promote the development of neuroscience tools and the study of neurological disorders. Each initiative was motivated by an assessment that scientific conditions were propitious for progress, rather than by contemporaneous trends in private investment, supporting their use as exogenous shifters to public research effort.

We find that diseases targeted by these federal initiatives receive a meaningful increase in NIH funding within the first year following the shock. This funding increase translates,

after approximately four years, into a significant rise in early-stage publicly-sponsored clinical trial activity—a lag that is informative in its own right, as it reveals how long it takes the public sector to translate research funds into clinical activity, and provides a benchmark for interpreting the private sector's response. Around the same time, privately-sponsored trial activity in treated disease areas begins to decline relative to untreated diseases, suggesting that private companies redirect resources away from diseases experiencing increased public investment roughly contemporaneously with—or shortly after—the public shock. This pattern points to the private sector acting as a substitute in the short run. In the medium-to-long run, however, private clinical trial activity trends upward for treated diseases five to six years following the public shock. This finding indicates that the private sector eventually acts as a complement, taking advantage of the knowledge generated by prior public investment and publicly-run trials.

Several normative questions remain open. These include, for example, whether the observed pattern of best responses is efficient (is sequential complementarity optimal, or should each sector instead specialize in areas of comparative advantage?); and what objective function the public sector ought to maximize (for instance, quality-adjusted lives saved subject to a budget constraint, or the wellbeing of the worst-off patients). We are currently exploring how to address at least some of these questions.

2. Theoretical Framework

Consider a game evolving over two periods ($t = 1, 2$) and involving two innovators ($i = u, v$). The superscript u stands for the public innovator, whereas the superscript v stands for the private innovator.

At time t, innovator i spends r_t^i dollars on research and d_t^i dollars on development. The probability of a successful innovation is given by

$$f_i(d_t^i, k_t^i)$$

where k_t^i is the stock of knowledge available to player i at time t. We assume that k_t^i evolves according to a knowledge accumulation equation as follows:

$$k_t^i = \delta\, k_{t-1}^i + r_{t-1}^i + \phi\, r_{t-1}^j$$

where δ is the depreciation rate, ϕ the knowledge spillover rate between the private and the public innovator, $i, j = u, v$, and $j \neq i$. We make the following assumption regarding the innovation process:

Assumption 1. *(a) Success is an absorbing state; (b) Success probabilities are independent across players.*

We assume that successful innovation by *any* player yields a payoff U to the public innovator. By contrast, the private innovator receives V if it is the *sole* innovator.

The payoff function (over the two periods) for the public innovator is given by

$$\begin{aligned}\mathcal{U} = U\left(1-\left(1-f_1^u\right)\left(1-f_1^v\right)\right) - r_1^u - d_1^u \\ + \beta\left(1-f_1^u\right)\left(1-f_1^v\right)\left(U\left(1-\left(1-f_2^u\right)\left(1-f_2^v\right)\right) - r_2^u - d_2^u\right)\end{aligned}$$

where, for simplicity, we write $f_t^u = f^u(d_t^u, k_t^u)$. Consider the first row in the above equation. The expression in brackets, $1-\left(1-f_1^u\right)\left(1-f_1^v\right)$, corresponds to the probability that *some* innovator is successful during the first period. This follows from our assumption that player U cares about innovation but not the innovator's identity. The expression $\left(1-f_1^u\right)\left(1-f_1^v\right)$, in the second row, corresponds to the probability that neither player was successful during the first period. Given that success is as absorbing state, the second-period payoff expression, $U\left(1-\left(1-f_2^u\right)\left(1-f_2^v\right)\right)-r_2^u-d_2^u$, is "discounted" both by the proper discount factor, β, and by the probability of no innovation during the first period.

The payoff function (over the two periods) for the private innovator is given by

$$\begin{aligned}\mathcal{V} = & V\left(1-f_1^u\right)\ f_1^v - r_1^v - d_1^v \\ & + \beta\left(1-f_2^u\right)\left(1-f_2^v\right)\left(V\left(1-f_1^u\right)\ f_1^v - r_2^v - d_2^v\right)\end{aligned}$$

where, for simplicity, we write $f_t^v = f^v(d_t^v, k_t^v)$. Consider the first row in the above equation. The expression in brackets, $\left(1-f_1^u\right)\ f_1^v$, corresponds to the probability that innovator v is the *sole* successful innovator during the first period. Similarly to the expression for $\mathcal{U}$,the second-period payoff expression is "discounted" both by the proper discount factor and by the probability of no innovation during the first period.

Our first result establishes that concurrent development decisions at $t=2$ are strategic substitutes.

Proposition 1. *$d(d_2^i)/d(d_2^j) < 0$, $i = u, v,\ i \neq j$.*

The proof of this and subsequent results can be found in the Appendix A. Proposition 1 is akin to strategic substitutability results in contest games. Consider, for example, the private innovator's second-period payoff conditional on success not having been attained during the first period. We have

$$\mathcal{V} = V\left(1-f_2^u\right)\ f_2^v - r_2^v - d_2^v$$

The higher the likelihood the public innovator innovates, that is, the higher f_2^u, the lower the effect of d^p (through f^u) on $\mathcal{V}$.

Next we show that, at $t=1$, player i's research and player j's development choices are strategic substitutes.

Proposition 2. *$d(r_1^i)/d(d_1^j) < 0$, $i = u, v,\ i \neq j$.*

Intuitively, a higher d_1^j implies a higher probability that innovation takes place today. This in turn lowers future impact of r_1^i, thus leading to a lower r_1^i.

Proposition 2 suggests that substitutability or complementarity patterns are not about the nature of the strategic variables (r vs r, r vs d) but rather about temporal relation between variables. In particular, Propositions 1 and 2 suggest that two contemporaneous innovation-effort variables will tend to be strategic substitutes.

Even though the exact mechanisms underlying Propositions 1 and 2 are different, they share this in common: when one player increases its innovation effort, the likelihood that such player succeeds is greater, which in turn reduces the marginal value of innovation effort by the other player.

We next specialize the model by making the following simplifying assumption:

Assumption 2. $f^i = \alpha_i \, f$

Our next result establishes the possibility of strategic complementarity in innovation efforts:

Proposition 3. *There exist* $\overline{\alpha}^u > 0$ *such that, if* $\alpha^u < \overline{\alpha}^u$*, then* $d(d_2^v)/d(r_1^u) > 0$.

Intuitively, through knowledge spillovers, an increase in r_1^i improves the effect of d_2^j. Technically, the proof of Proposition 3 works through a chain of derivatives. Unlike Propositions 1 and 2, Proposition 3 requires a technical condition regarding the value α_i. This is a sufficient, not necessary, condition. The point is that one must be careful with a possible indirect effect: an increase in r_1^i leads to an increase in d_2^i which, by Proposition 1, leads to decrease in d_2^j. Assumption 2 and the condition in Proposition 3 effectively bounds this effect.

