Exploit or Explore?
An Empirical Study of Resource Allocation in Research Labs

Ran Zhuo

Technology and Operations Department, Ross School of Business, University of Michigan, Ann Arbor, MI 48105,
ranzhuo@umich.edu

Balancing exploitation and exploration in resource allocation under incomplete information is a classic problem in operations management theory. Yet little research has empirically studied how and how well decision-makers make the exploitation-exploration tradeoff in a complex real-world situation. This paper empirically studies how a group of large publicly funded research labs traded off the exploitation of safe projects to maximize short-term productivity versus the exploration of high-variance projects to acquire information and improve long-term productivity. Using granular data on the allocation of almost one million input bundles to more than 300,000 research projects from 2000 to 2015, we model the resource allocation process as a multi-armed bandit and estimate a dynamic structural model to reveal how these labs balanced exploitation and exploration. We find the labs' decision model strongly resembles a simple Upper Confidence Bound (UCB) index. Estimates of the model's free parameters suggest that the labs explored extensively. Counterfactual simulations show that exploration substantially increased the labs' productivity—had they not explored, their output quantity would have decreased by 51%, and their citations would have decreased by 57%. Further simulations demonstrate that the labs' decision model outperformed popular alternative allocation models, including the Gittins Index, Thompson Sampling, and Explore-Then-Commit. Additionally, processes that promoted information utilization during allocation contributed to better outcomes. Had the labs not collected and performed data analytics on the information revealed during exploration, they would have saved 3% of funding but lowered output quantity by 7% and citations by 9%.

Key words: exploitation-exploration tradeoff, multi-armed bandit, resource allocation, innovation

1. Introduction

The exploitation-exploration tradeoff summarizes the fundamental dilemma in decision-making under incomplete information about different decisions' outcomes. Exploitation refers to taking advantage of known and proven options to gain immediate rewards or benefits, while exploration refers to experimenting with new methods, technologies, and strategies to gain information about new options' potential for improving long-term outcomes. Since seminal works of Thompson (1933) and Gittins (1979) on multi-armed bandits, the tradeoff has become a core concept in various fields such as economics (e.g. Bolton and Harris 1999), machine learning (e.g. Bubeck et al. 2012), and psychology (e.g. Cohen et al. 2007). In operations management theory, it plays a crucial role in resource allocation, process improvement, and product development (e.g. Caro and Gallien 2007, Papanastasiou et al. 2018, Zhalechian et al. 2022).

Yet, little research has studied empirically how decision-makers approach this tradeoff in a complex real-world setting. One challenge is obtaining data with sufficient granularity for decision-makers' choice sets. Another is the computational challenge of estimating dynamic choice models with large choice sets and long decision horizons. Workhorse recursive and simulation estimation methods (Pakes 1986, Rust 1987, Hotz and Miller 1993) do not work well for those problems due to the curse of dimensionality.

Employing novel data and proposing a new estimation approach, this paper studies empirically how a group of large publicly funded research labs balanced the exploitation of safe research projects for immediate productivity versus the exploration of new, high-variance projects to gain information about their potential productivity. This is interesting and important for several reasons. Firstly, this paper empirically tests in a multi-armed bandit setting theoretical intuition about the value of exploration and evaluates the level of sophistication in real-world decision-making. Secondly, this paper examines high-stake decision-making that involves large amounts of resources. The labs studied in this paper have allocated resources worth $1.3 billion. U.S. federal funding agencies distribute more than $50 billion of research funds each year (Boroush 2021) and largely face the same problem of how to allocate resource effectively among competing research programs (e.g. Jaffe 2002, Azoulay et al. 2011). Moreover, the managerial implications and methodological contributions of this paper are relevant to analyzing decision-making processes in many other aspects of operations management such as product assortment, service scheduling, hiring and supplier sourcing (similar in spirit to e.g. Bray et al. 2016, Li et al. 2020).

The empirical setting of this paper is a group of large structural biology labs funded by the Protein Structure Initiative (PSI), a $1.3 billion NIH program spanning 2000–2015. This is an ideal setting to study because it is realistically complex and affords highly granular data. Moreover, the labs thoroughly documented how they used a machine learning approach to analyze information observed during past allocations and to inform future allocation decisions, providing guidance for modeling. A few additional NIH policy features, such as restricting the pool of projects the labs could choose from, also make this setting particularly clean for model identification.

The main data is those labs' resource allocation and output across research projects, and the organization of this data allows the allocation problem to map nicely into the multi-armed bandit framework. The data contains, at the daily frequency, the input allocated to, and output produced by, each of the 335,553 research projects that those labs attempted over the sixteen-year horizon. A project in this setting is the determination of a specific protein molecule's structure, that is, the three-dimensional arrangement of atoms in that molecule. Each project is clearly defined and distinct from each other. The labs recorded input allocated to each project in discrete units. Each unit of input represents a distinct experimental trial on the project. The information revealed

during each *actual* allocation is the allocated trial's *observed* output, including whether the trial successfully produced a structure, and the number of citations and downloads of that structure. The labs made allocation decisions based on their *beliefs* about output from each *potential* allocation. To construct variables capturing the labs' posterior beliefs about output, a key model ingredient, the paper made a best-effort replication of the labs' machine learning approach of belief formation based on information revealed during past allocations.

To uncover how much the labs valued exploitation versus exploration, we estimate a dynamic structural model of the labs' decision-making process, assuming the labs used a simple heuristic index to approximate the complex underlying value function. This kind of index is well-studied in theory (e.g. Lai and Robbins 1985, Agrawal 1995) and in numerical simulations based on real data (e.g. Bietti et al. 2021, Cheung et al. 2022) and well-used in practice (e.g. Nguyen-Thanh et al. 2019, He et al. 2020) but has not been applied to estimation of empirical decision models. Under this model, in each period a lab first analyzes past information to form posterior beliefs about output from potential allocations. The lab then uses this posterior to compute a simple index to approximate the value function associated with allocating input to each project. The lab then allocates input to projects with the highest index values. The index approximation approach overcomes the curse of dimensionality because computation of the index only requires information available to the labs in the current period, in contrast to workhorse estimation methods, which tie the value function to the evolution of information in future periods. This approach builds upon a couple of earlier works that have studied bandit-like problems using structural estimation (Miller 1984, Erdem and Keane 1996, Crawford and Shum 2005, Dickstein 2021, Ganglmair et al. 2019, Jiang et al. 2022).

The main model modifies a well-used index called Upper Confidence Bound (UCB) (Auer et al. 2002). Intuitively, UCB captures "optimism in the face of uncertainty." Allocating input to a project that the lab has little information about has a high UCB. As the lab allocates more input to the project and has more information about its output, the UCB of further allocations decreases. To proxy for the unobserved factors during the allocation process, such as project-specific learning staying with individual researchers over time, we modify the baseline UCB index to incorporate an additional term that captures time-discounting of the value of older projects. As robustness checks, we also specify many alternative models with other types of indices discussed in the theoretical literature. These include the seminal Gittins (1979) index, which prescribes optimal decisions for some stylistic bandit problems.

We validate the main model in two stages. In the first stage, we estimate the free parameters in the model by maximizing the log likelihood of the observed allocation decisions. Identification of model parameters is based on revealed preferences and is very intuitive. We gauge model fit based

the log likelihood achieved at convergence. In the second stage, we use the estimated parameters from the model to forward simulate the labs' entire history of input allocation and output. We compare the patterns of input allocation and output from the simulated data to those from the actual data to determine whether the model could generate patterns similar to those in the actual data. We repeat this two-stage procedure for the alternative models to benchmark the fit of the main model against the alternatives.

We find the main model fits the data extremely well and captures the labs' decision-making. During the maximum likelihood stage, the main model by far has the smallest magnitude of log likelihood at convergence among the many alternative models tested. With the same number of parameters, its log likelihood at convergence is only 52% to 72% of the second best fit model across labs. During the simulation stage, the main model generates input allocation patterns and output that are very similar to those in the data. For all labs, the deviations of the simulated output from the actual output are within 10%. The alternative models fail to generate patterns matching the data as closely.

Based on this well-fitting model, we find that the labs explored extensively and that exploration had a large positive impact on the labs' productivity. Based on the estimates of the free parameters of the model, we are able to reject at a 95% confidence level for all labs the hypothesis that the labs did not explore and the hypothesis that unobserved factors did not impact the labs' incentives to explore a project over time. Counterfactual simulations show that exploration boosted the labs' productivity substantially. Had the labs not explored, they would have missed many low-hanging fruits and lowered their output quantity by 51% and their citations by 57%. The decrease is equivalent to forgoing at least $650 to $720 million of economic value. Moreover, the labs' decision model was more productive than alternative models including the Gittins index, Thompson sampling (Thompson 1933), and Explore-Then-Commit. Had the labs used those allocation models, their output quantity would have decreased by 14% to 43%.

A well-fitting model also enables analyzing the design of the allocation process, and we find that promoting the utilization of information revealed during exploration generated large returns. An initial pilot phase allowed the labs to explore freely and build a reasonable prior about different projects' productivity before they scaled up. Had the NIH cut the pilot, it would have saved 17% of funding ($220 million) but would have lowered the labs' output quantity by 23%, forgoing at least $290 million of economic value. The NIH also required the labs to collect data on their allocation decisions and outcomes and directed $40 million of funding to support data collection and analytics such as machine learning. Had the NIH cut this requirement and its funding support, it would have saved 3% of funding ($40 million) but would have lowered the labs' output quantity by 7% and citations by 9%, forgoing at least $90 to $110 million of economic value.

These findings bear significant managerial and policy implications. Firstly, they underscore the role of exploration and long-term vision in driving an organization's innovative productivity. Secondly, the results spotlight the value of "information" as a critical research output worthy of recognition. Presently, many organizations, including leading funding agencies, gauge innovation productivity solely by tangible outputs like publications and citations (see e.g. Balaban 2013, Lauer 2016). However, findings from this paper illuminate that even failures can yield invaluable information that improves subsequent resource allocation and overall productivity. Furthermore, to harness the full benefit of exploration, organizations should consider shaping their resource allocation process to emphasize information utilization, contemplating pilot programs, and investing in data analytics infrastructure. Lastly, theoretical insights from the multi-armed bandits literature have proven effective for resource allocation. Embracing these advanced methodologies is thus recommended for organizations.

The rest of this paper is organized as follows. Section 2 discusses the empirical setting. Section 3 describes the data. Section 4 builds the model. Section 5 describes model fitting. Section 6 shows model fitting results. Section 7 shows counterfactual results. Section 8 discusses managerial and policy implications and concludes.

2. Empirical Setting

The empirical setting of this paper is a group of large structural biology labs funded by the Protein Structure Initiative (PSI), a $1.3 billion NIH program from 2000–2015. This is an ideal setting to study for many reasons. Importantly, this setting captures the complexity of decision-making typical in resource allocation under incomplete information and affords highly granular data. Section 2.1 describes briefly the science of structural biology. Section 2.2 provides a more comprehensive explanation of the resource allocation process and the exploitation-exploration tradeoff in this context. It also describes how the labs used information observed during past allocations and data analytics to inform future decision-making, which provides guidance for modeling. Section 2.3 delves into the numerous policy features of the setting, which not only facilitate model simplification but also provide motivation for our policy analysis.

2.1. Scientific Background

A particularly clean feature of this setting is that research projects in structural biology are clearly defined and distinct from each other, making it easy to compile highly granular data that associates input and output with specific projects. A project in this setting is the determination of the three-dimensional structure of a protein molecule, and the molecule uniquely identifies the project. Each protein molecule consists of building blocks called amino acids. These amino acids are arranged into a chain, which folds up onto itself, creating a three-dimensional structure (Hill and Stein 2021).

Figure 1 shows one such structure, determined by one of the labs in the data. As is customary in the field of structural biology, the determined structure was considered published when it was deposited into a publicly accessible database known as the Protein Data Bank (PDB) and was assigned a unique identification number, which, in this case, is 2JUF. Structural biology projects are important basic research that could lead to valuable applied research. Knowing a protein's structure is often critical for developing drugs that target the protein. Protein structures have aided the development of therapeutics such as oncology drugs (Van Montfort and Workman 2017) and COVID-19 vaccines (Wrapp et al. 2020). Over a dozen Nobel Prizes have been awarded for advances in structural biology (Hill and Stein 2020).

Figure 1   Three-dimensional structure of a protein molecule
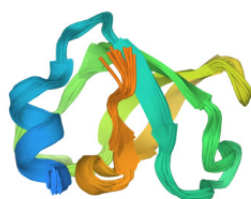
Note: The Northeast Structural Genomics Consortium (NESG) successfully determined this structure in 2007 (Kaustov et al. 2007). Its PDB identification number is 2JUF.

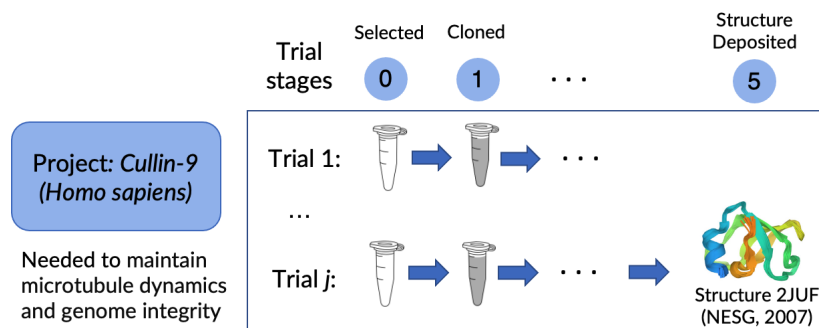**Figure 2   One experimental trial as an input bundle**

Image credits: The images of PCR tubes are from labicons.net.

Another clean feature of this setting is that structural biology labs allocate input to projects in discrete units, making it easy to track the amount of resources allocated to each project at any given point in time. A unit of input in this setting is a distinct experimental trial of the project. One can conceptualize a trial as being represented by an individual test tube containing the molecule, along with the combination of labor, capital, and materials dedicated to working on that specific test tube. Each trial proceeds in multiple sequential stages. For example, one of the stages involves

cloning the DNA sequence responsible for encoding the particular protein molecule. The outcome of each stage is either a success or a failure. If a trial succeeds in all stages, it produces a unit of output—the publication of the structure of the molecule. Figure 2 shows an example of an actual project in the data (protein Cullin-9 from Homo sapiens) and some example trials associated with that project. One of those trials successfully produced the structure shown in Figure 1. Trials are very risky. In the data, 98% of trials failed to produce output, and there were significant variations in outcomes across different stages, even for trials within a single project. As Chruszcz et al. (2008) put it, "...the success of any or all individual steps does not guarantee the success of the overall process... requires a significant amount of work and much luck..." Output is thus a random variable that structural biology labs have incomplete information about. To simplify the allocation problem, we assume that a trial in the data stopped if and only if the trial failed, which means the labs in our setting did not need to decide whether to kill an ongoing trial. This assumption is mostly supported by the patterns in the data.[1]

2.2. The Resource Allocation Process

**Figure 3   Lab's per period resource allocation decisions among projects**

---
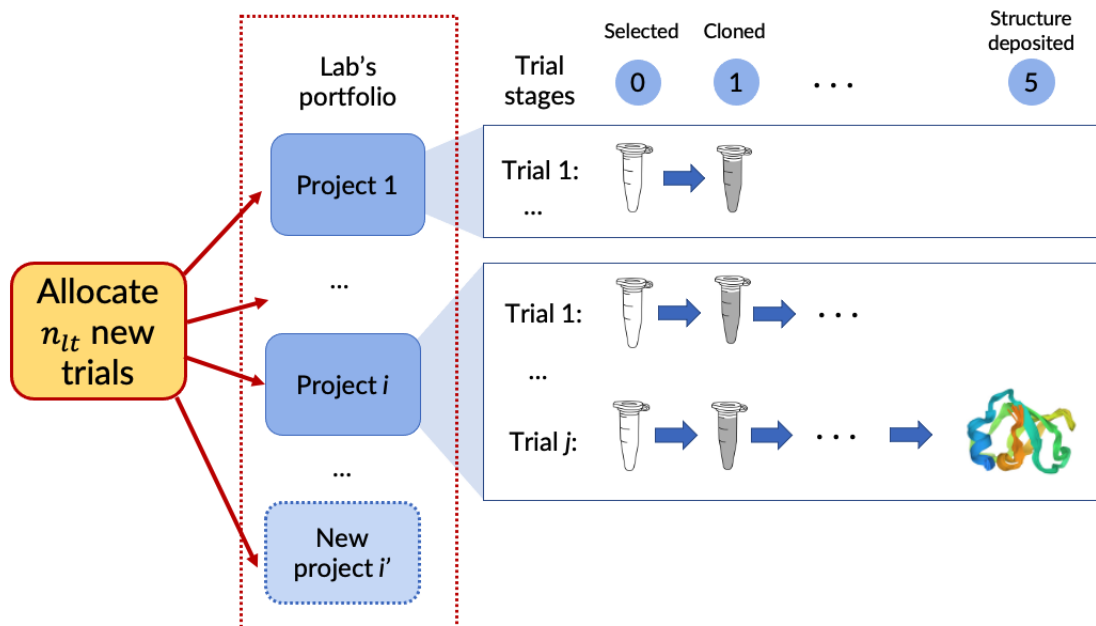
[1] For 15% of the trials, we observe the reasons for trial terminations. The reasons are mostly exogenous, e.g., "expression failed", "purification failed", "poor diffraction" and so on. In some cases, the trial termination reason is "duplicate target found" when the labs found a highly similar project was already determined. Upon further investigation, we find the ongoing trials in most of those cases were allowed to keep going. Some of those trials were even successful at all stages and produced duplicate structures.

This setting captures a realistically complex resource allocation problem over a long horizon. Figure 3 illustrates the decisions a structural biology lab faces. In each period, the lab decides to allocate units of input among a portfolio of research projects. The portfolio contains older projects the lab has worked on in the past and new projects that the lab or the NIH has identified for consideration. The lab's objective is to maximize over the horizon welfare from the output. The lab or the NIH determines the welfare from each unit of output according to some welfare weights unknown to us the researchers.

A key challenge in the lab's decision-making is that trial outcomes are very noisy and the lab has incomplete information about which trials are likely to produce output. For older projects that the lab has had trials on in the past or new projects that share similar physiochemical properties to older projects, the lab arguably has more information about their chances of success as the lab has observed past trial outcomes. For new projects that are novel and highly different, the lab has little information.

Figure 4   Machine learning analytics guides allocation decisions
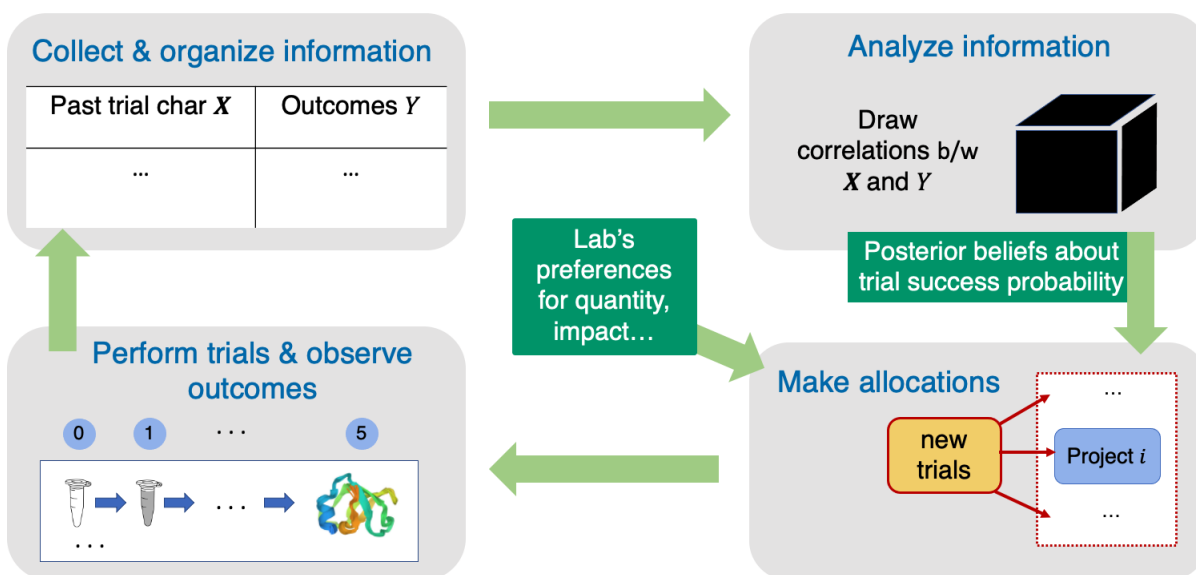
To improve efficiency of the resource allocation process, the labs in this setting used a machine learning approach to learn from the information they observed during past trial allocations to inform future allocation decision-making. They have thoroughly documented the process, providing guidance for modeling for this paper (Slabinski et al. 2007a,b, Jaroszewski et al. 2008, Price et al. 2009, Babnigg and Joachimiak 2010, Jahandideh et al. 2014). Figure 4 illustrates the process. As a lab performs trials and observes their outcomes, it collects and organizes the information about the characteristics and outcomes of its own trials and the trials performed by its peer labs. The lab

then analyzes that information, fitting machine learning models to draw correlations between trial characteristics and outcomes. The fitted models allow the lab to predict a new trial's probability of producing output based the trial's characteristics. That prediction forms the lab's posterior belief about the trial's output. The lab then allocates trials based on its posterior beliefs about different trials' output and its welfare weights. As the lab and its peer labs perform more trials and observe more information over time, the lab refits its machine learning models periodically.

The challenge of balancing exploration and exploitation naturally emerges in this resource allocation process. On one hand, the lab is motivated to concentrate on projects they are well-informed about, honing in on those with significant potential. On the other hand, there is a need for the lab to venture into unfamiliar projects to learn about their potential and uncover low-hanging fruits.

To illustrate the concept, let us consider two specific projects from the vast portfolio of projects considered by the Northeast Structural Genomics Consortium (NESG), one of the largest labs in our dataset, on May 30, 2009. Project A in Table 1 represents a project with a high exploitation value, while Project B represents a project with a high exploration value. The first row displays the lab's reasons for considering each project, as derived from the available labels and descriptions in the data. For instance, Project A was deemed valuable due to its human origin, biomedical significance, and disease relevance. The subsequent row indicates each project's trial history; Project A underwent eight trials before May 30, 2009, each with varying outcomes providing opportunities for the lab to learn about the project's potential, whereas Project B had not been trialed yet. The third row presents the predicted success probability for the next trial of both projects. These predictions come from a best-effort replica of the machine learning model the lab used, which was trained using all trial data available before May 30, 2009. Given that 98% of trials failed, Project A's 7% predicted success probability is considered high. The last row displays how similar each project was to the projects the lab and its peer labs attempted before May 30, 2009.[2] By construction, Project A had a 100% similarity because it had been attempted before this date. Project B shared a 58% similarity.

In this specific instance, the lab chose to proceed with a trial for Project B on May 30, 2009, bypassing Project A even though the latter was anticipated to have a higher likelihood of success and held biomedical significance. Despite the mere 0.1% success prediction for Project B, the trial was a success, resulting in the publication of a structure that became widely referenced and used. A crucial insight into the lab's decision-making process is the recognition that Project B was distinct from projects previously tackled by the lab and its peer labs. This distinction implied that the lab's prior experiences and the machine learning model, informed by past data, might not provide an

---

[2] Computed by comparing the sequence of amino acids of each molecule against all sequences attempted before May 30, 2009 using DIAMOND (Buchfink et al. (2015, 2021)). See Appendix A.4 for computational details.

Table 1   Exploit or explore? NESG's decision on May 30, 2009

	Project A		Project B
	Methyl-CpG-binding domain protein 4	**VS**	Malonyl-CoA decarboxylase
	(Homo sapiens)		(Cupriavidus metallidurans)
	involved in DNA repair		involved in fatty acid metabolism
Selection rationale	human molecule, biomedically important, related to diseases		novel
Previous trials	8 2 failed in stage 2 (expression) 3 failed in stage 3 (purification) 3 failed in stage 4 (crystalization)		0
Predicted prob of success next trial	6.92%		0.12%
Similarity to prev tried projects	100%		58%

Note: NESG allocated a trial to Project B on May 30, 2009. The trial was successful and produced the structure 4KS9 (Froese et al. 2013). In the initial six months after its addition to the PDB, the structure was downloaded 4,502 times. This surpassed 79% of the structures added between August 2007 and June 2013.

accurate prediction of the project's true likelihood of success. The lab's decision to explore Project B was influenced by factors beyond just the visible merits of the project and the success prediction rooted in historical data. Frequently in the data, we observe labs bypassing projects they seemed confident in, like Project A, to pursue those they had limited knowledge of, similar to Project B. This trend drives our use of modeling and structural estimation to uncover the extent to which labs prioritized exploration over exploitation.

2.3.   Key Policy Features

Several NIH policy features simplify this setting and mitigate worries about modeling and identification, one of them being restrictions on the labs' choices of projects. The NIH periodically drew families of novel molecules and solicited nominations of molecules from the biomedical research community for the labs to work on. The labs could also pursue projects of their own interest but had to communicate with the NIH about those projects well in advance (NIGMS 2007a,b, 2011b). Those processes limited the projects in the labs' choice sets, allowing us to construct choice sets that closely match the real ones.

Another feature is the collaborative nature of the grant program, which alleviates concerns that competition among labs might interfere with the balance between exploitation and exploration. The NIH funded the PSI labs through a collaborative U01 mechanism, rather than the competitive R01

mechanism. Throughout the program, the labs faced little competition for funding (and research questions due to the NIH assigning projects).[3]

The NIH also periodically evaluated the labs' productivity based on a set of metrics.[4] This evaluation process shaped the criteria that the labs deemed important, enabling us to focus solely on those metrics when modeling the observable aspects of the labs' preferences. Due to the NIH's active oversight of those labs' operations, it is plausible to disregard the principal-agent problem where the labs' preferences might deviate from the preferences of the funding agency. The evaluation metrics included the quantity of structures published, novelty, biomedical importance, human proteins, eukaryotic proteins,[5] and membrane proteins.[6]

In addition, an exogenous shift in NIH's preferences over the metrics in 2009 provides a nice robustness check. A well-fitting model of the labs' decision-making should capture this shift in its estimates of the labs' preferences. Before 2009, the NIH had a strong preference for the quantity of structures (NIGMS 2008a, 2011a). By mid-2008, the lack of emphasis on biomedically important projects had sparked heated debates in the community (Petsko 2007, Moore 2007); this prompted the NIH to boost its preference for biomedically important projects since 2009 (NIGMS 2007c, 2008b, 2009b,c). To facilitate the change, the NIH partnered with outside researchers to identify biomedical important projects and gave more attention to biomedical importance in its evaluation process.

The NIH also designed policies to facilitate the use of information in the resource allocation process, presenting valuable opportunities for policy analysis. One notable policy was a pilot phase from 2000 to 2004, during which the NIH refrained from setting production targets for the labs. The NIH's intention during the pilot was to enable the labs to engage in thorough exploration and tool development, ensuring that the labs would be well-equipped with valuable learnings when they scaled up production later on (NIGMS 2008a). The pilot consumed 17% of trials in the data, representing an approximate economic value of $220 million.[7] From 2005 onwards, the NIH set a

---

[3] A more typical lab faces additional tradeoffs due to competition. Hill and Stein (2021) found structural biology labs traded off quality of research and speed of publication, driven by the competitive race to be the first to publish. However, they also found the structural biology labs in our setting (labeled as " structural genomics (SG)" labs in their paper) were not as motivated by competition to sacrifice quality for speed.

[4] The NIH published statistics on those metrics online. The archived versions of those publications can be found by searching the urls http://targetdb.pdb.org/metrics/milestonestables.html and http://targetdb.pdb.org/Metrics/SummaryTable.html on the Internet Archive.

[5] Eukaryotes include all living organisms other than the eubacteria and archaebacteria. An eukaryote is an organism consisting of a cell or cells in which the genetic material is DNA in the form of chromosomes contained within a distinct nucleus.

[6] Membrane proteins are proteins found in the cell membrane. Membrane proteins are particularly hard for structure determination due to their physicochemical properties.

[7] This is by back-of-the-envelope calculations based on the average cost per trial. We multiply the total funding for the grant program ($1.269 billion) by 17% and round the result to the tens to obtain $220 million. Between 2000 and 2004,

production target of 200 structures per year for each large lab (NIGMS 2004, 2009a). Another key policy from the NIH required the labs to collect and share information among their peers regarding their trial allocation and outcomes. To facilitate this, the NIH invested $40 million in creating and maintaining an information system that not only archived and organized the data but also provided informatics tools for data analysis (NIGMS 2009d).[8]

3.   Data

The primary dataset used in this paper covers the allocation of 961,260 unique trials to 335,553 distinct projects by structural biology labs from 2000 to 2015. These allocation decisions were documented daily, noting whether a trial successfully produced a structure suitable for publication. The granular nature of this data is made possible due to the NIH's requirement for systematic information collection and sharing.

