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

Ran Zhuo

University of Michigan

NBER SI 2023

How do scientists decide which projects get resources?

How do scientists decide which projects get resources?

- Objective: max productivity (quantity, impact, ...) over some horizon
- Challenge: **incomplete information** about which projects will be productive

How do scientists decide which projects get resources?

- Objective: max productivity (quantity, impact, ...) over some horizon
- Challenge: **incomplete information** about which projects will be productive

Exploitation	**vs**	**Exploration**
prioritize projects that one has...		
good info about		**poor info** about
pursue safe projects		acquire information
max short-term productivity		improve long-term productivity

How do scientists decide which projects get resources?

- Objective: max productivity (quantity, impact, ...) over some horizon
- Challenge: **incomplete information** about which projects will be productive

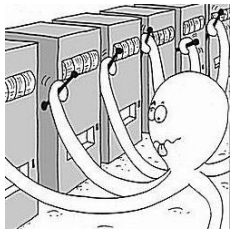
Exploitation	**vs**	**Exploration**
prioritize projects that one has...		
good info about		**poor info** about
pursue safe projects		acquire information
max short-term productivity		improve long-term productivity

My research empirically models and estimates how a group of large labs traded off exploitation and exploration in resource allocation under incomplete information

Why is this question interesting and important?

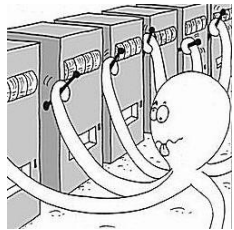
- Classic problem in theoretical literature, little empirical evidence
  - Multi-armed bandit
- Large stakes
  - Labs in sample spending $=$\$1.3B over 2000–2015
  - US spending on R&D $>$\$500B per year

Why is this question interesting and important?



- Classic problem in theoretical literature, little empirical evidence
  - Multi-armed bandit
- Large stakes
  - Labs in sample spending $=$\$1.3B over 2000–2015
  - US spending on R&D $>$\$500B per year

Why is this question interesting and important?



- Classic problem in theoretical literature, little empirical evidence
  - Multi-armed bandit
- Large stakes
  - Labs in sample spending $=$\$1.3B over 2000–2015
  - US spending on R&D $>$\$500B per year

Results

- Labs **explored extensively**
- Exploration had a **large positive impact** on their productivity
- **Policy counterfactuals**: Alternative allocation models? Effect of informatics?

Sections

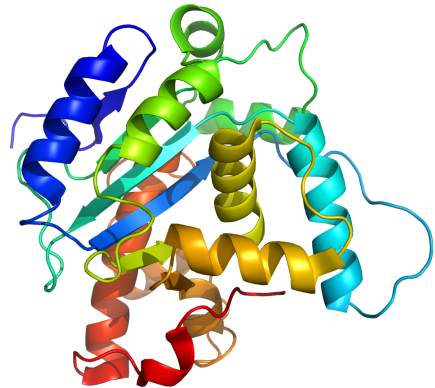
- **Setting**
- Model & Estimation
- Results

Setting enables studying allocation at an ultra-micro level

- **Structural biology labs**
  - Funded by $1.3B NIH Protein Structure Initiative (PSI)

Setting enables studying allocation at an ultra-micro level

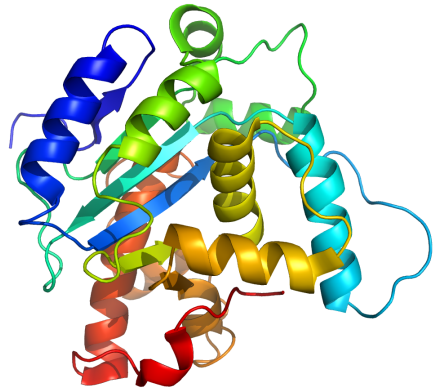
- **Structural biology labs**
  - Funded by $1.3B NIH Protein Structure Initiative (PSI)



Project = finding molecule's 3D image

Setting enables studying allocation at an ultra-micro level

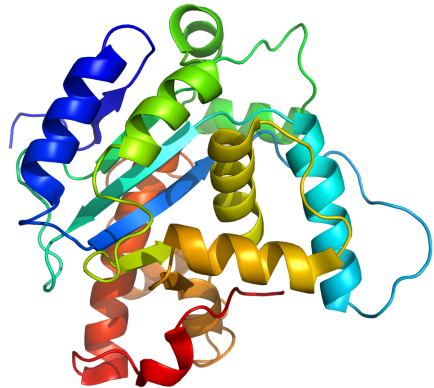
- **Structural biology labs**
  - Funded by $1.3B NIH Protein Structure Initiative (PSI)
- **Important basic research**
  - Lead to valuable applied research e.g. structure-based drug design of COVID vaccines



Project = finding molecule's 3D image

Setting enables studying allocation at an ultra-micro level

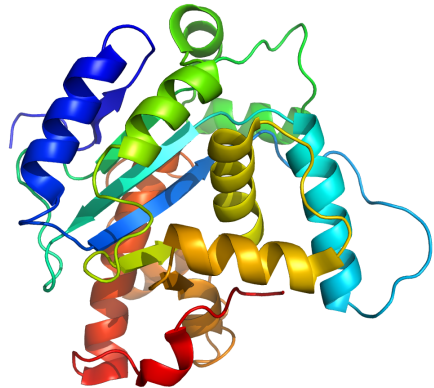
- **Structural biology labs**
  - Funded by $1.3B NIH Protein Structure Initiative (PSI)
- **Important basic research**
  - Lead to valuable applied research e.g. structure-based drug design of COVID vaccines
- **Highly granular data**
  - Clearly defined projects; discrete input bundles
  - Daily input allocation to $>$**300,000** projects, $\sim$**1 million** input bundles
  - Output from each allocation (structure Y/N, citations, downloads)



Project = finding molecule's 3D image

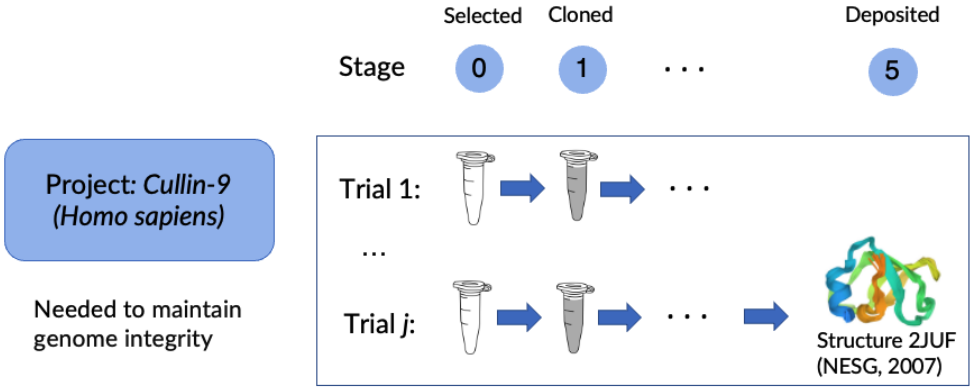
Setting enables studying allocation at an ultra-micro level