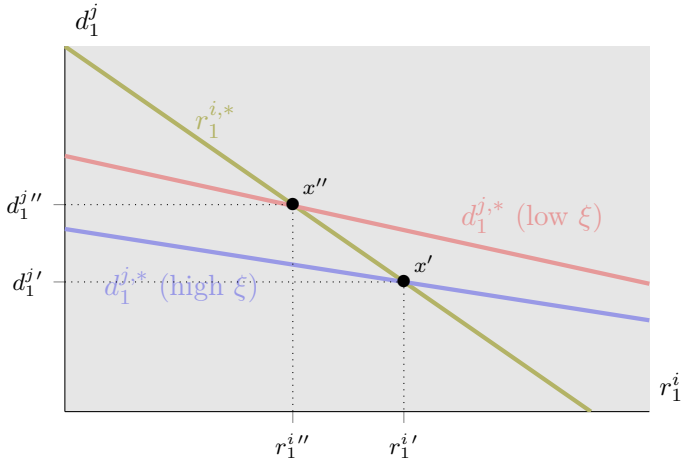
■ **Testable implications.** Propositions 1–3 suggest that innovation efforts can be either substitutes or complements. The strategic complementarity prediction is very much in line with the common wisdom regarding follow-up innovation (Azoulay et al., 2019). The strategic substitutability prediction, in turn, is very much in line with the literature on public spending crowding out (Wallsten, 2000). The gist of Propositions 1–3 is that strategic substitutability is more likely to take place when innovation effort choices are simultaneous, whereas strategic complementarity is more likely to take place when innovation effort choices are sequential.

Testing empirically for strategic complementarity (Proposition 3) can be thought of, intuitively, as regressing player i's innovation effort at a given time on player j's prior innovation effort. One complication is determining what "prior" means, that is, what the relevant time period for sequential choices is. Moreover, omitted variables may be persistent over time and thus may explain innovation efforts of both i and j on a certain domain over multiple periods.

Testing strategic substitutability involves the further challenge of endogeneity and, specifically, of simultaneity bias. Ideally, one would like to exploit a variable that shifts player i's best-response mapping but not player j's. This is illustrated in Figure 1. Suppose that there exists a variable ξ that shifts $d_1^{j,*}$ but not $r_1^{i,*}$. Specifically, suppose a lower value of ξ pushes $d_1^{j,*}$ upwards. Since $r_1^{i,*}$ remains constant, we should observe a shift in equilibrium from x' to x''. This corresponds to a higher value of d_1^j and a lower value of r_1^i. In terms of empirical analysis, including ξ as a regressor explaining, separately, d_1^j and r_1^i should lead to a negative estimated coefficient in the d regression and a positive coefficient in the r regression.

We present a careful discussion of our empirical strategy and the dataset we construct in the following sections.

Figure 1
Simultaneous choice of r_1^i and d_1^j: best-response curves.



3. Data

In this section we detail how we assemble a dataset that includes several measures for our main variables of interest (public and private innovation efforts), as well as for other important correlates.

We use NIH research funding as the main proxy for public innovation efforts, and privately-sponsored clinical trials to measure private innovation effort.

In addition to variation across years (our sample ranges from 2000 to 2019), we also categorize R&D by disease type (a total of 96 diseases). This cross-sectional variation, which plays a crucial role in our identification strategy, requires that we also assemble variables that might affect R&D levels, such as the existing body of knowledge (a proxy for k in our model) and the burden of the disease (a measure of the prize from successful innovation, which in our model we normalized to 1). We also include variables that potentially affect the private value from innovation but not the public value from innovation, such as the level of medical expenditures (a proxy for revenue potential) associated to each disease.

■ **NIH research funding.** We extract the universe of research grants awarded yearly by the NIH between 2000 and 2019 from the ExPORTER-RePORTER open-source platform. This platform provides detailed information for all research projects receiving funding, including their titles and abstracts, the awarded amount, and the designation year. Starting in 2008, following a Congress requirement designed to make NIH spending fully transparent, a computer-based algorithm assigns each project to one or to multiple category designations based on the research area, disease, or condition that the project targets.

We conduct a supervised machine learning exercise to expand this categorization back to years 2000-2007. Specifically, our aim is to assign each NIH funded project between 2000 and 2007 to one or multiple diseases from the same list of categories used by the NIH algorithm to classify projects in the subsequent years. To this end, we train a machine learning model on a random sub-sample of projects from 2008 onwards (for which we know the corresponding diseases assigned by the NIH algorithm), using project titles or, alternatively, abstracts.

We verify the precision of the model by comparing the predicted diseases against the true diseases in the residual validation sample of projects after 2008. We construct precision, recall, and F1 score, which are standard measures of model performance in this context. In our empirical analysis, we use the model predictions based on project titles rather than abstracts as a baseline, as they display a slightly higher F1 score on average (87%). We use predictions based on abstracts as a further validation of our results.[2]

We manually match each disease in the NIH list to the corresponding ICD10 code(s) using CMS codebooks and validating our matches against Sampat et al. (2013). We follow Sampat et al. (2013) and exclude fields of research, substances or risk factors that are not diseases, and thus cannot be mapped to diseases' features, such as burden, through standard ICD10 codes. We end up with around 100 well-defined and non-overlapping disease categories.

Furthermore, awarded projects are categorized as either NIH-solicited research or investigator-initiated research in our data. In the former, NIH solicits grant applications in specific research areas considered to be of high priority and, thus, is able to steer the direction of research more directly (Dresser, 2001, Hegde and Sampat, 2015). In the latter, researchers submit proposals to NIH based on their own interests, without specific restrictions or guidance from NIH.

■ **Privately sponsored clinical trials.** We collect information on past and ongoing clinical trials from ClinicalTrials.gov, an online database of clinical research studies submitted directly by sponsors or investigators, and maintained by the National Library of Medicine (NLM).[3] Information about each study is permanently available in the database, even after the study ends and even if the trial is unsuccessful. We focus on trials conducted in the U.S. and sponsored by the private sector, as opposed to public institutes or universities. For each clinical trial, sponsors are required to report the disease(s) or condition(s) under study using appropriate descriptors from the NLM's Medical Subject Headings (MeSH) controlled vocabulary thesaurus. We are able to match the MeSH terms to the corresponding ICD10 code(s) using publicly available crosswalks from NLM. As a result, we are able to map clinical trials to the disease categories from the NIH list used to classify NIH-funded research.[4]

■ **Scientific publications.** We use the lagged number of publications related to a certain disease as a proxy for the scientific opportunity for that disease. The underlying idea is that changes in the number of published articles in a certain area are likely to correlate with advances in that field (Hegde and Sampat, 2015). We collect research articles from the PubMed database, a free resource enabling the search of articles within the life sciences and biomedical literature. For each article, this database provides basic information about the title, journal and year of publication, as well as about the major MeSH term and all MeSH terms associated to the article's content. Using the PubMed Advanced Search, we can enter our MeSH terms of interest - i.e. those associated to our list of NIH disease categories - and

2. Additional details on our supervised learning procedure are available in Appendix B.
3. ClinicalTrials.gov was launched in 2000 as part of the FDA Modernization Act (1997). In 2007, the FDA Amendments Acts expanded the information and type of trials that investigators must submit to ClinicalTrials.gov by law.
4. In the rare cases where the match between MeSH terms and ICD10 codes was unsuccessful, we performed a fuzzy match directly between the MeSH terms and the NIH disease categories. We verified the outcome of this procedure manually and were successful in most cases.

retrieve the number of publications related to each MeSH term, for each year.