The primary dataset comprises 27 labs funded by the NIH and an additional 13 international labs that voluntarily shared their trial data. Through the shared information system, each lab had continuous access to the data of its peers.[9] The scale of operation varied significantly among labs in the dataset. Our analysis will primarily concentrate on the four largest labs,[10] which represented 71% of the projects and 85% of the trials documented. The remaining labs were significantly smaller, with diverse research focuses and funding amounts.

Our primary measure of output aggregates the outcomes of individual trials and counts the number of unique structures published by each lab. A trial successfully produces a structure for publication if it succeeds in all stages. We stress the importance of a structure's uniqueness, as a higher number of publications does not always equate to more value, especially when many structures are duplicates of previous ones. Such duplication arises because labs often allocated multiple trials to the same project concurrently, and sometimes, multiple trials succeeded. These

the NIH allocated approximately $370 million to the labs. However, a large part of this amount was earmarked for initial lab setups and purchases of long-lasting equipment. Neither the NIH nor the labs provide itemized breakdown of expenses. All dollar values are presented in 2015 dollars.

[8] Between 2010 and 2015, the information system was supported by NIH grants associated with FOA RFA-GM-10-004. Those grants were on average $2.6 million per year and had small variations across years. Between 2000 and 2009, the information system was hosted by the Research Collaboratory for Structural Bioinformatics (RCSB). The RCSB has been supported by grants from the NSF, the DOE, and six units of the NIH: the NIGMS, NLM, NCI, NCRR, NIBIB, and the NINDS (Chen et al. 2004). The RCSB does not detail the specific funding amount dedicated to the information system. We therefore assume a $2.6 million spending on the information system per year between 2000–2009. Multiplying $2.6 million with 16 years and rounding the result to the tens, we obtain $40 million. All dollar values are presented in 2015 dollars.

[9] Typically, a lab functioned as a consortium with sub-labs responsible for various trial stages. In total, 147 sub-labs are documented within this data.

[10] They are Joint Center for Structural Genomics (JCSG), Midwest Center for Structural Genomics (MCSG), New York Structural Genomics Research Consortium (NYSGRC), and Northeast Structural Genomics Consortium (NESG).

trials yielded structures of the same molecule, often with similar qualities. Hence, we only account for the first published structure from each project in measuring a lab's overall output. Our primary dataset includes 15,848 published structures, of which 10,501 are unique, accounting for 15% of global output of structures between 2000 and 2015. We supplement this primary output metric by tracking the number of citations and downloads each structure garners. These metrics offer insights into quality from different angles: citations in the academic literature reflect the amount of follow-on research directly built on the published structure, while downloads suggest other less tangible impacts, such as interest and the structure's role in advancing technologies in related areas, like DeepMind (2020).

Table 2 presents a list of the key variables along with their sources. While many of these variables are intuitive, detailed explanations of their construction are provided in Appendix A. In the remainder of this section, we will focus mainly on the construction of the variables that capture the labs' posterior beliefs about output.

3.1. Posterior Beliefs about Output

A key ingredient of this paper's model is the labs' posterior beliefs regarding the output from potential allocations at the time of making these allocations. We construct measures of these posterior beliefs by closely replicating the documented approach the labs used to form beliefs based on information they observed from previous allocations. The construction of these variables involves hundreds of variables sourced from various places. A comprehensive list of these variables can be found in Appendix A.5.

The four largest labs which we focus our analysis on formed and updated their posterior beliefs about whether a trial would succeed through supervised machine learning (Slabinski et al. 2007a,b, Jaroszewski et al. 2008, Price et al. 2009, Babnigg and Joachimiak 2010, Jahandideh et al. 2014). This involved using machine learning models to fit observed trial outcomes (successes or failures) on those trials' characteristics. The fitted model would be able to predict a new trial's probability of success based on the trial's characteristics. The labs periodically refitted the models when newer trial data became available to make updates. The labs started using this approach to form posterior beliefs as early as the beginning of 2005, when they had accumulated a considerable amount of trial data in the initial years of operation (Slabinski et al. 2007a,b). Over time and lab-wise, they used a variety of models, ranging from logistic regressions to support vector machine to random forest. Jahandideh et al. (2014) conducted a comparison study and found random forest worked best based on multiple metrics, including prediction accuracy. According to a discussion with a project coordinator from one of the labs, the accumulation of trial data played a more significant role in enhancing prediction quality than the selection of the machine learning model.

Table 2   Key variables and data sources

Variable	Variable Description	Source
Input allocation		
i	Identifier of a project or molecule; 335,553 unique projects	Berman et al. (2017)
j_i	Identifier of the jth trial of project i; 961,260 unique trials	Berman et al. (2017)
n_{lt}	Number of trials allocated at lab l on day t; mean= 35 trials per day per lab for the four largest labs	Berman et al. (2017)
NIH evaluation metrics		
$novel_i$	Binary $=1$ if the labs regarded project i as novel	Berman et al. (2017)
$prevStruct_{iy}$	Number of published structures in the same protein family as i prior to year y	EMBL-EBI (2021)
$biomed_i$	Binary $=1$ if the labs regarded project i as biomedically important	Berman et al. (2017)
$prevPub_{iy}$	Number of publications (not limited to structures) on molecule i prior to year y	UniProt Consortium (2021)
$human_i$	Maximal percentage identity of i to any human molecule	UniProt Consortium (2021)
$eukaryote_i$	Maximal percentage identity of i to any eukaryotic molecule	UniProt Consortium (2021)
$membrane_i$	Binary $=1$ if molecule i is a membrane protein	UniProt Consortium (2021)
Observed output		
Y_{ijt}	Binary $=1$ if trial j_i on day t successfully produced a structure; 961,260 observations	Berman et al. (2017)
Y_{ijkt}	Binary $=1$ if stage k of trial j_i on day t succeeded; 3,783,026 observations	Berman et al. (2017)
$citation_{iy}$	five-year citations and mentions of structure i published in year y	Varadi et al. (2020)
$download_{im}$	Number of downloads of structure i in month m (between Aug 2007 and Nov 2013)	wwPDB (2013)
Posterior beliefs about output		
$\hat{E}_{\bar{F}_t}(p_{ijt})$	Best-effort replication of the labs' posterior expectation of the probability of success of trial j_i on day t	Appendix B.1
$\widehat{Var}_{\bar{F}_t}(p_{ijt})$	Posterior variance of the probability of success of trial j_i on day t	Appendix B.1
$\hat{E}(citation_{iy})$	Expectation of five-year citations and mentions of project i published in year y	Appendix B.2
$\hat{E}(download_{iy})$	Expectation of five-year downloads of project i published in year y	Appendix B.3
Others		
$funding_{ly}$	Funding in dollars lab l received in year y	NIH (2019), NIH (2021)

Note: This table only summarizes key variables and data sources. For a full description of all data sources and variables, please see Appendix A.

To mimic how the labs updated their posterior for trial success probability, we fit a series of machine learning models over time between 2005 and 2015. To fit a model corresponding to the labs' posterior beliefs at time t, we use a training sample consisting of trial outcomes realized before t and those trials' characteristics. The characteristics cover the ones the labs used and fall under three categories:

- Physicochemical properties of the molecule based on scientific reasoning.
- Other characteristics of the project, such as novelty, biomedical importance, and the number of prior publications on the molecule.
- Past successes and failures of trials on the same project.

We use random forest to fit each model. Appendix B.1 shows the details.

The fitted models allow us to construct two variables, the posterior expectation of the probability of success $\widehat{E}_{\tilde{F}_t}(p_{ijt})$ and the posterior variance $\widehat{Var}_{\tilde{F}_t}(p_{ijt})$. Let the probability of success of a trial j_i at time t be p_{ijt}. The fitted model $\tilde{F}_t$ makes predictions about p_{ijt} based on the trial's characteristics, and these predictions constitute the posterior belief about p_{ijt}. The fitted model $\tilde{F}_t$ is a random forest consisting of an ensemble of submodels, each called a decision tree. Each decision tree fits the training sample independently and makes an independent prediction $\hat{p}_{ijt}^{ntree}$. The collection of these predictions forms the estimated posterior distribution. The mean of this distribution is the posterior expectation $\widehat{E}_{\tilde{F}_t}(p_{ijt})$, and the variance is the posterior variance $\widehat{Var}_{\tilde{F}_t}(p_{ijt})$.

Our best-effort replication may deviate from the labs' actual posterior beliefs, but we do not expect the deviations to bias estimation results of the allocation model. We note a list of possible deviations in Appendix B.1. One example is that the labs did not always predict $\widehat{Var}_{\tilde{F}_t}(p_{ijt})$. When they did, the variable took the form of comparing predictions from multiple models side by side (Slabinski et al. (2007a,b), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)). It is reasonable to believe the labs knew the value of analyzing the variation in predictions from different models, though they did not formally use a variable to represent that variation. $\widehat{E}_{\tilde{F}_t}(p_{ijt})$ and $\widehat{Var}_{\tilde{F}_t}(p_{ijt})$ therefore contain errors in the sense they may differ from the labs' actual posterior beliefs. But as long as the errors are random or at least not correlated with the actual allocation decisions, they should not bias the estimates of the allocation model. The errors may be correlated with the actual allocations if we fail to include in the training sample variables the labs actually used to form beliefs or make allocations. To minimize this risk, we include in our training samples all variables the labs ever mentioned using and all NIH evaluation metrics.

For the posterior beliefs about citations and downloads, we make the simplifying assumption that the labs' posterior beliefs stayed the same in all periods and were the same as the ground truth. This is equivalent to assuming that the labs only had incomplete information about the possibility of production and had perfect information about the kind of output produced if production does

happen. Given that the projects were very well-defined, this assumption is reasonable. Another evidence in support of this assumption is that, unlike the probability of success of a trial, the labs did not form and update their posterior beliefs about the number of citations and downloads of a publication in a systematic way.

We use ridge regressions to model the posterior expectations about citations $\hat{E}(citation_{iy})$ and about downloads $\hat{E}(download_{iy})$. For these models to closely match the ground truth, they must be capable of making minimal prediction errors outside the training sample. Appendices B.2 and B.3 show details of model fitting and visual evidence of out-of-sample model fit in cross validation.

4. Model

We introduce an estimable model of the labs' decision-making process, characterized by its use of a simple index to approximate the complex value function. Under this model, in each period, each lab first employs machine learning techniques to evaluate historical data, which helps them update their posterior beliefs about the output of different potential allocations. The lab then uses this posterior to compute a simple index to approximate the complex value function associated with allocating a trial to a given project. The lab then allocates trials to projects with the highest index values. The key strength of this model is its computational tractability. Additionally, it is realistically implementable, as labs could have feasibly used this approach in their actual decision-making.

We build intuition for this model as follows. First, we formally define the objective function of the labs' resource allocation problem. We then explain why the standard model of the labs' decision-making process is computationally intractable for estimation; this model assumes that the labs optimally solved the objective function using backward induction. Following this, we explain how index approximation differs from the standard model and why it is more tractable. We wrap up the discussion by explaining the theoretical foundation for index approximation. We define the main index approximation model as well as several alternative models for robustness checks at the end of this section.

4.1. Objective Function

To formally define the labs' objective function, we begin by introducing some notation. Let C_{lt} denote a lab l's choice set on day t. This choice set includes additional trials of existing projects in the lab's portfolio as well as initial trials of new projects. For example, on a day t when the lab has capacity to allocate $n_{lt} = 2$ new trials, and if the lab has already had three trials of project i up to day $t-1$, then the fourth and fifth trials of project i would be part of the choice set for day t. Similarly, for another project i' the lab has received from the NIH and but has not tried up to day $t-1$, the first and second trials of project i' would be included in the choice set for day t.

a_{ijt} denotes the allocation decision. $a_{ijt} \in \{0,1\}$ where 1 represents allocating trial j_i to project i on day t and 0 represents not allocating j_i on day t. The vector $\boldsymbol{a}_{lt}$ has its length equal to the cardinality of the choice set C_{lt} and represents the action on each of the project-trials in the choice set.

Y_{ijt} denotes the outcome of trial j_i on day t. $Y_{ijt} = 1$ if the trial succeeded and produced a structure for publication. Y_{ijkt} denotes the outcome of stage k of trial j_i on day t.

Ω_t denotes the labs' information set on day t. Because the labs had the ability to observe each others' allocations and output at all times, they all possessed the same information set on any given day. Ω_t includes the actions $\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_{t-1}$ and outcomes and stage-specific outcomes $\boldsymbol{Y}_1, ..., \boldsymbol{Y}_{t-1}$ that have been observed from all labs prior to day t.

p_{ijt} denotes the probability of success of trial j_i on day t. The labs and we, the researchers, have incomplete information about p_{ijt}. Let the prior distribution of p_{ijt} be $F_t(p_{ijt})$. The labs formed and updated posterior $\tilde{F}_t(p_{ijt}|\Omega_t)$ using supervised machine learning, and we made a best-effort replication of the labs' posterior offline in Section 3.1.

$\pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})$ denotes the payoff (in welfare units) from project-trial j_i. Whenever $a_{ijt} = 0$, $\pi_{ijt}(\cdot) = 0$ because a trial not allocated does not pay off. When $a_{ijt} = 1$, $\pi_{ijt}(\cdot)$ also depends on p_{ijt}, the probability of success of the trial, and $\boldsymbol{\theta}_{Xl}$, the welfare weights predetermined by the lab. One can think of $\pi_{ijt}(a_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})$ as a utility function that translates output into welfare according to the welfare weights $\boldsymbol{\theta}_{Xl}$. Take, for instance, a trial j_i expected to yield a novel and biomedically important structure if successful. This structure represents a single unit of output. Assume the lab assigns a welfare weight of 5 to a novel structure and a weight of 3 to a biomedically important structure; therefore, the total welfare the lab would receive from the success of this trial is 8. Each lab knew its welfare weights but we, the researchers, do not, so we will estimate $\boldsymbol{\theta}_{Xl}$ based on each lab's observed allocations. We describe the estimation procedure in Section 5.1.

A lab's objective is to maximize the posterior expected payoff over the finite horizon by choosing a sequence of allocation decisions. The objective function is as follows:

$$\max_{\boldsymbol{a}_{l1}(\Omega_1),...,\boldsymbol{a}_{lT}(\Omega_T)} \sum_{t=1}^{T} \sum_{j_i \in C_{lt}(\Omega_t)} \int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}_t(p_{ijt}|\Omega_t),$$
$$\text{subject to} \sum_{j_i \in C_{lt}(\Omega_t)} a_{ijt} = n_{lt}, \tag{1}$$
$$\text{and for all } j_i < j_i' \in C_{lt}(\Omega_t), \text{ if } a_{ijt} = 0 \text{ then } a_{ij't} = 0.$$

The first constraint is the capacity constraint. The second constraint rules out allocating a fifth trial to a project when the fourth trial has not been allocated.

Equation (1) characterizes a finite-horizon multi-armed bandit. Allocation problems character-ized by objective functions like equation (1) are called multi-armed bandits because of a classic

example of this type of problem. Imagine you are playing a five-armed bandit slot machine and you have ten opportunities to pull. In what way should you allocate your pulls across arms to maximize the expected total payoff if you have little prior information about the payoff of each arm? Intuitively, you want to use a few pulls to sample payoffs from different arms to learn which arms are promising then use the remaining opportunities to pull those arms. In our setting, each project is an arm and each trial of a project is a pull of an arm.

4.2.   Intractability of Backward Induction

In empirical structural estimation, the standard way of modeling how agents choose actions in a problem like equation (1) is to assume they solve the problem optimally through backward induction. The first step of this solution method involves specifying the value function for various actions using a Bellman equation. The value function of a particular state equates to the value attained by equation (1) at that state, following the sequence of optimal actions $\boldsymbol{a}_{l1}^*, \boldsymbol{a}_{l2}^*, ..., \boldsymbol{a}_{lT}^*$. In our context, the state variable is the information set Ω_t.

The Bellman equation has two additively separable components. One component represents the posterior expected payoff in the current period, which captures the value of exploitation. The other component represents the continuation value, which captures the value of exploration. We can define the action-specific value function $V_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl})$ of project-trial j_i on day t as follows

$$V_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}) = \underbrace{\int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})\, d\tilde{F}_t(p_{ijt}|\Omega_t)}_{\text{posterior expected payoff}} + \underbrace{E_{\Omega_{t+1}'}\big[\max_{\boldsymbol{a}_{l,t+1}} V_{i,j',t+1}(\Omega_{t+1}', \boldsymbol{a}_{l,t+1}; \boldsymbol{\theta}_{Xl})|\Omega_t, \boldsymbol{a}_{lt}\big]}_{\text{continuation value}}.$$
(2)

The value function $V_{ijt}(\Omega_t; \boldsymbol{\theta}_{Xl})$ of project-trial j_i on day t is equal to $V_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl})$ evaluated at the optimal $\boldsymbol{a}_{l1}^*, \boldsymbol{a}_{l2}^*, ..., \boldsymbol{a}_{lT}^*$.

Given that the continuation value integrates over the future evolutions of the state, solving the Bellman equation using backward induction to derive the optimal sequence of actions $\boldsymbol{a}_{l1}^*, \boldsymbol{a}_{l2}^*, ..., \boldsymbol{a}_{lT}^*$ is computationally intractable in our setting due to the curse of dimensionality. The state variable Ω_t includes all previous actions and outcomes. On certain days, the possible actions can number in the millions. With a time horizon spanning thousands of days, the total number of possible states at the end of the time horizon becomes astronomically large.

4.3.   Index Approximation

We present an alternative way of modeling how the labs chose actions, that is, we assume they used a simple index to approximate the complex value function. Index approximation overcomes the curse of dimensionality in empirical structural estimation because the approximated continuation

value does not integrate over the future evolutions of the state. Let $V_{ijt}^A(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_l)$ represent the lab's approximation to the action-specific value function in equation (2),

$$V_{ijt}^A(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_l) = \underbrace{\int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}_t(p_{ijt}|\Omega_t)}_{\text{as before}} + \underbrace{B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Bl})}_{\text{approximation to continuation value}} . \tag{3}$$

$V_{ijt}^A(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_l)$ has two additively separable components in direct parallel to the Bellman equation. The first component represents the posterior expected payoff in the current period and is identical to that in equation (2). The second component is a heuristic function $B_{ijt}(\cdot)$ that we assume the lab used to approximate the value of exploring of a project, and this heuristic is based only on information available to the lab at time t. Theoretical literature on multi-armed bandits would call $B_{ijt}(\cdot)$ the "exploration bonus."

Because the computation of the index approximation utilizes only the currently available information, the lab can directly solve equation (3), eliminating the need for backward induction. Assume the lab chose actions based on the approximate value function, acting as if the approximate value function was the true value function, it would simply choose actions $\boldsymbol{a}_{lt}^{A*}$ that would maximize the sum of the approximate values on each day t:

$$\begin{aligned} \boldsymbol{a}_{lt}^{A*} = \arg\max_{\boldsymbol{a}_{lt}} \quad & \sum_{j_i \in C_{lt}(\Omega_t)} V_{ijt}^A(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_l), \\ \text{subject to} \quad & \sum_{j_i \in C_{lt}(\Omega_t)} a_{ijt} = n_{lt}, \\ & \text{and for all } j_i < j_i' \in C_{lt}(\Omega_t), \text{ if } a_{ijt} = 0 \text{ then } a_{ij't} = 0. \end{aligned} \tag{4}$$

By adopting some very intuitive functional form assumptions about the payoff function $\pi_{ijt}(\cdot)$ (see Appendix C.1 for details), we can simplify $V_{ijt}^A(\cdot)$ to depend solely on a_{ijt}, rather than the entire $\boldsymbol{a}_{lt}$ vector, and we can make the second constraint always hold at the solution. Under those assumptions, $\boldsymbol{a}_{lt}^{A*}$ effectively becomes synonymous with an index rule: the lab computes index $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$ for each project-trial in C_{lt}, and $\boldsymbol{a}_{lt}^{A*}$ simply involves allocating resources to the n_{lt} trials with the highest index values.

There is a substantial body of literature on multi-armed bandits that theoretically supports the use of an index to approximate the value function in computationally challenging settings. Theoretical literature has found optimal or nearly optimal indices for some bandits. In the case of the standard infinite-horizon discounted multi-armed bandit,[11] the optimal index is the Gittins

---

[11] Defined by the objective function

$$\max_{\boldsymbol{a}_{l1}(\Omega_1), \boldsymbol{a}_{l2}(\Omega_2), \dots} \sum_{t=1}^{\infty} \beta^t \sum_{j_i \in C_{lt}(\Omega_t)} \int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}^{(i)}(p_{ijt}|\Omega_t^{(i)}), \text{subject to} \sum_{j_i \in C_{lt}(\Omega_t)} a_{ijt} = 1 \text{ for all } t. \tag{5}$$

index (Gittins and Jones 1979, Gittins 1979). In the case of the standard finite-horizon multi-armed bandit defined by the objective function

$$
\max_{\boldsymbol{a}_{l1}(\Omega_1),\ldots,\boldsymbol{a}_{lT}(\Omega_T)} \sum_{t=1}^{T} \sum_{j_i \in C_{lt}(\Omega_t)} \int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}^{(i)}(p_{ijt}|\Omega_t^{(i)}),
$$
$$
\text{subject to} \sum_{j_i \in C_{lt}(\Omega_t)} a_{ijt} = 1 \text{ for all } t, \tag{6}
$$

Lai and Robbins (1985) and Lai (1987) show that index rules do not provide exact solutions but are asymptotically optimal as the number of periods T goes to infinity, and have nearly optimal performance from both the Bayesian and frequentist viewpoints for moderate and small T. These nearly optimal index rules can be interpreted as the upper confidence bounds (UCB) for the posterior expected payoffs. The findings from their research catalyzed the development of a body of literature focused on indices that serve as approximations to the UCB (Agrawal 1995, Bubeck et al. 2012, Cappe et al. 2013). The simplest form of a UCB index is just a function based on the number of previous pulls of the arm (Auer et al. 2002).

Equation (1) differs from the standard finite-horizon multi-armed bandit in a few ways, and, as a result, theoretical literature has yet to find an optimal index for it, but the literature has also shown indices may be good heuristics in those cases. The first difference is that equation (1) does not assume stationarity of payoffs. In equation (1), $\tilde{F}_t(\cdot)$ is indexed with the t subscript, reflecting the continuous evolution of science that could change the payoff of a trial independent of states and actions. Non-stationary bandits are called "restless bandits" (Whittle 1988). Theoretical research (Ortner et al. 2012, Lattimore and Szepesvári 2020) suggests indices are generally suboptimal for such bandits. Nonetheless, under specific assumptions about changes in $\tilde{F}_t(\cdot)$, like abrupt changes in unknown periods, Garivier and Moulines (2011) show that UCB-like indices can match the lower bound on regret up to a logarithmic factor.

Another difference in equation (1) is its more complex action space. While equation (6) restricts to allocating one trial to one project in each period, equation (1) permits multiple trials either to the same or different projects per period if $n_{lt} > 1$ and n_{lt} can vary across periods. Such bandits, where multiple arms can be pulled each period and reveal their payoffs, are known as "combinatorial semi-bandits." Research indicates that indices are effective for stochastic combinatorial semi-bandits (Kveton et al. 2015, Chen et al. 2016, Wang and Chen 2018, Chen et al. 2021), though this is contingent on strong assumptions regarding the action space and the payoff function.

Moreover, equation (1) differs by not assuming independence among arms. It computes payoffs using the posterior distribution $\tilde{F}_t(p_{ijt}|\Omega_t)$, based on the complete information set at t, while equation (6) uses $\tilde{F}^{(i)}(p_{ijt}|\Omega_t^{(i)})$, where the information set is specific to each project i, preventing learning spillovers from one project's trials to another. Given many research projects share

physicochemical characteristics, their payoffs might be correlated. In cases where agents have such contextual data to predict payoffs, the bandits are called "contextual bandits" (Woodroofe 1979, Langford and Zhang 2007). Recent studies indicate that UCB-like indices in such settings can achieve a nearly optimal regret guarantee (Guan and Jiang 2018, Zhou et al. 2020). These indices formulate models correlating contextual features with outcomes, aiding in predicting each pull's UCB payoff.

Indices are also widely studied in numerical simulations with real data and widely used in various real-world settings like inventory control, online advertising, and dynamic pricing (e.g. Nguyen-Thanh et al. 2019, Jin et al. 2021). Due to the complexity of real-world bandit problems, the application of indices often lacks tight theoretical justification. Given that current theoretical research often cannot identify optimal solutions for complex bandit problems and instead relies on index-based heuristics, it seems unlikely that the labs in our setting always made optimal decisions. Hence, we contend that index approximation serves more than just computational convenience; it represents a credible empirical model for lab decision-making, as labs could realistically implement this straightforward approach in their decision processes.

4.4. Models of Index Approximation

Our main model adapts the widely-used UCB index from Auer et al. (2002), where $B_{ijt}(\cdot)$ is a convex decreasing function of the number of trials previously allocated to the project.[12] Intuitively, UCB approximates the upper confidence bound of the posterior expected payoff from an allocation. Allocating trials to a new project that the lab has poor information about has a high UCB. As the lab allocates more trials to the project and has better information about its productivity, the UCB of further allocations decreases. Our modification to Auer et al. (2002) is an additional term that captures time-discounting of the value of older projects attempted long ago. This term serves as a proxy for changes in the perceived value of exploring a project over time from the lab's perspective. For instance, if the lab experienced personnel changes and project-specific knowledge was linked to departing researchers, this could lead to a decreased inclination to revisit projects that were attempted a long time ago. The main model is as follows.

Main Model (UCB+D) *The labs set* $B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Bl}) = a_{ijt} \left[\sqrt{\frac{\theta_{B1,l}}{j}} + \theta_{B2,l} \cdot (t - t'_{i,t}) \right]$.

---

[12] According to Auer et al. (2002),

$$B_{ijt}(\cdot, a_{ijt} = 1) = \begin{cases} \infty & \text{if } j = 1 \\ \sqrt{\frac{2\ln(N_{lt})}{j-1}} & \text{if } j = 2, 3, 4... \end{cases} \qquad (7)$$

Here, N_{lt} represents the total number of pulls conducted by the agent up to that point. More recent implementations adopt a constant value $\theta_{B1,l}$ instead of $2\ln(N_{lt})$ (Lattimore and Szepesvári 2020). In our main model, we do not assign an infinite value to $B_{ijt}(\cdot)$ when $j = 1$. Doing so would result in an infinite V_{ijt}^A, complicating the estimation process through maximum likelihood and the identification of θ_{B1}.

The free parameter $\theta_{B1,l}$ captures the amount of exploration. A larger $\theta_{B1,l}$ shifts out $B_{ijt}(\cdot)$, which means the lab would value exploration more. If a lab did not value exploration at all, $\theta_{B1,l} = 0$. The free parameter $\theta_{B2,l}$ captures the amount of discounting of older projects. The variable $(t - t'_{i,t})$ measures the duration between the current period t and the period $t'_{i,t}$ in which the last trial of project i occurred.[13] If $\theta_{B2,l}$ has a negative coefficient estimate, it would suggest the lab discounted projects attempted long ago when deciding which projects to explore further. If the lab did not discount older projects, $\theta_{B2,l} \geq 0$.

We assess the fit of our main model by comparing it with several alternative index models as robustness checks. While these alternatives are the apparent choices, it is possible to consider other models as well. A variety of indices have been suggested by both theoretical research and practical applications for bandit problems, and we could in principle evaluate each of them. Discussions with NIH program officers and a project coordinator from one of the four largest labs indicated that their trial allocation was heuristic in nature. They adopted a "high-throughput" approach, initially assigning one trial to many projects and then adding more trials to projects that were either important or showed potential based on earlier trials. Therefore, simpler indices might effectively replicate the actual allocation decision process. Consequently, we opted to test only the most parsimonious yet well-researched indices from the theoretical literature.

The first alternative model, the greedy model, assumes the labs maximize the sum of posterior expected payoffs and completely ignore exploration. This model sets the exploration bonus $B_{ijt}(\cdot)$ to zero and does not have any free parameters in $B_{ijt}(\cdot)$. It is equivalent to restricting both $\theta_{B1,l}$ and $\theta_{B2,l}$ in the main model to zero. The second model, the Gittins index, prescribes optimal actions for the standard infinite-horizon discounted bandit discussed in Section 4.2. Since Equation (1) differs, this model assumes that the labs used the Gittins index without its optimality guarantee. Computing the exact Gittins index is difficult, so we use Brezzi and Lai (2002)'s approximation to the index. This approximation explicitly includes the posterior variance of payoff $Var(\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}))$ in $B_{ijt}(\cdot)$.[14] Like the first alternative model, this model does not have free parameters in $B_{ijt}(\cdot)$, except for $\boldsymbol{\theta}_{Xl}$. The third model, a simple UCB index akin to Auer et al. (2002), has one free parameter $\theta_{B1,l}$ in $B_{ijt}(\cdot)$ that needs to be estimated. The fourth model relaxes how posterior variance enters $B_{ijt}(\cdot)$ compared to the Gittins index. It also has one free parameter in $B_{ijt}(\cdot)$. The final model introduces time-discounting to the fourth model. Like the main model, it has two free parameters in addition to $\boldsymbol{\theta}_{Xl}$.