- **Structural biology labs**
  - Funded by $1.3B NIH Protein Structure Initiative (PSI)
- **Important basic research**
  - Lead to valuable applied research e.g. structure-based drug design of COVID vaccines
- **Highly granular data**
  - Clearly defined projects; discrete input bundles
  - Daily input allocation to $>$**300,000** projects, $\sim$**1 million** input bundles
  - Output from each allocation (structure Y/N, citations, downloads)
  - 4 large labs (71% of projects, 85% of input)
    - Funding & productivity of those labs



Project = finding molecule's 3D image

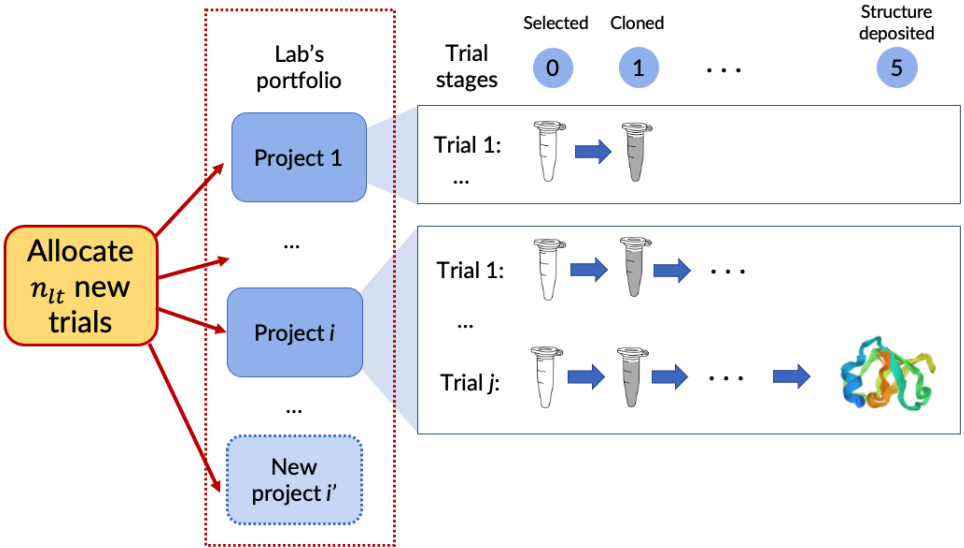
One input bundle = one experimental trial



Large variations in success/failure at every stage, even within project

"...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..." (Chruszcz et al., 2008)

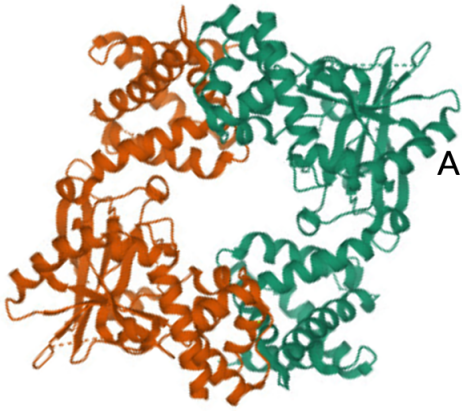
Lab allocates trials among many projects



Exploit or Explore? Let's be NESG on May 30, 2009...

	Project A Methyl-CpG-binding domain protein 4 (Homo sapiens) involved in DNA repair	**VS**	**Project B** Malonyl-CoA decarboxylase (Cupriavidus metallidurans) 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
ML pred prob of success next trial	0.0692		0.0012
Similarity to prev tried projects	100%		58%

Exploit or Explore? Let's be NESG on May 30, 2009...



Lab chose
Project B
A success on the first trial!

Structure 4KS9
Five-year citations=10
(Avg five-year citations=4.7)

Selection rationale

Previous trials

ML pred pro success nex

Similarity to tried projects

rboxylase
llidurans)
metabolism

NIH policies simplify the setting and motivate counterfactuals

Simplifications

- Assume away competition U01 collaborative grant, ▸ More evidence committee determined and assigned projects to labs ▸ Details
- Assume away principal-agent problem, predetermined preferences NIH closely monitored labs based on evaluation metrics ▸ Evidence

Many interesting policy features

- Strong emphasis on exploration of poorly covered knowledge space ▸ Evidence
- Support for informatics: databases, informaticians,... ($40M)

. . .

Sections

- Setting
- **Model & Estimation**
- Results

Sketch of a simple model

Labs faced a finite-horizon dynamic decision problem

For each period t in horizon $1, \ldots, T$:

1. Lab uses info from prev trials to retrain ML models and update posterior about future trial success probability $\boldsymbol{p}$

2. Based on this posterior & its preferences, lab determines the "value" of each trial

$$\text{value} = \underbrace{V_{exploit}}_{\text{current payoff}} + \underbrace{V_{explore}}_{\text{benefit for future}}$$

3. Lab chooses trials with the highest values up to capacity constraint

Estimation & model validation

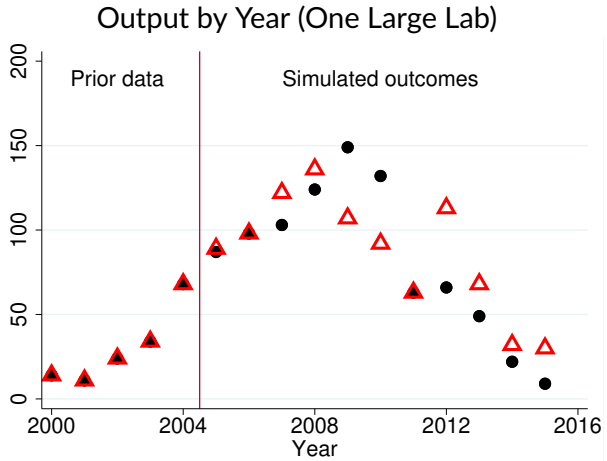
- **Estimation recovers the lab's perceived** $V_{exploit}(\cdot\,; \widehat{\boldsymbol{\theta}}, \boldsymbol{p})$ **and** $V_{explore}(\cdot\,; \widehat{\boldsymbol{\theta}}, \boldsymbol{p})$, methodological innovation for computational tractability ▸ Likelihood Function ▸ Intuition for Identification
- Use estimated model to **simulate labs' allocation, training of ML models, and output period-by-period and compare w/ actual data** ▸ Simulation Procedure

Sections

- Setting
- Model & Estimation
- **Results**

Model captures the labs' decisions: evidence from model validation simulations

▸ Additional Evidence



▸ Another lab

Model captures the labs' decisions: evidence from model validation simulations

▸ Additional Evidence

Simulated number of projects attempted, output quantity, citations, downloads within 10% different from actual output for all labs

Labs explored extensively—as they should!

Estimates reject no exploration

Estimates

- $\widehat{V}_{explore} >> 0$ for all large labs

Labs explored extensively—as they should!

