Matthew Effect, Research Productivity, and the Dynamic Allocation of NIH Grants

Y. J. Jeff Qiu*

November 7 2018

Job Market Paper

Click here for the most recent version

Abstract

Funding is important for research. However, research funding may suffer from the well-known Matthew Effect: the more researchers already have, the more they will be given. This paper develops an empirical framework to study how the National Institutes of Health (NIH) could allocate research funding in a dynamically optimal manner, especially in terms of balancing funds between young and veteran principal investigators (PIs). I first estimate a research production function and show that the dynamic effects of funding via learning-by-doing are of first-order importance. Using these estimates, I then develop a funding allocation model in which the planner (the NIH) maximizes the discounted sum of research output subject to a budget constraint. Because the planner's dynamic programming problem suffers from the curse of dimensionality, I adopt approximate dynamic programming methods from the operations research literature to allow computation. I provide three main results. First, a forward-looking policy with a discount factor of 0.9 funds 30% more young PIs than the myopic policy does, which translates to 5% more research output per year in the long run. Second, the NIH appears to be accounting for some intertemporal tradeoffs, but may still underfund young PIs: the discount factor that rationalizes the NIH's funding behavior is about 0.75. Finally, a temporary funding cut, similar to the one proposed by the current administration, would have a long-lasting effect on overall research output through its adverse impact on investment in young PIs.

*Department of Economics, Yale University. email: yinjia.qiu@yale.edu

I am indebted to my advisors Steve Berry, Penny Goldberg, Phil Haile, Michi Igami, and Olav Sorenson for their continuous guidance and encouragement. I am grateful to Warren Powell for suggestions on the Approximate Dynamic Programming algorithm. I also thank Daniel Ackerberg, Pierre Azoulay, Jin-Wook Chang, Luis Carvalho Monteiro, Giovanni Compiani, Yin Fang, Wayne Gao, Luong Hoang, Jordi Jaumandreu, Donghyuk Kim, Julia Lane, Masayuki Sawada, Suk-Joon Son, Fabian Schrey, Kosuke Uetake, Allen Vong, Bruce Weinberg, Jaya Wen, and Lawson Wong, for their valuable comments. All potential errors are mine.

1 Introduction

Academic scientific research is essential for economic growth, technological progress, and public health. In the United States, academic research in life sciences is heavily reliant on funding from the National Institutes of Health (NIH).[1] This paper studies how the NIH could allocate research funding in a dynamically optimal manner, especially in terms of balancing funds between young and veteran principal investigators (PIs). A myopically optimal allocation rule would fund the current "best" PIs.[2] Such a rule would favor veteran PIs, as they have the strongest track records, both in terms of past publications and past funding, and they have the experience in conducting research and managing labs. While experience is an essential component in the scientific research process, it can only be accumulated by doing research (i.e., through learning-by-doing). However, doing research requires funding. Therefore, underfunding young investigators could stall scientific advancement in the long run, as veterans eventually retire, and unfunded young investigators never got a chance to be veterans. This can be seen as a form of Merton's (1968) "Matthew Effect:" the more the researchers already have, the more they will be given.

Due to the learning-by-doing nature of scientific research production, the funding allocation problem is an inherently dynamic problem. Determining the optimal distribution of funds between novices and veterans is one of the most important questions decision makers face. Recognizing that young investigators are the future of scientific advancement, the NIH has implemented a series of policies to tackle this issue. With a wide range of policies aimed at improving the funding allocation system, it is important to understand how different funding rules affect scientific output and distribution of PIs over time.

This paper develops an empirical model to examine research funding allocation problems, and provide a benchmark for evaluating the effectiveness of different funding policies. I focus the analysis on academic cancer research. The empirical framework consists of two steps: the first step estimates a research production function using data, and the second step uses the estimates from the first step to formulate and solve the dynamic funding allocation problem. I then compare the solution with the current policy.

In the first step, I estimate a research production that takes funding and experience as inputs, recognizing that PIs also have unobserved productivities. I measure research output as a quality-weighted sum of publications, which captures one important feature of research output.[3] In addition, I model experience as a discounted accumulation of funding. As a

[1] The NIH submitted a $35 billion budget in 2016.

[2] In this paper, the unit of the analysis is at the research lab level, which is equivalent to the level of PIs, because a PI owns a research lab. Therefore, I use PI, lab, and investigator interchangeably throughout the paper.

[3] This measure neglects other interesting output measures such as the innovativeness of research; hence,

result, funding enters research production not only directly as an input but also indirectly via the experience channel.[4] Parameters in the production function, such as the elasticity of output with respect to funding, are difficult to estimate due to endogeneity — variable inputs are correlated with the productivity unobserved by the econometrician.[5] To address this endogeneity issue, I use instrumental variables motivated by variation in budgets set by the Congress. This variation includes the effect of policy stimulus, such as the American Recovery and Reinvestment Act in 2008 to 2010, and a series of policies implemented to support early-career investigators.

While variations in budget can be used as instruments, they may or may not be enough in their own for identification since inputs to the production function include experience, which is also endogenous. Thus, in addition to the instrumental variables, I impose an autoregressive structure on the unobserved productivity process, and employ techniques from the production function literature, as in Olley and Pakes (1996), Blundell and Bond (2000), and Ackerberg et al. (2015), to identify and estimate the model. The methods of these papers are often employed in the industrial organization and international trade literatures. Specifically, I assume that each period's innovation to a PI's unobserved productivity is independent of the PI's past inputs. Past observables of the PI can then be used as instruments for identification and estimation.

I construct a new PI-year level panel dataset from the detailed NIH cancer grant archive, which consists of over 4000 cancer research PIs from 2000 to 2014. I estimate the model using moments constructed from instrumental variables and orthogonality conditions implied by the dynamic panel structure. The model yields over-identifying restrictions, which allow me to test the validity of the moment conditions implied by my assumptions. The estimates of the production function imply that the elasticity of output with respect to funding is 0.21, output increases by 18% when experience doubles, experience depreciates at a rate of 40% per year, and the unobserved productivity at the PI level is persistent. These estimates are statistically significant, and the model passes the over-identification test. These estimates reveal that funding increases not only current research output, but also future research output, through the experience channel.

In the second step, I utilize the estimated production function to demonstrate the importance of dynamic funding policies through counterfactual simulation. I develop a funding allocation model in which the planner (or the NIH) maximizes the discounted sum of research output by choosing an allocation from a pool of investigators with different characteristics

analyses that examine the effect of funding allocation on the innovativeness of science is outside the scope of this paper. Hereafter, I use the term research output to refer to a quality weighted sum of publications.

[4]The research production function is similar to Griliches's (1979) Knowledge Capital model.

[5]Productivity in this paper refers to the Hicks-neutral total factor productivity (TFP).

to fund, subject to a budget constraint. Due to the dynamic impacts of funding, the planner needs to consider the impact of funding on experience and the persistence of productivity when deciding whom to fund. Therefore, the funder needs to choose a funding allocation that balances the marginal return in current research output with that of future research output.