---

[13] If $j = 1$, $(t - t'_{i,t}) = 0$.

[14]

$$Var(\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl})) = \int \left(\pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) - \int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}_t(p_{ijt}|\Omega_t) \right)^2 d\tilde{F}_t(p_{ijt}|\Omega_t). \tag{8}$$

Alternative Model 1 (Greedy) *The labs set* $B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}) = 0$.

Alternative Model 2 (Gittins) *The labs set* $B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}) = \psi(\cdot) \cdot \sqrt{Var(\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}))}$.[15]

Alternative Model 3 (UCB) *The labs set* $B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \theta_{B1,l}) = a_{ijt}\sqrt{\frac{\theta_{B1,l}}{j}}$.

Alternative Model 4 (FlexGittins) *The labs set*
$B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}, \theta_{B1,l}) = \theta_{B1,l} \cdot \psi(\cdot) \cdot \sqrt{Var(\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}))}$.

Alternative Model 5 (FlexGittins+D) *The labs set*
$B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}, \boldsymbol{\theta}_{Bl}) = \theta_{B1,l} \cdot \psi(\cdot) \cdot \sqrt{Var(\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}))} + a_{ijt} \cdot \theta_{B2,l} \cdot (t - t'_{i,t})$.

5. Model Fitting

We developed a two-stage method to estimate and validate the main model. In the first stage, we estimate the free parameters $\boldsymbol{\theta}_{Xl}$ and $\boldsymbol{\theta}_{Bl}$ in the main model by maximizing the likelihood of the observed allocation decisions. In the second stage, we use the estimated parameters $\widehat{\boldsymbol{\theta}}_{Xl}$ and $\widehat{\boldsymbol{\theta}}_{Bl}$ to forward simulate the labs' entire history of input allocation and output. We compare the patterns of input allocated and output from the simulated data to those from the actual data to assess model fit. We also repeat this two-stage procedure for each of the alternative models to benchmark the fit of the main model against the alternatives.

5.1. Estimation

In our model, we assume that the labs made decisions using the approximated value function. Therefore, the observed actions $\boldsymbol{a}_{l1}^o, \boldsymbol{a}_{l2}^o, ..., \boldsymbol{a}_{lT}^o$ correspond to the solutions $\boldsymbol{a}_{l1}^{A*}, \boldsymbol{a}_{l2}^{A*}, ..., \boldsymbol{a}_{lT}^{A*}$ in equation (4). From this, we can construct the likelihood of the observed allocations $P(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$ based on the form of $V_{ijt}^A(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$ in each model. This likelihood function is then used to estimate the model-specific free parameters $\boldsymbol{\theta}_l = \boldsymbol{\theta}_{Xl} \cup \boldsymbol{\theta}_{Bl}$.

Figure 5 presents a three-step overview of our estimation procedure. In the first step, we set the lab's information set Ω_t in each period as the one observed in actual data, which includes all previous actual allocations and outcomes up to each period t. In the second step, given Ω_t in each period, we make a best-effort replication of the lab's machine learning training process using Ω_t

---

[15] The function $\psi(\cdot)$ is defined as

$$\psi(s) = \begin{cases} \sqrt{s/2} & \text{if } s \le 0.2 \\ 0.49 - 0.11s^{-1/2} & \text{if } 0.2 < s \le 1 \\ 0.63 - 0.26s^{-1/2} & \text{if } 1 < s \le 5 \\ 0.77 - 0.58s^{-1/2} & \text{if } 5 < s \le 15 \\ \{2log(s) - log(log(s)) - log(16\pi)\}^{-1/2} & \text{if } s > 15, \end{cases} \quad (9)$$

where $s = \frac{Var(p_{ijt}|\Omega_t)}{-ln(\beta)E(p_{ijt}|\Omega_t)(1 - E(p_{ijt}|\Omega_t))}$. We set the discount factor $\beta = 0.95$.

Figure 5   **Overview of the estimation procedure**
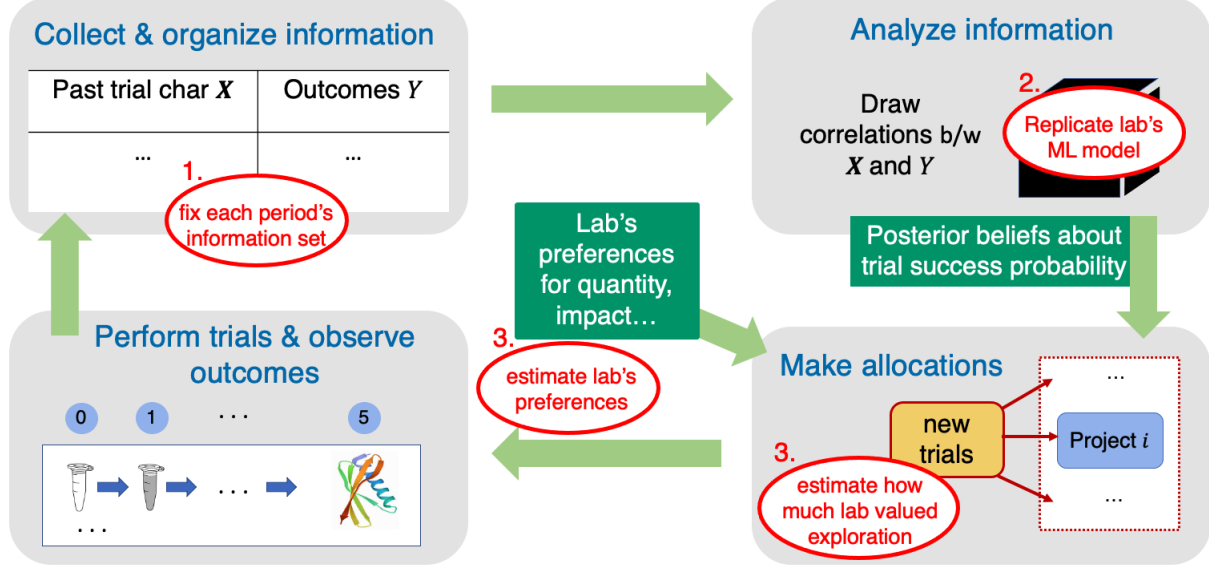
as the training data to form their posterior beliefs $\tilde{F}_t(p_{ijt}|\Omega_t)$. With this posterior, we formulate the index value of each project-trial $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$ as a function of $\boldsymbol{\theta}_l$.[16] We then express the likelihood of choosing the specific project-trial $P(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$ as a smooth, monotonically increasing function of the index value $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$.[17] In the third step, we construct the total likelihood function to be maximized. This likelihood function sums over the log likelihood of the observed action for each project-trial in choice sets $C_{l1}, C_{l2}, ..., C_{lT}$.[18] These likelihoods include $P(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ for trials actually allocated and $1 - P(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ for trials in those choice sets but which were *not* allocated. The total likelihood function is represented as:

$$\boldsymbol{\theta}_l^* = \arg\max_{\boldsymbol{\theta}_l} \sum_t^T \left(\underbrace{\sum_{j_i \in C_{lt}, \, a_{ijt}^o = 1} log(P(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l))}_{\text{actually allocated trials}} + \underbrace{\sum_{j_i \in C_{lt}, \, a_{ijt}^o = 0} log(1 - P(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l))}_{\text{actually not allocated trials}} \right).$$

(10)

We then search for the lab preference parameters $\boldsymbol{\theta}_{Xl}$ and free parameters in the exploration bonus term $\boldsymbol{\theta}_{Bl}$ that maximize the above likelihood function.

---

[16] We made some very intuitive functional form assumptions about the payoff function $\pi_{ijt}(\cdot)$, so that $V_{ijt}^A(\cdot)$ depends solely on a_{ijt}, rather than the full $\boldsymbol{a}_{lt}$ vector. These assumptions allow us to replace $V_{ijt}^A(\Omega_t, \boldsymbol{a}_{ijt}; \boldsymbol{\theta}_l)$ with $V_{ijt}^A(\Omega_t, a_{ijt}; \boldsymbol{\theta}_l)$. The functional form assumptions also ensure that the second constraint of equation (4) is satisfied at the solution, enabling us to ignore this constraint in our maximum likelihood estimation routine. Furthermore, the functional form is designed to accommodate scenarios where multiple trials can be allocated to a project in a single period, provided the lab's capacity $n_{lt} > 1$. It ensures the lab only receives welfare from the first successful trial on a project that yields a unique structure. Details on the functional form assumptions and derivations can be found in Appendix C.1.

[17] See Appendix C.2 for details of the formulation of the likelihood function $P(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$.

[18] See Appendix C.3 for how we constructed the choice sets.

Due to the simplicity of computing index approximations, estimation is feasible even though the choice sets contain millions of possible actions over thousands of days. A further trick to reduce computational burden is to compute in each iteration the log likelihoods for a random sample of the possible allocations in the choice sets, rather than for the full choice sets. The number of possible allocations in a full choice set could be large because n_{lt} could be large. Recall that the mean of n_{lt} for the four largest labs is 35. When $n_{lt} = 35$, the $(j+1)$th, $(j+2)$th,..., $(j+35)$th trials of every project in the lab's portfolio are in the full choice set. See Appendix C.3 for how we sampled from the choice sets.

The estimates include two kinds of parameters. $\hat{\boldsymbol{\theta}}_{Bl}$ are the free parameters in the exploration bonus term $B_{ijt}(\cdot)$, if there are any. $\hat{\boldsymbol{\theta}}_{Xl}$ are welfare weights that enter into the payoff function $\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl})$. We let $\boldsymbol{\theta}_{Xl}$ have eight parameters. Seven of these parameters correspond to the NIH evaluation metrics shown in Table 2. For example, $\theta_{biomed,l}$ corresponds to $biomed_i$ and captures the amount of welfare lab l receives when it publishes a biomedically important structure. One additional parameter $\theta_{quant,l}$ captures the baseline amount of welfare lab l receives per unique structure. See Appendix C.1 for the specification of the payoff function $\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl})$.

Identification of model parameters is very intuitive and is based on revealed preferences and variation of project-trial characteristics in the choice sets. Let us consider the identification of $\theta_{biomed,l}$. Suppose we observe a variety of biomedically important and less important projects in the choice sets, and we notice lab l choosing a large portion of biomedically important projects. Intuitively, this suggests the lab strongly valued this characteristic. Our likelihood function is specified such that a large positive $\theta_{biomed,l}$ increases the posterior expected payoff of biomedically important projects and, consequently, the index values $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$ of those projects relative to the distribution of index values in the lab's choice sets. In contrast, a large positive $\theta_{biomed,l}$ would decrease the index values of non-biomedically important projects relative to the distribution of index values in the lab's choice set. Therefore, a large positive $\theta_{biomed,l}$ increases the likelihood of allocating to biomedically important projects and decreases the likelihood of allocating to non-biomedically important projects, which maximizes the likelihood of the observed allocation decisions. Identification of other parameters follows the same intuition.

5.2. Simulation

In this second stage, we use the estimated parameters $\widehat{\boldsymbol{\theta}}_{Xl}$ and $\widehat{\boldsymbol{\theta}}_{Bl}$ of each model to forward simulate each lab's entire history of input allocation and output. This enables us to compare the patterns of input allocation and output from the simulated data with those from the actual data, thereby gauging the model's fit. The simulation procedure is as follows.

Algorithm 1: Forward simulation of trial allocations and outcomes

Initialization:

- **"Ground truth" output:** We train a flexible model F^* using Ω_{T+1}, the terminal information set including all trial outcomes and characteristics, to predict the "true" trial success probability. Unlike the posterior belief model $\tilde{F}_t$ from Section 3.1, which reflects how labs form their posterior beliefs and does not necessarily provide unbiased estimates of trial success probability, F^* needs to provide unbiased estimates. To correct potential biases and improve the labs' machine learning models, F^* diverges from $\tilde{F}_t$ in several aspects. The training details of F^* are outlined in Appendix B.1.

- **Project space:** F^* may not accurately predict the success probabilities of trials on projects that the labs never attempted, particularly those substantially different from the projects they did undertake. Therefore, we constrain our project space to include only projects the labs actually attempted. Despite this limitation, the labs retain ample opportunities for exploration and exploitation, considering that just 3% of the attempted projects led to successful structures.

- **Allocation model:** For model validation, we set the allocation model's parameters $\boldsymbol{\theta}'_{Xl}$ and $\boldsymbol{\theta}'_{Bl}$ to be equal to the parameters $\widehat{\boldsymbol{\theta}}_{Xl}$ and $\widehat{\boldsymbol{\theta}}_{Bl}$ as estimated under each model.

- **Prior allocations and output:** Allocations and outcomes before 2005 are used as prior data and not simulated.

For each day t between 2005 and 2015, do

- **Update posterior beliefs** Refit the lab's posterior belief model $\tilde{F}'_t(\Omega'_t)$, based on counterfactual information set Ω'_t which includes both actual and simulated actions and outcomes observed up to t.

- **Make allocations** Based on posterior $\tilde{F}'_t(\Omega'_t)$ and allocation model $\boldsymbol{\theta}'_{Xl}$ and $\boldsymbol{\theta}'_{Bl}$, compute $\widehat{VA}'_{ijt}(\Omega'_t, a_{ijt} = 1; \boldsymbol{\theta}'_{Xl}, \boldsymbol{\theta}'_{Bl})$ for trials in the counterfactual choice set C'_{lt}. Allocate trials according to $\widehat{VA}'_{ijt}$ up to lab capacity constraint n_{lt}.

- **Simulate outcomes** Generate the "true" trial success probability $p^*_{ijt}|F^*(\Omega_{T+1})$ for the allocated trials. Draw outcomes $Y'_{ijt} \sim Bernoulli(\hat{p}^*_{ijt})$ for these trials.

- **Update information set** Incorporate the actions and outcomes realized during this period into the information set, then proceed to the next period.

End

---

6. Model Fitting Results

6.1. Estimation Results

Table 3 illustrates that our main model, UCB+D, achieves significantly better fit than alternative models with minimal or no extra parameters, as shown by log likelihood comparisons. Notably, UCB+D's log likelihood is just half of FlexGittins+D's with the same parameter count for this particular lab. This result extends to other major labs as indicated in Appendix Tables D1–D3, where UCB+D's log likelihood is 28% to 41% smaller in magnitude than FlexGittins+D's. Even the basic UCB model, without discounting, either matches or surpasses FlexGittins+D in fit across labs, all while having fewer parameters. Moreover, Table 3 highlights the main model's superior

Table 3   **Comparison of model fits**

Model	Free parameters in $B_{ijt}(\cdot)$	Log likelihood	Avg $\hat{P}(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l)$ actually allocated trials	Avg $1 - \hat{P}(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l)$ actually not allocated trials
Greedy	0	-613,630	0.578	0.993
Gittins	0	-430,746	0.608	0.996
UCB	1	-204,228	0.763	0.998
FlexGittins	1	-412,112	0.614	0.996
FlexGittins+D	2	-230,048	0.710	0.998
UCB+D	2	-119,861	0.837	0.999

Note: Each model is estimated separately for the periods 2005–2008 and 2009–2015, reflecting the NIH's preference shift in 2009. The estimates are presented in Appendix Table D5. The total log likelihood for each model is the sum of the log likelihoods from both periods. To calculate the average $\hat{P}(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l)$ for the actually allocated trials, we sum the predicted likelihoods of all allocated trials across these periods and divide by their total number. Similarly, the average $1 - \hat{P}(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l)$ for the actually not allocated trials is the sum of the predicted likelihoods for all trials not allocated, divided by their count. The estimation uses data from 2005 to 2015 from NESG, one of the four major labs, encompassing 109,738 actual allocations. To manage the size of the choice sets, we employed random sampling (detailed in Appendix C.3), resulting in 32,410,947 trials in the sampled choice sets. Appendix Tables D1–D3 show estimation results from the other three large labs.

Table 4   **Key parameter estimates in the UCB+D model**

Parameter	2005–2008 (1)	2009–2015 (2)
θ_{B1}	158.28 [156.35, 160.13]	119.52 [118.15, 121.41]
θ_{B2}	-2.28 [-2.23, -2.30]	-4.71 [-4.66, -4.73]
θ_{biomed}	21.50 [21.32, 21.72]	52.86 [52.69, 52.93]

Note: This table presents estimates for key parameters in the UCB+D model, specifically for NESG, one of the four major labs. For the full set of estimates across various models for this lab, refer to Appendix Table D5. Estimates for the key parameters in the UCB+D model for other labs are in Appendix Table D6. The "2005–2008" column includes data from that period, featuring 59,261 actual allocations, with 5,628,158 trials in the choice sets post-random sampling (see Appendix C.3 for details). The "2009–2015" column covers data from these years, with 50,477 actual allocations and 26,782,789 trials in the choice sets after sampling. 95% confidence intervals, calculated using Chernozhukov and Hong (2003)'s MCMC approach, are provided in brackets. These confidence intervals are almost identical to those calculated using Chen et al. (2018)'s Procedure 1.

predictive accuracy for both actually allocated (83.7% likelihood on average) and actually not allocated project-trials (99.9% likelihood on average), the highest among all models. Similar fit is observed for other large labs in the appendix tables. We also validated the main model's fit out-of-sample: using odd-year data for fitting and even-year data for log likelihood comparisons. This approach, detailed in Appendix Table D4, shows comparable in-sample and out-of-sample fit across labs.

Table 4 presents the main model's key parameter estimates. Since these estimates reflect the lab's preferences in welfare units, they are hard to interpret on their own. Still, comparisons of these estimates to zero or across different periods offer insights. θ_{B1} captures the amount of exploration under the main model. If the lab did not explore, $\theta_{B1} = 0$. The estimates of θ_{B1} in Table 4 are large, positive, and significantly different from zero at the 95% confidence level, allowing us to confidently reject the hypothesis that the lab did not explore. Additionally, estimates of θ_{B2} are negative and significantly different from zero at the 95% confidence level, allowing us to reject the hypothesis that the lab did not discount older projects. Additional robustness is evident from the post-2009 increase in the θ_{biomed} estimate compared to pre-2009, supported by non-overlapping 95% confidence intervals. This aligns with the NIH's heightened emphasis on biomedical importance post-2009. This table focuses on one major lab, with Appendix Table D6 providing similar findings for other large labs.

6.2. Simulation Results

Table 5   Comparison of simulated outcomes in model validation

Model	Projects attempted	Unique structures	Citations	Downloads (millions)
Greedy	14,175	597	1,787	14.1
Gittins	14,980	631	1,898	14.9
UCB	59,164	1,052	3,236	24.5
FlexGittins	15,005	621	1,830	14.6
FlexGittins+D	17,382	638	1,892	15.0
UCB+D	59,947	1,097	3,376	25.6
Actual	59,953	1,053	3,502	24.5[†]

Note: Each simulation utilizes parameter estimates of the corresponding model. See Table D5 for those estimates. The results represent the average of three simulation runs for each model. The actual outcomes are displayed in the final row. These simulation outcomes are specific to NESG, one of the four major labs. Similar findings for the other three large labs are available in Appendix Table D7. [†]The download data are only available between 2007 and 2013, so the actual five-year downloads may not be available for some projects. We predicted the five-year downloads for each project based on the project's characteristics using the model described in Appendix B.3.

Figure 6   Simulated number of unique structures under different models

(a) Greedy

(b) Gittins

(c) UCB

(d) FlexGittins

(e) FlexGittins+D

(f) UCB+D
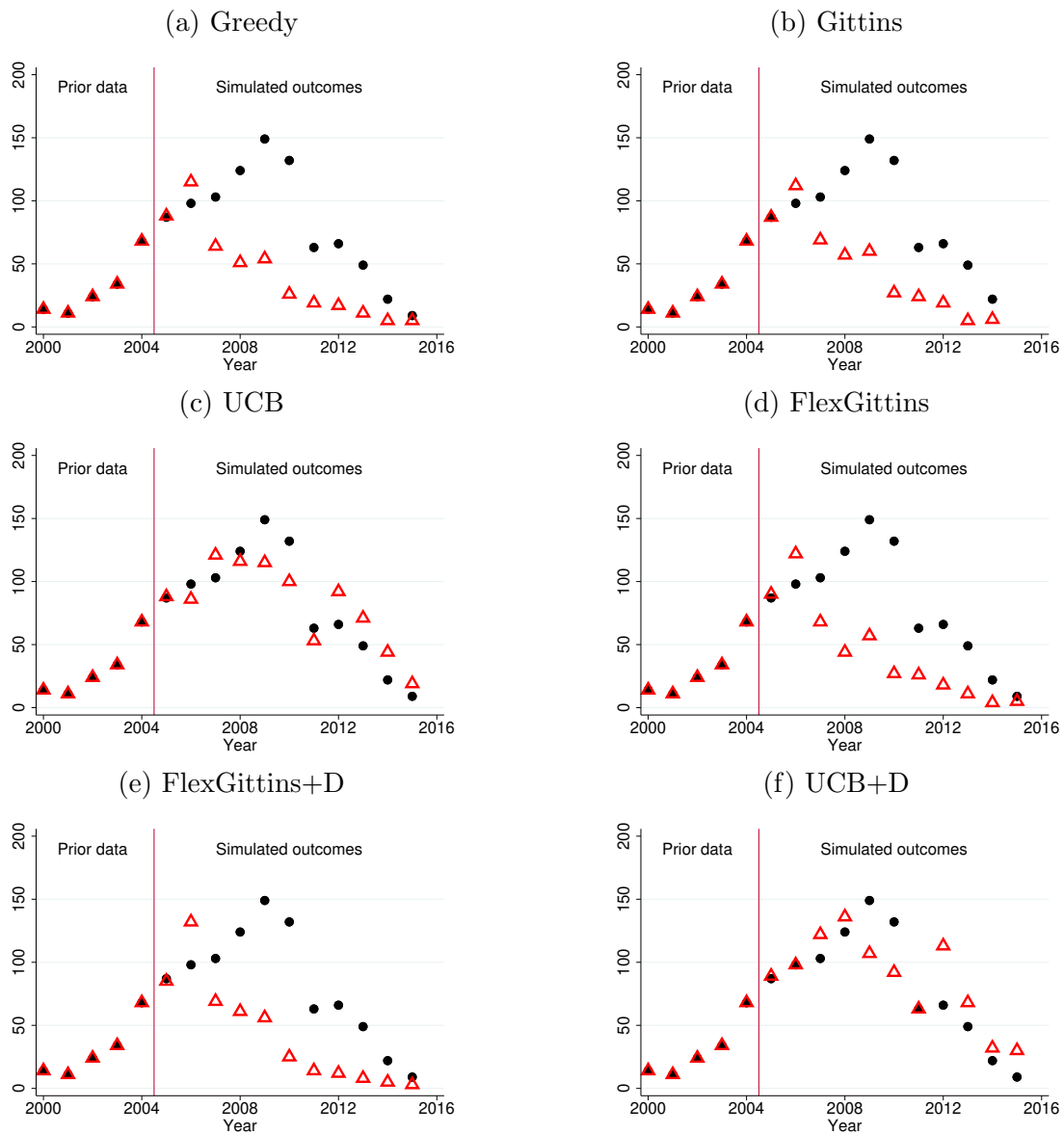
Note: Black dots represent the actual number of unique structures published by NESG each year, exhibiting a hump-shaped trend. Before 2009, the lab gradually built up its capacity. In 2009, upon learning that the grant program would conclude in 2015, it began to gradually reduce its capacity. Red triangles represent the simulated number of unique structures published by NESG each year. Each data series is derived from a single simulation of the corresponding model.

Figure 6 provides a detailed breakdown of simulated output for a major lab, showing that while the greedy and the Gittins index models diverge from real data, the UCB models closely mirror it. Additional visual evidence of the main model's ability to replicate patterns in real data is shown in Appendix Figure D2, which displays the main model's simulated distribution of the number of

trials per project closely matching the actual distribution. Appendix Figure D3 demonstrates that the characteristics of counterfactually allocated trials in the main model's simulation closely match those of the actually allocated trials and accurately reflect the drastic shift in the lab's preferences in 2009. To further validate the main model's robustness, we conducted out-of-sample simulations by fitting the main model with allocation decisions from odd years only and then simulating the entire allocation history and outputs for each lab, covering both odd and even years. Since even-year data were not used for model fitting, these results are considered out-of-sample. Appendix Table D8, comparing in-sample and out-of-sample simulations, reveals their close similarity across all labs.

Although we can never prove a model to be true, because the scientific method only allows us to falsify untrue ones, given the superior fit of the main model in both the estimation stage and the simulation stage, we have reasonable confidence that the main model captures the labs' actual allocation decision-making.

7. Counterfactual Policy Analyses

A well-fitting model of the lab's decision-making process enables us to examine policy counterfactuals through simulations. The simulation procedure for counterfactuals is largely identical to that used for model validation in Section 5.2. In all counterfactual scenarios, we set the welfare weights $\boldsymbol{\theta}'_{Xl}$ to $\hat{\boldsymbol{\theta}}_{Xl,UCB+D}$, those recovered from estimating the main model. The distinction between the counterfactual simulations and model validation lies in either how we set the parameters $\boldsymbol{\theta}'_{Bl}$ in the exploration bonus term or in our assumptions about the type of information Ω'_t that is available to the labs in the counterfactual scenarios. Intuitively, these counterfactuals preserve the lab's estimated preferences for projects and investigate how modifying only the lab's model of exploration or form of information utilization impacts their productivity.

7.1. Alternative Models of Exploration

The first counterfactual examines the extent of productivity gains in labs through exploration. This involves simulating a scenario where the labs do not explore and comparing the outcomes with the baseline outcomes from the main model. To mimic the effect of no exploration, we set $\boldsymbol{\theta}'_{Bl}$ to zero. Intuitively, setting $\boldsymbol{\theta}'_{Bl}$ to zero reduces the exploration bonus term $B_{ijt}(\cdot)$ to zero for all trials in the labs' choice sets. Thus, the labs make allocations as if they were following the greedy model—allocating only to trials they believe would yield the highest posterior expected payoff. The first row of Table 6 displays the counterfactual results for one large lab, while Appendix Table D9 shows results for other labs, which are qualitatively similar. Based on the simulation results, we conclude exploration was extensive and had a large positive impact on the labs' productivity. For the lab displayed in Table 6, no exploration would result in the lab attempting 79% fewer projects,

Table 6   Counterfactual outcomes, alternative models of exploration

Counterfactual model	Projects attempted	Unique structures	Citations	Downloads (millions)
Greedy (no exploration)	12,421	394	1,179	9.2
	(-79%)	(-64%)	(-65%)	(-64%)
Gittins	13,788	558	1,700	13.1
	(-77%)	(-49%)	(-50%)	(-49%)
Thompson sampling	19,340	661	2,016	15.6
	(-68%)	(-40%)	(-40%)	(-39%)
Explore-Then-Commit	59,953	790	2,418	18.4
	$(+0\%)$	(-28%)	(-28%)	(-28%)
Baseline model	59,947	1,097	3,376	25.6

Note: Each simulation uses $\hat{\boldsymbol{\theta}}_{Xl}$ from the parameter estimates of the UCB+D model, as shown in columns (1) and (2) of Table D5. The results are averaged from three simulations of the model. Outputs from the baseline model are identical to those in the second last row of Table 5. Parentheses indicate percentage differences compared to the baseline model. This table presents results from NESG, one of the four large labs. See Appendix Table D9 for results from the other labs.

producing 64% fewer unique structures, and 65% fewer citations. Summing up results from all four large labs, no exploration would result in the labs attempting 82% fewer projects, producing 51% fewer unique publications, and 57% fewer citations. Applying back-of-the-envelope calculations with some simple assumptions, this equates to a loss of at least \$650 to \$720 million of economic value.[19]

We then examine whether adopting several popular alternative allocation models could have enhanced the labs' productivity. Given the wide array of of alternative models available, our focus is on those that we deem ex ante "feasible." This means that these models do not include free parameters in the exploration bonus term $B_{ijt}(\cdot)$. Implementing a sophisticated allocation model with numerous free parameters presents challenges in tuning these parameters. Without ex post data on actual outcomes, it is difficult to determine whether increasing or decreasing parameter values would yield better outcomes. Therefore, models with free parameters in the $B_{ijt}(\cdot)$ term are of limited value for making ex ante policy recommendations. In contrast, ex ante feasible models, which do not encounter this issue, can be readily employed even without extensive knowledge of

---

[19] Assuming that the smaller labs under the \$1.3 billion NIH program would experience the same percentage decrease in output as the larger labs if they did not explore, and considering the economic value of each unit of output to be uniform, a 51% decrease in output would equate to a loss of at least \$650 million in economic value, provided the NIH program had a nonnegative economic return. Similarly, if these smaller labs had the same percentage decrease in citations as a result of not exploring, and the economic value of each citation was uniform, then the 57% decrease in citations would correspond to forgoing at least \$720 million in economic value. Results are rounded to the nearest tens.

the specific setting. These models act as useful benchmarks for assessing the effectiveness of the labs' allocation model.