■ **Disease burden.** We proxy the burden of a disease with its mortality rate in a certain year, in the spirit of Cutler et al. (2012) and Dossi (2024). We use data on total deaths by disease-year from administrative records from the Center for Disease Control and Prevention (CDC).[5] We recognize that mortality is not the only dimension of disease burden, but we consider it the most salient measure. Additionally, we focus on mortality because comprehensive data on disease incidence (unlike mortality data) are not as consistently available across disease areas. We include diseases with an average number of deaths per year equal to or above a minimum threshold (e.g., 40 deaths per year) and experiment with alternative thresholds to ensure the robustness of our results.

■ **Medical expenditure.** We use the Medical Expenditure Panel Survey (MEPS) to obtain information on the total annual medical expenditure by disease and, thus, to construct our potential market size measure. MEPS is among the most comprehensive source of data on the cost and use of health care in the U.S. We focus on the Household Component data, which are based on questionnaires fielded to individual household members and their medical providers. We match files on population characteristics and medical conditions to obtain, for each surveyed individual over time, the main medical conditions diagnosed (ICD10 code) and the corresponding costs that insurers and patient incur to cover treatment. We weight each patient using the survey weights provided by the data provider, to make each individual representative within the US population. We then average across individuals sharing the same condition in the same year, to obtain a measure for the overall costs that insurers and patients incur annually to cover treatments for each separate disease.

4. Empirical Analysis

Propositions 1–3 suggest that private and public innovation efforts can be either substitutes or complements, depending on whether innovation effort choices are simultaneous or sequential. In this section, we test these propositions empirically in the context of biomedical R&D.

We face two main challenges. First, measuring the public sector's best response to the private sector's investment choices (and viceversa) presents a well-known endogeneity problem: for example, omitted variables might shift both variables at the same time or we might risk simultaneity bias. Second, time periods are well defined in a theoretical context, but less so when dealing with empirics. To deal with the first challenge, we consider plausibly exogenous shocks to private (resp. public) investment that hit certain disease areas in specific years. We then employ a staggered event-study design to quantify the differential evolution of public (resp. private) investment on 'treated' diseases relative to untreated ones over several years. Estimating an event-study is also useful to address the second challenge, as it offers a model-free approach to identify differential investment choices over time, without defining 'simultaneous' or 'sequential' choices ex-ante. In other words, we let the data reveal what the 'short' vs 'long' term responses of the various actors are, depending on the specifics

5. Specifically, we use data from the "Underlying Cause-of-Death" files, which report the disease or injury which initiated the train of events leading directly to death.

of the technological and regulatory environment (e.g., how long it takes to translate research funds into early-stage clinical trials).

In what follows, we will first describe the plausibly exogenous shocks we employ in our analysis, as well as the event study regression we estimate. We will then turn to present our main results: (i) the public sector's estimated response to the private sector's investment decisions, and (ii) the private sector's estimated response to the public sector's investment choices.

■ **Exogenous shocks** To construct plausibly exogenous shocks to the *private* sector's clinical trials conducted on specific disease areas, we use the granting of FDA Accelerated Approval status to drugs targeting those areas. Established in 1992, the FDA's Accelerated Approval Program[6] allows faster approval of drugs provided they address conditions that are life-threatening or serious, fill an unmet medical need where limited effective treatments exist, and show evidence of effectiveness through surrogate endpoints or early clinical data. Thus, receiving this designation at a given point in time signals to the broader private sector that the underlying science is sufficiently mature and the regulatory pathway sufficiently favorable to justify launching or accelerating late-stage trials in that disease. We therefore expect disease areas in which one or more drugs receive Accelerated Approval status at t to exhibit a subsequent increase in the probability of becoming the focus of new late-stage private clinical trial activity. Importantly, the timing of FDA designations is driven by pre-clinical and early-phase scientific evidence and regulatory assessments, and is plausibly orthogonal to the broader economic or strategic determinants of firm-level trial portfolios.

As plausibly exogenous shocks to the *public* sector's allocation of research funds to particular disease areas, we exploit a sequence of high-profile federal initiatives and executive actions that directed public resources toward specific scientific domains. Each of those initiatives was motivated by an assessment that the conditions for scientific progress were favorable rather than by contemporaneous trends in private investment. The federal initiatives we consider are: (i) the Executive Order 13505, signed by President Obama on March 2009, lifting prior presidential restrictions on NIH's authority to fund human embryonic stem cell research;[7] (ii) the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, announced by President Obama in April 2013 to solicitate research aimed at advancing neuroscience tools and addressing neurological disorders;[8] (iii) the Cancer Moonshot, launched in January 2016 by President Obama with the establishment of a specific Task Force of experts, aimed at accelerating progress in the prevention, diagnosis, and treatment of specific forms of cancer.[9]

■ **Baseline specification** We employ a staggered event-study design to estimate the public

6. Source: `https://www.fda.gov/drugs/nda-and-bla-approvals/accelerated-approval-program`
7. The NIH Director was instructed to develop new guidelines governing the support and conduct of responsible stem cell research, with final guidance to be issued within 120 days. This created a discrete, policy-driven increase in the pool of NIH-eligible projects in stem-cell-related disease areas. Source: `https://obamawhitehouse.archives.gov/the-press-office/fact-sheet-presidential-executive-order`
8. Source: `https://obamawhitehouse.archives.gov/the-press-office/2013/04/02/fact-sheet-brain-initiative`
9. Source: `https://obamawhitehouse.archives.gov/the-press-office/2016/02/01/fact-sheet-investing-national-cancer-moonshot`

sector's best response to the private sector's investment choices (and viceversa). Moreover, our goal is to distinguish a short term ("simultaneous") response from a medium-to-long term ("sequential") response.[10] The main equation for estimation is:[11]

$$\Delta Y_{i,t} = \alpha_t + \sum_{\tau \neq \tau_{-1}} \beta_\tau YearRelativeToShock_{t,\tau} * TargetDisease_i + \pi Ci, t + \epsilon_{i,t} \quad (1)$$

Our unbalanced panel features around 100 diseases i and 20 years t. We employ the plausibly exogenous shocks described in the previous section to define the dummy variable "Target-Disease", which is equal to 1 for disease areas targeted by at least one policy shock and 0 otherwise. For example, when studying the public sector's response to private investment choices, we set the dummy equal to 1 for all disease areas whose drugs were granted Accelerated Approval status by the FDA.[12]

The event study we consider is staggered, as different diseases are treated by the shock at different points in time, if at all. For example, the FDA Accelerated Approval status might be granted to two drugs treating two different diseases in two different years. We define τ_0 as the period in which the shock is announced and realized. At each point in time t, we measure the temporal distance of each treated disease from τ_0: thus, "YearRelativeToShock" is 1 if t is exactly τ years away from the shock realization and 0 otherwise.