I show that the funding allocation problem can be formulated as a single-agent dynamic programming problem where the state vector is the PI-population distribution by their characteristics. However, the funding allocation model is difficult to solve due to the three curses of dimensionality — high-dimensional state space, action space, and probability transition of state variables. My benchmark model has over 1000 state variables representing PI characteristics, so standard techniques in the dynamic programming literature, such as brute-force value function iteration, are infeasible. I show that the funding allocation problem can be formulated as a dynamic resource allocation model that is well-known in the operations research literature (e.g. Topaloglu and Powell (2006)). Since dynamic resource allocation models have a concave structure in their value functions, I can use a range of techniques from the approximate dynamic programming literature to overcome the curses of dimensionality (e.g. Powell (2011)).[6] Specifically, I use a concave piecewise-linear functional form to approximate the value function and provide a feasible algorithm. This algorithm is similar to policy-iteration algorithms and can be used to compute the solution even with a high-dimensional state space.

I use the funding allocation model to evaluate counterfactual analyses of different funding policies. The analyses provide three main results. First, there are substantial tradeoffs between forward-looking policies and the myopic policy. A forward looking policy with a discount factor of 0.9 produces 1% less current research output in the first four years but 5% more research output per year thereafter as compared to the myopic policy. This difference is driven by the investment in productive novice investigators under the forward-looking policy — the forward-looking policy funds 30% more young PIs than the myopic policy does. These additional PIs eventually build up their experience and produce higher research output in the future. As a result, there are more productive PIs in the system on average, which translate to more research output in the long run.

Second, the actual funding behavior of the NIH appears to account for some intertemporal tradeoffs. The current NIH policy funds 18% more young PIs than the myopic policy, and the discount factor that rationalizes the NIH's funding behavior is about 0.75. The NIH, however, may still be underfunding novice PIs, as it is funding 8% fewer young PIs as

[6]This method is also known as neurodynamic programming and reinforcement learning in the literature. See e.g. Bertsekas and Tsitsiklis (1996) and Powell (2010).

compared to an optimal forward-looking policy with a discount factor of 0.9. Nevertheless, the observed funding behavior appears to be in line with the NIH's mission to fund the younger generation.

Finally, I employ the model to study the impact of a temporary funding reduction — a policy similar to the one proposed by the current administration. I simulate a stylized budget cut policy in which the funding budget is reduced by 20% for the first four years. The results show that the aggregate research output drops by 10% during the four-year budget cut period. The reduction in output remains significant even fifteen years after a return to the baseline level of funding. Furthermore, such a budget cut would mostly affect young and mid-career PIs. This analysis suggests that a budget cut, even if a temporary one, can lead to significant long-lasting reductions in aggregate research output.

1.1 Related Literature

This paper contributes to the literature on computational economics, the economics of science, the impact of funding on research, production function estimation, and funding allocation.

This paper uses methods developed in computational economics and operations research. Dynamic programming problems are ubiquitous in economics and operations research, but the curse of dimensionality due to large state space remains a bottleneck for many applications. A series of papers in the computational economics literature using value function approximation to solve high-dimensional dynamic programming problems include Keane and Wolpin (1994), Rust (1997), Benitez-Silva et al. (2000), Pakes and McGuire (2001), Crawford and Shum (2005), and Brumm and Scheidegger (2017). While these papers make advancements on the high-dimensional state space problem, their applications are to models with low-dimensional action spaces. I contribute to the literature by showing that methods in the operations research and approximate dynamic programming literature (e.g. Powell (2011)) can be useful to tackle problems with both high-dimensional state spaces and action spaces.

The empirical methods in this paper contribute to the estimation of the impact of funding on research output. Earlier works, including Adams and Griliches (1998), are mostly descriptive. Using NIH administrative data, Jacob and Lefgren (2011a; 2011b) focus on a regression discontinuity design to identify the impact of funding on future output. Their (2011a) paper finds a limited impact of R01 grants on research output for the group of investigators in the vicinity of the funding cutoff; their (2011b) paper finds a significant impact of the NIH post-doctoral training grant on research output. Recent work utilizes the instrumental variable approach. Tabakovic and Wollmann (2016) use unexpected NCAA tournament results as an instrument for funding, and finds the elasticity of output with respect to funding to be 0.2. Azoulay et al. (2017) exploit the NIH funding rules for investigation on different sciences and diseases together with variation in the quality of applications to construct an instrument, and find a large impact of funding on patent outcome. I contribute to this literature by using frontier methods from the production function literature (e.g. Olley and Pakes (1996), Blundell and Bond (2000), Ackerberg et al. (2015), Ackerberg (2016), and Ackerberg and Hahn (2015)) for identification and estimation. The results in this paper complement the existing work that utilizes the instrumental variable approach.

In addition, the research production function model developed in this paper explicitly incorporates funding dynamics into research output, where as the previous literature has neglected them. My work shows that the dynamic effect is of first-order importance. The present production model is related to the research production function literature for cases where dynamic inputs are important. Related work include Griliches (1979), Benkard (2000), and Doraszelski and Jaumandreu (2013).

This paper also contributes to the literature on science policy. The aging of the scientific workforce has been a concern in the scientific community (e.g. Freeman and Van Reenen (2008), Jones (2010), Jones and Weinberg (2011), Azoulay et al. (2011), and Levin and Stephan (1991)).[7] The empirical framework developed in this paper allows characterization of the optimal share of funds to young PIs when the funder maximize the discount sum of research output. The funding allocation model developed here therefore provides empirical assessment of a version of the classic exploration versus exploitation tradeoff in the organizational learning literature (e.g. March (1991)). Thus, the general framework developed in this paper can also be applied to other settings.

1.2 Overview

This paper is organized as follows. Section 2 discusses the setup of the institution and data construction. Section 3 presents the dynamic research allocation model. Section 4 presents the research production function, identification and estimation strategies, as well as their results. Sections 5 illustrates the computational details. Sections 6 shows results from counterfactual funding allocation exercises. Section 7 concludes.

[7]See also Fortin and Currie (2013), Fang and Casadevall (2016), and Cook et al. (2015) for recent debates on how to allocate research funds.

2 The NIH Funding System and Data

The present paper makes use of detailed grant level data from the NIH. This dataset includes information of grants from 1985 to the present, and contains funding amount, the principal investigator's information, the review section to which the grant was sent, budget period, abstract, etc. Only successful applicants are observed in this dataset. Grants are divided into different activities. For example, the most common grant is R01, which is awarded to PIs for conducting research projects in an area that represents their specific interest. It is well known that most life science labs depend on R01 grants to stay active in academic research Stephan (2012).[8]

In life sciences, an academic research lab is managed by a PI, typically a professor at a university, who receives research funding from governmental agencies such as the NIH for operation. The funding received by PIs is spent on acquiring input resources such as capital and labor (postdocs, graduate students, and other scientists), which PIs employ to produce research output (publications, patents, clinical studies, etc). Although we can in principle study all disciplines in the life sciences, the current paper focuses on one particular biological science: academic cancer research. I follow approximately 4000 cancer research PIs identified by recipients of R01 grants from the National Cancer Institute (NCI),[9] from 2000 to 2014.[10]

I focus only on established academic cancer research PIs, and it is assumed that the PIs must be recipients of NCI R01 grants prior to production and that the PIs maintain their NIH grant recipient status during the period of production.[11] Academic life scientists in the US will need to win an R01 grant in order to carry out their own research agenda.[12] That is, upon receiving an R01 grant, PIs act similarly to a CEO of a firm, managing resources (funding) and giving directions to lab members to produce research output. By focusing on PIs, we effectively treat each lab as an independent firm that takes in variable inputs to produce research output.