The first alternative model we consider is the Gittins Index, which we have previously discussed. To simulate this counterfactual, we configure the $B_{ijt}(\cdot)$ term as described in Alternative Model 2. This model does not have free parameters other than $\boldsymbol{\theta}'_{Xl}$. We use the estimated welfare weights, $\hat{\boldsymbol{\theta}}_{Xl,UCB+D}$, from the main model as $\boldsymbol{\theta}'_{Xl}$. The Gittins model generates more structures and citations than the greedy model, indicating it is preferable to a scenario with no exploration. However, its performance significantly lags behind the actual output of the labs. For the lab presented in Table 6, this model results in 49% fewer unique structures and 50% fewer citations compared to the baseline UCB+D model. Overall, when considering results from all four large labs, this model produces 43% fewer unique structures and 54% fewer citations

The second alternative model we consider is Thompson sampling, another prominent model in the multi-armed bandit literature (Thompson 1933). Unlike models that have an exploration bonus $B_{ijt}(\cdot)$, Thompson sampling approximates the value function by directly drawing p_{ijt} from the posterior distribution:

$$V_{ijt}^A(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_l) = \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}^{DRAW} \sim \tilde{F}_t(\Omega_t); \boldsymbol{\theta}_{Xl}). \tag{11}$$

The idea is to enable the lab to continue allocating resources to projects likely to have high payoffs, while shifting away from those with lower payoff potential (Russo et al. 2017). Although Thompson sampling outperforms the Gittins Index in our setting, it does not match the actual output of the labs. It results in 40% fewer unique structures and citations compared to the baseline UCB+D model. When aggregating results from all four large labs, Thompson sampling yields 32% fewer unique structures and 29% fewer citations. A notable issue in our context is the significantly lower number of projects attempted under Thompson sampling compared to actual lab activity. This could be due to the influence of the prior. As Lattimore and Szepesvári (2020) point out, the prior can crucially affect Thompson sampling's performance. If the prior underestimates the quality of an arm, then Thompson sampling may never play that arm with high probability and no data is ever observed. In our context, the prior, based on a limited range of initial trials, might not accurately predict the success of vastly different projects, suggesting a need for models capable of exploring projects with low prior success probabilities to overcome poor prior assumptions.

We present a third alternative model, "Explore-Then-Commit," which may address the issue of poor prior in Thompson sampling. As the name implies, Explore-Then-Commit initially explores each new project a fixed number of times, then commits to projects with the highest posterior expected payoffs. In our simulations, we have the labs try each new project once as they arise

before committing to projects with the highest posterior expected payoffs. This model, by design, attempts all projects in the given choice sets. It outperforms both the Gittins model and Thompson sampling. Summing up results from all labs, Explore-Then-Commit produces 26% more structures and 11% more citations than Thompson sampling. Despite its simplicity, this model's output in terms of the number of structures and citations is closest to the actual output.

A remarkable result from the above counterfactuals is that the labs did quite well in balancing exploration and exploitation. None of these popular alternative models surpassed the labs' actual output. Even the Explore-Then-Commit model falls short, as the projects' outcomes are noisy and trying every project once is ineffective for acquiring sufficient information. For the lab displayed in Table 6, the Explore-Then-Commit model produces 28% fewer unique structures and citations compared to the baseline UCB+D model. Summing up results from all four large labs, Explore-Then-Commit yields 14% fewer unique structures and 21% fewer citations. This indicates that having in-depth experience in the specific project area, as the labs did, is crucial when choosing an appropriate allocation model. While we have found some ad hoc tuning of the estimated parameters of the baseline UCB+D model occasionally led to marginally better outcomes in our simulations, we acknowledge that such adjustments have limited value for making ex ante policy recommendations.

7.2. Alternative Forms of Information Utilization

We next explore the impact of processes that aided in gathering and utilizing information during allocation, such as a pilot phase where the NIH refrained from setting immediate production targets for the labs. This involves simulating a counterfactual scenario where the labs skip the pilot phase from 2000 to 2004. In this counterfactual, we start with a flat prior in 2005, rather than one informed by data accumulated during the pilot. The remaining simulation steps are unchanged, employing the estimated UCB+D model for allocation decisions.

Table 7   Counterfactual outcomes, alternative forms of information utilization

Counterfactual model	Projects attempted	Unique structures	Citations	Downloads (millions)
UCB+D, no piloting	48,340 (−19%)	832 (−24%)	2,544 (−25%)	20.1 (−22%)
UCB+D, no analytics	59,904 (−0%)	939 (−14%)	2,841 (−16%)	22.6 (−12%)
Baseline model	59,947	1,097	3,376	25.6

Note: Table shows results from NESG, one of the four large labs. See Appendix Table D10 for results from other labs. The rest of the notes of Table 6 apply.

Figure 7   Simulated number of unique structures under alternative forms of information utilization

(a) No piloting

(b) No analytics
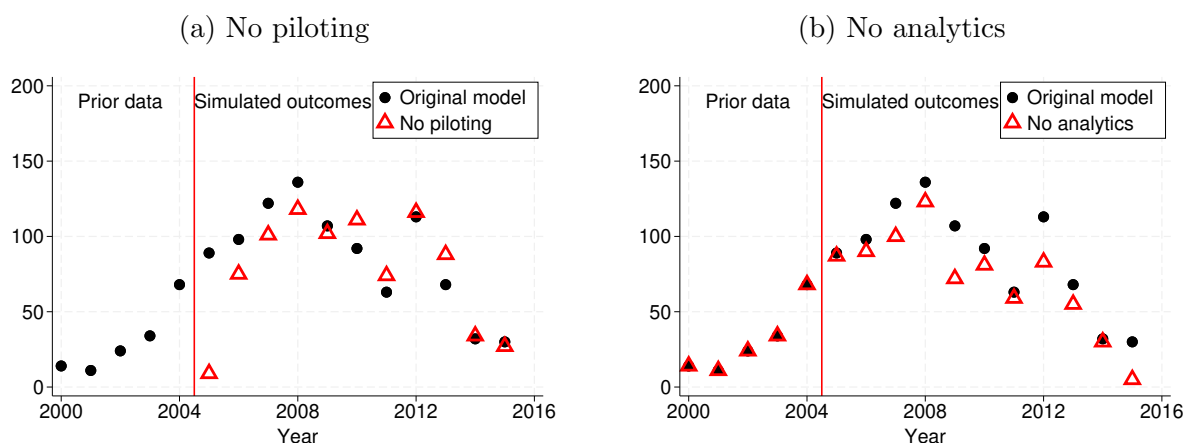
Note: Each data series is derived from a single simulation of the corresponding model for NESG.

We found that piloting had a large positive impact on the labs' productivity through improved allocation. For the lab shown in Table 7, not conducting the pilot resulted in 24% fewer unique structures and 25% fewer citations. When considering results from all four large labs, as presented in Appendix Table D10, the absence of a pilot phase led to a 23% reduction in unique structures and a 26% decrease in citations. This decline is not solely due to the mechanical loss of output that would have been produced during the pilot period. The output in 2005 and subsequent years also diminished because a less informative prior led to misallocation of resources towards less productive projects, particularly in the early years. Figure 7a illustrates this decrease for one large lab. As depicted, the output in 2005, immediately following the start of the simulation, was notably low due to insufficient information for guiding allocation. The output in subsequent years gradually improved as the lab accumulated more information and developed better posterior predictions. In later years, the output aligns more closely with that from the baseline model. If we compare only the output from this counterfactual to the baseline model from 2005 onwards, not conducting the pilot resulted in 13% fewer unique structures and 12% fewer citations during that period.[20]

We also examined the impact of continued data collection and analytics. For this, we simulated a counterfactual scenario where the labs do not collect information about trial outcomes nor update their posterior beliefs using machine learning from 2005 onwards. In this simulation, the labs rely on the prior formed at the end of the pilot phase as their posterior beliefs for the entire period of 2005–2015. We do not refit the machine learning models during each simulation. The remaining simulation procedure stays the same. We found that the continued collection of data and analytics enhanced the labs' productivity through improved allocation as well. In the case of the lab shown

---

[20] See Appendix Table D11 for simulated outcomes of the baseline model in 2005 and beyond.

in Table 7, the absence of ongoing data collection and analytics led to a 14% decrease in unique structures and a 16% reduction in citations. When aggregating results from all four large labs, as presented in Appendix Table D10, this led to a decrease of 7% in unique structures and 9% in citations. Figure 7b illustrates this decline for one large lab. As indicated in the figure, the output for all years is lower compared to the baseline model, a decline attributed solely to the misallocation of resources to less productive projects.

Applying back-of-the-envelope calculations with some simple assumptions,[21] these results suggest that processes facilitating the collection and use of information yielded significant returns. Had the NIH eliminated the pilot phase, it would have saved 17% of its input, or around $220 million. However, this would have resulted in 23% less output quantity and 26% fewer citations. This reduction in output would be equivalent to forgoing at least $290 to $330 million in economic value, making the return to piloting at least 30%. If the NIH had ceased requiring ongoing data collection and analysis, it would have saved the $40 million needed to support these activities. However, this would also have led to 7% less output quantity and 9% fewer citations. Using the same assumptions as before, this decrease in output would be equivalent to forgoing at least $90 to $110 million in economic value. The return on continued data collection and analysis is therefore at least twofold.

8.   Managerial Implications and Future Research

Extensive literature has recognized the importance of the exploitation-exploration tradeoff in operations management. Yet, there has been limited empirical study on how organizations navigate this tradeoff in complex real-world settings. This paper overcomes empirical challenges by utilizing novel data and a new estimation method. We examine how a group of large, publicly-funded research labs managed the exploitation-exploration tradeoff in allocating resources across a vast portfolio of research projects. We found that a simple model captures the labs' decision-making process remarkably well. This model embodies extensive exploration and strongly resembles the simple UCB index of Auer et al. (2002). Its only modification is the addition of a variable to capture time-discounting of older projects. During the estimation stage, we found this model to be the best fit for the data with minimal additional parameters compared to the many alternative models we tested. In the simulation stage, we observed that this model generates input allocation patterns and outputs that are very similar to those in the actual data. Based on this well-fitting model, we found that extensive exploration significantly boosted the productivity of the labs studied. Moreover, enhancing information acquisition and utilization through methods such as piloting and data analytics helped to unlock the full benefit of exploration.

---

[21] Same as those in Footnote 19.

These findings have important implications for both management and policy. Firstly, they underscore the critical role of exploration in enhancing an organization's productivity, especially when the outcomes of various decisions are unknown. This is particularly relevant for innovative and creative industries, where the dilemma between adhering to tried-and-tested ideas and venturing into uncharted territories is ever-present. If these findings are indicative of broader industry trends, then exploration plays a significant role in boosting societal innovation productivity by optimizing resource allocation within organizations, making it a key factor for consideration by both managers and policymakers.

Secondly, these findings highlight information as a crucial output in decision problems involving the exploitation-exploration tradeoff. In the innovation context, many organizations, including leading funding agencies, measure innovation productivity solely through tangible outcomes such as publications and citations. However, this research suggests that even unsuccessful efforts can yield valuable insights that enhance future resource allocation and overall innovation productivity. To maximize the benefits of exploration, organizations are advised to adjust their resource allocation processes to effectively utilize information. This includes considering pilot programs and investing in data analytics infrastructure for the ongoing collection and analysis of information gathered during exploratory activities. Regarding policy implications, it could be suggested that funding agencies require grantees of public funds to document and publicly share details of their research endeavors, successful or otherwise, thereby amplifying the societal gains from information acquired by individual organizations and researchers.

Moreover, these findings demonstrate the practical relevance of theoretical insights from multi-armed bandit literature. These empirical results reinforce the theoretical emphasis on the importance of exploration in bandit problems and showcase the effectiveness of allocation strategies that resemble cutting-edge bandit algorithms. Therefore, organizations are advised to consider applying these advanced methodologies in their operations management practices.

This research opens the door to various future research areas. One promising direction involves exploring how different elements of the allocation process design can enhance productivity in innovative and creative industries. Many questions remain unanswered, and our real-world findings invite a reevaluation of many established practices in these sectors. For instance, funding agencies often award smaller, early-career grants to numerous emerging researchers and later prioritize track records for more substantial, follow-up funding as researchers' careers advance. This decision-making approach could be better understood and potentially improved in light of insights from theoretical research. Such understanding will influence future policy directions and how publicly funded research organizations manage their research and development activities.

Furthermore, the methodology developed in this paper can be applied to study a broad range of complex, real-world operational challenges involving the exploitation-exploration tradeoff, from product development to supplier sourcing. Our index approximation estimation approach addresses the significant computational challenges associated with traditional estimation methods, which stem from the curse of dimensionality. This advancement enables researchers to empirically analyze decision-making in highly complex dynamic choice problems where information is the state variable and decision-makers encounter millions of choices per period across thousands of periods.

References

Agrawal R (1995) Sample mean based index policies with O(log n) regret for the multi-armed bandit problem. *Advances in Applied Probability* 27(4):1054–1078.

Auer P, Cesa-Bianchi N, Fischer P (2002) Finite-time analysis of the multiarmed bandit problem. *Machine Learning* 47(2):235–256.

Azoulay P, Graff Zivin JS, Manso G (2011) Incentives and creativity: Evidence from the academic life sciences. *RAND Journal of Economics* 42(3):527–554.

Babnigg G, Joachimiak A (2010) Predicting protein crystallization propensity from protein sequence. *Journal of Structural and Functional Genomics* 11(1):71–80.

Balaban RS (2013) Evaluation of scientific productivity and excellence in the NHLBI Division of Intramural Research. *Journal of General Physiology* 142(3):177–178.

Berman HM, Gabanyi MJ, Kouranov A, Micallef DI, Westbrook J, Protein Structure Initiative network of investigators (2017) Protein Structure Initiative—TargetTrack 2000-2017—all data files. URL `https://zenodo.org/record/821654`, accessed on June 25, 2019.

Bietti A, Agarwal A, Langford J (2021) A contextual bandit bake-off. *The Journal of Machine Learning Research* 22(1):5928–5976.

Bolton P, Harris C (1999) Strategic experimentation. *Econometrica* 67(2):349–374.

Boroush MA (2021) New data on U.S. R&D: Summary statistics from the 2019–20 edition of national patterns of R&D resources URL `https://ncses.nsf.gov/pubs/nsf22314`, accessed on May 29, 2022.

Bray RL, Coviello D, Ichino A, Persico N (2016) Multitasking, multiarmed bandits, and the Italian judiciary. *Manufacturing & Service Operations Management* 18(4):545–558.

Brezzi M, Lai TL (2002) Optimal learning and experimentation in bandit problems. *Journal of Economic Dynamics and Control* 27(1):87–108.

Bubeck S, Cesa-Bianchi N, et al. (2012) Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning* 5(1):1–122.

Buchfink B, Reuter K, Drost HG (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods* 18(4):366–368.

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. *Nature Methods* 12(1):59–60.

Cappe O, Garivier A, Maillard OA, Munos R, Stoltz G, et al. (2013) Kullback–Leibler upper confidence bounds for optimal sequential allocation. *Annals of Statistics* 41(3):1516–1541.

Caro F, Gallien J (2007) Dynamic assortment with demand learning for seasonal consumer goods. *Management science* 53(2):276–292.

Chen L, Oughtred R, Berman HM, Westbrook J (2004) TargetDB: A target registration database for structural genomics projects. *Bioinformatics* 20(16):2860–2862.

Chen W, Hu W, Li F, Li J, Liu Y, Lu P (2016) Combinatorial multi-armed bandit with general reward functions. *Advances in Neural Information Processing Systems* 29.

Chen W, Wang L, Zhao H, Zheng K (2021) Combinatorial semi-bandit in the non-stationary environment. *Uncertainty in Artificial Intelligence*, 865–875 (Proceedings of Machine Learning Research).

Chen X, Christensen TM, Tamer E (2018) Monte Carlo confidence sets for identified sets. *Econometrica* 86(6):1965–2018.

Chernozhukov V, Hong H (2003) An MCMC approach to classical estimation. *Journal of Econometrics* 115(2):293–346.

Cheung WC, Simchi-Levi D, Zhu R (2022) Hedging the drift: Learning to optimize under nonstationarity. *Management Science* 68(3):1696–1713.

Chruszcz M, Wlodawer A, Minor W (2008) Determination of protein structures–A series of fortunate events. *Biophysical Journal* 95(1):1–9.

Cohen JD, McClure SM, Yu AJ (2007) Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences* 362(1481):933–942.

Crawford GS, Shum M (2005) Uncertainty and learning in pharmaceutical demand. *Econometrica* 73(4):1137–1173.

DeepMind (2020) AlphaFold: Using AI for scientific discovery URL `https://deepmind.com/blog/article/AlphaFold-Using-AI-for-scientific-discovery`, accessed on April 15, 2021.

Dickstein MJ (2021) Efficient provision of experience goods: Evidence from antidepressant choice Working Paper.

EMBL-EBI (2021) EMBL-EBI Pfam to PDB mapping. URL `http://ftp.ebi.ac.uk/pub/databases/Pfam/mappings/`, accessed on Mar 13, 2021.

Erdem T, Keane MP (1996) Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science* 15(1):1–20.

Froese DS, Forouhar F, Tran TH, Vollmar M, Kim YS, Lew S, Neely H, Seetharaman J, Shen Y, Xiao R, et al. (2013) Crystal structures of malonyl-coenzyme A decarboxylase provide insights into its catalytic mechanism and disease-causing mutations. *Structure* 21(7):1182–1192.

Ganglmair B, Simcoe T, Tarantino E (2019) Learning when to quit: An empirical model of experimentation in standards development Working Paper.

Garivier A, Moulines E (2011) On upper-confidence bound policies for switching bandit problems. *International Conference on Algorithmic Learning Theory*, 174–188.

Gittins JC (1979) Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society: Series B (Methodological)* 41(2):148–164.

Gittins JC, Jones DM (1979) A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika* 66(3):561–565.

Guan M, Jiang H (2018) Nonparametric stochastic contextual bandits. *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32 (1).

He X, An B, Li Y, Chen H, Guo Q, Li X, Wang Z (2020) Contextual user browsing bandits for large-scale online mobile recommendation. *Proceedings of the 14th ACM Conference on Recommender Systems*, 63–72.

Hill R, Stein C (2020) Scooped! Estimating rewards for priority in science Working Paper.

Hill R, Stein C (2021) Race to the bottom: Competition and quality in science Working Paper.

Hotz VJ, Miller RA (1993) Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies* 60(3):497–529.

Jaffe AB (2002) Building programme evaluation into the design of public research-support programmes. *Oxford Review of Economic Policy* 18(1):22–34.

Jahandideh S, Jaroszewski L, Godzik A (2014) Improving the chances of successful protein structure determination with a random forest classifier. *Acta Crystallographica Section D: Biological Crystallography* 70(3):627–635.

Jaroszewski L, Slabinski L, Wooley J, Deacon AM, Lesley SA, Wilson IA, Godzik A (2008) Genome pool strategy for structural coverage of protein families. *Structure* 16(11):1659–1667.

Jiang Z, Huang Y, Beil DR (2022) The role of feedback in dynamic crowdsourcing contests: A structural empirical analysis. *Management Science* 68(7):4858–4877.

Jin R, Simchi-Levi D, Wang L, Wang X, Yang S (2021) Shrinking the upper confidence bound: A dynamic product selection problem for urban warehouses. *Management Science* 67(8):4756–4771.

Kaustov L, Liao J, Lemak S, Duan S, Muhandiram R, Karra M, Srisailam S, Sundstrom M, Weigelt J, Edwards A, Dhe-Paganon S, Arrowsmith C (2007) NMR solution structure of PARC CPH domain. URL https://www.rcsb.org/structure/2JUF, accessed on November 1, 2023.

Kveton B, Wen Z, Ashkan A, Szepesvari C (2015) Tight regret bounds for stochastic combinatorial semi-bandits. *Artificial Intelligence and Statistics*, 535–543 (Proceedings of Machine Learning Research).

Lai TL (1987) Adaptive treatment allocation and the multi-armed bandit problem. *Annals of Statistics* 15(1):1091–1114.

Lai TL, Robbins H (1985) Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics* 6(1):4–22.

Langford J, Zhang T (2007) The epoch-greedy algorithm for contextual multi-armed bandits. *Advances in Neural Information Processing Systems* 20(1):96–1.

Lattimore T, Szepesvári C (2020) *Bandit Algorithms* (Cambridge: Cambridge University Press).

Lauer M (2016) Citations per dollar as a measure of productivity. URL `https://nexus.od.nih.gov/all/2016/04/28/citations-per-dollar/`, accessed on October 24, 2023.

Li D, Raymond LR, Bergman P (2020) Hiring as exploration Working Paper.

Miller RA (1984) Job matching and occupational choice. *Journal of Political Economy* 92(6):1086–1120.

Moore PB (2007) Let's call the whole thing off: Some thoughts on the protein structure initiative. *Structure* 15(11):1350–1352.

Nguyen-Thanh N, Marinca D, Khawam K, Rohde D, Vasile F, Lohan ES, Martin S, Quadri D (2019) Recommendation system-based upper confidence bound for online advertising. *arXiv preprint arXiv:1909.04190* .

NIGMS (2004) Large-scale centers for the Protein Structure Initiative. URL `https://grants.nih.gov/grants/guide/rfa-files/RFA-GM-05-001.html`, accessed on May 31, 2022.

NIGMS (2007a) Frequently asked questions. URL `http://www.nigms.nih.gov/Initiatives/PSI/Background/FAQs.htm`, accessed on March 1, 2020.

NIGMS (2007b) Frequently asked questions for PSI-2. URL `http://www.nigms.nih.gov/Initiatives/PSI/Background/FAQsPSI-2.htm`, accessed on March 1, 2020.

NIGMS (2007c) Request for information: To solicit input for the assessment of the Protein Structure Initiative. URL `http://grants.nih.gov/grants/guide/notice-files/NOT-GM-07-108.html`, accessed the Internet Archive capture from June 25, 2011.

NIGMS (2008a) Protein Structure Initiative (Pilot Phase) fact sheet. URL `http://www.nigms.nih.gov/Initiatives/PSI/Background/PilotFacts.htm`, accessed the Internet Archive capture from Oct 1, 2008.

NIGMS (2008b) Report of the future structural genomics initiatives meeting URL `http://www.nigms.nih.gov/News/Reports/FutureSGMeeting102008.htm`, accessed the Internet Archive capture from Dec 5, 2010.

NIGMS (2009a) Centers for high-throughput structure determination (U54). URL `https://grants.nih.gov/grants/guide/rfa-files/rfa-gm-10-005.html`, accessed on May 31, 2022.

NIGMS (2009b) Concept clearance: High-throughput structural biology URL `https://www.nigms.nih.gov/News/Reports/council_concept_clearance_2009`, accessed the Internet Archive capture from May 23, 2009.

NIGMS (2009c) Protein Structure Initiative Advisory Committee (PSIAC): Recommendations for the future of the PSI URL `http://www.nigms.nih.gov/News/Reports/PSIAC_Future_2009.htm`, accessed the Internet Archive capture from July 6, 2010.

NIGMS (2009d) PSI:biology knowledgebase (U01). URL `https://grants.nih.gov/grants/guide/rfa-files/rfa-gm-10-004.html`, accessed on May 31, 2022.

NIGMS (2011a) PSI Production Phase fact sheet. URL `http://www.nigms.nih.gov/Research/FeaturedPrograms/PSI/Background/PSI2FactSheet.htm`, accessed the Internet Archive capture from June 27, 2011.

NIGMS (2011b) PSI:Biology FOA frequently asked questions. URL `https://www.nigms.nih.gov/research/specificareas/PSI/PSI_biology/background/pages/PSIbiology_faqs.aspx`, accessed on March 1, 2020.

NIH (2019) NIGMS funding opportunities search. URL `https://www.nigms.nih.gov/grants/Pages/Funding.aspx?expired=1`, accessed on Feb 29, 2020.

NIH (2021) NIH RePORT advanced projects search. URL `https://reporter.nih.gov/`, accessed on Feb 29, 2020.

Ortner R, Ryabko D, Auer P, Munos R (2012) Regret bounds for restless Markov bandits. *International Conference on Algorithmic Learning Theory*, 214–228.

Pakes A (1986) Patents as options: Some estimates of the value of holding European patent stocks. *Econometrica* 755–784.

Papanastasiou Y, Bimpikis K, Savva N (2018) Crowdsourcing exploration. *Management Science* 64(4):1727–1746.

Petsko GA (2007) An idea whose time has gone. *Genome Biology* 8(6):1–3.

Price WN, Chen Y, Handelman SK, Neely H, Manor P, Karlin R, Nair R, Liu J, Baran M, Everett J, et al. (2009) Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nature Biotechnology* 27(1):51–57.

Russo D, Van Roy B, Kazerouni A, Osband I, Wen Z (2017) A tutorial on Thompson sampling. *arXiv preprint arXiv:1707.02038* .

Rust J (1987) Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica* 999–1033.

Slabinski L, Jaroszewski L, Rodrigues AP, Rychlewski L, Wilson IA, Lesley SA, Godzik A (2007a) The challenge of protein structure determination–lessons from structural genomics. *Protein Science* 16(11):2472–2482.

Slabinski L, Jaroszewski L, Rychlewski L, Wilson IA, Lesley SA, Godzik A (2007b) XtalPred: A web server for prediction of protein crystallizability. *Bioinformatics* 23(24):3403–3405.

Thompson WR (1933) On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika* 25(3/4):285–294.

UniProt Consortium (2021) UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Research* 49(D1):D480–D489.

Van Montfort RL, Workman P (2017) Structure-based drug design: Aiming for a perfect fit. *Essays in Biochemistry* 61(5):431–437.

Varadi M, Berrisford J, Deshpande M, Nair SS, Gutmanas A, Armstrong D, Pravda L, Al-Lazikani B, Anyango S, Barton GJ, et al. (2020) PDBe-KB: a community-driven resource for structural and functional annotations. *Nucleic Acids Research* 48(D1):D344–D353.

Wang S, Chen W (2018) Thompson sampling for combinatorial semi-bandits. *International Conference on Machine Learning*, 5114–5122 (Proceedings of Machine Learning Research).

Whittle P (1988) Restless bandits: Activity allocation in a changing world. *Journal of Applied Probability* 25(A):287–298.

Woodroofe M (1979) A one-armed bandit problem with a concomitant variable. *Journal of the American Statistical Association* 74(368):799–806.

Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS, McLellan JS (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 367(6483):1260–1263.

wwPDB (2013) PDB structure download statistics. URL `http://www.wwpdb.org/downloadStats.php`, accessed the Internet Archive capture from July 9, 2014.

Zhalechian M, Keyvanshokooh E, Shi C, Van Oyen MP (2022) Online resource allocation with personalized learning. *Operations Research* 70(4):2138–2161.

Zhou D, Li L, Gu Q (2020) Neural contextual bandits with UCB-based exploration. *International Conference on Machine Learning*, 11492–11502 (Proceedings of Machine Learning Research).

Appendix A: Data and Variable Construction

A.1. Project Rationale/NIH Evaluation Metrics

Several variables capture a lab's observable rationale for allocating a trial to a project, and these variables correspond to the NIH's evaluation metrics for the labs' productivity, among which a key metric is the novelty of the project. The variable $novel_i$ is a binary and is equal to 1 if the labs cited novelty as a reason to allocate trials to project i in the TargetTrack information system. Another key NIH evaluation metric is the biomedical importance of the project. The variable $biomed_i$ is a binary and is equal to 1 if the labs cited biomedical importance as a reason to allocate trials to project i in the information system.The TargetTrack information system contains textual descriptions of why labs allocated trials to a project. The relevant fields are populated for 84% of projects at the four largest labs. Construction of $novel_i$ and $biomed_i$ is based on keywords in those descriptions. The following paragraphs describe the variable construction process.

First, we use keywords to identify projects that were novel and/or biomedically important. TargetTrack contains a variable called *targetCategoryList* where labs give projects categorical labels such as "biomedical," "structural coverage,"[22] and so on. It also contains a text field called *targetRationale* where labs give textual descriptions of projects' rationales. Whenever *targetCategoryList* and *targetRationale* contain the following keywords, we set $novel_i$ equal to 1:

big,[23] coverage of protein universe, diversity, first structure of class, low sequence identity, mega,[24] metagenomic, new fold, no structural information, no structure, numer of homologs,[25] pfam, remote homologs, structural coverage, structural template for unsolved, structure coverage, unsolved families, without any solved structures, without structure.