We add year-specific fixed effects α_t, that absorb annual investment trends common to all diseases; and disease-year specific control variables Ci, t, that might affect changes in investment strategies over time, including changes in mortality rates, scientific opportunity, or revenue potential associated to a certain disease.

Our outcome variable is the standardized year-on-year change in investment activity for disease i at t relative to the prior year. Employing year-on-year changes allows to get rid of unobservable time-invariant heterogeneity across diseases that could motivate differential investment effort. For each event-study we run, the definition of Y depends on (i) the direction we are testing (public response to private activity or viceversa), and (ii) whether we are documenting the "first stage" relationship between the exogenous shocks and the endogenous variable, or the proper best response of interest. For example, when studying the public sector's response to private investment choices, in the first stage regression, Y is the standardized year-on-year change in private clinical trials, as we expect the private sector's clinical activity to increase almost mechanically for treated diseases receiving the FDA Accelerated Approval status. After establishing the first stage, we turn to our main outcome of interest, namely the degree of public sector's investments on the same disease areas; to this end, we define Y as the standardized year-on-year change in public funding allocation.

Given our theoretical predictions, we expect the public sector to respond in a diametrically opposite direction to the private sector in the short run—acting as a substitute—and in the same direction in the medium-to-long run—acting as a complement—following the plausibly exogenous shocks. Estimating a set of β_τ coefficients allows us to verify the theory

10. Our approach is similar in spirit to Clemens and Gottlieb (2014), even though their focus is very different, i.e. physician-induced demand in the context of a Medicare consolidation reform.
11. Standard errors are clustered at the disease level.
12. The same disease may be treated multiple times over time. For the small number of diseases where this occurs, there is a risk of a disease being classified as newly treated when in fact it had first received treatment several years earlier. In our robustness checks, we exclude these diseases from the sample.

empirically, as they capture changes in public sector investment in diseases affected by plausibly exogenous shocks to private innovation efforts, relative to untreated disease areas[13] and at each point in time. We use the year preceding each shock, τ_{-1}, as the reference period. A symmetric argument applies when studying the private sector's best response to public investment.

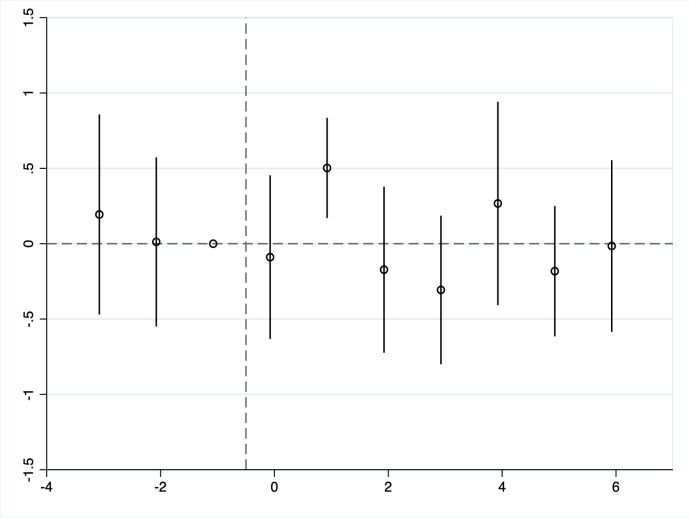
4.1. Results: how does the *public* sector respond to the private sector's investment decisions?

■ **First stage: private sector's response to exogenous shocks.** We begin by testing whether the proposed policy shocks to private investment induce an actual acceleration of the private sector's clinical trial activity. To this end, we run our main empirical specification using, as the main outcome of interest, the standardized year-on-year change in the number of privately-sponsored clinical trials starting at t relative to $t-1$ for disease i. As previously detailed, treated diseases are those in which one or more drugs receive the FDA Accelerated Approval status at time 0; time 0 is defined differently across diseases, as the FDA Accelerated Approval status is given at different times to different diseases (staggered adoption). For each year before and after the shock realization, we plot the estimated effect of the private shock on the clinical trial activity of the treated diseases relative to the untreated.

The event-study estimates confirm that the proposed shock contributes positively to the private sector's clinical trial activity, especially when it comes to the number of initiated late-stage clinical trials (Figure 2). Specifically, receiving FDA Accelerated Approval status induces a 0.5 standard deviation increase of the typical year-on-year change in clinical trial activity. The increase takes place in the year right after the status is granted, thus suggesting that the induced boost in clinical trials is concentrated around the set of drugs granted Accelerated Approval status, with limited spillovers to other drugs in the same diseases over time. The absence of pre-trends is reassuring, as it confirms that the proposed shock targets diseases that were substantially similar to the untreated ones in terms of private clinical trial activity, after controlling for key disease characteristics such as burden, profitability, or research opportunity.

13. We use never-treated diseases as control units, in the spirit of Sun and Abraham (2021).

Figure 2
Estimated effect of "private shocks" on private clinical trial activity.



Notes: This figure displays coefficients (β_τ) and confidence intervals from estimating an event-study as in Equation (1). The outcome variable is the standardized year-on-year change in the number of late-stage clinical trials for disease i, initiated at t, and sponsored by the private sector. The outcome variable is regressed on disease-specific shocks resulting from the granting of FDA Accelerated Approval status for one or more drugs targeting those conditions, interacted with indicator variables for each year before and after the shock. The event study design is staggered, as different diseases are treated at different points in time. The year before the shock ($\tau = -1$) is taken as the reference point. Controls include year fixed effects and disease-year specific variables, such as changes in mortality rates, research fertility, and revenue potential. Standard errors are clustered at the disease level.

■ **Main outcome: public sector's best response.** After establishing that the proposed private sector's shock had a substantial and positive impact on private clinical trial activity, we now turn to study the public sector's response. We use the same private shock to distinguish between treated and untreated diseases: the underlying idea is that disease affected by the plausibly exogenous private shock are more likely to experience a boost in private investment that is uncorrelated with other endogenous investment drivers, and thus more suitable to quantify the public sector's response correctly.

We estimate (1) using a set of outcome variables describing the year-on-year change of public sector's innovation effort. Our goal is to distinguish between a short term response and a medium-long term response, where 'short' and 'medium-to-long' terms are data-driven.

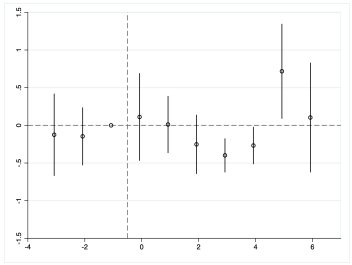
Our results are displayed in Figure 3. We find that treated diseases are more likely to experience a decline in public funding allocation from NIH between two and four years following the private shock (panel A). This funding reduction focuses specifically on new research projects (panel B) and early-stage explorative research (panel C),[14] both experiencing a drop of around 0.3 standard deviations two years after the private shock. Since the actual increase in private clinical trial activity occurs one year after the shock (see Figure 2), this means that NIH starts acting as a "substitute" immediately after the private boost to clinical trials. We also find implications of NIH budget cuts on real-world outputs a few

14. Explorative research is typically less demanding in terms of budget allocation than standard R01 research funds.

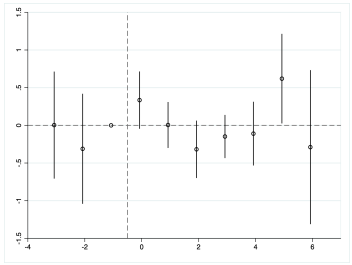
years later: the number of early-stage clinical trials funded exclusively by public money decreases around three years following the private shock (panel D); similarly, publications around targeted diseases decline (panel E).