Estimates reject no exploration

▸ Estimates

- $\widehat{V}_{explore} >> 0$ for all large labs

Effect of exploration?

- Compare w/ no exploration in counterfactual

Labs explored extensively—as they should!

Estimates reject no exploration

▸ Estimates

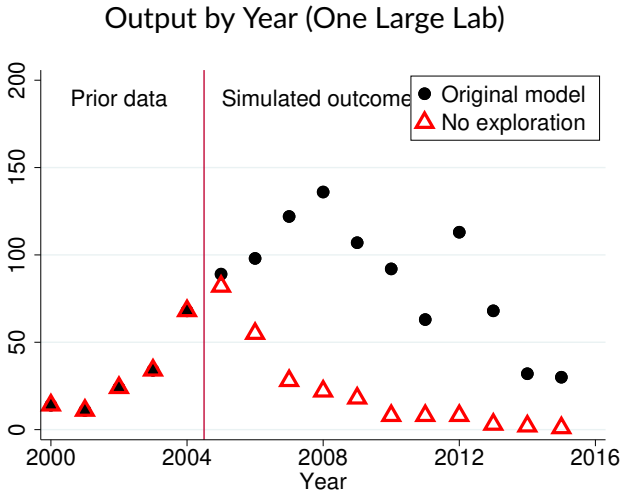
- $\widehat{V}_{explore} >> 0$ for all large labs

Effect of exploration?

- Compare w/ no exploration in counterfactual

Miss low-hanging fruits, inefficient allocation

- Output quantity $\downarrow 51\%$, citations $\downarrow 57\%$ across labs



Many policy counterfactuals: one example

What if no informatics?

- Save $40M (3% of funding)

Many policy counterfactuals: one example

What if no informatics?

- Save $40M (3% of funding)
- No machine learning in counterfactual

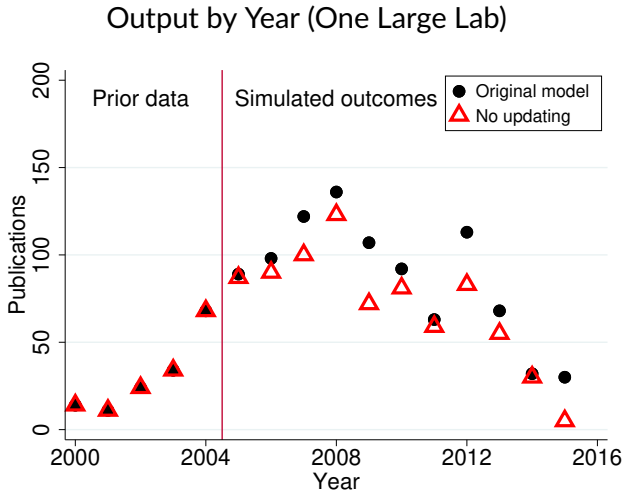
Many policy counterfactuals: one example

What if no informatics?

- Save $40M (3% of funding)
- No machine learning in counterfactual

Still find low-hanging fruits, but less inefficient allocation

- Output quantity ↓ 7%, citations ↓ 9% across labs



Discussion—why do these results matter?

- **Exploration improves long-term productivity** and should be encouraged
- **"Information" is an important research output** that improves future allocation and should be rewarded
- Policy relevance: PSI and beyond

Discussion—why do these results matter?

- **Exploration improves long-term productivity** and should be encouraged
- **"Information" is an important research output** that improves future allocation and should be rewarded
- Policy relevance: PSI and beyond

Some bold proposals

- Open, queryable database of research experience (what has been done, successes AND failures)
- Big data analytics on project potential
- Reward for research experience vs research successes
- More ideas?

Thank You

Literature (science of science)

Back to Contributions

- **Resource allocation under uncertainty & incomplete information:** Arrow (1962); Roberts & Weitzman (1981); Bergemann & Hege (2005)...
- **Idea development, experimentation & learning:** Cohen & Levinthal (1989); Henderson & Cockburn (1996); Azoulay et al. (2011); Ederer & Manso (2013); Manso (2016); Krieger (2021); Ganglmair et al. (2019); Khmelnitskaya (2021); Lane et al. (2022); Nagaraj et al. (2022)...
- **Impact of innovation policy & institutions:** Jaffe (2002); Furman & Stern (2011); Azoulay (2012); Cantoni & Yuchtman (2014); Lane et al. (2015); Azoulay et al. (2019); Myers (2020)...

Literature (single-agent dynamics and multi-armed bandit)

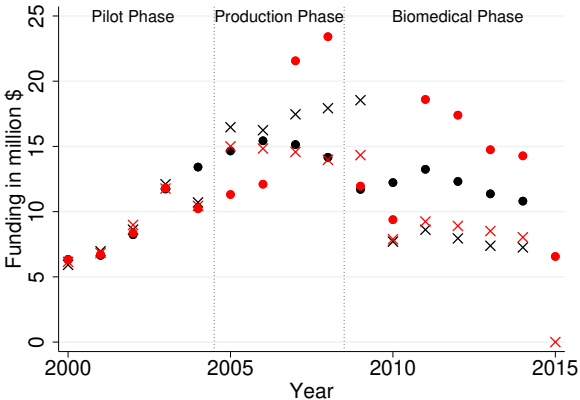
Back to Contributions

- **Recursive & simulation methods:** Pakes (1986); Rust (1987); Hotz & Miller (1993); Hotz et al. (1994); Rust (1994); Timmins (2002); Aguirregabiria & Mira (2010)...
- **Theoretical lit on MAB and dynamic allocation indices:** Gittins (1979); Weitzman (1979); Lai & Robbins (1985); Whittle (1988); Bergemann & Valimaki (1996); Bolton & Harris (1999); Auer et al. (2002); Keller et al. (2005); Bubeck & Cesa-Bianchi (2012); Russo et al. (2017)...
- **Empirical analyses of bandit-like single-agent problems:** Miller (1984); Erdem & Keane (1996); Crawford & Shum (2005); Dickstein (2018); Li et al. (2020); Caria et al. (2020)...