Whenever *targetCategoryList* and *targetRationale* contain the following keywords, we set $biomedical_i$ equal to 1:

activator, adhension, antibiotic, binding, biochemistry, biological interest, biomedical, cascade, catalyze, cell development, community nominated, communit-nominated,[26] community-nominated, community request, conserved, disease, coronavirus, drug, drug development, drug target, effector, enzyme, essential, function, functional studies, functional, gpcr, high value, hig-value,[27] hiv, homeostasis, host, immune, immunity, infection, infectious, inhibitor, interaction, interact, legionella, medical school, metabolism, mitochondria, model system, operon, parkinsons, partnership, pathogen, pathology, pathway, phosphatase, pneumonia, protein family of high biological importance, reagent, receptor, resistance, resistant, salmonella, school of

---

[22] "Structural coverage" means the project is in part of the structure space with no or few published structures.

[23] BIG and MEGA domain families were defined by the PSI-2 Target Selection Committee as having high value for extensive coverage. These familes contained hundreds to tens of thousands of members and many subfamilies which could not be modeled well due to a lack of structural coverage.

[24] Same as above.

[25] This typo occurs in the raw data.

[26] Same as above.

[27] Same as above.

medicine, secret, sensor, shen lab, shen_lab, shen_selection, stem cell, substrate, syndrome, synthesis, t-cell, t cell, therapeutic, thorson lab, toxoplasma, transcription, transport, tuberculosis, tumor, university, vaccine, vibrio, virulence, virulent.

Second, we use labs' selection protocols of projects for additional information. TargetTrack contains a field where labs describe the protocols they used to conduct each stage of the trials. One type of protocol is the selection protocol. For example, 15 projects were selected because of the protocol "TSel_101," which states "These proteins are important for cell development." We read the descriptions associated with each selection protocol and manually classified whether each protocol was "novel" and/or "biomedical."[28] Then we set $novel_i$ equal to 1 if the project was selected due to a "novel" protocol. We set $biomedical_i$ equal to 1 if the project was selected due to a "biomedical" protocol.

Lastly, TargetTrack has a field that contains a list of reference IDs of each molecule in large-scale bioinformatics databases.[29] These reference ids may yield additional information. Whenever the list of reference ids contains BIG and MEGA reference ids,[30] We set $novel_i$ equal to 1.

When the labs cited a project i as being novel, they often emphasized that there were no or few already published structures in the same protein family as i. We therefore construct $prevStruct_{iy}$, a continuous variable that captures the changes over year (subscripted with the letter y) of the number of published structures in the same protein family as i. To construct this variable, we first pull from UniProt the list of protein families $pfam_i$ associated with molecule i. We then obtain a mapping of each protein family to its associated structures from EMBL-EBI (2021) and the structures' publication dates (we take the structure's deposition date to the PDB as the publication date) from Varadi et al. (2020). Merging the datasets results in $prevStruct_{iy}$. If i is associated with multiple protein families, we take the average of the number of already published structures in each protein family associated with i.

As an additional proxy for the biomedical importance of a molecule, we look into the number of publications related to the molecule in UniProt, including structures and other types of publications. We construct $prevPub_{iy}$, a continuous variable that captures the changes over year of the number of publications on molecule i.

Additional NIH evaluation metrics correspond to whether the project was related to human beings, eukaryotes,[31] and the cell membrane. The variable $human_i$ captures how similar molecule i is to any molecules from human beings. When a lab worked on a "human" molecule, often the molecule was actually from bacteria but was very similar to a molecule from human beings and was much easier than the human molecule. Therefore, the right construction for $human_i$ is molecule i's degree of similarity to human molecules rather than being a human molecule itself. We learned this from a conversation with an NIH program officer in charge of the grant program. To construct this variable, we search each molecule i against all UniProt protein

---

[28] The manual classification is available upon request.

[29] These reference ids include, but are not limited to, the molecule's id in the Protein Data Bank (PDB), UniProt, and the National Center for Biotechnology Information (NCBI) database.

[30] See footnote 23.

[31] See definition in footnote 5.

sequences in the Homo sapiens (human) species (UniProt (2021d)). From the search results, we take the maximal percentage identity of i to any human molecule as the variable $human_i$. Due to potentially large number of search results, the search algorithm DIAMOND (Buchfink et al. (2015, 2021)) by default cuts off results at $evalue = 0.001$. $evalue$ is a well-understood metric for search quality in this field. If there are no search results meeting the cutoff, we let $human_i = 0$. For details on how to do the DIAMOND search, please see Appendix A.4.

The variable $eukaryote_i$ likewise captures how similar molecule i is to any molecules from eukaryotes. To construct this variable, we search each molecule i against all UniProt protein sequences in the Eukaryota superkingdom (UniProt (2021c)). From the search results, we take the maximal percentage identity of i to any eukaryotic molecule as the variable $eukaryote_i$. Due to potentially large number of search results, the search algorithm DIAMOND (Buchfink et al. (2015, 2021)) by default cuts off results at $evalue = 0.001$. If there are no search results meeting the cutoff, we let $eukaryote_i = 0$.

The variable $membrane_i$ is a binary and is equal to 1 if molecule i is related to the cell membrane. We set this binary variable $= 1$ if project i's UniProt information contains the word "membrane."

A.2. Observed output

The main measure of output is the number of unique structures published. A trial produced a publication if it succeeded in all stages. We use a binary Y_{ijt} to denote the outcome of trial j_i on day t. $Y_{ijt} = 1$ if the trial succeeded and produced a structure for publication. Only 1.6% of trials in the data succeeded. We also observe the outcomes of all stages of each trial and use a binary Y_{ijkt} to denote the outcome of stage k of trial j_i on day t. $k = 0, 1, 2, 3, 4$. $Y_{ij0t} = 1$ if DNA was successfully cloned. Y_{ij1t} is only defined when $Y_{ij0t} = 1$ and is equal to 1 if protein was successfully expressed. Y_{ij2t} is only defined when $Y_{ij0t} = 1$ and $Y_{ij1t} = 1$ and is equal to 1 if protein was successfully purified. Y_{ij3t} is only defined when $Y_{ij0t}, Y_{ij1t}, Y_{ij2t} = 1$ and is equal to 1 if protein was successfully prepared for studying its structure (through X-ray crystallography or NMR or cryo-EM). Y_{ij4t} is only defined when all previous stages were successful and is equal to 1 if the structure was successfully produced and deposited to the Protein Data Bank (PDB) for publication. We use these intermediate outcomes to construct the labs' posterior beliefs about output.[32]

We supplement the main output measure with the number of citations and downloads of each published project. The variable $citation_{iy}$ captures the five-year citations and mentions of a project i published in year y. When multiple structures of the same project were published, we take the mean of five-year citations and year of publication of those structures. We do not take the sum of five-year citations because those structures were often cited together. The variable $download_{im}$ captures the number of downloads of a project i in month m between August 2007 and November 2013. This variable is not directly useful because downloads have strong lifecycle trends—downloads were high when a structure just became published then fell over time— and comparing different publications at different stages of their lifecycles is meaningless. In Section 3.1, we construct a variable corresponding to the five-year downloads to make comparison across projects meaningful. For that variable, when multiple structures of the same project were published, we take the mean of five-year

---

[32] See Section 3.1 and Appendix B.

Figure A1   Observed output of the four largest labs

(a) Number of publications

(b) Citations and mentions
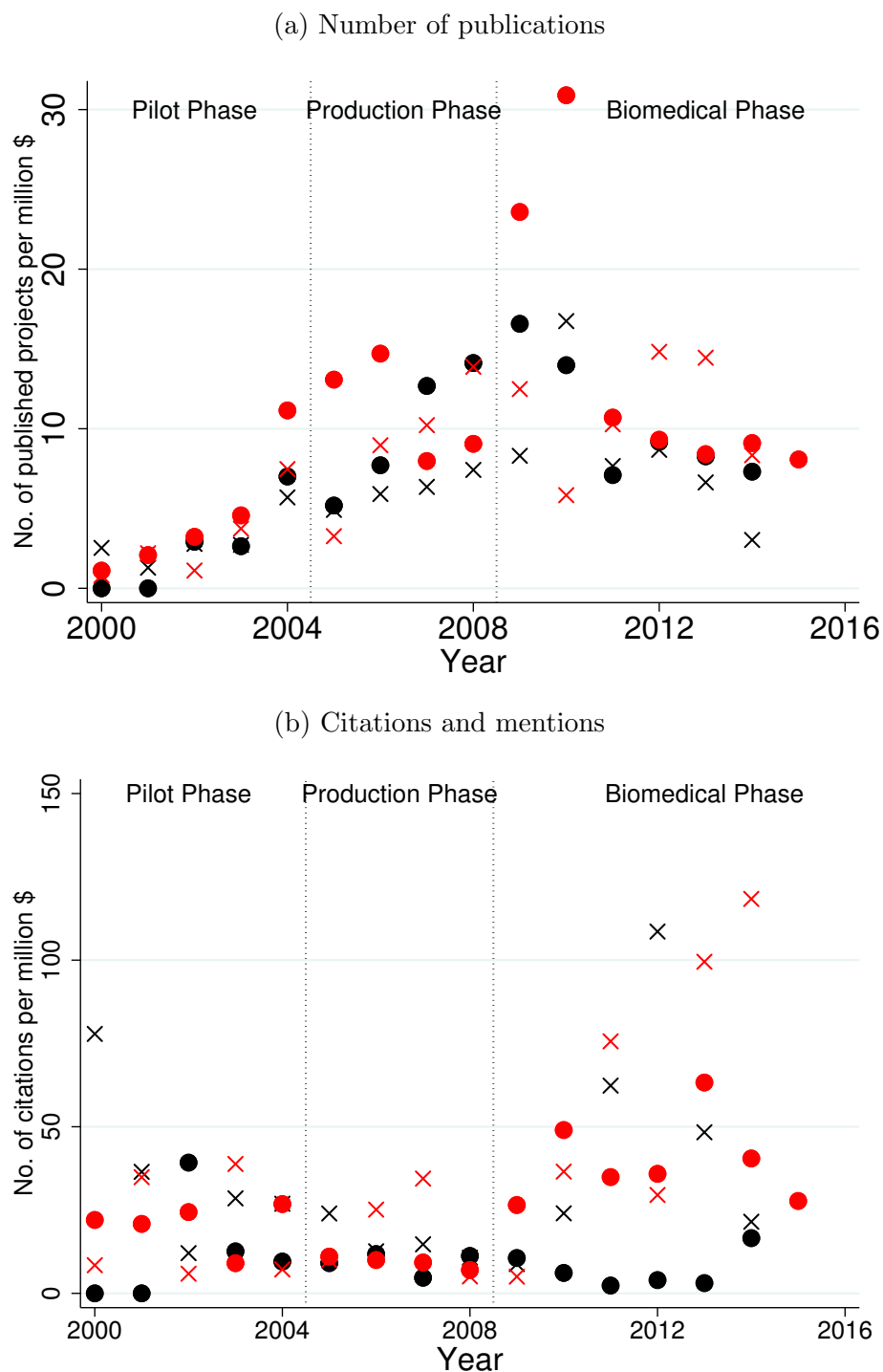
Note: Each series represents one lab. Panel A1a shows the number of published structures on unique molecules in a given year divided by the lab's funding in millions in that year. Panel A1b shows the number of five-year citations and mentions the published structures in that year generated, divided by the lab's funding in millions in that year.

Table A1   Funding opportunity announcements (FOA) tied to PSI

Id	Title	Year
RFA-GM-99-009	PILOT PROJECTS FOR THE PROTEIN STRUCTURE INITIATIVE (STRUCTURAL GENOMICS)	1999
PA-99-116	PROTEIN STRUCTURE INITIATIVE (STRUCTURAL GENOMICS)	1999
PA-99-117	PROTEIN STRUCTURE INITIATIVE (STRUCTURAL GENOMICS) – SBIR/STTR	1999
RFA-GM-00-006	PILOT PROJECTS FOR THE PROTEIN STRUCTURE INITIATIVE (STRUCTURAL GENOMICS)	2000
RFA-GM-05-001	LARGE-SCALE CENTERS FOR THE PROTEIN STRUCTURE INITIATIVE	2004
RFA-GM-05-002	SPECIALIZED CENTERS FOR THE PROTEIN STRUCTURE INITIATIVE	2004
RFA-GM-06-004	Structural Genomics Knowledgebase (U01)	2006
RFA-GM-10-004	PSI:Biology Knowledgebase (U01)	2009
RFA-GM-10-005	Centers for High-Throughput Structure Determination (U54)	2009
RFA-GM-10-006	Centers for Membrane Protein Structure Determination (U54)	2009
RFA-GM-10-007	Consortia for High-Throughput-Enabled Structural Biology Partnerships (U01)	2009
PAR-10-214	High-Throughput-Enabled Structural Biology Research (U01)	2010
PAR-11-176	High-Throughput-Enabled Structural Biology Partnerships (U01)	2011

downloads and year of publication of those structures. we do not take the sum of five-year downloads because those structures were likely downloaded together as they were often cited together. Figure A1 shows the observed output for the four largest labs.

A.3.   Lab Funding

The variable $funding_{ly}$ captures the total amount of funds lab l received from the NIH in year y. Funding information comes from two sources. First, the NIH released a series of funding opportunity announcements (FOAs) directly tied to the grant program (NIH 2019), which allows us to search directly all grants associated with those FOAs on NIH RePORT database (NIH 2021). Table A1 shows the full list of FOAs. Second, labs sometimes received supplementary funds from the NIH so we also perform a direct search of the labs' names and abbreviations using RePORT's advanced search functionality to obtain data on each labs' supplementary funding. The search term we used was (quotation marks included):

"[lab full name]" OR "[lab abbreviation]"

We then aggregate each lab's sum of research grants by year from the search results. Figure A2 shows the level of funding for the four largest labs.

A.4.   Matching Projects to UniProt Molecule Information

As a preliminary to using the UniProt data, we match projects from the TargetTrack information system to their molecule information on UniProt through two methods.

Figure A2　　Funding of the four largest labs
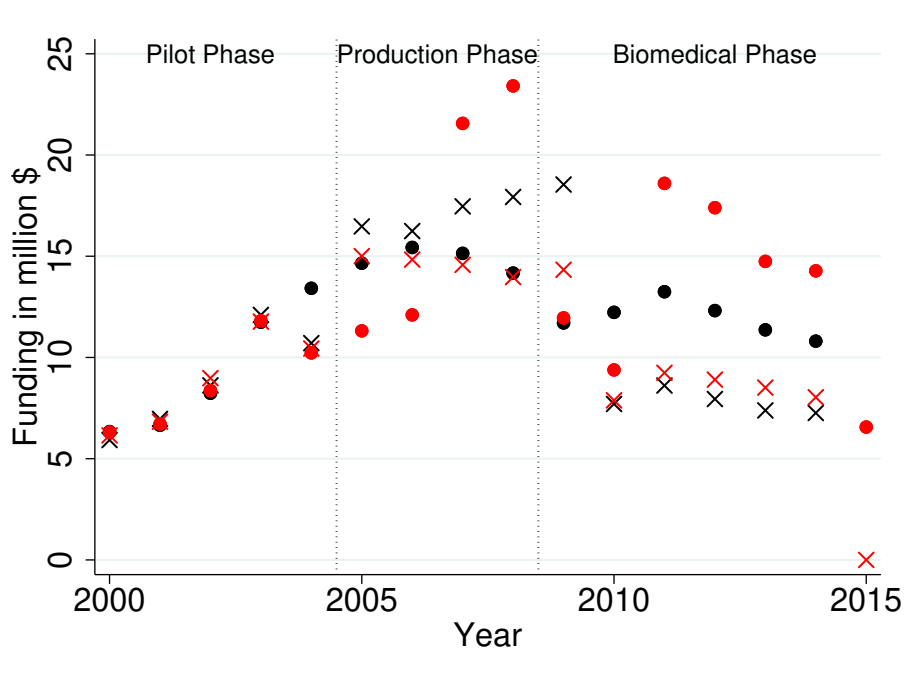
Note: Each series represents one lab.

TargetTrack has a field containing a list of reference ids of each molecule i in large-scale bioinformatics databases. These reference ids include, but are not limited to, the molecule's id in the Protein Data Bank (PDB), UniProt, and the National Center for Biotechnology Information (NCBI) database. When the UniProt id of the molecule is available in this field, the mapping is direct. We also use the following id types, which easily convert into UniProt molecule id through UniProt's ID Mapping service (Huang et al. 2011, UniProt 2021a):

- *PDB_ID*: a molecule's id in the Protein Data Bank (PDB), a database for 3D structures.
- *P_REFSEQ_AC*: a molecule's id in NCBI's RefSeq protein database.
- *EMBL*: a molecule's corresponding gene's id in European Molecular Biology Laboratory (EMBL)/GenBank/DNA Data Bank of Japan (DDBJ) CDS database.
- *P_ENTREZGENEID*: a molecule's corresponding gene's id in GeneID (Entrez Gene) database.
- *P_GI*: a molecule's GI number assigned by NCBI.

When the first method fails to find a match (usually due to an entirely missing reference id field or obsolete records in the relevant databases), we use a second method: directly searching the molecule's sequence of amino acids against all protein sequences in UniProt.[33] We perform this search using DIAMOND (Buchfink et al. 2015, 2021), a very fast algorithm for searching similar sequences. The diamond command we used was:

```
diamond blastp -d [database name] -q [input sequences in .fasta]
-o [output in .csv] -f 6 qseqid qlen sseqid slen evalue bitscore pident length
```

---

[33] Downloadable in .fasta format at https://www.uniprot.org/downloads.

```
-b4.0 --top 5
```

It produces search results with the following variables:

- *qseqid*: query sequence's identifier (the full sequence in this case).
- *qlen*: query sequence's length.
- *sseqid*: search result's UniProt id.
- *slen*: search result's length.
- *evalue*: the number of expected hits of similar quality that could be found just by chance in a random database of the same size. E-value is a commonly used measure for the degree of similarity between the query sequence and the search result.
- *bitscore*: the required size of a sequence database in which the current match could be found just by chance. Bit score does not depend on the size of the database and is a common alternative measure for the degree of similarity between the query sequence and the search result.
- *pident*: percentage of identical matches between the query sequence and the search result over the alignment length.
- *length*: the alignment length between the query sequence and the search result.

If the query sequence's best match search result, determined by the e-value, a standard metric for assessing sequence similarity, has at least 95% *pident* and the alignment length *length* is at least 67% of both *qlen* and *slen*, we map the query sequence to the result sequence's UniProt id.

We were able to match 262,984 (78.4%) of the 335,553 projects to their UniProt entries through the id mapping method and match an additional 58,593 (17.5%) projects through the direct search. Overall, we were able to map 321,577 (95.8%) projects to their UniProt entries. We then used UniProt's programmatic access for individual entries (UniProt (2021b)) to pull each molecule's information from UniProt. We successfully pulled this information for 319,986 (95.4%) projects.

A.5. Data Glossary

This paper uses hundreds of project characteristics extracted from a variety of sources. This data glossary offers a comprehensive view of these variables.

* Variable is included in the characteristics $\boldsymbol{X}_{ijt}$ in training $\tilde{F}_t(\Omega_t)$.

† Variable is included in the characteristics $\boldsymbol{X}_{ijt}$ in training $F^*(\Omega_T)$.

‡ Variable is included in the characteristics $\boldsymbol{X}_{ijt}$ in training $ridge((\boldsymbol{X}, citation)_T)$.

§ Variable is included in the characteristics $\boldsymbol{X}_{ijt}$ in training $ridge((\boldsymbol{X}, \Delta download)_T)$.

Please see Appendix B for these models.

Table A2: Data glossary

Variables	Description

[4 cap letters then 6 digits]$_i$*†	Amino acid attributes from the AAindex database (Kawashima et al. (2007)). Each attribute had an identifier that had four capital letters followed by six digits. We started with the 567 attributes in AAindex1, and then normalized and clustered them to a set of around 30 attribute classes as in Babnigg and Joachimiak (2010). We used scikit-learn's implementation of affinity propagation clustering, which automatically picked 34 clusters. We then kept the cluster center of each class. For each cluster center attribute, we calculated the local average value, the local minimum, and the local maximum of the sum of the attribute in a seven-amino acid sliding window for molecule i as in Babnigg and Joachimiak (2010). This resulted in 102 variables.
[consortium abbreviation]$_{ijt}$*†‡ §	Binary variable $= 1$ if trial j_i was conducted by the given consortium at time t. Only consortia with more than 70 observations of projects in the TargetTrack database have their corresponding variables. 36 variables in total.
[gene]$_i$†‡ §	Binary variable $= 1$ if molecule i is coded for by the given gene. From UniProt. 48,548 variables in total. Most variables are very sparse. For † we only include genes that have occurred more than 200 times in the data. For ‡ and § We only include genes that are associated with at least one molecule whose structure was successfully published in the data.
[keyword]$_i$†‡ §	Binary variable $= 1$ if molecule i is associated with the given keyword in UniProt. Examples of keywords include "Alzheimer disease," "Antioxidant," "RNA-binding," "Viral envelope protein." 1,053 variables in total. Most variables are very sparse. For † we only include keywords that have occurred more than 200 times in the data and remove the keyword "3D-structure" because this is the outcome. For ‡ and § we only include keywords that are associated with at least one molecule whose structure was successfully published in the data.
[superkingdom-phylum]$_i$*†‡ §	Binary variable $= 1$ if molecule i comes from an organism in the specific superkingdom and phylum. From UniProt. Due to the large number of species molecules in TargetTrack represent, we do not go down the UniProt taxonomy below phylum. 81 variables in total.
$aminoAcid\_[X]_i$*†	Counts the number of times amino acid "X" is in molecule i. 20 variables for each of amino acids A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Contents of certain amino acids are linked to more successes of trials (Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).
$aminoAcidPercent\_[X]_i$*†	Calculate the amino acid "X" content in molecule i in percentages. 20 variables for each of amino acids A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Contents of certain amino acids are linked to more successes of trials (Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).
$biomedical_i$*†‡ §	Binary variable $= 1$ if project i was biomedically important. See Appendix A.1 for variable construction.
$citation_{iy}$	five-year citations and mentions of project i published in year y, from PDBe (Varadi et al. (2020)). When multiple structures were published on molecule i, we take the mean values of the five-year citations and year of publication.

$download_{im}$	Number of downloads of published project i in month m across the three major structure databases in the world, from wwPDB (wwPDB (2013)). Available for Aug 2007–Nov 2013.
$\hat{E}(citation_{iy})$	Expected five-year citations and mentions of project i published in year y. See Appendix B.2 for construction.
$\hat{E}(download_{iy})$	Expected five-year downloads of project i published in year y. See Appendix B.3 for construction.
$\hat{E}_{\tilde{F}_t}(p_{ijt})$	Best-effort replication of the labs' posterior expectation of the probability of success of trial j_i on day t. See Appendix B.1 for construction.
$eukaryote_i$*†‡ §	Maximal percentage identity of molecule i to any eukaryotic molecule. To construct this variable, we search each molecule i against all UniProt protein sequences in the Eukaryota superkingdom (UniProt (2021c)). From the search results, we take the maximal percentage identity of i to any eukaryotic molecule as the variable $eukaryote_i$. Due to potentially large number of search results, the search algorithm DIAMOND (Buchfink et al. (2015, 2021)) by default cuts off results at $evalue = 0.001$. $evalue$ is a well-understood metric for search quality in this field. If there are no search results meeting the cutoff, we let $eukaryote_i = 0$.
$exposedAminoAcid\_[X]_i$*†	Counts the number of times amino acid "X" is on the predicted exposed surface of molecule i. 20 variables for each of amino acids A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y. Exposed surface was predicted using the NetSurfP (Klausen et al. (2019)) program with the cutoff of relative solvent accessibility (rsa) > 0.25. Contents of certain amino acids on the exposed surface of the molecule are linked to more successes of trials (Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).
$exposedAminoAcidPercent\_[X]_i$*†	Calculate the amino acid "X" content on the predicted exposed surface of molecule i in percentages. 20 variables for each of amino acids A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y. Exposed surface was predicted using the NetSurfP (Klausen et al. (2019)) program with the cutoff of relative solvent accessibility (rsa) > 0.25. Contents of certain amino acids on the exposed surface of the molecule are linked to more successes of trials (Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).
$extinctCoeffReduced_i$*†	Molar extinction coefficient of molecule i with reduced cysteines. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Slabinski et al. (2007b) used the extinction coefficient as a feature to predict project success.
$extinctCoeffOxidized_i$*†	Molar extinction coefficient of molecule i with disulfid bridges. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Slabinski et al. (2007b) used the extinction coefficient as a feature to predict project success.
$funding_{ly}$	Total sum of research grants consortium l received from NIH in year y. See Appendix A.3 for variable construction.
$gaps_i$*†	The average number of insertions in molecule i's alignment compared to homologs in UniProt protein sequences. Computed by searching sequence i against UniProt protein sequences using DIAMOND (Buchfink et al. (2015, 2021)). The output variable *gaps* captures this value. Insertions were included as a feature in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Jahandideh et al. (2014).

$gapOpen_i$*†	The average number of insertion openings in the alignment compared to homologs in UniProt protein sequences. Computed by searching sequence i against UniProt protein sequences using DIAMOND (Buchfink et al. (2015, 2021)). The output variable $gapOpen$ captures this value. Insertions were included as a feature in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Jahandideh et al. (2014).
$gravyIndex_i$*†	Grand average of hydropathicity index (GRAVY) of molecule i, used to represent the hydrophobicity value of a molecule. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Hydrophobicity is a key determinant of success of trials (Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).
$hasPrevSuccess_{ijkt}$*	Binary variable $=1$ if at least one previous trial on molecule i successfully completed stage k before date t.
$hasPrevFailure_{ijkt}$*	Binary variable $=1$ if at least one previous trial on molecule i failed at stage k before date t.
$human_i$*†‡ §	Maximal percentage identity of molecule i to any human molecule. To construct this variable, we search each molecule i against all UniProt protein sequences in the Homo sapiens (human) species (UniProt (2021d)). From the search results, we take the maximal percentage identity of i to any human molecule as the variable $human_i$. Due to potentially large number of search results, the search algorithm DIAMOND (Buchfink et al. (2015, 2021)) by default cuts off results at $evalue = 0.001$. $evalue$ is a well-understood metric for search quality in this field. If there are no search results meeting the cutoff, we let $human_i = 0$.
$instabilityIndex_i$*†	Instability index of molecule i, which is an estimate of the stability of the protein in a test tube. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Instability Index was included as a feature in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Jahandideh et al. (2014).
$isoelectricPoint_i$*†	Isoelectric point of molecule i. Calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module. Isoelectric point is a key determinant of success of trials (Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).
$membrane_i$*†‡ §	Binary variable $=1$ if project i's UniProt information contains the word "membrane."
$molecularWeight_i$*†‡ §	Molecular weight of molecule i, calculated using Biopython's ProteinAnalysis function from Bio.SeqUtils.ProtParam module.
$novel_i$*†‡ §	Binary variable $=1$ if project i was novel. See Appendix A.1 for variable construction.
$p^*_{ij,k-1,t_{k-1}}$†	The predicted probability of success of stage $k-1$ of project-trial j_i that started in period t_{k-1}. If $k=0$, this variable is set to 1.
$percentCoil_i$*†	Predicted percentage of coil secondary structure in molecule i. Predicted using the NetSurfP (Klausen et al. (2019)) program. Secondary structure features were used in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Jahandideh et al. (2014).

$percentCoiledCoil_i$*†   Percentage of coiled-coil regions in molecule i from UniProt. Coiled-coil regions were used in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014).

$percentDisordered_i$*†   Predicted percentage of disordered region in molecule i. Predicted using the NetSurfP (Klausen et al. (2019)) program. Disordered region was used as a feature in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014).

$percentDisorderedUniprot_i$*†   Percentage of disordered region in molecule i from UniProt. Disordered region was used as a feature in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014).

$percentExposed_i$*†   Predicted percentage of amino acids on the exposed surface of molecule i. Exposed surface was predicted using the NetSurfP (Klausen et al. (2019)) program with the cutoff of relative solvent accessibility (rsa) > 0.25. Extent of the exposed surface of the molecule are linked to more successes of trials (Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014)).

$percentHelix_i$*†   Predicted percentage of helix secondary structure in molecule i. Predicted using the NetSurfP (Klausen et al. (2019)) program. Secondary structure features were used in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014).

$percentLowComplexity_i$*†   Predicted percent low-complexity regions in molecule i. Computed using the SEG program (Wootton (1994)). Low-complexity regions were used as features in Slabinski et al. (2007a,b), Jaroszewski et al. (2008).

$percentSignalPeptide_i$*†   Percentage of signal peptide in molecule i. From UniProt. Slabinski et al. (2007a,b), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014) state molecules containing signal peptides have very low chances of success.