What about NIH response in the medium-to-long term? Our event-study suggests that public funding allocation patters change for treated diseases around five years following the private shock: NIH research funds are allocated primarily to those diseases (panel A), especially for new projects (panel B) and early-stage research (panel C). This finding suggests that the public sector acts as a complement in the medium-to-long run, possibly exploiting the knowledge accumulated through the acceleration of the private sector's trial activity in prior years.

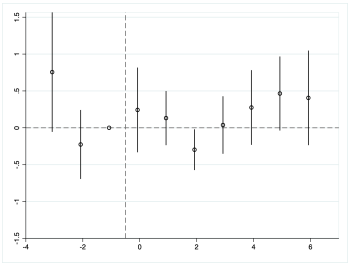
Figure 3
Estimated effect of "private shocks" on public sector innovation effort (several measures)



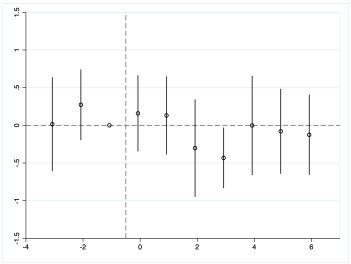
(a) NIH research funds allocated



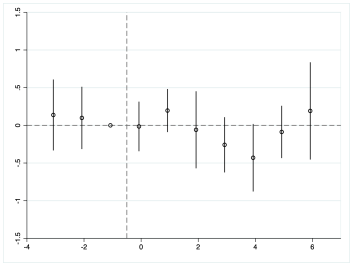
(b) NIH research funds allocated, new project only



(c) NIH research funds allocated, early-stage explorative projects only



(d) Publicly-funded clinical trials, early stages



(e) Publications

Notes: This figure displays coefficients (β_τ) and confidence intervals from estimating five separate event-studies as in Equation (1). The outcome variables are five measures for public sector innovation efforts, detailed under each panel. Each outcome variable is regressed on disease-specific shocks resulting from the granting of FDA Accelerated Approval status for one or more drugs targeting those conditions, interacted with indicator variables for each year before and after the shock. The event study design is staggered, as different diseases are treated at different points in time. The year before the shock ($\tau = -1$) is taken as the reference point. Controls include year fixed effects and disease-year specific variables, such as changes in mortality rates, research fertility, and revenue potential. Standard errors are clustered at the disease level.

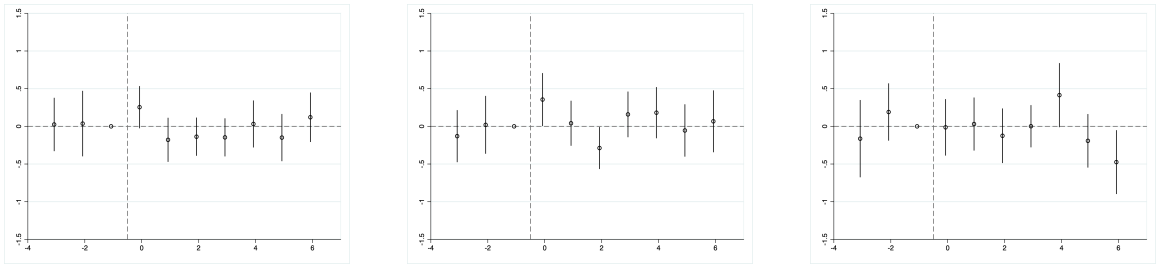
4.2. Results: how does the *private* sector respond to the public sector's investment decisions?

■ **First stage: public sector's response to exogenous shocks.** We test empirically whether the proposed policy shocks to public investment generate a boost to public funding allocations and related publicly-funded clinical trials. To this end, we run our main empirical specification using, as the main outcome of interest, a set of proxies for the public sector's innovation effort. Specifically, for disease i at time t, we use the standardized year-on-year change in: (i) public funds allocated by NIH to research project proposals (in dollars); (ii) NIH public funds allocated specifically to R01 grants, i.e. the most common and substantial public financial support available for biomedical researchers; (iii) the number of publicly-sponsored early-stage clinical trials started. As previously detailed, treated diseases are those targeted by a federal initiative or executive action directing public resources toward specific domains at time 0. Time 0 is defined differently across diseases, as several federal initiatives took place during our time-span of interest and targeted different disease areas (staggered adoption). For each year before and after the shock realization, we plot the estimated impact of the public shock on the public sector's innovation effort measures for the treated diseases relative to the untreated ones.

The event-study estimates shows that the proposed public shock has a significant and positive impact on public innovation effort on treated diseases (Figure 4). Diseases targeted by a federal initiative or executive action are associated with an approximately 0.3 standard deviation higher year-on-year change in NIH funding allocation relative to untreated diseases (panel A). The same is true if we restrict to R01 grants, the most common grant type available to mature biomedical researchers (panel B). The effect is immediate as the shock realizes, suggesting that NIH allocates its budget according to the new government provisions. Reassuringly, we don't observe any significant pre-trend; this confirms that the proposed shock targeted diseases that were substantially similar to the untreated ones in terms of NIH fund allocations, after controlling for key disease characteristics such as burden, profitability, or research opportunity.

The shocks to NIH funds propagates to publicly-funded clinical trials: in panel C, we show that the year-on-year variation in the number of early-stage public trials is 0.4 standard deviation higher than usual for targeted diseases relative to untreated ones, approximately four years after the public shock to NIH funds. This result is useful in two ways: first, it confirms that our proposed public sector shock had a real impact on innovation activity; second, through a model-free and purely data-driven approach, it reveals how long it typically takes the public sector to translate research funds into early-stage trials—a time-span that we can use to interpret the private sector's best response, to which we turn next.

Figure 4
Estimated effect of "public shocks" on public research investment (several measures)



(a) NIH research funds allocated

(b) NIH research funds allocated, R01 grants only

(c) Publicly-funded clinical trials, early stages

Notes: This figure shows coefficients (β_τ) and confidence intervals from estimating three event-studies as in Equation (1). The outcome variables are three measures for public research effort, as detailed under each panel. Each outcome variable is regressed on disease-specific shocks resulting from high-profile federal initiatives directing public resources toward those conditions, interacted with indicator variables for each year before and after the shock. The event study design is staggered, as different diseases are treated at different points in time. The year before the shock ($\tau = -1$) is taken as the reference point. Controls include year fixed effects and disease-year specific variables, such as changes in mortality rates, research fertility, and revenue potential. Standard errors are clustered at the disease level.