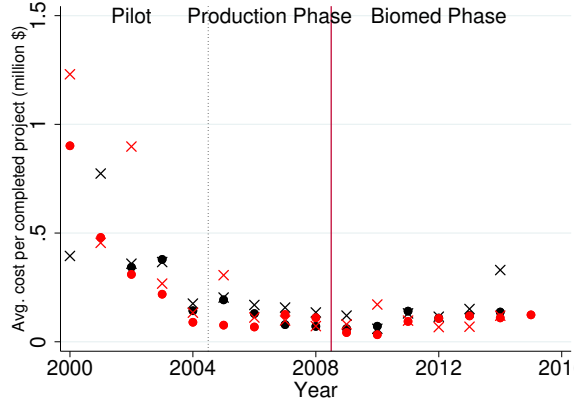
Funding & productivity of 4 large labs

Back to Empirical Setting

Funding (million $)



Productivity (cost per structure)

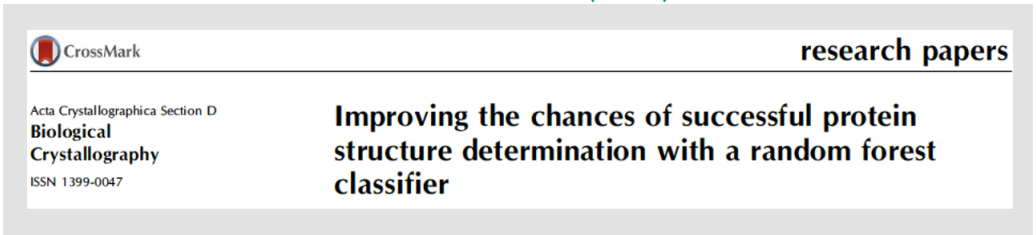


Labs using ML to form posterior

Back to Empirical Setting

- Labs published a series of journal articles describing their ML models (Slabinski et al. (2007a,b); Jaroszewski et al. (2008); Price Ii et al. (2009a,b); Babnigg & Joachimiak (2010); Jahandideh et al. (2014)

Jahandideh et al. (2014)



CrossMark

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Improving the chances of successful protein structure determination with a random forest classifier

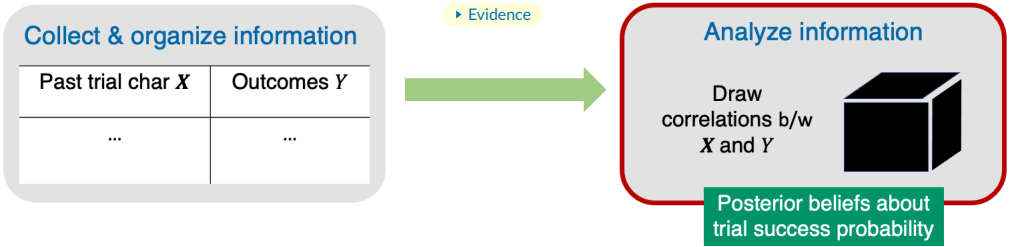
How did the labs use info to guide decision-making?

Collect & organize information

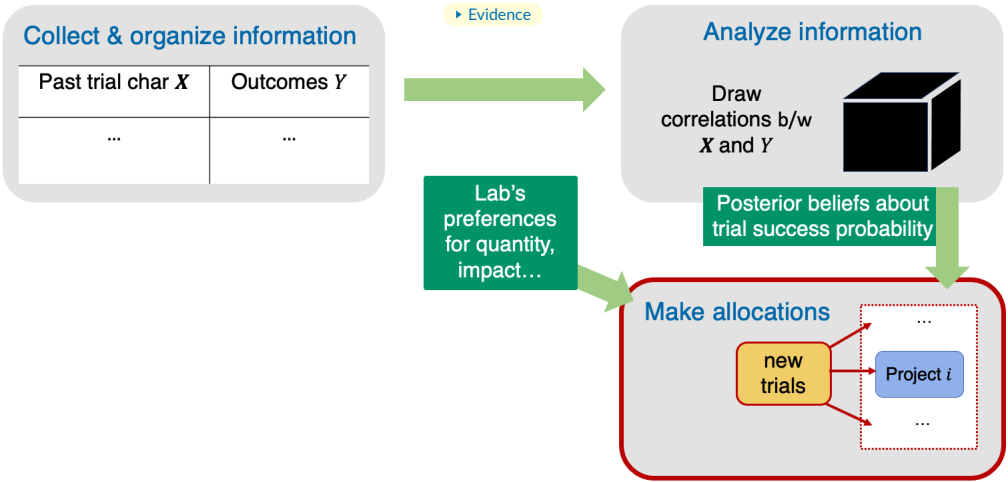
Past trial char X	Outcomes Y
...	...

▸ Evidence

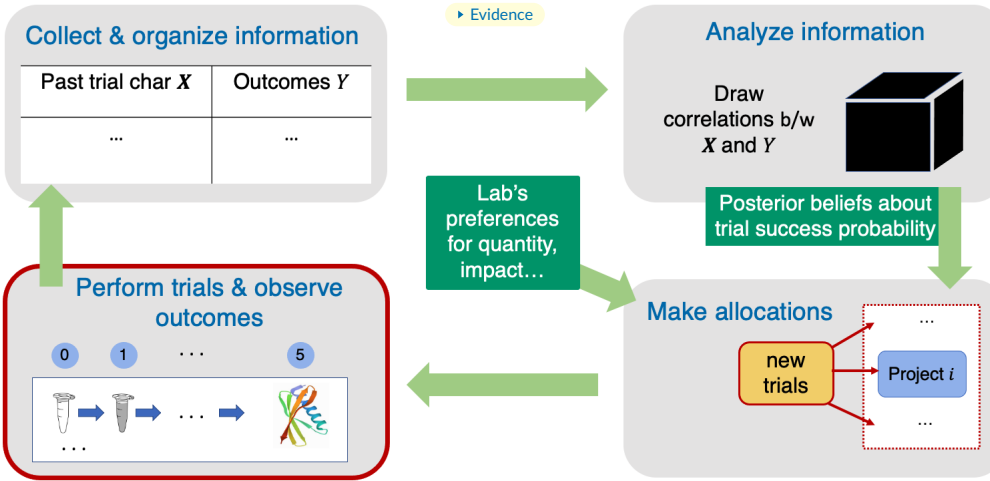
How did the labs use info to guide decision-making?



How did the labs use info to guide decision-making?



How did the labs use info to guide decision-making?



NIH evaluation metrics

Back to NIH Involvement

During production phase

Updated October 1, 2008

I.1.A	Number of Novel Experimental PSI-2 Structures	**1253**
I.1.B	Number of Distinct Experimental PSI-2 Structures with Nonredundant Sequences	**1762**
I.1.D	Total Number of Experimental PSI-2 Structures	**1983**
I.1.E	Numbers of Experimentally Determined Distinct Residues	**396764**
	Numbers of Experimentally Determined Novel Residues	**276296**
I.2.J	Number of Experimental Structures of Human Proteins	**71**
I.2.K	Number of Experimental Structures of Eukaryotic Proteins	**206**
I.2.M	Number of Experimental Structures of Membrane Proteins	**10**
I.2.N	Number of Experimental Structures Determined at the Atomic Level using X-ray Crystallography	**1801**
	Number of Experimental Structures Determined at the Atomic Level using NMR Methods	**182**