There are several reasons I have chosen to base my analysis on cancer research. First, by focusing on one particular science, the analysis is more tractable, and comparisons between different PIs are more valid, as different types of life sciences may have different research production processes. My choice thereby mitigates potential confounding of two different production functions.[13] Second, NIH funding accounts for the major share of funding to can-

[8]See Stephan (1996) and Azoulay et al. (2011) for in-depth surveys of the funding agency.

[9]The budget for NCI was $5 billion in 2016.

[10]I collect all grants that these cancer research PIs received from the NIH in a given year and aggregate the funding amount to the lab-year level. These include grants received from different institutional centers.

[11]Thus, in effect, only the most relevant cancer research PIs are considered in my sample.

[12]See Stephan (2012).

[13]This, of course, assumes that cancer research PIs with different disease specializations have the same

cer research PIs. Because of the prominence of funding from the NIH, it is approximately valid for me to assume that my sample of PIs is representative of all PIs in this sector. By extension, I assume that the NIH is the sole source of funding, and that PIs are unable to operate without its funding.[14] Furthermore, R01 grant applications are reviewed at the institutional center level, and NCI is the largest institute in NIH. Finally, cancer is one of the deadliest diseases that threaten public health. Therefore, it is important to understanding the academic cancer research production function and how to allocate funding on this discipline.[15]

While there are many types of research grants in the NIH, the present paper focuses on R01 cancer research grant funding decisions. There are several reasons to base the analysis on R01 grants. First, R01 grants are the most important and established grants in the life sciences. They account for more than 40% of the entire NIH budget, and the median R01 grant provides about $300,000 a year to a PI for three to five years, and accounts for about 65% (median) of the annual funding of a research lab (Stephan (2012)). Second, the NIH designates budgets for different types of grants, and grant allocation decisions are at the grant type level. For example, P01 grant applicants do not compete with R01 applicants.[16] Third, R01 grants are unsolicited grants, and there is less of a strategic component involved. Fourth, almost all of R01 grants are awarded to single PIs. Hereafter, I refer to a grant as an R01 grant.

A typical grant lasts for four to five years and allocates funds to PIs in 12-month budget periods. Hence, it is reasonable to define the unit of observation at the PI-year level. I construct the funding amount F_{it} of PI i at time t by aggregating the budget allocated to PI i in year t.[17] The NIH also contains data on all publications that cited the grant number,

production function. We could relax this assumption by disaggregating the data to the PI-disease-year level, and estimate the production function for different diseases separately. However, this disaggregation opens up an avenue of problems. First, it is difficult to associate PIs with certain disease specialities. Second, even if we are able to identify PIs with certain disease specializations, since most PIs have multiple and different specializations over time, we need further assumptions to associate these PIs with the proper disease speciality. Third, there are many diseases within the cancer category, so we face the additional problem of the choice of granularity of disease aggregation. Finally, the identification strategy utilized in this paper crucially relies on the panel structure, which is data demanding. Therefore, a finer level of aggregation implies that less data is available.

[14] This neglects funding from the private sector and grants from the Howard Hughes Medical Institute (HHMI). See Azoulay, Graff Zivin, and Manso (2011) for the difference between NIH and HHMI grants.

[15] Cancer has been the focus of several earlier studies in Economics because of its importance in research and health policy (see Budish, Roin, and Williams (2015) for example).

[16] P01 is another NIH grant, and it is typically awarded to a group of PIs. Therefore, a model for P01 funding allocation would need to account for team composition of PIs.

[17] In general, grant level data display the total cost of each grant, which contains direct and indirect costs. Indirect costs refer to funding amounts used to pay the grant recipient, i.e., overhead of the home institution of the PI. Direct costs refer to the actual amount of funding received by the PI. The funding variable F_{it} defined here represents the total cost. This implicitly assumes a constant proportion of overhead paid by

and I measure the research output of a PI at time t by a measure that represents the quality and contribution of publications written by the PI in year $t+1$ that cite the grant, which is calculated as a weighted sum.[18] Formally, I define the output of i at time t as

$$Y_{it} = \sum_{p \in \mathcal{P}_{i,t+1}} \frac{\varphi_p}{\#(p)}, \tag{1}$$

where $\mathcal{P}_{i,t+1}$ denotes the set of papers published by PI i in year $t+1$, φ_p is paper p's research output as measured by its quality, and $\#(p)$ denotes the number of PIs involved in paper p. The division of φ_p by $\#(p)$ captures the research output proration, which is often used to address publications with multiple coauthors, and is common in the economics of science literature. In contrast to the measure used in the literature, a slight nuance in (1) is that the paper output is prorated by the number of PIs rather than the number of coauthors, because our analysis is at the PI level.[19]

If $\varphi_p = 1$ for all p, then Y_{it} is the unweighted measure of research output, which simply aggregates the number of papers the PI published. This unweighted measure only captures the quantity of publication and ignores the quality of the papers. An alternative definition of output would be to measure φ_p by the number of citations paper p has. However, citations of publications take time, which may contaminate the more recent data, as we only have data from 2000 to 2014. This is the right-truncation problem in the literature. In this paper, I employ the impact factor of the journal in which paper p is published as a measure of φ_p, such that the resulting weighted output measure would not suffer from inconsistencies of quality measures over time. This measure, although unable to distinguish between the quality of two papers published in the same journal, is adequate in capturing quality differentiations across different journals, and avoiding the right-truncation problem. This is especially important for the dynamic panel setup.[20]

the PIs to different institutions.

[18]Life science publications are usually much shorter than social science ones; also, the review time of publications is generally much shorter for life sciences than for social sciences.

[19]This type of measurement is commonly used in the economics of science literature and the patent and innovation literature.

[20]Alternatively, we can measure output such as the number of patents that have cited the papers associated with a particular NIH grant. This measure is proposed and carefully constructed in Azoulay et al (2017). However, it takes a longer time to include all patents that cite publications associated with a grant. For instance, about 40% of NIH grants awarded in $1991-1995$ are associated with at least one patent 10 years after their approval (Azoulay et al (2017)). Using this patent output measure would potentially yield many zero outputs in the PI-year panel structure adopted in the present paper; therefore, this measure is unsuitable for the Cobb-Douglas production function model considered in the present paper.

3 The Allocation Problem

3.1 The Planner's Problem

I formulate the funding allocation problem of the planner (the NIH) as a single-agent dynamic programming (DP) problem. Time is discrete with finite horizon $t \in \mathcal{T} = \{1, 2, ..., T\}$. For this application, a period is assumed to be a year, and $T = 30$, which is the average life-cycle of a PI.[21] The planner receives a budget B_t and a pool of applicants $\mathcal{N}_t$ at the beginning of year t, and makes funding allocation decisions regarding $\mathcal{N}_t$.[22] Recall that I focus on R01 grant decisions, so that other types of NIH grants are treated together as a random variable.[23]

Each PI $i \in \mathcal{N}_t$ has five state variables: experience level E_{it}, total factor productivity (TFP) level ω_{it}, current Non-R01 funding amount F_{it}^{NonR01}, proposed R01 funding amount F_{it}^{R01}, and the associated R01 grant cycle C_{it}, which I denote by a characteristic vector $X_{it} = [E_{it}, \omega_{it}, F_{it}^{\text{NonR01}}, F_{it}^{\text{R01}}, C_{it}]'$. The experience state is the key dynamic to the problem, and it evolves endogenously as the planner's funding decision. In practice, a typical R01 grant provides funding for a PI annually for 4 to 5 years. A grant cycle state variable is defined as the grant cycle status when a PI applies for a grant, that is, the grant cycle status prior funding decision is made. TFP ω_{it} is another key state variable, it evolves stochastically according to an exogenous Markov process.