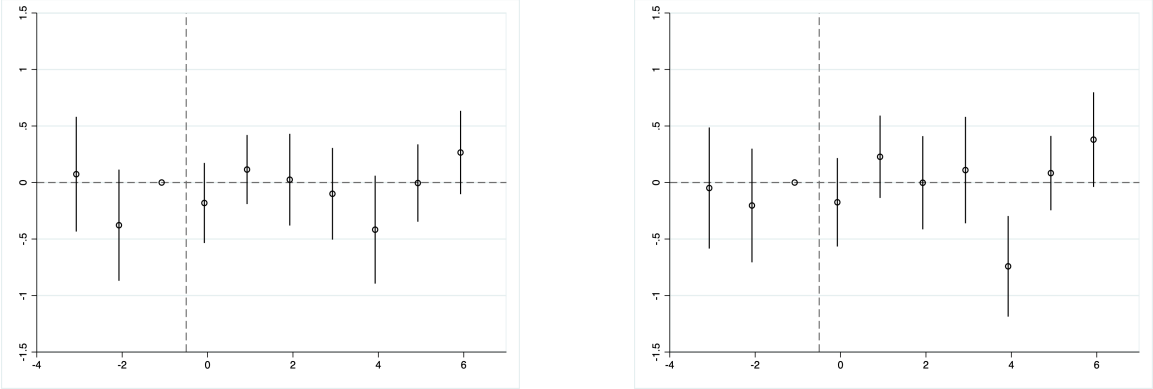
■ **Main outcome: private sector's best response** After showing that the proposed public sector's shock had a substantial and positive impact on public innovation efforts, we now move to analyze the private sector's response. Again, we use the public shock detailed above to define treated and untreated diseases. We estimate (1) using as main outcome variable the year-on-year change in the number of private sector's clinical trials, with a focus on early-stage and intermediate-stage trial activity. Similarly to before, our goal is to distinguish between a short term response and a medium-long term response, where 'short' and 'medium-to-long' terms are data-driven.

Our results are displayed in Figure 5. We find that treated diseases are more likely to experience a decline in privately-sponsored clinical trial activity around four years following the public shock. This is true for both early-stage (panel A) and intermediate-stage (panel B) clinical trials, each experiencing a drop of respectively 0.4 and 0.7 standard deviations. Note that the four-year time lag is supported by empirical evidence: in the prior section (see Figure 4, panel C), we established that the public sector took around 4 years to translate a boost in public funding to actual publicly-funded trials. A decline in private clinical trial activity around the same period suggests that the private sector's decision to divert resources away from diseases targeted by the public shock occurred more or less contemporaneously with that shock—or immediately thereafter, if we assume that the private sector translates investment decisions into clinical trials more quickly than the public sector. Taken together, these findings suggest that the private sector acts as a "substitute" immediately after the public funding increase on treated diseases.

What about the private sector's response in the medium-to-long term? Our event-study suggests that private clinical trial activity enters an upward-sloping trend for treated diseases around five-to-six years following the public shock. In other words, relatively more early-stage (panel A) and intermediate-stage (panel B) clinical trials targeting the treated diseases

are now sponsored by the private sector than prior to the public shock. This finding suggests that the private sector acts as a complement in the medium-to-long run, taking advantage of the public investments and publicly-run trials in the prior years.

Figure 5
Estimated effect of "public shocks" on private sector innovation effort (several measures)



(a) Privately-funded clinical trials, early stages

(b) Privately-funded clinical trials, intermediate stages

Notes: This figure shows coefficients (β_τ) and confidence intervals from estimating two event-studies as in Equation (1). The outcome variables are two measures for private research effort, as detailed under each panel. Each outcome variable is regressed on disease-specific shocks resulting from high-profile federal initiatives directing public resources toward those conditions, interacted with indicator variables for each year before and after the shock. The event study design is staggered, as different diseases are treated at different points in time. The year before the shock ($\tau = -1$) is taken as the reference point. Controls include year fixed effects and disease-year specific variables, such as changes in mortality rates, research fertility, and revenue potential. Standard errors are clustered at the disease level.

5. Discussion and conclusion

This paper argues that framing public and private innovation efforts as either complements or substitutes provides an incomplete picture of the strategic interaction between the two. We develop a theoretical framework showing that both patterns can arise within a single model, depending on the temporal structure of the interaction: simultaneous decisions generate strategic substitutability, while sequential decisions generate complementarity through knowledge spillovers. Testing these predictions in the context of U.S. biomedical research over 2000-2019, we find that both sectors act as substitutes in the short run following a positive (plausibly exogenous) shock to the other's activity, before transitioning to a complementary role in the medium-to-long run.

These findings carry direct implications for public policy. The short-run substitutability we document suggests that public funding increases may partially crowd out private investment (and viceversa) in the immediate aftermath of a shock, a pattern that naive cross-sectional estimates would likely miss. At the same time, the medium-to-long-run complementarity we uncover suggests that the full social return to public investment, for example, might materializes over longer horizons, as private actors build on publicly generated knowledge. Policymakers evaluating the effectiveness of research initiatives or innovation

subsidies should therefore account for these complex interactions.

Several important questions remain open. On the positive side, future work could explore whether the empirical patterns we document generalize beyond biomedical research to other innovation-intensive sectors, such as clean energy or defense. On the normative side, it remains unclear whether the observed best responses are efficient and what objective function the public sector ought to maximize. We view these as promising directions for future research.

A. Proofs

Proof of Proposition 1: At $t = 2$, conditional on no success at $t = 1$, the payoff functions are given by

$$\begin{aligned}\mathcal{U} &= U\left(1-\left(1-f_2^u\right)\left(1-f_2^v\right)\right) - r_2^u - d_2^u \\ \mathcal{V} &= V\left(1-f_2^u\right)\, f_2^v - r_2^v - d_2^v\end{aligned}$$

where, for simplicity, we write $f_t^i = f^i(d_t^i, k_t^i)$. Trivially, since we only have two periods, the optimal values of r_2^i are equal to zero, and so we have

$$\begin{aligned}\mathcal{U} &= U\left(1-\left(1-f_2^u\right)\left(1-f_2^v\right)\right) - d_2^u \\ \mathcal{V} &= V\left(1-f_1^u\right)\, f_1^v - d_2^v\end{aligned}$$

The first-order conditions with respect to d_2^i are given by

$$\begin{aligned}\left(1-f_2^v\right)\,\frac{df_2^u}{dd_2^u} &= 1/U \\ \left(1-f_2^u\right)\,\frac{df_2^v}{dd_2^v} &= 1/V\end{aligned}$$

By the implicit function theorem,

$$\frac{d(d_2^i)}{d(d_2^j)} = -\left(\frac{d(f_2^j)}{d(d_2^j)}\right)\left(\frac{d(f_2^i)}{d(d_2^i)}\right)\Bigg/\left(-\frac{d^2(f_2^i)}{d(d_2^i)^2}\right)$$

Since we assume an interior solution, $-d^2\left(f_2^i\right)/d(d_2^i)^2 > 0$. By assumption, f^i is increasing in d^i. It follows that $d(d_2^i)/d(d_2^j) < 0$. ■