NIH evaluation metrics

Back to NIH Involvement

During biomedical phase

Updated August 03, 2010

I.1.A	Number of Novel Experimental PSI-2 Structures	1985
I.1.B	Number of Distinct Experimental PSI-2 Structures with Nonredundant Sequences	3077
I.1.D	Total Number of Experimental PSI-2 Structures	3518
I.1.E	Numbers of Experimentally Determined Distinct Residues	704106
	Numbers of Experimentally Determined Novel Residues	446761
I.2.B	Number of Experimental Structures from Biomedical Theme Target Lists	1141
I.2.C	Number of Experimental Structures from Community Outreach Target Lists	406
I.2.J	Number of Experimental Structures of Human Proteins	130
I.2.K	Number of Experimental Structures of Eukaryotic Proteins	317
I.2.M	Number of Experimental Structures of Membrane Proteins	64
I.2.N	Number of Experimental Structures Determined at the Atomic Level using X-ray Crystallography	3162
	Number of Experimental Structures Determined at the Atomic Level using NMR Methods	356

Production target for the large labs

Back to NIH Involvement

This FOA (RFA-GM-10-005) specifically solicits applications to establish the **Centers for High-Throughput Structure Determination**. The centers must be able to provide capabilities for high-throughput structure determination on the order of those that have been developed during previous phases of the Protein Structure Initiative, e.g., ~ 200 structures per year deposited in the PDB. This rate of structure

NIH policy features

Back to NIH Involvement

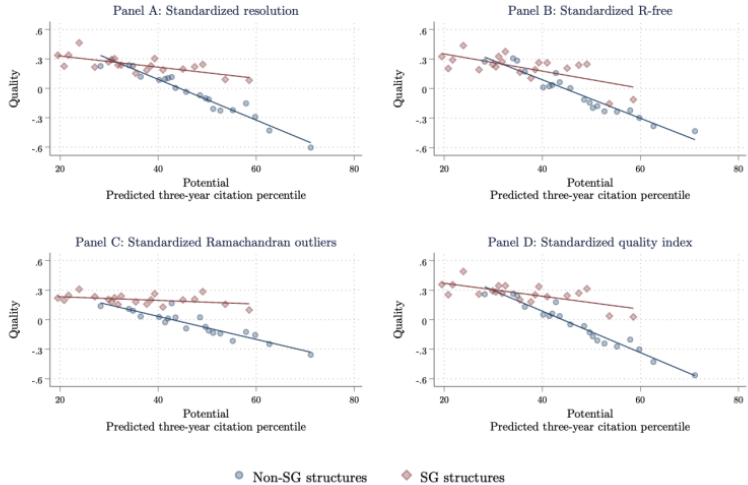
- Restrictions on choice sets
  - Centralized committee periodically drew families of molecules
  - Solicited nominations from research community
  - Reviewed projects proposed by labs
- Robustness check: exogenous shift in preferences in 2009
  - Pilot (2000–2004)
  - Production phase (2005–2008): publish a lot of unique structures
  - Biomedical phase (2009 onwards): publish a lot of unique structures and focus on biomedically important ones

Hill & Stein (2022)

Back to NIH Involvement

Labs in my sample (SG) not motivated by competition

Figure 11: The Effect of Potential on Quality by Structural Genomics Status



Emphasis on poorly covered knowledge space in structural biology

▸ Back to NIH Involvement

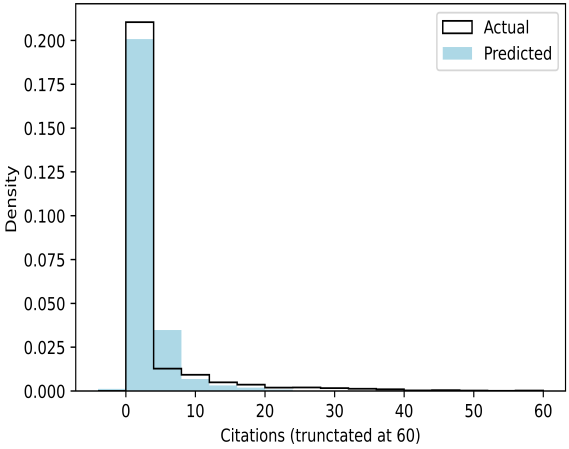
PURPOSE OF THIS RFA

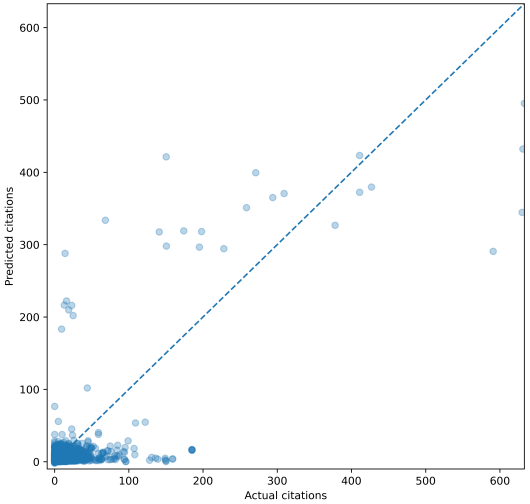
The National Institute of General Medical Sciences (NIGMS) encourages applications for cooperative agreements to support large-scale structural genomics research centers for the determination of unique protein structures. These centers will form one component of the Protein Structure Initiative (PSI) Research Network, the integrated second, or production, phase of the PSI (PSI-2). Each large-scale center must perform all tasks of structural genomics in a high-throughput operation to produce a large number of unique protein structures to meet the PSI-2 goals for structural coverage of sequenced genes. Each large-scale center must also develop technologies and methodologies that will make the production and structural determination of proteins less expensive, more efficient, and more likely to be successful.

Prediction of 5-year citations

To Simulation Procedure

- Fit a ridge reg of 5-year citations of actual pubs on their characteristics
- Hyperparameters chosen with cross validation

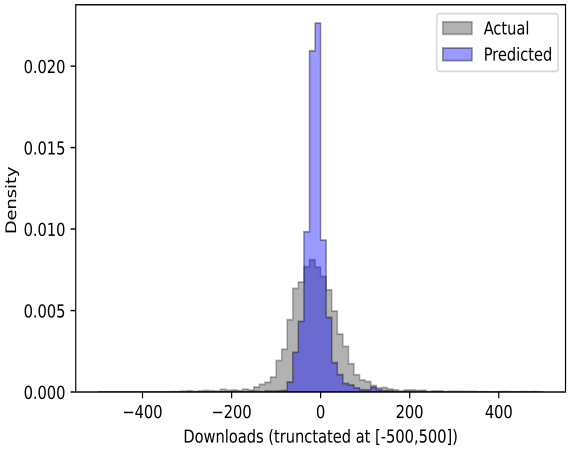


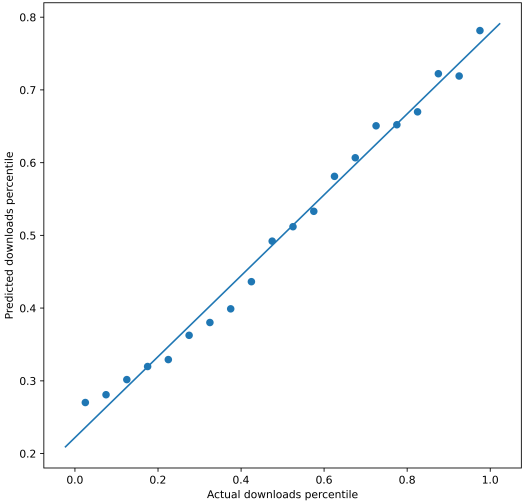


Prediction of 5-year downloads

▸ To Simulation Procedure

- Fit a ridge reg of 5-year detrended downloads of actual pubs on their characteristics
- Hyperparameter chosen with cross validation