Since this paper does not explicitly model Non-R01 funding decisions, I treat the Non-R01 funding variable as a random variable (a state variable), and formulate the state transition as follows:

$$F_{it}^{\text{NonR01}} \sim \mathbb{P}[.|X_{i,t-1}], \tag{2}$$

where F_{it}^{NonR01} is drawn from the conditional distribution $\mathbb{P}[\cdot|X_{i,t-1}]$, which can be estimated from data.

I assume that the planner maximizes the discounted sum of research output subject to a budget constraint. Therefore, the planner's decision is to decide whether to fund each applicant in the set $\mathcal{N}_t$. Importantly, the NIH does not choose how much to fund, but whether to fund the PI. Equivalently, the control variables of the planner's allocation problem can be formulated as to choose a subset $\mathcal{N}_t^1 \subset \mathcal{N}_t$ to fund. The set of PIs that are funded, $\mathcal{N}_t^1$,

[21] I formulate the problem as a finite horizon problem where T is large, and this formulation is an approximation of the infinite horizon problem in a steady-state environment. See Chapter 10.10 in Powell (2011) for a detailed discussion on using finite horizon approximations for steady-state applications. The results are similar with $T = 25$ and $T = 30$.

[22] I assume that the NIH cannot save, and it will spend all B_t at year t. Furthermore, residual funding, if there are any, cannot carry over to the subsequent year.

[23] Modeling the funding decisions of all NIH grants is complex, and is beyond the scope of this paper.

proceeds to year $t+1$. Furthermore, due to the presence of the grant cycle, PIs that are not funded but are not at the terminal grant cycle, denoted by the set $\mathcal{N}_t^0 \subset \mathcal{N}_t \setminus \mathcal{N}_t^1$, also proceed to year $t+1$. Formally, define $\hat{\mathcal{N}}_t := \mathcal{N}_t^0 \cup \mathcal{N}_t^1$ — PIs in $\hat{\mathcal{N}}_t$ proceed to the subsequent year $t+1$ with their characteristic vector transitioning from $\mathbf{X}_t$ to $\mathbf{X}_{t+1}$ based on a known transitional probability $\mathbb{P}(\cdot\,|X = X_{it})$. I refer to the set $\hat{\mathcal{N}}_t$ in year $t+1$ as the incumbent set, and it is denoted by $\mathcal{N}_{t+1}^{\text{incumbent}}$. In addition, there is a set of entrant PIs $\mathcal{N}_{t+1}^{\text{entrant}}$ that apply for funding with characteristic vector $\mathbf{X}$ drawn from a known distribution at year $t+1$. Therefore, the set of available PIs at period $t+1$ is $\mathcal{N}_{t+1} := \mathcal{N}_{t+1}^{\text{incumbent}} \cup \mathcal{N}_{t+1}^{\text{entrant}}$. Each PI $i \in \hat{\mathcal{N}}_t$ produces output Y_{it} at year t according to the

$$Y_{it} = h(F_{it}^{\text{NonR01}}, F_{it}^{\text{R01}}, E_{it}, \omega_{it}), \tag{3}$$

where $h(\cdot)$ is a production function that needs to be estimated.[24] The production function (3) is assumed to be known to the NIH. Figure 1 displays the timeline of the funding allocation model.

The DP model as outlined above can be formulated as the following Bellman equations, for all $t \in \mathcal{T} = \{1, 2, ..., T\}$

$$V_t(\mathcal{N}_t, B_t) \quad = \quad \max_{\hat{\mathcal{N}}_t \subset \mathcal{N}_t} \sum_{i \in \hat{\mathcal{N}}_t} Y_{it} + \beta \mathbb{E}\left[V_{t+1}(\mathcal{N}_{t+1}, B_{t+1}) | \hat{\mathcal{N}}_t, B_t\right] \tag{4}$$

$$\text{such that} \quad \sum_{i \in \mathcal{N}_t^1} F_{it}^{\text{R01}} \leq B_t, \tag{5}$$

where β is the discount factor, $\mathbb{E}\left[V_{t+1}(\mathcal{N}_{t+1}, B_{t+1}) | \hat{\mathcal{N}}_t, B_t\right]$ is the expected value of the NIH conditional on the allocation today, and (5) is the budget constraint.

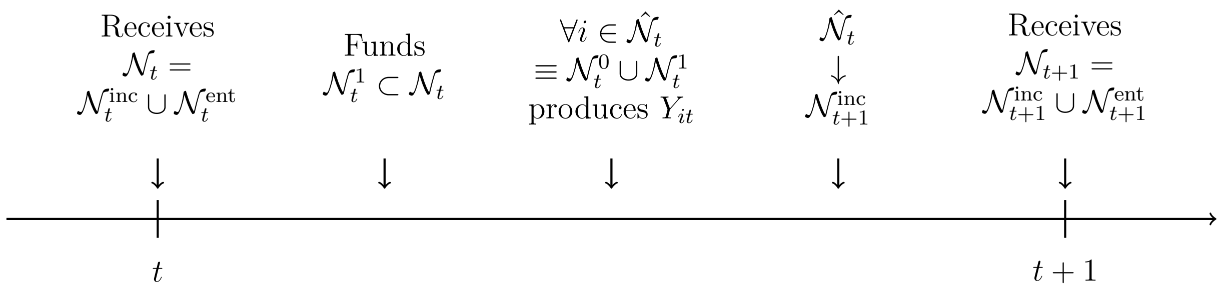
3.2 The Discount Factor

The discount factor β plays a crucial role in determining the funding allocation decisions to $\mathcal{N}_t$. One interpretation of the discount factor is, by convention, the time preference of the NIH. On one side of the spectrum, when β is relatively small, NIH puts a larger weight on current research outputs. In contrast, when β is large, the weights between current and future outputs are relatively balanced. Different discount factors would fund drastically different sets of investigators, and hence, different aggregate research output in the short-term and long-term.

For example, $\beta = 0$ corresponds to a myopic policy, which I will refer to as "the myopic policy" hereafter. In this case, the policy assigns no weight on continuation research output,

[24] See Section 4.1 for a detailed discussion of identification and estimation of the production function.

Figure 1: Timeline of the Funding Allocation Model



Notes: $\mathcal{N}_t^{inc}$ stands for a set of incumbents at year t; $\mathcal{N}_t^{ent}$ stands for a set of entrants at year t. $\mathcal{N}_t^1$ stands for a set of funded PIs; $\mathcal{N}_t^0$ stands for a set of unfunded PIs but have active grants. This timeline depicts the NIH's allocation problem. The planner (or the NIH) receives a set of applicants $\mathcal{N}_t$ which consist of a combination of incumbents and entrants, at the beginning of year t. The NIH then selects a set $\mathcal{N}_t^1$ to fund. Those who get funded or have active grants produce output, and proceed to the subsequent year $t+1$ as incumbents. The process continues.

and the optimal allocation is to fund labs that produces the most output for that given period. These are the labs with the highest marginal product of funding. Recall that the marginal product of funding for production function (3) is a function of both experience and productivity. Since this policy neglects the dynamic impact of funding through the experience channel, it tends to favor PIs with more experience, because high levels of experience can compensate for the low levels of TFP.