Proof of Proposition 2: At $t = 1$, the payoff functions are given by

$$\begin{aligned}\mathcal{U} = {} & U\left(1-\left(1-f_1^u\right)\left(1-f_1^v\right)\right) - r_1^u - d_1^u \\ & + \beta\left(1-f_1^u\right)\left(1-f_1^v\right)\left(U\left(1-\left(1-f_2^u\right)\left(1-f_2^v\right)\right) - r_2^u - d_2^u\right) \\ \mathcal{V} = {} & V\left(1-f_1^u\right)\, f_1^v - r_1^v - d_1^v \\ & + \beta\left(1-f_2^u\right)\left(1-f_2^v\right)\left(V\left(1-f_1^u\right)\, f_1^v - r_2^v - d_2^v\right)\end{aligned}$$

The first-order conditions with respect to r_1^i are given by

$$\begin{aligned}\beta\left(1-f_1^u\right)\left(1-f_1^v\right)\left(\left(\frac{\partial f_2^u}{\partial(d_2^u)}\,\frac{dd_2^u}{dk_2^u}\right)+\left(\frac{\partial f_2^u}{\partial k_2^u}\,\frac{\partial k_2^u}{\partial r_1^u}\right)\right) &= 1/U \\ \left(1-f_2^u\right)\,\frac{df_2^v}{dd_2^v} &= 1/V\end{aligned}$$

By the implicit function theorem,

$$\frac{d(d_2^i)}{d(d_2^j)} = -\left(\frac{d(f_2^j)}{d(d_2^j)}\right)\left(\frac{d(f_2^i)}{d(d_2^i)}\right)\Bigg/\left(-\frac{d^2(f_2^i)}{d(d_2^i)^2}\right)$$

Since we assume an interior solution, $-d^2(f_2^i)/d(d_2^i)^2 > 0$. By assumption, f^i is increasing in d^i. It follows that $d(d_2^i)/d(d_2^j) < 0$. ■

Proof of Proposition 3: Suppose that $\alpha^u = 0$. This implies that $d_t^u = 0$. At $t = 2$, conditional on no success at $t = 1$, $\mathcal{V}$'s payoff is

$$\mathcal{V} = V\left(1 - f_1^u\right) f_1^v - r_2^v - d_2^v$$

The first-order condition with respect to d_2^v is given by

$$V\frac{d(f_2^v)}{d(d_2^v)} = 1$$

By the implicit function theorem,

$$\frac{d(d_2^v)}{d(r_1^u)} = \left(\frac{\partial^2(f_2^v)}{\partial(d_2^v)\,\partial(r_1^u)}\right)\Bigg/\left(-\frac{d^2(f_2^v)}{d(d_2^v)^2}\right)$$

Since we assume an interior solution, $-d^2(f_2^v)/d(d_2^v)^2 > 0$. By assumption, $\partial^2 f/\partial d\,\partial k > 0$. Since $d(k_{t+1}^i)/d(r_t^j) > 0$, we conclude that $d(d_2^v)/d(r_1^u)$, as stated in the proposition. The result then follows by continuity. ■

B. Supervised Machine Learning Procedure

Since 2008, NIH has used a machine-based algorithm (the Research, Condition, and Disease Categorization system, RCDC) to classify funded research projects by disease area. This classification draws on project titles, abstracts, and key sections, matching extracted concepts to predefined disease categories using a thesaurus curated and validated by NIH experts.

Projects funded prior to 2008 are not classified under this system. To extend disease classifications to the 2000-2007 period, we implement a supervised machine learning model trained on post-2008 RCDC-labeled data. Our goal is to assign earlier projects to one or more of the same disease categories used by RCDC. We begin by restricting the sample to U.S.-based, extramural NIH projects and verify the stability of disease categories over time. Following prior literature, we exclude non- disease research areas (e.g., risk factors, chemical agents), and 17 high-level "supercategories" that overlap almost entirely with more specific disease types. This yields a consistent set of around 100 non-overlapping disease categories. We clean each project's title and abstract, removing stop words and non-discriminating terms (e.g., "and," "health").

We split the post-2008 RCDC-labeled data into a training set (80%) and validation set (20%). Using the FastText library, we train a multi-label classification model to assign

projects to one or more disease categories. FastText is well-suited to our task because it supports multi-label predictions and handles noisy, high-dimensional text efficiently. We test different classification models by varying the prediction threshold (i.e. the minimum probability required to assign a disease label) and the input (i.e., we use either projects' titles or abstracts to train the model).

Model performance is evaluated on the validation sample using standard metrics:

- Precision (P): Correct predictions among predicted labels
- Recall (R): Correct predictions among true labels
- F1 Score: Harmonic mean of P and R, i.e.

$$F1 = \frac{2PR}{P+R}$$

We find that predictions based on abstracts perform similarly to those based on titles alone for intermediate prediction threshold levels. We use the latter as our main model specification, as all research projects are assigned a title, while some lack an abstract. Applying this model to pre-2008 projects, we generate disease classifications for the full 2000-2019 period.

This produces a consistent 20-year panel of NIH-funded research projects, each mapped to specific disease areas. We describe in details all the steps we followed in the following subsections.

B.1. Training and Validation

Our objective was to retrospectively classify NIH-funded projects prior to 2008 using titles and abstracts. We began by preprocessing and cleaning the titles and abstracts of all NIH projects. We then split the data from 2008 onwards into a training set (80%) and a validation set (20%). We trained a multi-label classification model using FastText, a machine learning tool developed by Meta AI, well-suited for short-text classification and multiple-label assignments.

B.2. Model Performance Test 1: Out-of-Sample Validation

We use the validation sample to benchmark the model on out-of-sample data. For each project, we compare the predicted disease labels to the true NIH-assigned ones. Performance is measured using:

- Precision: Share of predicted labels that are correct.
- Recall: Share of true labels that are successfully predicted.
- F1 Score: Harmonic mean of precision and recall.

We vary:

- Input type: Using either titles or abstracts to train the model for predictions.
- Prediction threshold: The minimum probability required to assign a disease label (e.g., 0.1, 0.5, 0.8).

Key results (see Tables S2-S4):

- Precision increases with tighter thresholds (fewer, more confident predictions).
- Recall increases with looser thresholds (more inclusive predictions).
- The F1 score captures the Precision-Recall trade-off and is maximized at intermediate threshold levels.
- For intermediate threshold levels, title- and abstract-based models perform similarly in terms of F1. For more extreme threshold levels, titles outperform abstracts.

B.3. Model Performance Test 2: Temporal Backcasting

Our core use case is to use post-2008 data to predict disease categories for pre-2008 projects. To simulate this, we train models on 2014-2019 data and validate on 2008-2013 projects. As in Test 1, we evaluate precision, recall, and F1 scores, comparing abstract- and title-based models across thresholds (see Tables S5-S6). Results are consistent:

- Higher thresholds improve precision, lower ones improve recall.
- F1 scores favor intermediate threshold levels.
- Abstracts slightly outperform titles here, likely due to richer detail, although this was not obvious a priori.