Theoretical foundation of index approximation

Back to Index Approximation

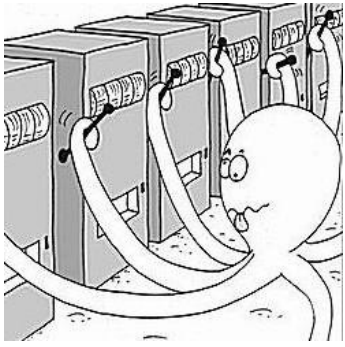
- The labs' allocation problem is a finite-horizon multi-armed bandit (MAB)
  - Project = arm
  - Trial = pull
- Some MABs have optimal index solutions
  - Compute an index value for each choice
  - Optimal action is picking the choice w/ the highest index
  - E.g. Gittins Index for infinite-horizon discounted MAB

Objective function

- The labs' problem does not have an optimal index solution
  - Nonstationarity, correlated choices, complex action space
  - But index could work reasonably well

Details

  - Lack of alternatives to index



Index rules could work reasonably well

Back to Setting Back to Index Approximation

- Finite horizon $\sum_{t=1}^{T}$ instead of $\sum_{t=1}^{\infty} \beta^t$
  - Index rules are asymptotically optimal as $T \to \infty$, nearly optimal (both in the Bayesian and frequentist viewpoints) for small T (Lai& Robbins (1985); Lai (1987))
- Nonstationarity $\tilde{F}_t$ instead of $\tilde{F}$
  - Restless bandits (Whittle (1988)), index rules suboptimal in the general case (Ortner et al. (2012))
  - Assume specific form of change in $\tilde{F}_t$, UCB-like index approximations could match the lower bound on regret up to a logarithmic factor (Garivier & Moulines (2011)
- Correlated choices $\tilde{F}_t(p_{ijt}|\Omega_t)$ instead of $\tilde{F}^{(i)}(p_{ijt}|\Omega_t^{(i)})$
  - Contextual bandits (Woodroofe (1979); Langford & Zhang (2007))
  - UCB-like index approximations could match the lower bound on regret up to a logarithmic factor (Guan & Jiang (2018); Zhou et al. (2020)), first build models to correlate the contextual characteristics with the observed outcomes, then use the models to predict the UCB
- Multiple choices in each period $\sum_{j_i \in C_{lt}} a_{ijt} = n_{lt}$, all pulled arms reveal payoffs
  - Combinatorial semi-bandits, need strong functional form assumptions on action space and payoff function for index approximations to work well (Kveton et al. (2015); Chen et al. (2016); Wang & Chen (2018))

Practical applications of index approximations

Back to Index Approximation

- Recommender systems: Netflix Link
- Dynamic pricing: Boston Globe Link
- Games: AlphaGo Link
- Self-driving cars: Tesla Link

Likelihood function – definitions

▸ Back to Model ▸ Back to Estimation

Define variables:

- a_{ijt} = 1 if allocate the jth trial to project i on day t, = 0 if not
- Ω_t information set on day t, includes allocations and outcomes observed before t
- θ_X lab's welfare weights on different evaluation metrics, unknown to us

Objective function: fix input, max welfare-weighted output over the horizon by choosing $\boldsymbol{a}_1,\ldots,\boldsymbol{a}_T$

Value function:

$$V_{ijt}(\Omega_t, \boldsymbol{a}_t; \boldsymbol{\theta}_X) = \underbrace{\pi_{ijt}(\Omega_t, \boldsymbol{a}_t; \boldsymbol{\theta}_X)}_{\text{posterior expected output}} + \underbrace{\mathbb{E}_{\Omega'_{t+1}}\left[\max_{\boldsymbol{a}_{l,t+1}} V_{i,j',t+1}(\Omega'_{t+1}, \boldsymbol{a}_{l,t+1} | \Omega_t, \boldsymbol{a}_{lt}) \right]}_{\text{continuation value}} \quad (1)$$

Try backward induction: infeasible to compute the value for every state of Ω_t

Likelihood function – index approximation

Back to Model | Back to Estimation

- True value is intractable to compute, both for labs and for us
- Indices are well-studied in theory and well-used in practice for multi-armed bandit
  - Theoretical foundation
  - Regret analysis
  - Many applications
- Assume lab approximated the perceived value w/

$$V^{A}_{ijt}(\Omega_t, \boldsymbol{a}_t; \boldsymbol{\theta}) = \underbrace{\pi_{ijt}(\Omega_t, \boldsymbol{a}_t; \boldsymbol{\theta}_X)}_{\text{as before}} + \underbrace{B_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Bl})}_{\text{exploration bonus}} \tag{2}$$

- $B_{ijt}(\cdot)$ does not integrate over future states $\Omega_{t+1}, \Omega_{t+2}, \ldots$ easily computable

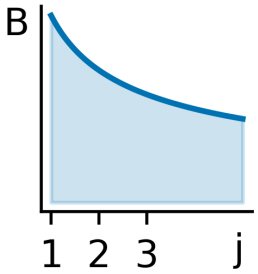
Main model modifies a well-used index

Back to Model Back to Estimation

Upper Confidence Bound (UCB)

$$B_{ijt}(\cdot) = \sqrt{\frac{\theta_{B1}}{j}}$$

- Auer et al. (2002) Detail
- Larger θ_{B1} = more exploration



Modification approximates for unobservables

Back to Model Back to Estimation

UCB + Discounting $B_{ijt}(\cdot) = \sqrt{\frac{\theta_{B1}}{j}} + \theta_{B2} \cdot (t - t'_{i,t})$

- Auer et al. (2002) Detail
- Larger θ_{B1} = more exploration

- $(t - t'_{i,t})$ = duration b/w current day & last previous allocation to project i
- Negative θ_{B2} = lab discounted older projects
- Proxy for learning staying with one individual, forgetting...