$percentStrand_i$*†   Predicted percentage of strand secondary structure in molecule i. Predicted using the NetSurfP (Klausen et al. (2019)) program. Secondary structure features were used in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014).

$percentTransmembraneHelices_i$*†   Percentage of transmembrane helices in molecule i. From UniProt. Transmembrane helices were used as a feature in Slabinski et al. (2007a,b), Jaroszewski et al. (2008), Price et al. (2009), Babnigg and Joachimiak (2010), Jahandideh et al. (2014) .

$pfam_i$ A list of protein families associated with molecule i, from UniProt (UniProt (2021b)).

$phase1_{ijkt}$†   Binary variable $= 1$ if stage k of project-trial j_i started in phase 1 of PSI (pilot phase). We let this variable be 1 if the stage start year is before or in 2005. Phase 1 ended in 2004. However, based on Figures D1 and A1, one can clearly see that 2004 and 2005 are transition periods: output quantity jumped up in 2004. We therefore let 2004 and 2005 be part of both Phase 1 and Phase 2.

$phase2_{ijkt}$†	Binary variable $= 1$ if stage k of project-trial j_i started in phase 2 of PSI (production phase). We let this variable be 1 if the stage start year is between 2004 and 2010. Phase 2 is between 2005 and 2008. However, based on Figures D1 and A1, one can clearly see that 2009 and 2010 are transition periods: output quantity stayed high but citations reversed the trend. We therefore let 2009 and 2010 be part of both Phase 2 and Phase 3.
$phase3_{ijkt}$†	Binary variable $= 1$ if stage k of project-trial j_i started in phase 3 of PSI (biomedical phase). We let this variable be 1 if the stage start year is 2009 and beyond.
$prevPub_{iy}$*†‡ §	Number of publications on molecule i by the start of year y, from UniProt (UniProt (2021b)).
$prevStruct_{iy}$*†‡ §	Number of already published structures in the same protein families associated with molecule i by the start of year y. To construct this variable, we first pull from UniProt the list of protein families $pfam_i$ associated with molecule i. We then obtain a mapping of each protein family to its associated structures from EMBL-EBI (2021) and the structures' publication dates (we take the structure's deposition date to the PDB as the publication date) from Varadi et al. (2020). Merging the datasets results in $prevStruct_{iy}$. If i is associated with multiple protein families, we take the average of the number of already published structures in each protein family associated with i.
$prevSuccesses_{ijkt}$*	Number of previous trials on molecule i that have successfully completed stage k before date t.
$prevTrials_{ijkt}$*	Number of previous trials on molecule i that have reached stage k before date t.
$refId_i$	A list of reference ids of molecule i in TargetTrack, used to map i to its information in UniProt.
seq_i	Sequence representation of molecule i's amino acids, unique identifier of project i.
$seqLength_i$*†‡ §	The number of amino acids in molecule i.
$simPrevProj_{it}$	The maximal degree of similarity between project i and all previously attempted projects at time t, measured by the bit score (see Appendix A.4 for the definition of bit score). Computed by searching sequence i against all sequences attempted before time t using DIAMOND (Buchfink et al. (2015, 2021)). The maximum of the output variable $bitscore$ among research results was used as $simPrevProj_{it}$.
$surfaceRuggedness_i$*†	Surface ruggedness of molecule i, defined by the total accessible surface of molecule i divided by the accessible surface predicted based on molecular mass. The total accessible surface of the molecule i is calculated by summing the predicted absolute solvent accessibility of each amino acid from NetSurfP (Klausen et al. (2019)). The accessible surface predicted based on molecular mass is calculated using the formula $6.3(molecularMass)^{0.73}$ (Miller et al. (1987)). Jahandideh et al. (2014) used this variable as a feature.
$trialId_{ij}$	Trial id of project-trial j_i, unique identifier of trial j_i in TargetTrack.
$\widehat{Var}_{\tilde{F}_t}(p_{ijt})$	Best-effort replication of the posterior variance of the labs' beliefs about the probability of success of trial j_i on day t. See Appendix B.1 for construction.

Y_{ijt}	Binary variable $= 1$ if trial j_i on date t was successful.
Y_{ijkt}	Binary variable $= 1$ if intermediate stage k of trial j_i on date t was successful. $Y_{ij0t} = 1$ if DNA was successfully cloned. Y_{ij1t} is only defined when $Y_{ij0t} = 1$ and is equal to 1 if protein was successfully expressed. Y_{ij2t} is only defined when $Y_{ij0t} = 1$ and $Y_{ij1t} = 1$ and is equal to 1 if protein was successfully purified. Y_{ij3t} is only defined when $Y_{ij0t}, Y_{ij1t}, Y_{ij2t} = 1$ and is equal to 1 if protein was successfully crystalized for X-ray crystallography or prepared for NMR or cryo-EM. Y_{ij4t} is only defined when all previous stages were successful and is equal to 1 if the structure was successfully produced and deposited to the Protein Data Bank (PDB) for publication.

Appendix B: Constructing Posteriors

B.1.   Random Forest for Trial Success Probabilities

There are two kinds of models of trial success probability in this paper. The first one is a model that captures how labs formed posterior beliefs. It does not have to produce an unbiased estimate of the true probability of success of a trial. However, it has to produce an unbiased estimate of the labs' perceived posterior beliefs about the probability of success. Our implementation of $\tilde{F}_t$ closely follows the machine learning approach the labs described in published journal articles.

The second one is a model of the true data generating process of trial success probability F^*, which is used in simulating counterfactual outcomes. Estimating F^* is different from estimating the posterior using $\tilde{F}_t$ because F^* needs to produce an unbiased estimate of the true probability of success of a trial. As such, our implementation of F^* deviates from $\tilde{F}_t$ in several ways to correct the potential bias of and improve upon the machine learning systems the labs described.

In this appendix, we first explain our implementation of $\tilde{F}_t$; then we move on to discuss how our implementation of F^* deviates from that of $\tilde{F}_t$.

B.1.1.   Implementation of $\tilde{F}_t$ Our implementation of $\tilde{F}_t$ fits stage-specific models to account for information embodied in outcomes of intermediate stages. Recall that each trial j_i has multiple sequential stages and the overall probability of success of j_i is equal to the product of the probabilities of success for all sequential stages $p_{ijt} = \prod_{k=0}^{4} p_{ijkt}$. The intermediate outcomes Y_{ijk} for stages $k = 0, 1, \ldots$ up to the point when the overall trial failed/succeeded provides information for future trials' potential.

For a given quarter $q(t)$ and each of the stages $k = 0, 1, 2, 3, 4$, we let the information set $\Omega_{k,q(t)}$ consist of project-trial outcomes realized before quarter $q(t)$ at stage k and these project-trials' characteristics. Following when the labs started to use machine learning to form posterior, we let $q(t)$ to be between 2005 and 2015. For $q(t) = 2005Q1$, we use the trial outcomes realized before 2005 and these trials' characteristics as the initial information set Ω_{2005Q1}. Project-trial characteristics $\boldsymbol{X}_{ijkt}$ we use for training $\tilde{F}_t$ and prediction of labs' posterior beliefs fall under three categories:[34]

• Physicochemical properties of molecule i based on scientific reasoning. These variables were identified by the series of journal articles the labs published and were quite similar across labs and time.

---

[34] Please see Appendix A.5 for the full list of variables used.

- Other characteristics of project i, for examples, novelty, biomedical importance, and the number of prior publications on molecule i.

- Past successes and failures of project i at stage k.

Then, for the given quarter $q(t)$ and each of the stages $k = 0, 1, 2, 3, 4$, we fit a random forest model $\tilde{F}_{k,q(t)}(\Omega_{k,q(t)})$ using `RandomForestClassifier` from python package `scikit-learn`. Random forest is an *ensemble*[35] machine learning method. The algorithm constructs a large number of decision trees at training time. Each decision tree is a learning model that aims to find the project-trial characteristics predictive of success/failure in the training set. When it comes to prediction, the trained random forest classifier $\tilde{F}_{k,q(t)}(\Omega_{k,q(t)})$ would pool individual trees and average predicted values of $\{\hat{p}_{ijkt}^{ntree}\}$ from individual trees as the final output. As Jahandideh et al. (2014) did, we set the number of trees in the random forest equal to 1000.

Decision trees and random forests are known for often overfitting without regularization. To avoid overfitting, we regularize by restricting the hyperparameters *max_depth*,[36] *min_samples_leaf*,[37] *max_features*,[38] and *min_samples_split*.[39] We perform model selection with a grid search of the combinations of the four hyperparameters.[40] For each hyperparameter combination, we evaluate the model with five-fold cross validation using `scikit-learn`'s `cross_validate` function. In each iteration of the cross-validation, the function fits a random forest on four out of five cross-validation folds and then computes the cross-validation score by comparing the model's predictions with the actual data from the remaining fold. We use the average log likelihood (`log_loss` scoring in `scikit-learn`) as the cross-validation scoring method. We choose the hyperparameter combination that maximizes the average log likelihood in cross-validation.

After training the models $\tilde{F}_{k,q(t)}$ for $k = 0, 1, 2, 3, 4$ for a given $q(t)$, we predict $\widehat{E}_{\tilde{F}_{q(t)}}(p_{ijt})$ and $\widehat{Var}_{\tilde{F}_{q(t)}}(p_{ijt})$ for each project-trial in the choice set C_{lt} at decision time t as follows. We first collect the predictions $\{\hat{p}_{ijkt}^{ntree}\}$ from the 1000 individual decision trees in $\tilde{F}_{k,q(t)}(\Omega_{k,q(t)})$, and then compute $\hat{p}_{ijt}^{ntree} = \prod_{k=0}^{4} \hat{p}_{ijkt}^{ntree}$. There are 1000 values in the set $\{\hat{p}_{ijt}^{ntree}\}$. We let

$$\widehat{E}_{\tilde{F}_{q(t)}}(p_{ijt}) = \bar{p}_{success,ijt}^{ntree}, \tag{EC.1}$$

$$\widehat{Var}_{\tilde{F}_{q(t)}}(p_{ijt}) = s^2(p_{success,ijt}^{ntree}) \tag{EC.2}$$

---

[35] Ensemble methods use multiple learning models to obtain better predictive performance than could be obtained from any of the constituent learning models alone.

[36] This hyperparameter determines the maximum depth of each decision tree.

[37] This hyperparameter determines the minimum number of observations a node in the decision tree must have before it can be split.

[38] This hyperparameter determines the maximum number of features to consider when looking for the best split.

[39] This hyperparameter determines the minimum number of observations required to split a node.

[40] To reduce computational burden, we do not perform model selection for all $\tilde{F}_{k,q(t)}$. Rather, for each $k = 0, ..., 4$, we construct $\Omega_{k,T}$ using all outcomes at stage k and only perform model selection for $\tilde{F}_{k,T}$ on this full training set. We then use the selected hyperparameters to train the models $\tilde{F}_{k,q(t)}$ where $q(t) = 2005Q1, 2005Q2, ..., 2015Q4$. The set of *max_depth* used in grid search is $[int(log(sample\_size, 2)), 2 \cdot int(log(sample\_size, 2)), 3 \cdot int(log(sample\_size, 2)), 4 \cdot int(log(sample\_size, 2))]$. The set of *min_samples_leaf* used in grid search is $[1, 2, 4]$. The set of *max_features* used in grid search is $[0.1, 0.2, 0.3, 0.4]$ of the total number of features. The set of *min_samples_split* used in grid search is $[8, 16, 32, 64, 128]$.

Although we extensively reference the labs' implementations of machine learning systems when we implement $\tilde{F}_t$, our estimate of the posterior is not a perfect replica of the labs' posteriors. We note why replicating perfectly the labs' posterior beliefs would be difficult and where our implementation corresponds to and deviates from the labs' learning and updating process below:

• We include in $\boldsymbol{X}_{ijt}$ the set of physicochemical properties of molecules identified in the labs' published journal articles (Slabinski et al. 2007a,b, Jaroszewski et al. 2008, Price et al. 2009, Babnigg and Joachimiak 2010, Jahandideh et al. 2014). This set is the union of the sets of such properties in different articles (to minimize the risk of selection on unobservables) and is fixed for all labs and time periods in our implementation. In contrast, though similar across lab and time, the set of physicochemical properties the labs used in training and prediction still varied. It is impossible to capture all of these potential variations during the labs' long operational history (some may not have been recorded by the published articles).

• The construction of some variables in $\boldsymbol{X}_{ijt}$ requires using software packages that are constantly being updated or have become obsolete. We make our best effort to construct variables using methods as close to the labs' original approach as possible (see Appendix A.5).

• The $\boldsymbol{X}_{ijt}$ in our implementation includes past trial outcomes of projects while the labs' implementations did not explicitly include those characteristics. Still, it is reasonable to believe that researchers working on a project would update their beliefs on the potential of the project upon seeing a trial success/failure.

• We use random forest as the model of posterior updating for all labs and time periods. In contrast, the machine learning models the labs used in training and prediction varied across lab and across time. It is impossible to capture all of these potential variations during the labs' long operational history (some may not have been recorded by the published articles).

• We set the frequency of "updating" and refitting models at the quarterly interval. In contrast, the labs' actual belief updating frequency is not clearly documented. We use the quarterly interval because training models at a finer interval, such as at the daily frequency, places large computational and storage burden. The day-to-day change of the information set Ω was also relatively small. Therefore, to improve computational tractability, we coarsen the frequency of refitting new models to quarterly.

• Our model predicts the overall potential of success of a trial while the labs' implementations focused on predicting the potential of success of bottleneck stages of a trial. That is, for stages where success rates were usually reasonable (for example cloning the DNA), the labs were often not explicitly reliant on something as rigorous as supervised machine learning systems to form and update beliefs, while they were explicitly reliant on such systems for predicting the potential of success in crystallizing a molecule and studying its structure through X-ray crystallography.

• The output the labs' systems produced may not exactly be $\widehat{E}_{\tilde{F}_{q(t)}}(p_{ijt})$ and $\widehat{Var}_{\tilde{F}_{q(t)}}(p_{ijt})$. For example, the model in Slabinski et al. (2007b) predicted the probability of success as an intermediate outcome. The final output was an integer score between 1 and 5, where 1 represents "optimal" and 5 represents "very difficult." The labs' systems did not always predict $\widehat{Var}_{\tilde{F}_{q(t)}}(p_{ijt})$. When they did, the measure took the form of comparing predictions from multiple models side by side (Slabinski et al. 2007a,b, Babnigg and Joachimiak 2010, Jahandideh et al. 2014). It is reasonable to believe that labs had some understanding that

predictions from different models (or submodels of an ensemble model) differed, and looking at how those predictions varied was valuable, though they did not percolate the idea down to form an additional metric just to measure that variation. This seems consistent with the notation that the labs used heuristics to guide their exploration of high-variance projects.

B.1.2.   Implementation of F^*   The implementation of F^* is almost identical to that of $\tilde{F}_t$ except for a few deviations. First of all, a new model $\tilde{F}_{k,q(t)}$ (for stages $k = 0,..,4$) is trained for every quarter $q(t)$ between 2005Q1 and 2015Q4, incorporating new trial outcomes realized in each quarter. In contrast, F_k^* (for stages $k = 0,...,4$) is trained only on the full information set Ω_T. Ω_T covers the characteristics and outcomes of all trials in my trial allocations and outcomes dataset in the entire sample period.

Second, F^* uses additional covariates to correct the potential bias of $\tilde{F}_t$ in predicting trial success probabilities. The model $\tilde{F}_t$ may be biased in predicting trial success probabilities because it does not account for the propensity of observing a specific stage of a trial. To see this, think about the probability of success of stage 1 of a trial. We observe stage 1 of a trial only if stage 0 of the trial was successful. If the probabilities of success of stages 0 and 1 are positively correlated, then we are more likely to observe stage 1 of trials that are more likely to succeed in stage 1. Therefore, models trained with the observed data on stage 1 would produce prediction results that are positively biased. Correcting this bias is simple if we assume that the selection into observing a given stage is only based on observable characteristics of trials: we can use the predicted probability of success of the previous stage as the propensity score of observing the given stage. As such, we include $p^*_{ij,k-1,t_{k-1}}$, the predicted probability of success of stage $k-1$ of trial j_i that started in period t_{k-1}, as a covariate when I train F_k^*. For stage $k = 0$, we set this variable to 1. The labs' published articles offer no discussion about this source of bias, so we do not include this variable in training $\tilde{F}_t$.

Another difference between F^* and $\tilde{F}_t$ is that F^* does not include variables on previous trial outcomes as covariates. In simulations, all previous trial outcomes of a project are simulated and should not shift the project's true probability of success; therefore, the simulated counterfactual outcome of a trial should not be based on the simulated previous outcomes.

To further improve the predictive power of F^*, we include in F^* project-trial characteristics the labs did not use in their machine learning systems. We include keywords and genes associated with the molecule i. We also include three phase-specific dummy variables to capture the effects of different phases of the grant program on the probabilities of success, as we learned during our conversations with NIH program officers that the labs underwent retooling corresponding to the changes of phases.

B.2.   Ridge Regression for Citations

Let the model be $ridge(\boldsymbol{X}, citation)_T)$, where the training set $(\boldsymbol{X}, citation)_T$ represents the characteristics and $citation_{iy}$ of all published projects in our data. The goal of this model is to predict $E(citation_{iy}|\boldsymbol{X}_{iy}, ridge((\boldsymbol{X}, citation)_T))$, the expected number of five-year citations a project i published in year y would generate conditional on the project's characteristics $\boldsymbol{X}_{iy}$.

The number of characteristics that could potentially predict higher citations is very large. Characteristics ranging from the organism the molecule i is from to the gene that expresses molecule i could all contribute to the biomedical significance and research interests on molecule i. The number of characteristics is on the

order of hundreds, most of which are very sparse, while we only have 10,424 observations.[41] This calls for regularization to avoid overfitting.

We use a ridge regression from the python package `scikit-learn`. The project characteristics $\boldsymbol{X}_{iy}$ included for model fitting are shown in Appendix A.5. We choose the regularization hyperparameters using cross-validation with the `RidgeCV` function provided by the `scikit-learn` package.

We standardize the outcome variable $citation_{iy}$ by subtracting away the lab mean value of this variable and then dividing by the lab standard deviation as publications from different labs had large variations in citation numbers. When the model makes a prediction, we multiply the predicted value with the lab standard deviation and add the lab mean to get the predicted citations.

We assess model fit by comparing the actual citations with their out-of-sample predicted citations under five-fold cross validation. Figure B1 shows the distribution of the predicted citations correctly captures a high proportion of zero values in the actual data. Figure B2 shows a scatterplot of predicted citations against the actual citations. A linear regression of the predicted citations on the actual citations without constant shows an $R^2 = 0.580$.

Figure B1   Distributions of actual citations and predicted citations
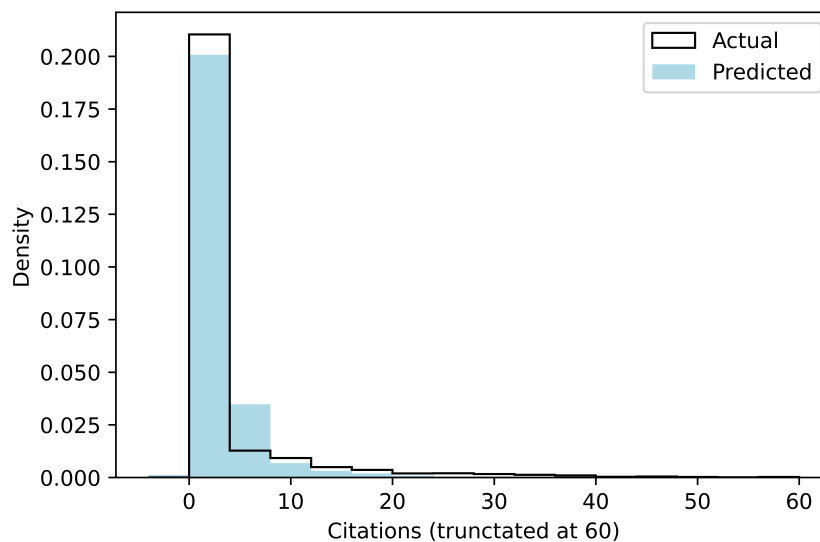
Note: Distributions are truncated at 60. Bin width is 4.

The predicted citations $\hat{E}(citation_{iy}|\boldsymbol{X}_{iy}, Ridge((\boldsymbol{X}, citation)_T))$ has a y subscript because some of the important characteristics vary with time, for example, the number of publications on molecule i prior to the year of publication of the structure. In Section B.4, since we do not know which exact year the structure of

---

[41] The total number of published projects in our data is 10,501. 77 published projects in our data did not give the PDB ids of their publications so we were not able to map their citation information.

Figure B2   **Scatterplot of predicted citations against actual citations**
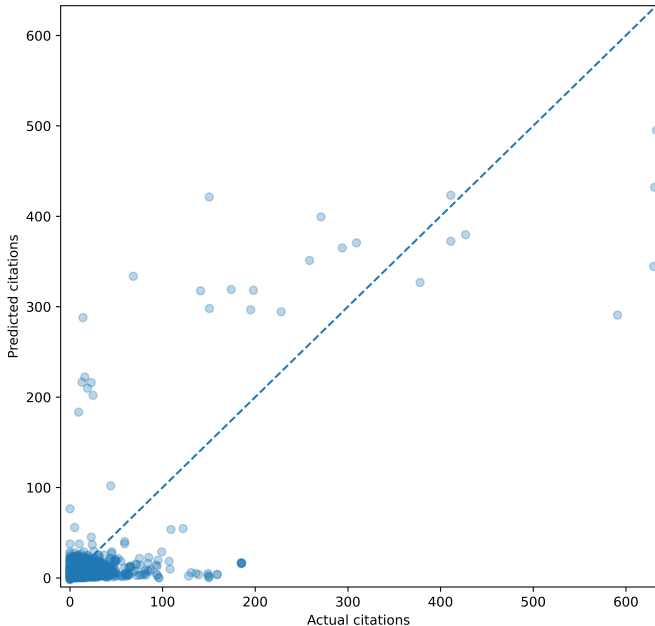
each trial in the choice set would have been published if the trial was allocated, we remove the y subscript by averaging the predicted citations for each molecule across years so that

$$\hat{E}(citation_i) = \frac{1}{16} \sum_{y=2000}^{2015} \hat{E}(citation_{iy}|\boldsymbol{X}_{iy}, ridge((\boldsymbol{X}, citation)_T)). \tag{EC.3}$$

In simulations, we simulate both the outcome (success or failure) of an allocated trial and the date that outcome is realized. In that case, the y subscript is preserved for the predicted citations of the publication.

B.3. Ridge Regression for Downloads

Before we start predicting downloads, the raw download data needs to be transformed because the number of downloads shows a strong time trend over the lifecycle of a publication. The raw download data we obtained consists of monthly downloads on the structure level between Aug 2007 and Nov 2013 for all publicly available structures human beings know of. The number of structure-month observations is 5,484,800. As some structures in a given month observed were published a long time ago while some just got published, comparing these structures' raw download counts in a month would be misleading. As shown in Figure B3, downloads peak within a month since the publication of a structure and then sharply decline over the following months until reaching some steady level in approximately two years.

We perform a transformation of the raw download data to detrend it as follows. Let $pubAge(i,m)$ be the age of publication i (in months) in month m. To detrend, we first compute $\overline{download_{pubAge(i,m)}}$, the mean downloads of structures that have been published for $pubAge(i,m)$ months.[42] We then compute how much the number of downloads structure i had in month m deviates from this mean, $\widetilde{download}_{im} = download_{im} -$

---

[42] We pool observations with $pubAge > 25$ months in computing this mean as the mean number of downloads flattens by 25 months since publication.

$\overline{download_{pubAge(i,m)}}$. We then define $\Delta download_i$, the average deviation of structure i's monthly downloads from the mean download trend, by the average of $\widetilde{download_{im}}$, in other words $\Delta download_i = \overline{\widetilde{download_{im}}}$. I treat the variable $\Delta download_i$ as the outcome variable. If there are multiple structures on the same project i, we take the mean of their average deviations as $\Delta download_i$. We then match the download data with the project-trial characteristics of completed projects in our data.

Let the model be $ridge((\boldsymbol{X}, \Delta download)_T)$, where the training set $(\boldsymbol{X}, \Delta download)_T$ represents the characteristics and $\Delta download_i$ of all published projects in my data. The goal of this model is to predict $E(\Delta download_{iy} | \boldsymbol{X}_{iy}, ridge((\boldsymbol{X}, \Delta download)_T))$, the expected average deviation of the structure's monthly downloads from the mean download trend for a project i published in year y, conditional on the project's characteristics $\boldsymbol{X}_{iy}$. There is a y subscript because some of the important characteristics vary with time, for example, the number of publications on molecule i prior to the year of publication of the structure.

The number of characteristics that could potentially predict higher downloads is very large. Characteristics ranging from those of molecule i's organism to those of molecule i's gene could all contribute to the level of interest on molecule i. The number of characteristics is on the order of hundreds, most of which are very sparse, while we only have 10,424 observations. This calls for regularization to avoid overfitting.

We use a ridge regression from the python package `scikit-learn`. The project characteristics $\boldsymbol{X}_{iy}$ included for model fitting are shown in Appendix A.5. We choose the regularization hyper-parameters using cross-validation with the `RidgeCV` function provided by the `scikit-learn` package.

We assess model fit by comparing the actual $\Delta download_{iy}$ with their out-of-sample predicted values under five-fold cross validation. Figure B4 shows a comparison of the distributions. As ridge regression shrinks all regression coefficients towards zero, the distribution of the predicted values is narrower. Still, the predicted values capture the rank order of the actual data well. Figure B5 shows a scatterplot. The plot shows a relationship quite close to the line $y = x$. Figure B6 shows a binned scatterplot.

In Section B.4, we perform additional transformations on the predicted value $\hat{E}(\Delta download_{iy} | \boldsymbol{X}_{iy}, ridge((\boldsymbol{X}, \Delta download)_T))$. First, the predicted value has a y subscript because some of the important characteristics vary with time, for example, the number of publications on molecule i prior to the year of publication of the structure. Since we do not know which exact year the structure of each trial in the choice set would have been published if the trial was allocated, we remove the y subscript by averaging the predicted values for each molecule across years so that

$$\hat{E}(\Delta download_i) = \frac{1}{16} \sum_{y=2000}^{2015} \hat{E}(\Delta download_{iy} | \boldsymbol{X}_{iy}, ridge((\boldsymbol{X}, \Delta download)_T)). \qquad \text{(EC.4)}$$

Second, the variable $\hat{E}(\Delta download_i)$ predicts the average deviation of monthly downloads from a trend and is difficult to interpret. We therefore transform this variable to a prediction of five-year downloads $\hat{E}(download_i)$ by using the following formula:
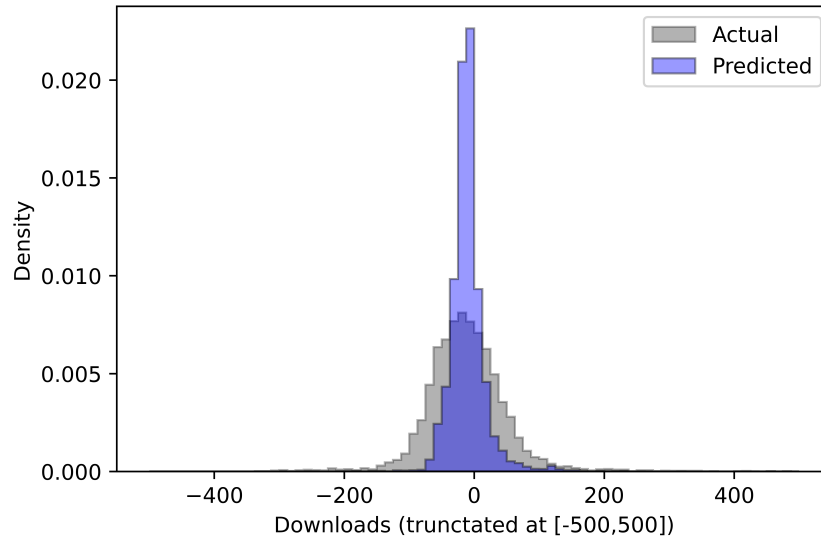
$$\hat{E}(download_i) = \sum_{pubAge=0}^{59} \overline{download_{pubAge}} + 60 \times \hat{E}(\Delta download_i), \qquad \text{(EC.5)}$$

Figure B3   Average downloads per structure in months since publication
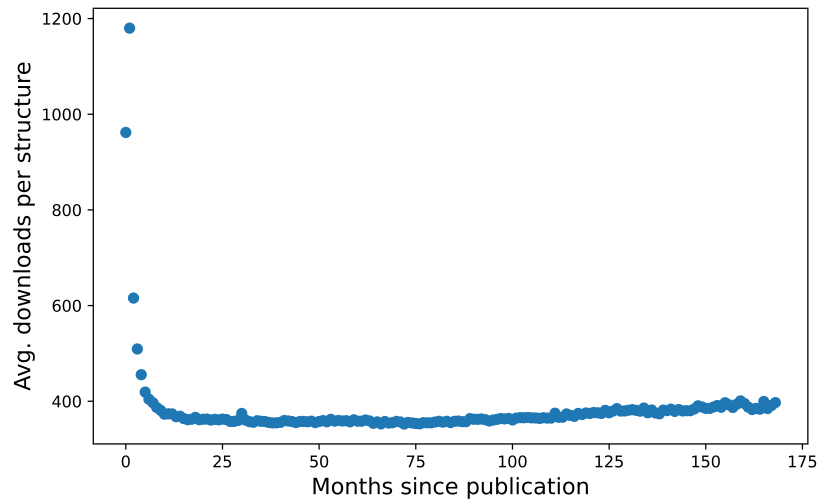

Note: The plot is based on 5,484,800 structure-month observations of download counts between Aug 2007 and Nov 2013. Each blue dot aggregates in this data the average monthly downloads for structures published m months ago.