On the other hand, when $\beta \in (0, 1)$, NIH would make funding decisions in a way that accounts for future research output. In this case, since productivity is persistent (as shown in Section 4), the NIH would also evaluate the importance of funding on future research output through the experience channel. More precisely, in addition to how much the funded PIs can produce in the current period, the NIH would also assess how much the funding in the current period could help the funded PIs to climb the experience ladder, and thereby assessing how much the expected research output these PIs can produce in the future. As a result, forward looking policies may fund PIs that have lower marginal product of funding in the current period but high expected research outputs in the future. These policies would give higher weights to productive young investigators, where the weights depend on the value of the discount factor.

4 Estimation of Funding Allocation Model Primitives

4.1 Production Function

A PI's production function is assumed to follow the Cobb-Douglass form, and takes funding and experience as variable inputs to produce research output:

$$\ln Y_{it} = \alpha_0 + \alpha_1 \ln F_{it} + \alpha_2 \ln E_{it} + \alpha_3' Z_{it} + \omega_{it} + \varepsilon_{it}, \qquad (6)$$

where the subscript it indicates PI i at year t, output Y_{it} is defined by (1), F_{it} and E_{it} are funding and experience variable inputs for the PI respectively, Z_{it} is a vector of the PI's observed characteristics such as a time trend and school effects, ω_{it} is the total factor productivity (TFP) of the PI, which is observed by the PI and the NIH but not observed by the econometrician, and ε_{it} is the measurement error of output.[25] The key parameters of interest are α_1, the elasticity of output with respect to funding, and α_2, the elasticity of output with respect to experience, which describes the effect of learning-by-doing.[26] Note that, although (4) distinguishes F_{it}^{R01} and F_{it}^{NonR01}, the production function (6) treats funding PI i's funding at year t as a sum of R01 funding and non-R01 funding. That is, $F_{it} = F_{it}^{\text{R01}} + F_{it}^{\text{NonR01}}$.

There are a few differences between (6) and the standard setup of production function in the production function literature. First, I do not include labor input because funding input already encapsulates it. Possible measures of labor input may include the number of postdocs, Ph.D. students, undergraduate students, and technicians in the research lab. Generally, we do not have this information in the NIH dataset. In academic science in the United States, PIs need to provide financial support for postdocs, Ph.D. students, and other labor input using their grants. This practice may be very different for academic science in Asia, for example, where postdocs salaries and Ph.D. student stipends are paid by the PIs' home institutions in some cases. However, given that we only focus on NIH grant recipients, this assumption is innocuous. Second, we have experience, E_{it}, as a variable input. It can be viewed as a state variable that determines research output. The effect of this variable is the key dynamic channel in research production. It is included to capture the important

[25] ε_{it} is included to acknowledge the fact that output defined by (1) is susceptible to large measurement error.

[26] Learning rate refers to percentage increase in research output when experience doubles, or $(2^{\alpha_2} - 1) \times 100\%$. PIs that do not produce any papers over the entire sample period are dropped. Observations with no NIH funding are dropped. The output measure, Y_{it}, is transformed by incrementing it by one in order to allow for PIs that produce zero output for certain years. This transformation, used by many researchers (e.g., Tabakovic and Wollmann (2016)), is a normalization allowing for minimum output for PIs to remain in the sample.

feature of research, especially in life science, that research output is a function of extensive trial and error.

4.1.1 Experience

I model the experience component of (6) similar to the knowledge models in Griliches (1979) and Levin and Stephan (1991), and the learning models in the learning-by-doing and organizational forgetting literature, e.g., Benkard (2000).[27] The key feature of experience is that it captures the accumulation of relevant research capital over time. I model the experience dynamic as follows:

$$E_{it} = \delta E_{i,t-1} + R_{i,t-1}, \text{ with } E_{i\tau_i} = \theta_i, \tag{7}$$

where $R_{i,t-1}$ is a measure of research capital of PI i at time t, τ_i indicates the entry year of PI i, θ_i is the initial condition, and δ is the rate of experience depreciation (or the forgetting parameter). In this model, a PI can improve its experience only by acquiring research capital, and experience depreciates at a constant rate $(1-\delta)$. The constant rate of depreciation is specified in an attempt to capture worker turnover and depreciation of relevant knowledge, etc. The model (7) refers to the experience process in a general learning model. Note that δ is unknown, and it will be estimated.

In the present paper, a PI is modeled as a firm, and the experience process refers to research capital accumulation of a lab (e.g. Griliches (1979)). The experience process (7) encapsulates both scientific knowledge of the PI and research experiences of lab members. As a result, the depreciation rate $(1-\delta)$ in (7) refers to depreciation of relevant knowledge of the PI as well as employment effects on a micro-level such as lab member turnover and training of new lab members. Since scientific research is labor intensive and turnover rate is high, we can expect a higher depreciation rate for model (7) than for knowledge processes in the literature.

In this paper, I measure $R_{i,t-1}$ using $F_{i,t-1}$, the funding amount i received in year $t-1$. In this specification, experience improves by how much funding the PI received in the previous year. When $\theta_i = 0$, $E_{it} = \sum_{j=1}^{t-1} \delta^{j-1} F_{i,t-j}$, which is a measure of i's discounted accumulation of funding. In general, $R_{i,t-1}$ could be measured as the number of grants won, number of papers published, etc.

Although similar, the research production model studied in the present paper, described by (6) and (7), is conceptually different from the learning model in the learning-by-doing literature. In the learning models studied in the literature (e.g. Benkard (2000)), experience

[27] Knowledge plays the same role as experience in the production process, and I use the experience model terminology throughout this paper. I use the terms experience and knowledge interchangeably.

refers to discounted cumulative past output. Therefore, in these models, experience only accumulates if the firm produces output. Scientific research production, however, is a function of trial and error. Scientists typically only publish successful research, and thus failed experiments are often left unacknowledged, even though failed experiments are a crucial component of and pathway to experience accumulation.

For instance, suppose we take some measure of research output, say the number of publications, as a measure of research capital in (7), then the contribution of failed experiments to the experience dynamic is immediately omitted, despite its fundamental importance to scientific research. Therefore, research capital measured by output would substantially understate the actual experience.

This paper focuses on the discounted accumulation of funding as a measure of experience as the benchmark model for two reasons.[28] First, research capital measured by funding can better capture the trial and error aspect of science experience dynamic. Second, one drawback of the output definition (1) is that it may be sensitive to different time lag definitions. For example, (1) with production function (6) assumes away the impact of funding at $t-2$ on research output at t. Experience dynamic measured as discounted cumulative funding, however, mitigates this problem because past funding is also an input to current research output. Hence, past funding has an indirect effect on current research output.[29] Nevertheless, experience as measured by the cumulative number of publications is considered for comparison purposes.

4.1.2 Identification

Identification and estimation of the model (6) suffers from common problems associated with estimation of production function. Namely, variable inputs are determined with knowledge of ω_{it}, which is unobserved by the econometrician. Therefore, F_{it} and E_{it} are correlated with the unobserved variable ω_{it}.

Under specification (6), shocks to funding can be good instruments for this model.[30] In this paper, I use variation in the aggregate NIH budget over time as an instrument for the funding variable in (6). The exogenous variation of this instrument arise from various

[28] This measure omits the contribution of failed grant applications to experience. However, we are interested in the effect of experience on research output and not the effect of experience on writing grant proposal. Therefore, the contribution of failed grant applications may not understate the actual experience for research production.