Likelihood function

Back to Estimation

- Assume lab treats whether to allocate each trial as a two-armed bandit (choose a trial if its V^A_{ijt} is greater than a threshold value)
  for computational tractability

- Let threshold $V^{n_{lt}}_{lt}(\boldsymbol{\theta}) = n_{lt}$th largest value of V^A_{ijt} on day t, $\epsilon \overset{\text{iid}}{\sim}$ Type-I EV

$$
\begin{aligned}
Pr(a^o_{ijt} = 1; \boldsymbol{\theta}) &= Pr\{ V^A_{ijt}(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}) + \epsilon_{it} > V^{n_{lt}}_{lt}(\boldsymbol{\theta}) + \epsilon_{lt} \} \\
&= \frac{exp(V^A_{ijt}(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}))}{exp(V^A_{ijt}(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta})) + exp(V^{n_{lt}}_{lt}(\boldsymbol{\theta}))} \qquad (3)
\end{aligned}
$$

- Total log likelihood adds up
  - Log $Pr(a^o_{ijt} = 1; \boldsymbol{\theta})$ for all actually allocated trials in $C_{l1}, ..., C_{lT}$
  - Log $Pr(a^o_{ijt} = 0; \boldsymbol{\theta})$ for all actually not allocated trials in $C_{l1}, ..., C_{lT}$

Functional form of $\pi_{ijt}(\cdot)$

Back to Estimation

$$\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; \boldsymbol{\theta}_{Xl}) = \int \underbrace{r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl})}_{\text{welfare upon payoff}} \cdot \underbrace{q(\boldsymbol{a}_{lt}, p_{ijt})}_{\substack{\text{posterior belief of} \\ \text{prob of payoff}}} d \underbrace{\tilde{F}_t(p_{ijt}|\Omega_t)}_{\substack{\text{posterior belief of} \\ \text{prob of success}}}, \tag{4}$$

where

$$r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) = 1 \cdot \theta_{quant,l} + biomed_i \cdot \theta_{biomed,l} + ... + human_i \cdot \theta_{human,l} \tag{5}$$

and

$$q(\boldsymbol{a}_{lt}, p_{ijt}) = a_{ijt}(1 - p_{ijt})^m p_{ijt}. \tag{6}$$

- r includes 8 variables corresponding to NIH evaluation metrics of labs' productivity
- m is the number of ongoing trials of project i before j_i
- Duplicated structure does not receive additional payoff $\Leftrightarrow$ a success of trial j_i pays off only if all of the already ongoing trials of the project fail $\Leftrightarrow (1 - p_{ijt})^m p_{ijt}$

Functional form of $\pi_{ijt}(\cdot)$

Back to Estimation

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

$$= a_{ijt} \cdot r(\boldsymbol{X}_{it}; \boldsymbol{\theta}_{Xl}) \underbrace{\int \overbrace{[(1-p_{ijt})^m p_{ijt}]}^{\text{let it be } M_{ijt}} \; d\tilde{F}_t(p_{ijt}|\Omega_t)}_{\text{estimated offline}}. \tag{7}$$

- M_{ijt} only depends on p_{ijt}, estimated offline
- $\pi_{ijt}(\Omega_t, \boldsymbol{a}_{lt}; ; \boldsymbol{\theta}_{Xl})$ is specified in a way that for all $j_i < j_i'$ in choice set, $V^A_{ijt}(\Omega_t, a_{ijt} = 1; \boldsymbol{\theta}_l) + \epsilon_{it} \geq V^A_{ij't}(\Omega_t, a_{ij't} = 1; \boldsymbol{\theta}_l) + \epsilon_{it}$, because the two terms only differ by $E_{\tilde{F}_t}(M_{ijt}) \geq E_{\tilde{F}_t}(M_{ij't})$
- So one would always choose the 4th trial before choosing the 5th trial of a project in simulation

Intuition for Identification

Back to Estimation

Search for $\hat{\boldsymbol{\theta}}_X$ and $\hat{\boldsymbol{\theta}}_B$ so $\widehat{V}^A_{ijt}(\,\cdot\,;\boldsymbol{\theta})$ maximizes likelihood of observed actions $\boldsymbol{a}^o_1, \ldots, \boldsymbol{a}^o_T$

- Specified the likelihood of choosing trial j of project i as a smooth, monotonically increasing function of $\widehat{V}^A_{ijt}(\,\cdot\,;\boldsymbol{\theta})$ Likelihood Function

- Intuition for identification: variation in biomedically importance of choices identifies $\hat{\theta}_{bio}$
  - Observe many trials on biomedically important projects $\Leftrightarrow$ large θ_{bio}
  - Want a large $\hat{\theta}_{bio}$ to maximize the log likelihood
  - Biomed projects should have large V^A_{ijt} and have a high likelihood of being chosen
  - Non-biomed projects should have small V^A_{ijt} and have a high likelihood of not being chosen

Auer et al. (2002)

Back to Main Model

$$B_{ijt}(\cdot, a_{ijt} = 1) = \begin{cases} \infty & \text{if } j = 1 \\ \sqrt{\frac{2\ln(N_{lt})}{j-1}} & \text{if } j = 2, 3, 4... \end{cases} \tag{8}$$

- N_{lt} is the units of resources allocated so far, more recent implementation uses a fixed value θ_{B1} rather than $2\ln(N_{lt})$ (see Lattimore & Szepesvari (2020))
- I do not use an infinite value for $B_{ijt}(\cdot)$ when $j = 1$. If I do, V_{ijt}^{A} would be infinite and cause problems in estimation via MLE and in identifying θ_{B1}.

Model Validation Simulations

▸ To Estimation ▸ To Counterfactual

Use $\hat{\theta}_X$ and $\hat{\theta}_B$ to simulate labs' history of allocation and output and compare w/ actual data

Initialization

- **Ground truth output:** use all trial data Ω_{T+1} to fit a flexible model F^* to predict trials' "true" productivity ▸ Prediction of Prob of Success ▸ Prediction of Citations ▸ Prediction of Downloads
- **Prior:** trial allocation and outcomes realized before 2005, not simulated

For period t in 2005–2015 do:

- **Update posterior:** use data before t to **refit** labs' posterior belief $\tilde{F}'_t(\Omega'_t)$ ▸ Details
- **Make allocations:** based on posterior $\tilde{F}'_t(\Omega'_t)$ and allocation model $\hat{\theta}_X$ and $\hat{\theta}_B$, compute $V^{A'}_{ijt}$ for trials in choice set; allocate trials according to $V^{A'}_{ijt}$ up to capacity constraint
- **Simulate trial outcomes:** use $F^*(\Omega_{T+1})$ to simulate output of the allocated trials
- **Update information set** Ω'_{t+1}

Modeling of labs' posterior

▸ Back to Simulations ▸ To "True" DGP

- Fit random forests of trial outcomes on characteristics
- Model updated every quarter as newer data incorporated in information set
- Characteristics include every variable the labs mentioned ever using and NIH evaluation metrics to minimize OVB and selection on unobservables
- Each random forest consists of 1,000 decision trees (as in Jahandideh et al. (2014)), the average of the predictions of different trees is the posterior mean, the variance of the predictions of different trees is the posterior variance
- Best-effort replication, not perfect:
  - Labs changed the models and variables used over time, some of which may be uncaptured by the published articles
  - Some variable constructions relied on obsolete software packages
  - Some models predicted different outcomes, e.g. 1-5 scores of likelihood of success

  ...
- My estimate of $\tilde{F}_t$ contains errors (different from the labs' actual posterior beliefs), as long as the errors are not correlated w/ allocation choices, should not bias the results