Figure B4   Distributions of predicted $\hat{E}(\Delta download_{iy})$ against actual $\Delta download_{iy}$

where $\sum_{pubAge=0}^{59} \overline{download_{pubAge}} = 23960.11$ is computed based on the full download data with 5,484,800 structure-month observations.

In simulations, we simulate both the outcome (success or failure) of an allocated trial and the date that outcome is realized. In that case, the y subscript is preserved for the predicted average deviation of monthly downloads of the published projects. Therefore the predicted five-year downloads becomes

Figure B5 **Scatterplot of predicted** $\hat{E}(\Delta download_{iy})$ **against actual** $\Delta download_{iy}$
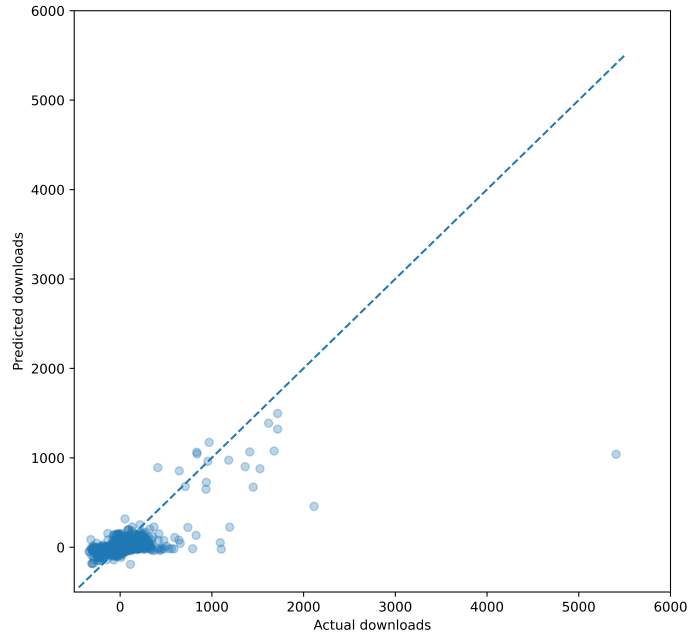
Figure B6 **Binned scatterplot of predicted** $\hat{E}(\Delta download_{iy})$ **against actual** $\Delta download_{iy}$
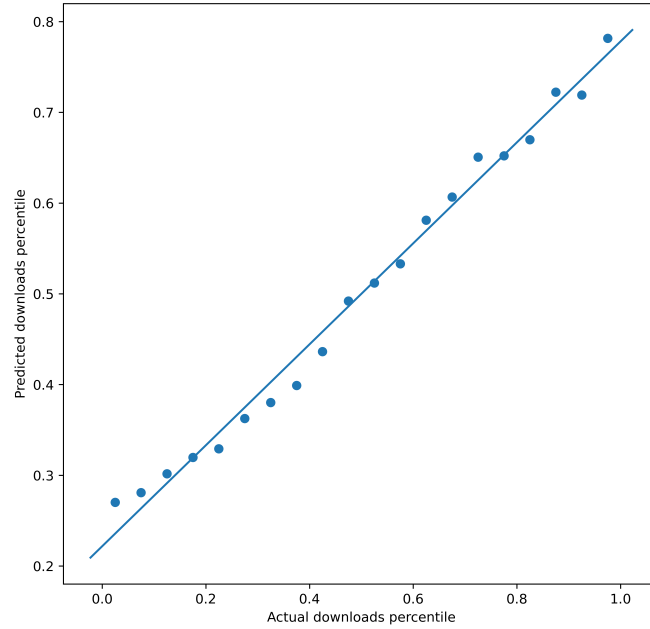
$$\hat{E}(download_{iy}) = \sum_{pubAge=0}^{59} \overline{download_{pubAge}} + 60 \times \hat{E}(\Delta download_{iy}|\boldsymbol{X}_{iy}, ridge((\boldsymbol{X}, \Delta download)_T)). \quad \text{(EC.6)}$$

Figure B7 **Estimated posterior beliefs about output**
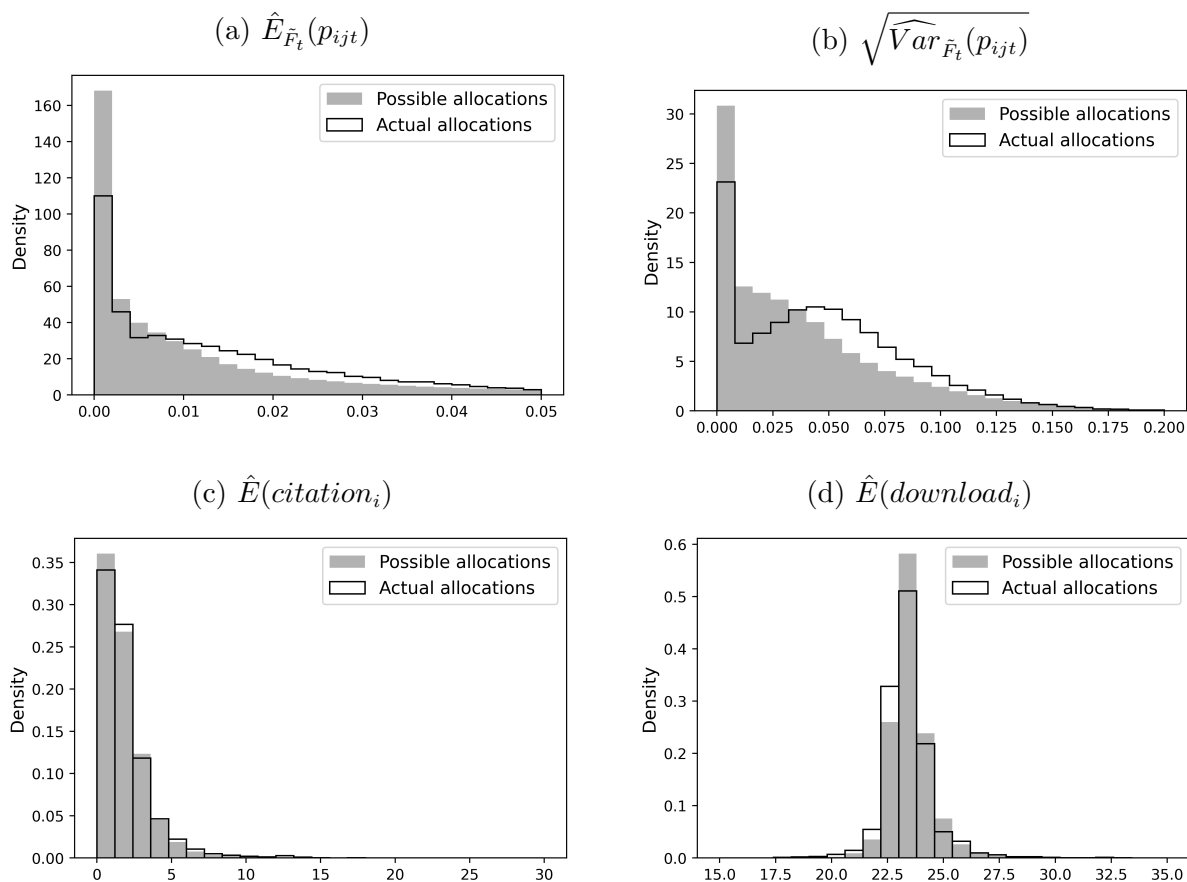
(a) $\hat{E}_{\tilde{F}_t}(p_{ijt})$

(b) $\sqrt{\widehat{Var}_{\tilde{F}_t}(p_{ijt})}$

(c) $\hat{E}(citation_i)$

(d) $\hat{E}(download_i)$

Note: Each panel shows two distributions. The outlined bars show the distribution of the variable for trials actually allocated at the four largest labs. The size of this data is 714,736. The grey unoutlined bars show the distribution of the variable for a random sample of all possible trials in those labs' choice sets. The size of this data is 105,533,539. See Appendix C.3 about sampling from choice sets. Panel B7a shows the distributions of the posterior expected probability of success of a trial. Panel B7b shows the posterior standard deviation of the probability of success of a trial. Panel B7c shows the expected five-year citations and mentions of a trial upon publication. Panel B7d shows the expected five-year downloads (in thousands) of a trial upon publication. An observation from a lab is inversely weighted by the total number of observations from this lab so each lab contributes equally to the aggregate distribution. The number of bins is coarse and equal to 25 to avoid depicting very tall bars near zero. The distributions are truncated to their respective ranges of x-axis values.

B.4. Estimated Posterior Beliefs

Figure B7 shows the estimated posterior beliefs and provides descriptive evidence that the labs explored high-variance projects in resource allocation and, in doing so, forwent some opportunities to exploit projects with high posterior expected output.

Figure B7a shows the distribution of $\hat{E}_{\tilde{F}_t}(p_{ijt})$ for trials the four largest labs actually allocated versus that for a random sample of trials in those labs' choice sets. While the distribution of $\hat{E}_{\tilde{F}_t}(p_{ijt})$ for the actual allocations is less right-skewed than that of the possible allocations, a considerable proportion of actual

allocations had extremely low posterior expected probabilities of success. This is unlikely due to exhaustion of good possible allocations with high posterior expected probabilities of success. The actual allocations accounted for only 0.7% of the random sample of the possible allocations.

Figure B7a shows the distribution of $\hat{E}_{\tilde{F}_t}(p_{ijt})$ for trials the four largest labs actually allocated versus that for a random sample of trials in those labs' choice sets. While the distribution of $\hat{E}_{\tilde{F}_t}(p_{ijt})$ for the actual allocations is less right-skewed than that of the possible allocations, a considerable proportion of actual allocations had extremely low posterior expected probabilities of success. This is unlikely due to exhaustion of good possible allocations with high posterior expected probabilities of success. The actual allocations accounted for only 0.7% of the random sample of the possible allocations.

Figure B7b shows the distribution of $\sqrt{\widehat{Var}_{\tilde{F}_t}(p_{ijt})}$. The figure shows that on average the actual allocations had considerably more posterior variance in the probability of success than the possible allocations, providing evidence that the labs allocated a considerable amount of resources to high-variance projects.

Figures B7c and B7d show that the distributions of $\hat{E}(citation_i)$ and $\hat{E}(download_i)$. Both show the trials the labs allocated to were not very different from the trials the labs passed in terms of expected citations and downloads they would generate upon publication. This suggests choosing high-variance projects was associated with not being chiefly motivated by citations and downloads, providing evidence that the labs allocated to high-variance projects for the sake of exploration.

Appendix C: Additional Details on Modeling and Estimation Procedure

C.1. Functional Form of Payoff Function $\pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})$

To begin, recall that the payoff $\pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})$ of project-trial j_i at t given action $\boldsymbol{a}_{lt}$ has a probability distribution depending on p_{ijt}. The lab or we, the researchers, do not perfectly know p_{ijt}, but previous outcomes of trials reveal information about it so one can form a posterior $\tilde{F}_t(p_{ijt}|\Omega_t)$. Integrating $\pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})$ over the posterior, we obtain the posterior expected payoff $\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl})$.

We define a function $q(\boldsymbol{a}_{lt}, p_{ijt})$ that maps the probability of success to the probability of *payoff* of project-trial j_i at t given actions. When a project-trial is not allocated on day t, it does not pay off even though it may have a nonzero probability of success. Moreover, the labs often simultaneously allocated multiple trials to the same project. A successful trial j_i of project i should only receive payoff if the simultaneous trials $(j-m)_i, (j-m+1)_i, ..., (j-1)_i$ fail, because we have assumed only the first successful trial/structure of a project produces welfare. Each trial j_i is a Bernoulli trial with probability p_{ijt}. We can express $q(\boldsymbol{a}_{lt}, p_{ijt})$ as follows:

$$q(\boldsymbol{a}_{lt}, p_{ijt}) = a_{ijt}(1 - p_{ijt})^m p_{ijt}, \tag{EC.7}$$

where trials $(j-m)_i, (j-m+1)_i, ..., (j-1)_i$ are in the choice set C_{lt} and trial $(j-m)_i$ is the smallest-numbered trial of project i in C_{lt}.[43] When $m = 0$, trial j_i is the smallest-numbered trial of project i in C_{lt} and its probability of payoff is simply $a_{ijt}p_{ijt}$.

---

[43] $q(\boldsymbol{a}_{lt}, p_{ijt}) = a_{ijt}(1 - p_{i,j-m,t})...(1 - p_{i,j-1,t})p_{ijt}$. As all trials on day t share the same information set Ω_t, the posteriors for $p_{i,j-m,t}, ..., p_{ijt}$ are the same.

We then specify a deterministic reward function, which captures the amount of payoff a lab will get when a trial pays off. We let the reward function $r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl})$ be a function of project i's characteristics $\boldsymbol{X}_{it}$ on day t. $\boldsymbol{\theta}_{Xl}$ are the welfare weights on $\boldsymbol{X}_{it}$ and are to be estimated. To reduce the number of parameters, we restrict $\boldsymbol{X}_{it}$ to correspond to the set of NIH evaluation metrics. $\boldsymbol{X}_{it}$ includes a constant 1 to capture preference for quantity; $novel_i$ and $prevStruct_{iy}$ to capture preference for novelty; $biomed_i$ and $prevPub_{iy}$ to capture preference for biomedical importance; and $human_i$, $eukaryote_i$, and $membrane_i$ to capture preferences for human, eukaryotic, and membrane proteins, respectively. We let $r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl})$ have a simple linear form

$$r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) = 1 \cdot \theta_{quant,l} + biomed_i \cdot \theta_{biomed,l} + \ldots + membrane_i \cdot \theta_{membrane,l}. \tag{EC.8}$$

Whenever a trial pays off, the lab receives a baseline amount $\theta_{quant,l}$ plus additional amounts depending on the other characteristics of the project.

We can then break the posterior expected payoff into a few pieces:

$$\int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}_t(p_{ijt}|\Omega_t) = \int r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) \cdot q(\boldsymbol{a}_{lt}, p_{ijt}) \, d\tilde{F}_t(p_{ijt}|\Omega_t)$$

$$\overset{\text{let it be } M_{ijt}}{= a_{ijt} \cdot r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) \underbrace{\int \overbrace{[(1-p_{ijt})^m p_{ijt}]}^{} \, d\tilde{F}_t(p_{ijt}|\Omega_t)}_{\text{estimated offline}}. \tag{EC.9}$$

Notice that M_{ijt} only depends on p_{ijt}. Since we have estimated $\tilde{F}_t(p_{ijt}|\Omega_t)$ offline (see Appendix B.1), we can estimate $E_{\tilde{F}_t}(M_{ijt})$ and $Var_{\tilde{F}_t}(M_{ijt})$ offline as well.[44]

Also notice that given the breakdown of the posterior expected payoff in equation (EC.9), V_{ijt}^A does not depend on the full action vector $\boldsymbol{a}_{lt}$. It only depends on the action a_{ijt}. Plugging in equation (EC.9) into V_{ijt}^A in equation (3) and evaluating it at $a_{ijt} = 1$, we obtain:

$$V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) = \int \pi_{ijt}(a_{ijt} = 1, p_{ijt}; \boldsymbol{\theta}_{Xl}) \, d\tilde{F}_t(p_{ijt}|\Omega_t) + B_{ijt}(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_{Bl})$$

$$= r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) E_{\tilde{F}_t}(M_{ijt}) + B_{ijt}(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_{Bl}). \tag{EC.12}$$

For the main model, $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) = r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) E_{\tilde{F}_t}(M_{ijt}) + \sqrt{\frac{\theta_{B1,l}}{j}} + \theta_{B2,l} \cdot (t - t'_{i,t})$. For alternative model 1, where $B_{ijt}(\cdot) = 0$, this reduces to $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) = r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) E_{\tilde{F}_t}(M_{ijt})$. For alternative model 2, $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) = r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) E_{\tilde{F}_t}(M_{ijt}) + \psi(\cdot) r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) Var_{\tilde{F}_t}(M_{ijt})$. Evaluating V_{ijt}^A at $a_{ijt} = 0$, we obtain $V_{ijt}^A(\Omega_t, a_{ijt} = 0; \boldsymbol{\theta}_l) = 0$ for all models.

Moreover, with this functional form, the second constraint in equation (4) will always be guaranteed by the solution. Notice that in most models including the main model, for all $j_i < j'_i \in C_{lt}$, $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) \geq V_{ij't}^A(\Omega_t, a_{ij't} = 1; \boldsymbol{\theta}_l)$ because the two terms only differ by $E_{\tilde{F}_t}(M_{ijt}) \geq E_{\tilde{F}_t}(M_{ij't})$. For models based on the Gittins index, the two terms also differ by $Var_{\tilde{F}_t}(M_{ijt}) \geq Var_{\tilde{F}_t}(M_{ij't})$ and the constraint continues to be satisfied. In equation (EC.13), we add an ϵ_{it} to both terms and the constraint continues to be satisfied.

---

[44] Each trial j_i is a Bernoulli trial with probability of success p_{ijt},

$$E_{\tilde{F}_{q(t)}}(M_{ijt}) = E_{\tilde{F}_{q(t)}}((1-p_{ijt})^m p_{ijt}) = [1 - E_{\tilde{F}_{q(t)}}(p_{ijt})]^m \cdot E_{\tilde{F}_{q(t)}}(p_{ijt}), \tag{EC.10}$$

$$Var_{\tilde{F}_{q(t)}}(M_{ijt}) = Var_{\tilde{F}_{q(t)}}((1-p_{ijt})^m p_{ijt})$$

$$= \{Var_{\tilde{F}_{q(t)}}(p_{ijt}) + [E_{\tilde{F}_{q(t)}}(p_{ijt})]^2\} \times \{Var_{\tilde{F}_{q(t)}}(1-p_{ijt}) + [E_{\tilde{F}_{q(t)}}(1-p_{ijt})]^2\}^m \tag{EC.11}$$

$$- [E_{\tilde{F}_{q(t)}}(p_{ijt})]^2 \times \{[E_{\tilde{F}_{q(t)}}(1-p_{ijt})]^2\}^m.$$

Plugging in $\hat{E}_{\tilde{F}_{q(t)}}(p_{ijt})$ and $\widehat{Var}_{\tilde{F}_{q(t)}}(p_{ijt})$ into the above equations, one obtains $\hat{E}_{\tilde{F}_{q(t)}}(M_{ijt})$ and $\widehat{Var}_{\tilde{F}_{q(t)}}(M_{ijt})$.

C.2. Likelihood Function

Our model assumes that the labs chose actions based on the approximate value function; therefore the observed actions $\boldsymbol{a}_{l1}^o, \boldsymbol{a}_{l2}^o, ..., \boldsymbol{a}_{lT}^o$ are the solutions $\boldsymbol{a}_{l1}^{A*}, \boldsymbol{a}_{l2}^{A*}, ..., \boldsymbol{a}_{lT}^{A*}$ to equation (4). We can use the likelihood for the observed allocations $P(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$ to estimate $\boldsymbol{\theta}_l$.

To form $P(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$, we start by rewriting the allocation problem in equation (4) as follows

$$
\begin{aligned}
\boldsymbol{a}_{lt}^o = \arg\max_{\boldsymbol{a}_{lt}} \sum_{j_i \in C_{lt}} V_{ijt}^A(\Omega_t, a_{ijt}; \boldsymbol{\theta}_l) + \epsilon_{it}, \\
\text{subject to } \sum_{j_i \in C_{lt}} a_{ijt} = n_{lt}.
\end{aligned}
\tag{EC.13}
$$

We replaced $\boldsymbol{a}_{lt}^{A*}$ with $\boldsymbol{a}_{lt}^o$ in equation (EC.13) to reflect that the observed allocations are the optimal actions based on the approximate value function. We introduced an additive error term ϵ_{it}. We assume all $\epsilon \overset{iid}{\sim}$ Type I Extreme Value and capture the unobservables in decision-making. This will help us to write $P(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$ as a closed-form function of $\boldsymbol{\theta}_l$. We also made some very intuitive functional form assumptions about the payoff function $\pi_{ijt}(\cdot)$, so that $V_{ijt}^A(\cdot)$ is only a function of a_{ijt}, and not a function of the full $\boldsymbol{a}_{lt}$ vector. These assumptions allow us to replace $V_{ijt}^A(\Omega_t, \boldsymbol{a}_{ijt}; \boldsymbol{\theta}_l)$ with $V_{ijt}^A(\Omega_t, a_{ijt}; \boldsymbol{\theta}_l)$. The functional form assumptions also guarantee that the second constraint of equation (4) holds at the solution, allowing us to drop the second constraint for equation (EC.13). See Appendix C.1 for details of the functional form assumptions and derivations.

The solution $\boldsymbol{a}_{lt}^o$ to equation (EC.13) is equivalent to the following index rule: one first computes index $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) + \epsilon_{it}$ for each project-trial in C_{lt} and then allocates to the n_{lt} trials with the highest index values.

For computational tractability, we assume the lab made iid decisions whether to allocate each trial, using a threshold rule. The lab allocated a trial whenever the trial's index value was greater than the threshold on that day. Doing so allows us to express the likelihood $P(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$ with a simple closed form. Let $V_{lt}^{n_{lt}}(\boldsymbol{\theta}_l)$ denote the n_{lt}th highest value of $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)$ on day t. Define the threshold value which is equal to $V_{lt}^{n_{lt}}(\boldsymbol{\theta}_l)$ plus an error term ϵ_{lt}:[45]

$$
threshold_{lt} = V_{lt}^{n_{lt}}(\boldsymbol{\theta}_l) + \epsilon_{lt}.
\tag{EC.14}
$$

The likelihood of observing $a_{ijt}^o = 1$ is equal to the likelihood of $V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) + \epsilon_{it}$ being greater than this threshold, so the likelihood $P(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ is

$$
\begin{aligned}
P(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l) &= P(V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) + \epsilon_{it} > V_{lt}^{n_{lt}}(\boldsymbol{\theta}_l) + \epsilon_{lt}) \\
&= \frac{exp(V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l))}{exp(V_{ijt}^A(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l)) + exp(V_{lt}^{n_{lt}}(\boldsymbol{\theta}_l))}.
\end{aligned}
\tag{EC.15}
$$

---

[45] Note that this value is not necessarily the n_{lt}th highest index value on day t due to the iid error terms. In principle, one can obtain the n_{lt}th highest index value on day t with simulation draws of ϵ's and use that as the threshold. Doing so would be computationally more challenging than our approach.

The total likelihood function sums over the log likelihood of the observed action for each project-trial in choice sets $C_{l1}, C_{l2}, ..., C_{lT}$. These likelihoods include $P(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l)$ for trials actually allocated and $1 - P(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l)$ for trials in those choice sets but which were *not* allocated:

$$\boldsymbol{\theta}^*_l = \arg\max_{\boldsymbol{\theta}_l} \sum_t^T \left(\underbrace{\sum_{j_i \in C_{lt},\, a^o_{ijt}=1} log(P(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l))}_{\text{actually allocated trials}} + \underbrace{\sum_{j_i \in C_{lt},\, a^o_{ijt}=0} log(1 - P(\Omega_t, a^o_{ijt} = 1; \boldsymbol{\theta}_l))}_{\text{actually not allocated trials}} \right). \quad \text{(EC.16)}$$

One can then estimate $\boldsymbol{\theta}_l$ by maximizing the above likelihood function. Due to the simplicity of computing index approximations, estimation is feasible even though the choice sets contain millions of possible actions over thousands of days. A further trick to reduce computational burden is to compute in each iteration the log likelihoods for a random sample of the possible allocations in the choice sets, rather than for the full choice sets. The number of possible allocations in a full choice set could be large because n_{lt} could be large. Recall that the mean of n_{lt} for the four largest labs is 35. When $n_{lt} = 35$, the $(j+1)$th, $(j+2)$th,..., $(j+35)$th trials of every project in the lab's portfolio are in the full choice set. See Appendix C.3 for how we sampled from the choice sets.

C.3. Specifying Choice Set C_{lt}

As discussed in Section 2, the major labs in our setting received new projects through three ways with close NIH involvement: 1) a centralized planning committee periodically assigned families of novel molecules; 2) the biomedical research community nominated projects; and 3) the labs determined projects of their own interest, which they reported to the NIH well in advance. These processes placed limits on the new projects the labs could plausibly consider when they made trial allocations. These limits allow us to considerably reduce the size of C_{lt}.

We restrict C_{lt} to include only the following project-trials. For an older project i that the lab has attempted up until trial j_i in period $t' < t$, we include trials $(j+1)_i, ..., (j+n_t)_i$ in C_{lt}. For a new project i' that the lab has not attempted until t but attempts within the next six months, we include trials $1_{i'}, ..., (n_t)_{i'}$ in C_{lt}. One can also consider using alternative windows for the new projects, such as projects attempted within the next three months or nine months. Doing so changes the magnitudes of the estimates, but all qualitative results are the same as when using six months as the window. Likewise, adding some new projects that the lab could have considered but never actually attempted could change the estimates, but the qualitative results should stay the same.

We further reduce the sizes of the choice sets used in estimation by taking random subsamples of C_{lt}. The reason is related to computation. Labs sometimes allocated hundreds of project-trials on a day and they usually had tens of thousands of projects in their portfolios. The sizes of some choice sets C_{lt} could be on the order of millions. Given that we have thousands of periods, if we use the full choice sets C_{lt}, we need to compute log likelihood for billions of choices in each iteration of maximum likelihood. This would result in a very large memory burden and slow computation. Moreover, it is not necessary to include every possible choice in C_{lt} to consistently estimate $\boldsymbol{\theta}_l$. A random subsample of the choices on each side of the threshold value would be sufficient. Due to the sheer number of projects in each lab's portfolio, the set of actual

Table C1 Trials included in the reduced choice set for estimation

Project	Trial	Actually allocated on day t?	Notes
	$(j+1)_i$	Y	Trials $(j+1)_i$ through $(j+m-1)_i$ were actually allocated on day t, include all in C_{lt}^R. When no trial was allocated on day t, m equals 1.
	$\vdots$	$\vdots$	
	$(j+m-1)_i$	Y	
i	$(j+m)_i$	N	
	$\vdots$	$\vdots$	Trials $(j+m)_i$ through $(j+n_{lt})_i$ were not actually allocated on day t, include one random trial $(j+r)_i$ in C_{lt}^R.
	$(j+r)_i$	N	
	$\vdots$	$\vdots$	
	$(j+n_{lt})_i$	N	

Notes: For an older project that has been attempted before t, j equals the number of trials allocated to the project before t. For a new project, $j=0$. Trials in black are included in the reduced choice set C_{lt}^R. Trials in grey are in the choice set C_{lt} but are excluded from C_{lt}^R to reduce computational burden.

trials is much smaller than the set of not-allocated trials. We therefore reduce the sizes of the choice sets for estimation by taking random subsamples of the latter on the level of project and date. Let the reduced choice sets be $C_{lt}^R \subset C_{lt}$. Table C1 shows the project-trials in C_{lt}^R after random sampling from C_{lt}.

C.4. Additional Notes on the Simulation Procedure

In the "simulate outcomes" step of our simulation procedure, we allocate to the top n_{lt} trials having the greatest values of $\widehat{VA}'_{ijt} + \epsilon'_{it}$, where ϵ'_{it} are iid draws from a Type-I Extreme Value distribution. As detailed in Appendix Section C.2, these iid errors are introduced to encapsulate any remaining unobserved variables in the labs' decision-making process. These errors also allow us to express $P(\Omega_t, \boldsymbol{a}_{lt}^o; \boldsymbol{\theta}_l)$ as a closed-form function of $\boldsymbol{\theta}_l$, facilitating the estimation process.