[29] Using funding as a measure of research capital, the present paper uses several measures of θ_i. The default measure employs the first year R01 funding as a measure of θ_i. The main results are robust to other measures of θ_i, including first year NIH funding, medium NIH funding, and a constant normalization.

[30] Recent works that estimate production functions with instrumental variables include Benkard (2000), Bloom et al. (2013), Tabakovic and Wollmann (2016), and Azoulay et al. (2017).

changes in budgets set by the Congress. There have been three key changes in the NIH budget time series over the past twenty years. First, the NIH budget doubled from 1998 to 2003 - the nominal budget rose from $25.7. billion in 1998 to $43.2 billion in 2003 [31]. This was a result of efforts made by the Clinton Administration and the Bush Administration to further scientific advancement (Macilwain (1998) and Kaiser (2002)). Second, there is a steady decline in the NIH budget (in real term) after the budget doubling. Third, the fiscal stimulus package - American Recovery and Reinvestment Act (ARRA), implemented in 2008 in response to the financial crisis - resulted in another exogenous increase to the NIH budget. The nominal budget increase from $29.6 billion to $35.5 billion in 2008 and $30.9 to $36.7 billion in 2009. Importantly, the changes in budget were determined by the Congress. Therefore, these changes are unlikely to be associated with the unobserved TFP of PIs.

Furthermore, there were policies aimed at supporting researchers at different career stages. Recognizing that the scientific workforce is aging, and the lack of junior investigators may be detrimental for scientific advancement, the former NIH director Elias Zerhouni implemented a series of polices to support younger researchers. Some of these measures included creating special grants only awarded to young researchers, and giving priorities to early-career researchers during peer reviewed panel (Zerhouni (2006)). These policies have helped maintained a more balanced growth for early-career investigators, but it have been shown to affect the mid-career investigators unfavorably (Collins (2017)). These changes in this period of time can be interpreted as exogenous budget changes at different career stages due to the taste of the policy maker. Therefore, it is unlikely to be correlated with the unobserved TFP. Combining these observations, I construct a budget instrument using percentage change in aggregate budget at different career-stages. The key idea is that variation of budget change is set by the congress, and there are cross-sectional variation at the career stage level due to taste of the NIH director.

The exogenous budget variation provides one instrumental variable for the model. However, it under-identifies the production function (6) because the model has more endogenous variables than the number of instruments.[32] To identify the model, therefore, it is necessary to impose additional restrictions on the unobserved TFP structure.

Following the frontier methods in the literature (e.g. Olley and Pakes (1996), Ackerberg et al. (2015)), I impose a structure along with timing assumptions on the unobserved TFP ω_{it} to tackle the endogeneity problem. Precisely, I assume that the TFP evolves over time

[31] See the annual report by Johnson (2018).

[32] While we can use the lagged budget changes as additional instruments, the estimation results suggest that these are rather weak, i.e. , standard errors are high.

according to an exogenous first-order Markov process

$$\omega_{it} = \Psi(\omega_{i,t-1}) + \xi_{it}, \tag{8}$$

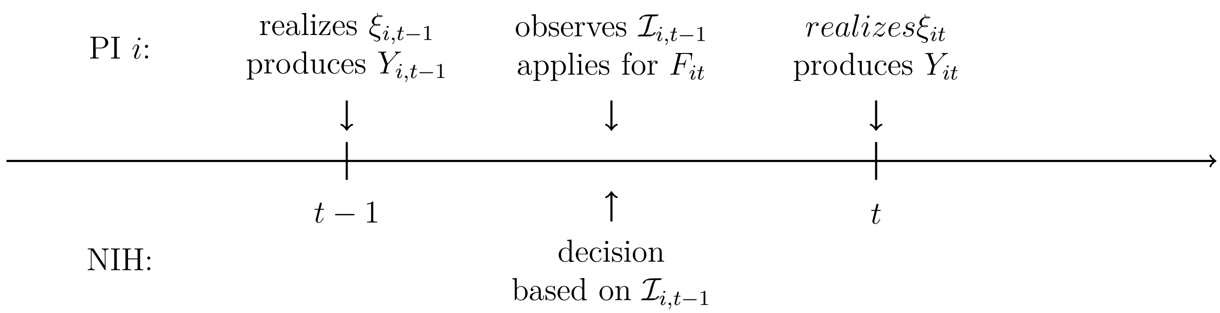
where $\Psi(\cdot)$ is a nonparametric function, and ξ_{it} is an innovation shock to the TFP centered at the origin. ξ_{it} is assumed to not be forecastable with the orthogonality condition

$$\xi_{it} \perp \mathcal{I}_{i,t-1} \tag{9}$$

where $\mathcal{I}_{i,t-1}$ is the information set of PI i up to year $t-1$, a set of characteristics observable to the NIH and the PI.

The timing assumptions on the production process implied by (8) and (9) are the following: PIs observe $\mathcal{I}_{i,t-1}$ and apply for research funding with proposed amount F_{it} at the end of year $t-1$, then the NIH observes $\mathcal{I}_{i,t-1}$ for all $i \in \mathcal{N}_t$ where $\mathcal{N}_t$ is a set of all PIs applying for the grant, and makes decisions about whether to fund each PIs. After the funding decision, PIs move on to year t and produce research output according to technology (6).

Figure 2: Timing Assumption: Strong



Notes: This stronger timing structure assumes that the funding decision at year t is determined at year $t-1$, and the innovation shock is realized after the funding decision. Since ξ_{it} is not forecastable, it is independent of the variable F_{it}. Therefore, $\mathcal{I}_{i,t-1}$ includes F_{it}. This structure allows the funding decision at period t to depend on ω_{it} through the knowledge of $\omega_{i,t-1}$.

The strength of the timing assumptions depend on the time of the realization of the innovation shock ξ_{it}. On the one hand, determination of funding F_{it} can be thought of as determination of capital under the traditional production function framework. Figure 2 depicts such timing structure. Specifically, this timing structure assumes that funding decision at year t is determined at year $t-1$, and that the innovation shock is realized after the funding decision. Since ξ_{it} cannot be predicted, it is independent of the variable F_{it}. Therefore, $\mathcal{I}_{i,t-1}$ includes F_{it}. This structure allows the funding decision at year t to depend on ω_{it} through the knowledge of $\omega_{i,t-1}$.[33]

This timing assumptions and the TFP structure (8) and (9) serve as an important identification condition in the literature.[34] This identification condition relies on the knowledge of the innovation shock ξ_{it} at the time decisions are made. It is common that uncertainties enter the research production. For example, the success of novel research ideas can go in either direction (Stephan (1996; 2012)), implying that the decision makers do not have perfect information of ω_{it} at $t-1$. This assumption implies that the best predictor for ω_{it} is $\Psi(\omega_{i,t-1})$. Since the NIH grant review process takes more than half a year, the inclusion of F_{it} in $\mathcal{I}_{i,t-1}$ may be valid.[35]

This timing assumption would yield more precise estimates, but this assumption would be violated if the PI possesses information about ξ_{it} before applying for grants, or of the NIH has knowledge about ξ_{it} when making its decision at year t. For example, suppose a PI writes a grant application listing a set of working papers at the end of year $t-1$. The NIH would make funding decisions based on the quality of these working papers, and these working papers are published after the funding decision.[36] The timing structure in this case would likely violate the strong timing assumption depicted in Figure 2. Note, however, output at year t is defined as the quality weighted sum of papers published at year $t+1$, see (1). If these working papers are published at year t, then the timing structure defined in this paper would mitigate this concern. However, if these working papers are published at year $t+1$, then ξ_{it} would be correlated with the decision at year t, and the strong timing assumption depicted in Figure 2 would be invalid.[37]