F_t^* vs $\tilde{F}_t$ ▸ Back to Simulation Procedure

F_t^* almost identical to $\tilde{F}_t$, except

- F_t^* corrects the selection bias of the labs' models in predicting the probability of success by conditioning on a propensity score of observing a specific trial stage. There is no evidence the labs corrected this bias.
  - E.g. 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.
  - Assuming 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 stages as the propensity score of observing the given stage
- F_t^* does not fit/predict on previous outcomes of trials on the project. In simulations, all trial outcomes are simulated, should not shift the "true" probability of success
- F_t^* includes additional variables for better fit: keywords & genes associated w/ the molecule, dummies for different NIH policy phases

UCB+D captures the labs' decisions: other evidence

Back to Main Slides

Estimation:

- UCB+D model has the best fit among many alternative models
- With the same number of params, log likelihood of UCB+D model is 52%—72% of that of the second best fit model across labs
- For trials the labs actually allocated, UCB+D model predicts on average a 70%—84% likelihood of allocating those trials

  For trials the labs did not allocate, UCB+D model predicts on average a <0.5% likelihood of allocating those trials

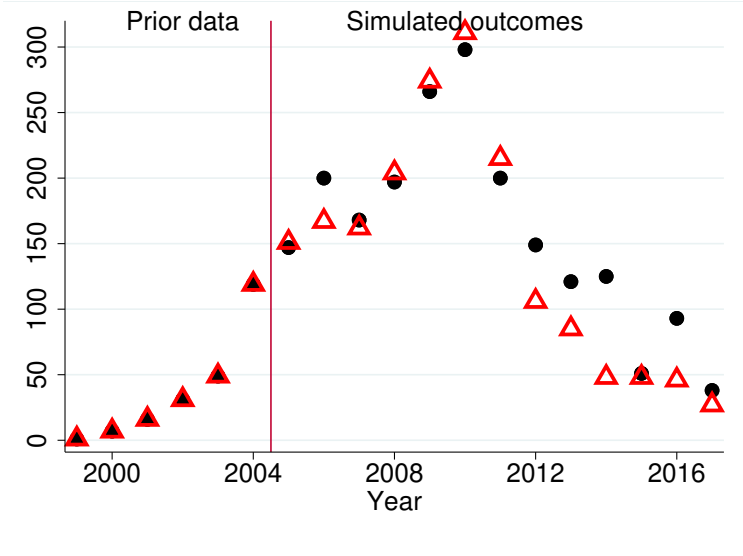
  far better than any alternative model

Model validation simulation:

- Simulated number of projects attempted, output quantity, citations, downloads within 10% different from actual output for all labs

Output by year (UCB+D, another lab)

Back to Main Slides



UCB+D coefficients reject no exploration

Back to Main Slides

- No exploration: $\theta_{B1} = 0$
  Exploration: $\theta_{B1} > 0$
- Lab discounted older projects: $\theta_{B2} < 0$
- Lab preferred biomed projects more since 2009: $\theta_{biomed,before} < \theta_{biomed,after}$

Hard to interpret magnitudes, study effects → counterfactuals

Variable	2005–2008 (1)	2009–2015 (2)
θ_{B1}	158.3 [156.4, 160.1]	119.5 [118.1, 121.4]
θ_{B2}	-2.28 [-2.23, -2.30]	-4.71 [-4.66, -4.73]
θ_{biomed}	21.5 [21.3, 21.7]	52.9 [52.7,52.9]

Results from one large lab Other labs 95% CI based on Chernozhukov & Hong (2003)

UCB+D coefficients reject no exploration

Back to Main Slides

Variable	2005–2008 (1)	2009–2015 (2)
θ_{B1}	273.3 [266.2, 284.8]	127.1 [126.3, 127.8]
θ_{B2}	-3.91 [-3.86, -3.98]	-3.76 [-3.75, -3.76]
θ_{biomed}	12.65 [12.61, 12.72]	84.79 [84.77,84.81]

Results from MCSG. 95% CI based on Chernozhukov & Hong (2003), almost identical to those based on Chen et al. (2018).

UCB+D coefficients reject no exploration

Back to Main Slides

Variable	2005–2008 (1)	2009–2015 (2)
θ_{B1}	61.6 [61.2, 61.9]	115.6 [114.2,116.9]
θ_{B2}	-2.94 [-2.93, -2.96]	-3.93 [-3.91, -3.95]
θ_{biomed}	33.0 [32.9, 33.1]	89.4 [89.1,89.8]

Results from NYSGRC. 95% CI based on Chernozhukov & Hong (2003), almost identical to those based on Chen et al. (2018).

UCB+D coefficients reject no exploration

Back to Main Slides

Variable	2005–2008 (1)	2009–2015 (2)
θ_{B1}	558.2 [551.1,573.4]	1001.9 [977.6,1028.0]
θ_{B2}	-247.9 [-247.6, -248.5]	-104.7 [-102.9, -106.6]
θ_{biomed}	-34.3 [-34.6,-34.1]	105.1 [104.7,105.8]

Results from JCSG. 95% CI based on Chernozhukov & Hong (2003), almost identical to those based on Chen et al. (2018).

Objective function for a simple MAB

Back to Index Solution · Back to Counterfactuals

... where Gittins Index is the optimal solution

$$\max_{\boldsymbol{a}_{l1},\boldsymbol{a}_{l2},\ldots} \sum_{t=1}^{\infty} \beta^t \sum_{j_i \in C_{lt}} \int \pi_{ijt}(\boldsymbol{a}_{lt}, p_{ijt}; \boldsymbol{\theta}_{Xl})\; d\tilde{F}^{(i)}(p_{ijt}|\Omega_t^{(i)}), \text{ subject to } \sum_{j_i \in C_{lt}} a_{ijt} = 1 \text{ for all } t \qquad (9)$$

- Infinite horizon, discounted $\sum_{t=1}^{\infty} \beta^t$ vs $\sum_{t=1}^{T}$
- Stationarity $\tilde{F}$ vs $\tilde{F}_t$
- Independent choices $\tilde{F}^{(i)}(p_{ijt}|\Omega_t^{(i)})$ vs $\tilde{F}_t(p_{ijt}|\Omega_t)$
- One choice in each period $\sum_{j_i \in C_{lt}} a_{ijt} = 1$ vs $\sum_{j_i \in C_{lt}} a_{ijt} = n_{lt}$

Form of $\psi(\cdot)$

Back to Counterfactuals

Computing the exact Gittins Index is hard. Use Brezzi & Lai (2002)'s approximation to the index. The function $\psi(\cdot)$ is defined as

$$\psi(s) = \begin{cases} \sqrt{s/2} & \text{if } s \leq 0.2 \\ 0.49 - 0.11 s^{-1/2} & \text{if } 0.2 < s \leq 1 \\ 0.63 - 0.26 s^{-1/2} & \text{if } 1 < s \leq 5 \\ 0.77 - 0.58 s^{-1/2} & \text{if } 5 < s \leq 15 \\ \{2log(s) - log(log(s)) - log(16\pi)\}^{-1/2} & \text{if } s > 15, \end{cases} \tag{10}$$

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