B.4. Model Selection and Deployment

We select the title-based model for primary use due to consistent title availability across all grants. Both title- and abstract-based models perform similarly at optimal thresholds. We apply this model to classify NIH-funded projects from 2000-2007, providing consistent categorization across our full study period (2000-2019). Abstract-based classifications are reserved for robustness checks.

C. Additional Tables

Table S1
Summary statistics of the key variables of interest.

				Percentile				
	Count	Mean	SD	5th	25th	50th	75th	95th
Key Variables								
NIH-initiated funding (million $)		51.4	150.2	0.1	2.0	10.1	42.1	188.8
Privately sponsored clinical trials (count)		25.8	42.3	0.0	2.0	7.0	33.0	109.0
Controls								
Black deaths		2,101.1	5,666.4	5.0	29.0	169.0	1,208.0	10,108.0
White deaths		16,518.1	47,979.4	46.0	310.0	1,112.0	7,868.0	88,119.0
Male deaths		9,554.2	28,118.4	20.0	186.0	727.0	5,826.0	37,764.0
Female deaths		9,580.7	27,406.5	32.0	182.0	654.0	4,805.0	59,977.0
Publications		5,391.5	10,628.5	114.0	426.0	1,463.0	5,098.0	22,767.0
Potential market size (million $)		839.9	1,853.9	5.6	44.0	169.3	763.6	3,658.0
Sample Overview								
Year	20							
Numbers of diseases	96							

Table S2
Model performance, test 1: Precision.

Year	P(A-0.1)	P(A-0.5)	P(A-0.8)	P(T-0.1)	P(T-0.5)	P(T-0.8)
2008	0.58	0.89	0.95	0.73	0.89	0.93
2009	0.65	0.91	0.96	0.78	0.91	0.95
2010	0.62	0.91	0.96	0.75	0.89	0.93
2011	0.60	0.89	0.95	0.73	0.88	0.93
2012	0.60	0.90	0.96	0.73	0.88	0.93
2013	0.59	0.89	0.95	0.73	0.88	0.93
2014	0.58	0.89	0.95	0.72	0.88	0.93
2015	0.59	0.89	0.95	0.73	0.89	0.93
2016	0.59	0.89	0.96	0.73	0.88	0.93
2017	0.59	0.89	0.95	0.73	0.89	0.94
2018	0.62	0.90	0.96	0.74	0.88	0.92
2019	0.62	0.90	0.96	0.74	0.88	0.93

Table S3
Model performance, test 1: Recall.

Year	R(A-0.1)	R(A-0.5)	R(A-0.8)	R(T-0.1)	R(T-0.5)	R(T-0.8)
2008	0.79	0.57	0.43	0.71	0.61	0.54
2009	0.84	0.65	0.51	0.75	0.66	0.60
2010	0.82	0.63	0.50	0.70	0.60	0.54
2011	0.82	0.61	0.48	0.70	0.60	0.53
2012	0.82	0.62	0.49	0.70	0.60	0.53
2013	0.80	0.59	0.45	0.70	0.60	0.53
2014	0.80	0.59	0.46	0.70	0.59	0.53
2015	0.81	0.59	0.46	0.70	0.59	0.53
2016	0.80	0.60	0.47	0.70	0.60	0.53
2017	0.81	0.62	0.50	0.72	0.61	0.56
2018	0.82	0.63	0.52	0.72	0.62	0.56
2019	0.84	0.66	0.54	0.73	0.64	0.58

Table S4
Model performance, test 1: F1 Score.

Year	F1(A-0.1)	F1(A-0.5)	F1(A-0.8)	F1(T-0.1)	F1(T-0.5)	F1(T-0.8)
2008	0.67	0.69	0.60	0.72	0.72	0.68
2009	0.73	0.76	0.67	0.76	0.77	0.73
2010	0.71	0.75	0.66	0.72	0.72	0.69
2011	0.69	0.73	0.64	0.71	0.72	0.68
2012	0.70	0.73	0.65	0.71	0.71	0.67
2013	0.68	0.71	0.61	0.72	0.71	0.67
2014	0.67	0.71	0.62	0.71	0.71	0.68
2015	0.68	0.71	0.62	0.72	0.71	0.68
2016	0.68	0.71	0.63	0.71	0.71	0.68
2017	0.69	0.73	0.66	0.72	0.72	0.69
2018	0.70	0.75	0.67	0.73	0.73	0.69
2019	0.71	0.76	0.69	0.74	0.74	0.71

Table S5
Model performance, test 2: F1 Score, based on research projects' abstracts.

Threshold	Precision	Recall	F1
0.1	0.66	0.88	0.75
0.3	0.75	0.83	0.79
0.5	0.79	0.79	0.79
0.7	0.81	0.75	0.78
0.9	0.84	0.68	0.75

Table S6
Model performance, test 2: F1 Score, based on research projects' titles.

Threshold	Precision	Recall	F1
0.1	0.68	0.74	0.71
0.3	0.75	0.70	0.72
0.5	0.77	0.68	0.72
0.7	0.80	0.65	0.72
0.9	0.82	0.61	0.70

References

Azoulay, P., Zivin, J. S. G., Li, D., and Sampat, B. N. (2019). Public R&D investments and private-sector patenting: Evidence from NIH funding rules. *The Review of Economic Studies*, 86(1):pp. 117–152.

Chakravarthy, R., Cotter, K., DiMasi, J., Milne, C.-P., and Wendel, N. (2016). Public-and private-sector contributions to the research and development of the most transformational drugs in the past 25 years: from theory to therapy. *Therapeutic Innovation & Regulatory Science*, 50(6):759–768.

Clemens, J. and Gottlieb, J. D. (2014). Do physicians' financial incentives affect medical treatment and patient health? *American Economic Review*, 104(4):1320–1349.

Cohen, W. M., Nelson, R. R., and Walsh, J. P. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 48(1):1–23.

Cutler, D. M., Meara, E., and Richards-Shubik, S. (2012). Induced innovation and social inequality: evidence from infant medical care. *Journal of Human Resources*, 47(2):456–492.

Dossi, G. (2024). Race and science. Working paper.

Dresser, R. (2001). *When science offers salvation: Patient advocacy and research ethics*. Oxford University Press.

Hegde, D. and Sampat, B. (2015). Can private money buy public science? disease group lobbying and federal funding for biomedical research. *Management Science*, 61(10):2281–2298.

Kingma, B. R. (1989). An accurate measurement of the crowd-out effect, income effect, and price effect for charitable contributions. *Journal of Political Economy*, 97(5):1197–1207.

Li, D., Azoulay, P., and Sampat, B. N. (2017). The applied value of public investments in biomedical research. *Science*, 356(6333):78–81.

Lichtenberg, F. R. (2001). The allocation of publicly funded biomedical research. In *Medical Care Output and Productivity*, pages 565–590. University of Chicago Press.

Sampat, B. N., Buterbaugh, K., and Perl, M. (2013). New evidence on the allocation of nih funds across diseases. *The Milbank Quarterly*, 91(1):163–185.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Toole, A. A. (2007). Does public scientific research complement private investment in research and development in the pharmaceutical industry? *Journal of Law and Economics*, 50(1):81–104.

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