With the aid of the budget instrumental variable described above, however, this assumption may be weakened. Figure 3 depicts an alternative timing assumption that is weaker than the timing assumptions depicted in Figure 2. Specifically, this timing structure assumes that funding decision at year t is made at year $t-1$ with the knowledge of the innovation shock ξ_{it}. This assumption implies that F_{it} is correlated with ξ_{it}. Therefore, F_{it} is not a valid instrument. However, $F_{i,t-1}$ is realized without the knowledge of ξ_{it}, and is hence uncorrelated with ξ_{it} and can be used as an instrument.[38] This weaker timing assumption, however,

[33] This timing assumption, as used in the literature (e.g., Levinsohn and Petrin (2003), Doraszelski and Jaumandreu (2013), and Ackerberg et al. (2015)), guarantees that $\mathbb{E}[\xi_{it}|\mathcal{I}_{i,t-1},\text{exit}_{it}=0]=\mathbb{E}[\xi_{it}|\mathcal{I}_{i,t-1}]=0$. Without this assumption, a selection model would be needed.

[34] This type of timing assumptions has also been used in the demand estimation literature (e.g. Lee (2013)). See Ackerberg and Hahn (2015) for recent results on nonparametric identification of production function using these timing assumptions.

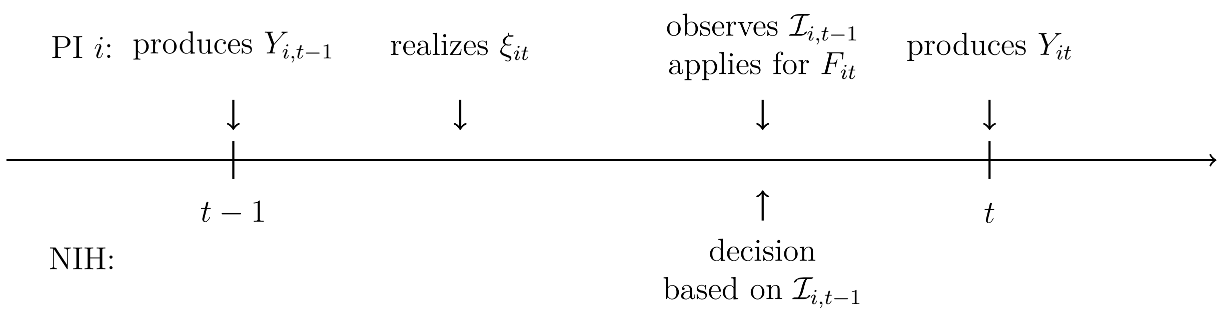
[35] See https://grants.nih.gov/grants/how-to-apply-application-guide/due-dates-and-submission-policies/due-dates.htm

[36] See e.g. Li and Agha (2015), and Li (2017).

[37] The timing assumption depicted in Figure 2 is strong, but it is commonly used in the determination of capital in the production function literature.

[38] While this weaker timing assumption allows for correlation between F_{it} and ξ_{it}, it rules out correlation

Figure 3: Timing Assumption: Weak



Notes: This weaker timing structure assumes that funding decision at year t is determined at year $t-1$ but with the knowledge of innovation shock ξ_{it}. This assumption implies that F_{it} is correlated with ξ_{it}. Therefore, F_{it} is not a valid instrument. However, $F_{i,t-1}$ is realized without the knowledge of ξ_{it}, and is thus uncorrelated with ξ_{it}, which can be used as an instrument.

cannot identify the coefficient α_1 in (6). It is thus necessary to include an additional instrument that is correlated with F_{it} but not correlated with ω_{it}. The budget instrument can be utilized as an instrument. In summary, it is possible to identify the model using a weaker timing assumption, but one additional instrument is required in this case. In this paper, I estimate the model using both timing assumptions as a robustness check.

4.1.3 Estimation

I employ the dynamic panel approach to identify and estimate the parameters in (6) .[39] Specifically, I specify a first-order autoregressive (AR1) productivity structure

$$\omega_{it} = \rho\omega_{i,t-1} + \xi_{it}, \tag{10}$$

where ρ captures the persistency of the productivity process.[40] By ρ-differencing of (6), conditions (10) and (9) yield the following moment conditions

between F_{it} and $\xi_{i,t+1}$.

[39]In general, there are two approaches in the production function literature: the proxy variable approach (e.g. OP, and ACF), and the dynamic panel approach (e.g. Blundell and Bond (2000)). The proxy variable approach requires more assumptions on the existence of a control function for the unobserved productivity, and restriction to unobserved heterogeneity, but it allows the unobserved first-order Markov productivity process to be nonparametric. The dynamic panel approach does not require the existence of a control function, but imposes a stronger assumption on the productivity process, which has to be linear. Both approaches have their own strengths and weaknesses. Nevertheless, both approaches rely crucially on the same timing assumption (9), and yield similar results (Ackerberg (2016)). Due to data limitations, I use the dynamic panel approach.

[40]In general, a linear productivity structure is needed. For instance, we could have the AR2 productivity structure, i.e. $\omega_{it} = \rho_1\omega_{i,t-1} + \rho_2\omega_{i,t-2} + \xi_{it}$.

$$
\begin{aligned}
\mathbb{E}\left[\xi_{it}+\varepsilon_{it}-\rho\varepsilon_{i,t-1}\left|\mathcal{I}_{i,t-1}\right.\right] &= 0 \\
\mathbb{E}\left[\left(\ln Y_{it}-\alpha_0-\alpha_1\ln F_{it}-\alpha_2\ln E_{it}-\alpha_3' X_{it}\right)\right. &- \\
\rho\left(\ln Y_{i,t-1}-\alpha_0-\alpha_1\ln F_{i,t-1}-\alpha_2\ln E_{i,t-1}-\alpha_3' X_{i,t-1}\right)\left|\mathcal{I}_{i,t-1}\right] &= 0. \qquad (11)
\end{aligned}
$$

To identify the parameters of interest, we need funding variables from at least three different years. The underlying identification condition requires that the past funding a PI received (conditional on observables) is not a perfect predictor for future funding of the PI. The sources of variation come from different types of smaller grants a PI can receive from the NIH, application luck (due to the taste of the study section), and variations in the NIH budget.

Parameters in (6) can be estimated with (11) using Generalized Method of Moments (GMM). The set of instruments for the moment condition (11) is all the information of PI i up to and including year $t-1$. The strength of the timing assumption (9) depends on whether F_{it} is included in $\mathcal{I}_{i,t-1}$. Under the stronger timing assumption, I include F_{it} as an instrument; under the weaker timing assumption, I exclude F_{it} as an instrument but combine the budget instrument described in Section 4.1.2 for identification and estimation. In all cases, we can use all observables in the information set of PI i at $t-1$, that is $F_{i,t-1}$ and further lags in funding and output prior to $t-1$.

One attractive feature of the dynamic panel model is that the moment conditions (11) yield over-identification restrictions, which allows us to test the validity of different instruments via Sargan-Hansen's J-test. This, in turn, tests the validity of different timing assumptions. Furthermore, the framework developed in the present paper extends naturally to different definitions of time aggregation. For instance, we can define time biennially instead of annually as in our baseline model.[41]

4.1.4 Production Function Estimation Results

Table 1 presents the estimates for the production function.[42] Columns (i) and (ii) present OLS estimates while columns (iii) and (iv) present fixed effect (FE) estimates with two deterministic experience measures: cumulative amount of funding and cumulative number of publications.[43] In all cases, the age of the labs, defined by number of years since their first R01, school ranking by the total funding received from the NIH, and time fixed effects

[41] An analysis using this level of aggregation shows similar results.