Appendix D: Additional Results

Figure D1   Observed output: number of structures and citations
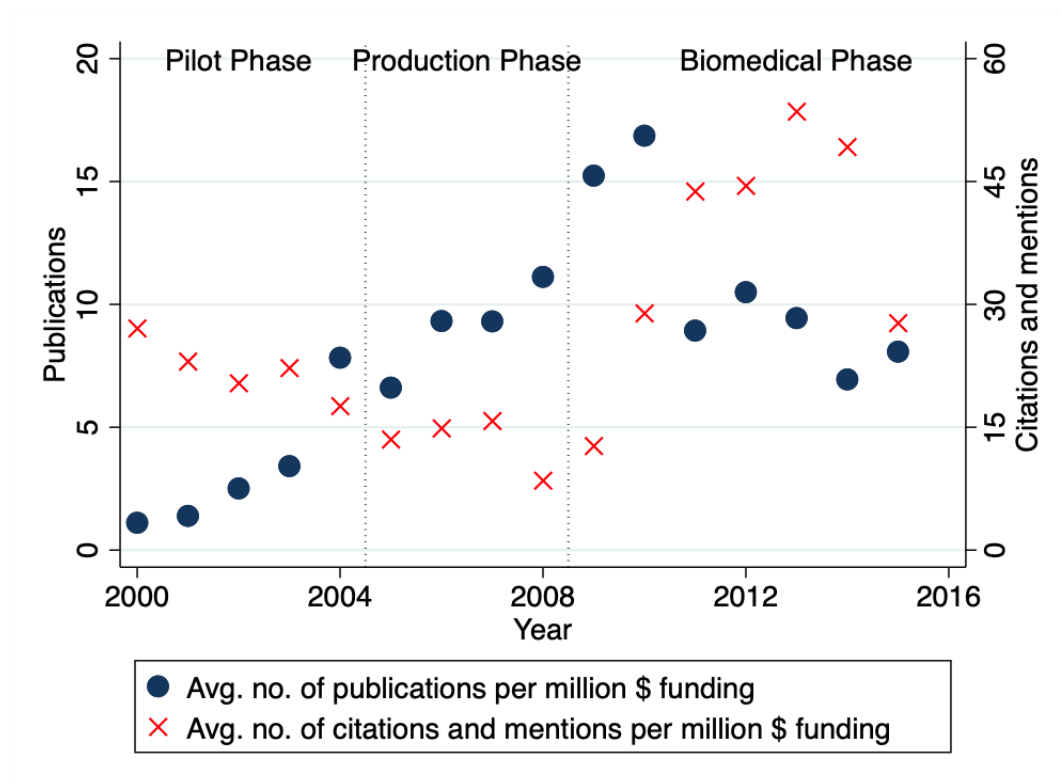
Note: The blue dots show the number of published structures on unique molecules in a given year divided by the lab consortium's funding in millions in that year. The red crosses show the number of 5-year citations and mentions the published structures in that year generated, divided by the lab consortium's funding in millions in that year. Each plot shows the average value across the four large lab in each year. The disaggregated values are in Figure A1.

Table D1   Comparison of model fits, JCSG

Model	Free parameters in $B_{ijt}(\cdot)$	Log likelihood	Avg $\hat{P}(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ actually allocated trials	Avg $1 - \hat{P}(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ actually not allocated trials
Greedy	0	-1,308,021	0.691	0.891
Gittins	0	-1,104,171	0.640	0.910
UCB	1	-544,118	0.854	0.981
FlexGittins	1	-1,056,772	0.640	0.917
FlexGittins+D	2	-349,531	0.644	0.981
UCB+D	2	-206,951	0.909	0.995

Note: The estimation uses data from 2005 to 2015 from JCSG, one of the four major labs. The number of trials actually allocated was 320,295. The number of trials in the choice sets after random sampling is 5,807,902. The rest of the notes of Table 3 apply.

Table D2   Comparison of model fits, MCSG

Model	Free parameters in $B_{ijt}(\cdot)$	Log likelihood	Avg $\hat{P}(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ actually allocated trials	1-Avg $\hat{P}(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ actually not allocated trials
Greedy	0	-636,089	0.644	0.994
Gittins	0	-473,481	0.669	0.996
UCB	1	-274,230	0.769	0.998
FlexGittins	1	-402,754	0.681	0.997
FlexGittins+D	2	-225,201	0.780	0.998
UCB+D	2	-157,987	0.850	0.999

Note: The estimation uses data from 2005 to 2015 from MCSG, one of the four major labs. The number of trials actually allocated was 141,059. The number of trials in the choice sets after random sampling is 39,136,035. The rest of the notes of Table 3 apply.

Table D3   Comparison of model fits, NYSGRC

Model	Free parameters in $B_{ijt}(\cdot)$	Log likelihood	Avg $\hat{P}(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ actually allocated trials	Avg $1 - \hat{P}(\Omega_t, a_{ijt}^o = 1; \boldsymbol{\theta}_l)$ actually not allocatd trials
Greedy	0	-677,382	0.538	0.990
Gittins	0	-550,827	0.561	0.992
UCB	1	-339,498	0.682	0.996
FlexGittins	1	-529,851	0.565	0.993
FlexGittins+D	2	-400,671	0.626	0.994
UCB+D	2	-288,162	0.700	0.996

Note: The estimation uses data from 2005 to 2015 from NYSGRC, one of the four major labs. The number of trials actually allocated was 139,276. The number of trials in the choice sets after random sampling is 23,883,552. The rest of the notes of Table 3 apply.

Table D4   Out-of-sample fit of UCB+D model

Lab	Sample	Avg Log likelihood	Avg $\hat{P}(\Omega_t, a_{ijt}^o = 1; \hat{\boldsymbol{\theta}}_l)$ actually allocated trials	Avg $1 - \hat{P}(\Omega_t, a_{ijt}^o = 1; \hat{\boldsymbol{\theta}}_l)$ actually not allocated trials
JCSG	in	-0.033	0.909	0.994
	out	-0.036	0.914	0.996
MCSG	in	-0.004	0.826	0.999
	out	-0.004	0.871	0.999
NESG	in	-0.004	0.817	0.999
	out	-0.003	0.850	0.999
NYSGRC	in	-0.013	0.656	0.996
	out	-0.011	0.734	0.997

Note: To compute these results, we first estimated the UCB+D model using only observed allocation decisions in odd years. We then computed the in-sample results using the estimates and the odd years' decisions. We computed the out-of-sample results using the estimates and even years' decisions. The rest of the notes of Table 3 apply.

Table D5: Parameter estimates, NESG

	UCB+D		Greedy		Gittins		UCB		FlexGittins		FlexGittins+D	
	(1) 2005–2008	(2) 2009–2015	(3) 2005–2008	(4) 2009–2015	(5) 2005–2008	(6) 2009–2015	(7) 2005–2008	(8) 2009–2015	2005–2008	2009–2015	2005–2008	2009–2015
θ_{B1}	158.3	119.5	—	—	—	—	210.5	188.8	1.6	1.4	1.4	0.8
	[156.4, 160.1]	[118.1,121.4]					[208.8, 212.5]	[188.5,189.2]	[1.6,1.6]	[1.4,1.4]	[1.4,1.4]	[0.8,0.8]
θ_{B2}	-2.3	-4.7									-6.5	-8.4
	[-2.3, -2.2]	[-4.7,-4.7]					—	—		—	[-6.4,-6.6]	[-8.4,-8.4]
θ_{quant}	105.6	95.9	288.9	232.6	167.5	162.3	106.1	146.1	150.0	149.5	138.4	143.5
	[105.5, 105.8]	[95.8,95.9]	[286.3, 291.4]	[232.4,232.9]	[167.1, 167.8]	[162.1,162.5]	[106.0, 106.3]	[146.1,146.2]	[149.3,151.0]	[149.4,149.8]	[137.9,139.1]	[143.3,143.7]
θ_{novel}	28.4	-23.3	202.6	20.7	67.6	-8.5	28.7	-39.8	48.4	-14.4	38.7	-7.7
	[27.8, 28.9]	[-23.4,-23.1]	[200.5, 204.3]	[20.5,20.8]	[67.3, 67.8]	[-8.6,-8.4]	[28.6, 28.9]	[-39.8,-39.8]	[47.9,49.1]	[-14.5,-14.3]	[38.1,39.4]	[-7.7,-7.7]
$\theta_{prevStructZ}$	18.1	20.1	9.8	56.5	3.8	7.9	22.3	47.6	3.0	5.2	-0.4	4.9
	[17.6, 18.6]	[20.1,20.3]	[9.4, 10.0]	[56.4,56.6]	[3.8, 3.9]	[7.8,8.0]	[22.1, 22.5]	[47.5,47.7]	[2.7,3.2]	[5.1,5.2]	[-0.7,-0.1]	[4.8,5.0]
θ_{biomed}	21.5	52.9	45.3	158.5	7.3	44.1	21.5	102.5	8.0	38.0	3.4	29.3
	[21.3, 21.7]	[52.7,52.9]	[43.1, 47.0]	[158.3,158.8]	[7.1, 7.5]	[44.0,44.2]	[21.4, 21.7]	[102.4,102.5]	[7.8,8.3]	[37.8,38.4]	[2.8,3.9]	[29.1,29.3]
$\theta_{prevPubZ}$	-9.2	-3.0	-14.5	-1.7	-6.9	-0.5	-9.4	-6.8	-6.9	-4.5	-3.8	-3.1
	[-9.3, -9.0]	[-3.1,-3.0]	[-14.7, -14.2]	[-1.7,-1.7]	[-7.0, -6.8]	[-0.5,-0.4]	[-9.6, -9.2]	[-6.8,-6.7]	[-7.2,-6.7]	[-4.5,-4.4]	[-3.9,-3.6]	[-3.1,-3.1]
θ_{human}	67.1	124.0	105.9	191.0	51.4	90.5	69.8	160.6	40.2	83.0	36.3	74.5
	[66.9, 67.3]	[123.9,124.0]	[104.8, 106.9]	[190.7,191.3]	[51.2, 51.6]	[90.4,90.6]	[69.7, 70.0]	[160.6,160.7]	[40.1,40.5]	[82.9,83.1]	[36.0,36.7]	[74.5,74.6]
$\theta_{eukaryote}$	-29.8	-35.9	67.7	61.6	-8.3	-26.9	-30.0	-0.3	-14.3	-33.7	-11.7	-33.4
	[-29.9, -29.7]	[-36.0,-35.8]	[67.1, 69.1]	[61.5,61.8]	[-8.7, -8.0]	[-27.0,-26.8]	[-30.4, -29.8]	[-0.3,-0.2]	[-14.4,-13.9]	[-33.7,-33.6]	[-12.0,-11.6]	[-33.4,-33.4]
$\theta_{membrane}$	58.3	13.1	96.5	52.3	35.2	28.4	58.1	22.7	31.2	24.8	23.5	22.1
	[58.0, 58.6]	[13.1,13.3]	[96.1, 97.2]	[52.1,52.4]	[35.1, 35.4]	[28.3,28.5]	[57.9, 58.5]	[22.7,22.8]	[31.1,31.4]	[24.8,24.9]	[23.1,23.9]	[22.1,22.2]

Note: Table displays the full estimates of parameters in different models for NESG, one of the four large labs. Results from other labs are available upon request. $prevStructZ_{iy}$ represents the number of standard deviations by which $prevStruct_{iy}$ differs from the yearly mean $\overline{prevStruct}_y$. $prevPubZ_{iy}$ represents the number of standard deviations by which $prevPub_{iy}$ differs from the yearly mean $\overline{prevPub}_y$. The rest of the notes of Table 4 apply.

Table D6: Key parameter estimates in the UCB+D model, other labs

Parameter	JCSG		MCSG		NYSGRC	
	2005–2008	2009–2015	2005–2008	2009–2015	2005–2008	2009–2015
	(1)	(2)	(3)	(4)	(5)	(6)
θ_{B1}	558.2	1001.9	273.3	127.1	61.6	115.6
	[551.1,573.4]	[977.6,1028.0]	[266.2,284.8]	[126.3,127.8]	[61.2,61.9]	[114.2,116.9]
θ_{B2}	-247.9	-104.7	-3.9	-3.8	-2.9	-3.9
	[-247.6, -248.5]	[-102.9, -106.6]	[-3.9,-3.9]	[-3.8,-3.8]	[-2.9,-3.0]	[-3.9,-3.9]
θ_{biomed}	-34.3	105.1	12.7	84.8	33.0	89.4
	[-34.6,-34.1]	[104.7,105.8]	[12.6,12.7]	[84.8,84.8]	[32.9,33.1]	[89.1,89.8]

Note: Table displays the estimates of the main parameters of interest in the UCB+D model for the other three large labs. The full set of parameter estimates in different models are available upon request. 95% confidence intervals, calculated using Chernozhukov and Hong (2003)'s MCMC approach, are provided in brackets. These confidence intervals are almost identical to those calculated using Chen et al. (2018)'s Procedure 1.
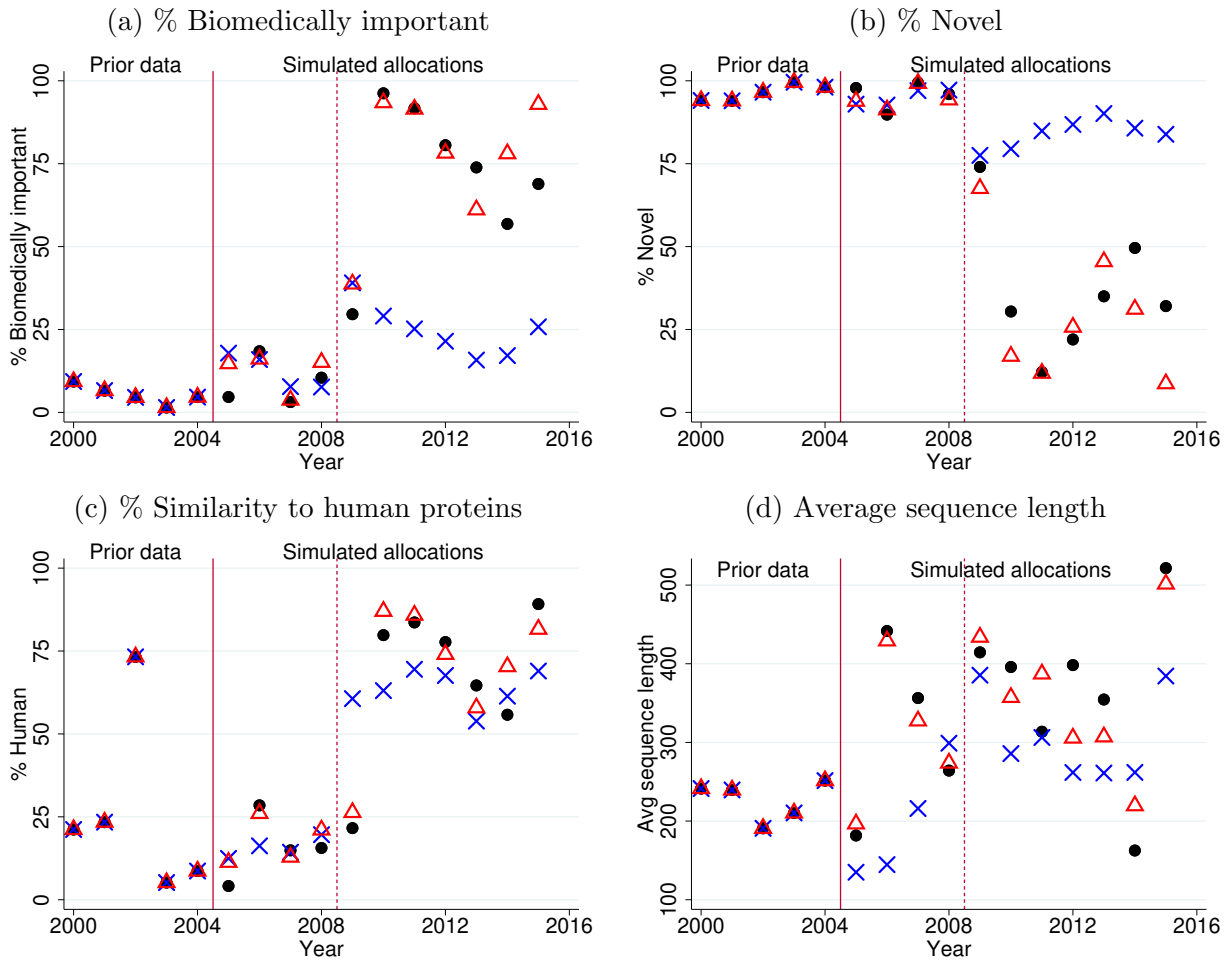
Table D7 **Simulated outcomes of UCB+D model, other labs**

Lab	Model	Projects attempted	Unique structures	Citations	Downloads (millions)
JCSG	UCB+D	40,881	1,607	1,495	38.4
	Actual	40,881	1,512	1,463	36.3†
MCSG	UCB+D	77,503	2,040	2,524	46.3
	Actual	78,740	2,276	3,145	50.0†
NYSGRC*	UCB+D	59,734	626	2,579	14.7
	Actual	59,734	617	2,575	14.4†

Note: *We encountered data issues related to NYSGRC. For this lab, over half of the trials that produced structures have missing or incorrectly ordered key stage dates (e.g., the publication of a structure dated earlier than previous stages). Therefore, we were unable to accurately simulate the dates and outcomes of different stages of trials. While these simulation results appear promising, they should be interpreted with caution. The rest of the notes of Table 5 apply.

Figure D2 **Simulated distributions of input allocation across projects**

(a) Greedy

(b) Gittins

(c) UCB

(d) FlexGittins

(e) FlexGittins+D

(f) UCB+D
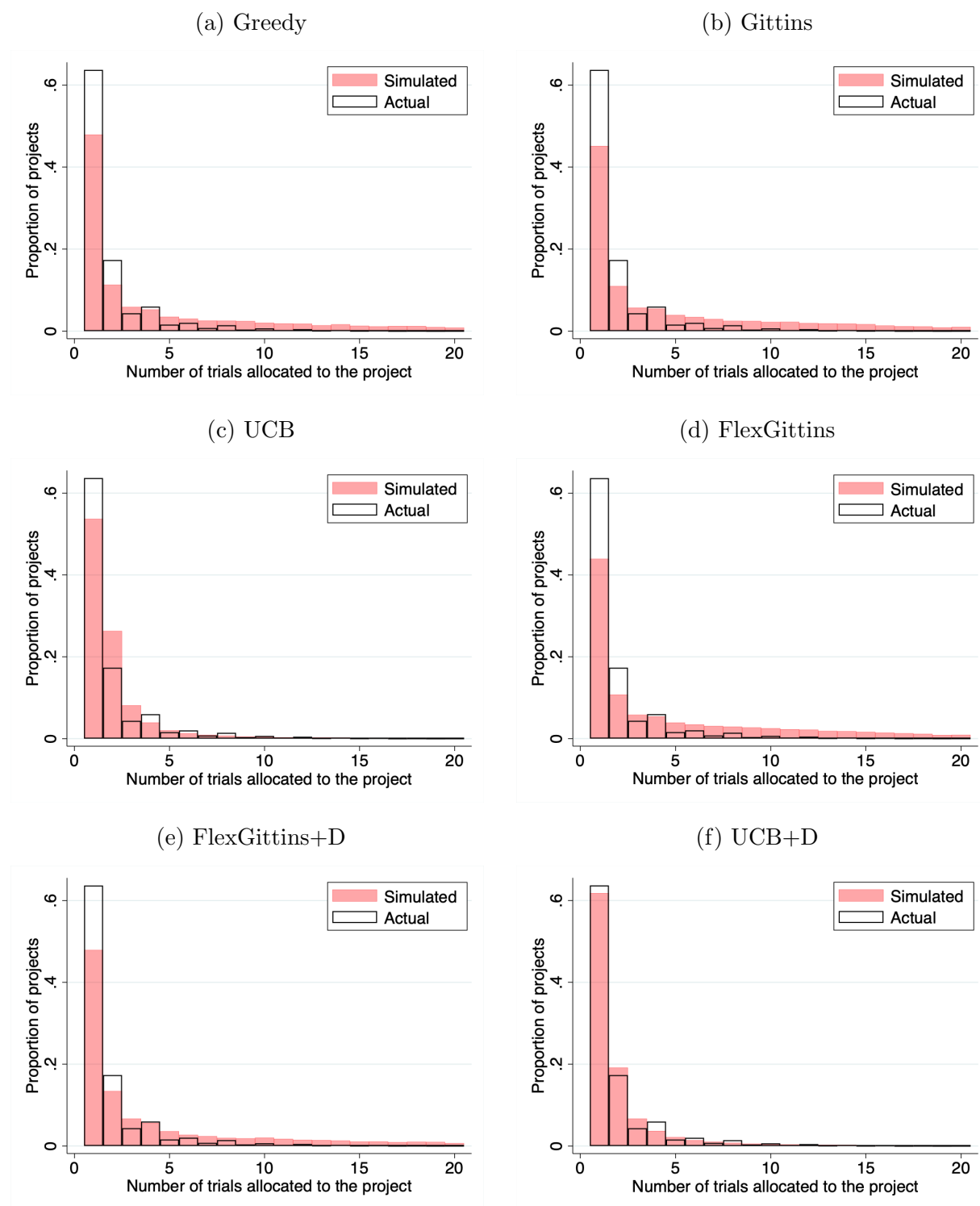
Note: Each panel shows two distributions. The outlined bars show the actual distribution of the number of trials across projects at NESG. The pink bars show the simulated distribution under the model. Each distribution is based on a single simulation of the corresponding model.

Figure D3 **Characteristics of trials allocated in simulations, greedy model versus UCB+D model**

(a) % Biomedically important

(b) % Novel

(c) % Similarity to human proteins

(d) Average sequence length

Note: Black dots represent characteristics of actually allocated trials at NESG. Blue crosses represent those of trials allocated under the greedy model. Red triangles represent those of trials allocated under the UCB+D model. Each series is based on a single simulation of the corresponding model. The first three panels show project characteristics that the NIH evaluated the lab's progress on. The fourth panel shows one project characteristic that the NIH did not evaluate on.

Table D8   Out-of-sample simulation results for UCB+D model

Lab	Model	Projects attempted	Unique structures	Citations	Downloads (millions)
JCSG	in	40,881	1,607	1,495	38.4
	out	40,881	1,621	1,481	38.6
	actual	40,881	1,512	1,463	36.3†
MCSG	in	77,503	2,040	2,524	46.3
	out	77,504	2,051	2,553	46.4
	actual	78,740	2,276	3,145	50.0†
NESG	in	59,947	1,097	3,376	25.6
	out	59,913	1,085	3,336	25.3
	actual	59,953	1,053	3,502	24.5†
NYSGRC*	in	59,734	626	2,579	14.7
	out	59,734	628	2,594	14.7
	actual	59,734	617	2,575	14.4†

Note: In-sample simulation results are identical to those for the UCB+D model in Tables 5 and D7. We simulate the out-of-sample results as follows. We first fit the UCB+D model for each lab with odd years of observed decisions. Using those estimated parameters, we simulate each lab's full input allocation history and output, in odd and even years. The rest of the notes of Table D7 apply.

Table D9   Counterfactual outcomes, alternative models of exploration, other labs

Lab	Counterfactual model	Projects attempted	Unique structures	Citations	Downloads (millions)
JCSG	Greedy (no exploration)	10,710	1,338	1,270	32.4
		(-74%)	(−17%)	(−15%)	(−16%)
	Gittins	11,296	1,310	1,219	31.7
		(-72%)	(−18%)	(−18%)	(−17%)
	Thompson sampling	11,756	1,310	1,257	31.8
		(-71%)	(−18%)	(−16%)	(−17%)
	Explore-Then-Commit	40,881	1,685	1,483	40.4
		(+0%)	(+5%)	(−0%)	(+5%)
	Baseline model	40,881	1,607	1,495	38.4
MCSG	Greedy (no exploration)	14,883	668	856	15.6
		(-81%)	(−67%)	(−66%)	(−66%)
	Gittins	17,859	870	1,106	20.3
		(-77%)	(−57%)	(−56%)	(−56%)
	Thompson sampling	25,826	1,090	1,347	25.4
		(-67%)	(−47%)	(−47%)	(−45%)
	Explore-Then-Commit	77,539	1,680	2,065	38.6
		(+0%)	(−18%)	(−18%)	(−17%)
	Baseline model	77,503	2,040	2,524	46.3
NYSGRC*	Greedy (no exploration)	5,203	247	1,026	5.9
		(-91%)	(−61%)	(−60%)	(−60%)
	Gittins	6,278	313	1,317	7.4
		(-89%)	(−50%)	(−49%)	(−50%)
	Thompson sampling	21,742	598	2,454	14.0
		(-64%)	(−4%)	(−5%)	(−5%)
	Explore-Then-Commit	59,734	454	1,885	10.7
		(+0%)	(−27%)	(−27%)	(−27%)
	Baseline model	59,734	626	2,579	14.7

Note: Each simulation uses $\hat{\boldsymbol{\theta}}_{Xl}$ from the parameter estimates of the UCB+D model for the corresponding lab, which are available upon request. Results are averaged from three simulations of the model. Outputs from the baseline model are identical to those for the UCB+D model in Appendix Table D7. Parentheses indicate percentage differences compared to the baseline model. This table shows results from the three other large labs. See Table 6 for results from NESG. *We encountered data issues related to NYSGRC. For this lab, over half of the trials that produced structures have missing or incorrectly ordered key stage dates (e.g., the publication of a structure dated earlier than previous stages). Therefore, we were unable to accurately simulate the dates and outcomes of different stages of trials. While these simulation results appear promising, they should be interpreted with caution.

Table D10   Counterfactual outcomes, alternative forms of information utilization, other labs

Lab	Counterfactual model	Projects attempted	Unique structures	Citations	Downloads (millions)
JCSG	UCB+D, no piloting	33,836	1,332	1,092	32.8
		(-17%)	(-17%)	(-27%)	(-15%)
	UCB+D, no analytics	40,881	1,523	1,392	37.8
		$(+0\%)$	(-5%)	(-7%)	(-2%)
	Baseline model	40,881	1,607	1,495	38.4
MCSG	UCB+D, no piloting	63,989	1,509	1,926	36.8
		(-17%)	(-26%)	(-24%)	(-21%)
	UCB+D, no analytics	77,425	1,920	2,416	46.7
		(-0%)	(-6%)	(-4%)	$(+0\%)$
	Baseline model	77,503	2,040	2,524	46.3
NYSGRC*	UCB+D, no piloting	55,637	440	1,820	10.4
		(-7%)	(-30%)	(-29%)	(-29%)
	UCB+D, no analytics	59,718	597	2,464	14.2
		(-0%)	(-5%)	(-4%)	(-3%)
	Baseline model	59,734	626	2,579	14.7

Note: Table shows results from the three other large labs. See Table 7 for results from NESG. The rest of the notes of Appendix Table D9 apply.

Table D11   Simulated outcomes of UCB+D model, all labs, 2005 and beyond

Lab	Model	Unique structures	Citations	Downloads (millions)
JCSG	UCB+D	1,458	1,290	34.7
	Actual	1,363	863	32.6[†]
MCSG	UCB+D	1,817	2,221	41.0
	Actual	2,053	2,489	44.7[†]
NESG	UCB+D	946	2,916	22.1
	Actual	902	2,257	21.0[†]
NYSGRC*	UCB+D	484	1,974	11.3
	Actual	475	1,881	11.0[†]

Note: This table displays the simulated outcomes of the main model for 2005 and beyond. The simulation runs are identical to those runs displayed in Tables 5 and D7. The rest of the notes of Tables 5 and D7 apply.

Additional References for the Appendices

Huang H, McGarvey PB, Suzek BE, Mazumder R, Zhang J, Chen Y, Wu CH (2011) A comprehensive protein-centric ID mapping service for molecular data integration. *Bioinformatics* 27(8):1190–1191.

Kawashima S, Pokarowski P, Pokarowska M, Kolinski A, Katayama T, Kanehisa M (2007) AAindex: Amino acid index database, progress report 2008. *Nucleic Acids Research* 36(suppl_1):D202–D205.

Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Sønderby CK, Sommer MOA, Winther O, Nielsen M, Petersen B, et al. (2019) NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. *Proteins: Structure, Function, and Bioinformatics* 87(6):520–527.

Miller S, Janin J, Lesk AM, Chothia C (1987) Interior and surface of monomeric proteins. *Journal of Molecular Biology* 196(3):641–656.

UniProt (2021a) Programmatic access—mapping database identifiers. URL `https://www.uniprot.org/help/api_idmapping`, accessed on Apr 20, 2021.

UniProt (2021b) Programmatic access—retrieving individual entries. URL `https://www.uniprot.org/help/api_retrieve_entries`, accessed on April 18, 2021.

UniProt (2021c) Taxonomy—Eukaryota. URL `https://www.uniprot.org/taxonomy/2759`, accessed on Dec 14, 2020.

UniProt (2021d) Taxonomy—Homo sapiens (human). URL `https://www.uniprot.org/taxonomy/9606`, accessed on Dec 14, 2020.

Wootton JC (1994) Non-globular domains in protein sequences: Automated segmentation using complexity measures. *Computers & Chemistry* 18(3):269–285.

Acknowledgments