[42] I use the sample from year 2000 to investigate cancer research in more recent years.

[43] These two measures refer to (7) with $\delta = 1$.

Table 1: Production Function Estimates

	Dependent variable: log weighted output					
	OLS		*Fixed Effect*		*Dynamic Panel*	
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Funding	0.393	0.310	0.182	0.181	0.193	0.207
	(0.006)	(0.005)	(0.008)	(0.008)	(0.025)	(0.041)
Experience: funding	0.156	-	0.103	-	0.234	0.248
	(0.007)		(0.010)		(0.067)	(0.061)
Experience: # publications	-	0.510	-	-0.764	-	-
		(0.005)		(0.014)		
δ	1	1	1	1	0.604	0.590
					(0.098)	(0.083)
Age	-0.023	-0.069	-	-	-0.021	-0.021
	(0.001)	(0.001)			(0.015)	(0.011)
School Ranking	-0.058	-0.045	-0.050	-0.051	-0.100	-0.097
	(0.004)	(0.003)	(0.011)	(0.011)	(0.264)	(0.218)
ρ	-	-	-	-	0.922	0.922
					(0.012)	(0.012)
Observations	49,837	49,837	43,442	43,442	37,305	37,305
GMM J test p-value	-	-	-	-	0.442	0.808
$F_{it} \in \mathcal{I}_{i,t-1}$	-	-	-	-	Yes	No
Budget Instrument	-	-	-	-	No	Yes

Notes: HAC standard errors in parantheses are added as controls.

OLS results show higher estimated elasticity of funding than FE results, which suggests the crucial endogeneity problem with variable inputs: higher unobserved productivity is associated with higher variable inputs; therefore, OLS estimates of the elasticity of funding is upward biased. One problem presented in column (iv) is that the coefficient of experience when measured in cumulative number of publications makes little sense. This suggests that defining experience as an accumulated number of papers is problematic. This is not surprising, however. First, the number of papers a PI produces is stable, and FE first differencing results in insufficient variation of the inputs. Second, using number of papers as research capital significantly underrepresents the contribution of failed experiments to research experience. Third, without controlling for past funding, the output measure defined in (1) is sensitive to the timing of lags. Finally, if experience input is subject to measurement error, the FE estimator would be downward biased (Griliches and Hausman (1986)). An FE model with experience specified as the cumulative amount of funding, however, yields robust and reasonable results.

Columns (v) and (vi) show the dynamic panel estimates with experience modeled as discounted cumulative funding. Time trend is used instead of time fixed effects. The main results in columns (v) and (vi) use two timing assumptions. Column (v) employs a stronger timing assumption by including F_{it} as an instrument, while column (vi) adopts a weaker timing assumption by excluding F_{it} but including the present and lagged budget change as instruments.[44] Other instruments for both column (v) and (vi) include X_{it}, $F_{i,t-1}$, $F_{i,t-2}$, and $Y_{i,t-2}$. Heterostedasticity Autocorrelation Consistent (HAC) standard errors are reported.

Columns (v) and (vi), despite having different timing assumptions, yield similar results. The estimated elasticity of funding is about 0.2, the implied learning rate is about 18%, the depreciation rate is about 0.4, and the autocorrelation coefficient ρ is 0.92. The coefficients of age and school ranking are similar throughout all estimates. I employ Sargan-Hansen's J-test to access the validity of the moments (11) implied by the dynamic panel model. The J-test cannot reject the null hypothesis that the moments in (11) are valid for all models estimated in columns (v) and (vi) at the 10% level.

The estimated elasticities of funding are higher than those obtained in the FE models and lower than those obtained using OLS. These estimates are comparable to those found in Tabakovic and Wollmann (2016) and higher than those found in the previous literature (e.g. Adams and Griliches (1998) and Jacob and Lefgren (2011b)). The high estimates of ρ suggest persistency of the productivity process. The estimated depreciation rate, however, implies that 60% of stock experience from the beginning of the year remains relevant at the

[44] I measure the budget instrument as the percentage change of budget allocated to labs of specific ages.

end of the year. The estimated depreciation rate may seem high. However, as pointed out earlier, the experience process contains the scientific knowledge of the PI, research experience of the lab members, and other forms of research capital. Thus, the depreciation rate does not merely measure the depreciation of scientific knowledge of the PI.

The high depreciation rate can be justified by the organizational structure of academic research labs and the nature of scientific research. First, lab members are typically Ph.D. students and postdocs. They are not permanent lab members, and may move between labs through different stages of their career. In addition, new lab members are not perfect substitutes for those who left since training takes time. Hence, the high depreciation rate can be explained by the high turnover rate of lab members. Second, scientific research production requires numerous scientific experiments. Some of the experimental designs are unique and may not be easily transferable. Lastly, as pointed out by Stephan (2012), certain fields progress quickly, so that scientists (both the PI and lab members) may not be able to keep up with the pace at which the discipline is changing over time.[45]

4.2 Other Model Primitives

This subsection discusses the model primitives of the DP problem including the state variables and their transition probabilities.

4.2.1 State Variables

To summarize, each PI has 5 state variables, Non-R01 funding, R01 funding, grant cycle, experience, and TFP. Since the key dynamics of the problem lies in experience and productivity, I utilize coarse discretization for the funding variables, experience, and productivity state variables. Table 2 summarizes the state variables and the transition probabilities.

R01 Funding and Grant Cycle

For R01 Funding, we focus on one-point support, taken as the median of R01 grant.

$$F_{it}^{\text{R01}} \in \{F_{\text{meidum}}^{\text{R01}}\} = 300$$

That is, I impose the restriction that a research PI can only propose one type of R01 grant, which lasts for four years.

[45] I ran regressions with the specifications $\ln Y_{it} = \alpha_0 + \sum_{j=1}^{J} \alpha_j \ln F_{i,t+1-j} + \varepsilon_{it}$ under different J. Overall, the coefficient α_j is smaller as j increases, implying that the contribution of lagged funding on current output gets smaller as we increase the lag.

Non-R01 funding

While we focus on R01 grant funding decisions, PIs often have Non-R01 grants, such as P01 and R21. To account for this factor, I allow for Non-R01 funding to have a three-point support

$$\begin{aligned} F_{it}^{\text{NonR01}} &\in \{0, F_{\text{low}}^{\text{NonR01}}, F_{\text{high}}^{\text{NonR01}}\} \\ &\in \{0, 250, 500\} \end{aligned}$$

where $F_{\text{low}}^{\text{NonR01}}$ is taken to be the median of Non-R01 funding, and $F_{\text{hight}}^{\text{NonR01}}$ is taken to be the upper quartile of Non-R01 funding.[46] The probability transition of this random variable is estimated from data nonparametrically.

Experience and Productivity

Experience is the key dynamic input for the funding allocation problem. The parameters estimated from Section 4.1 determine the impact of experience on research output as well as the transition of the experience state variable over time. The experience process can be well approximated if we have enough discrete experience states. Due to computational complexity, however, I limit the number of experience states to be $d_E = 7$, and focus on the following discrete support

$$\begin{aligned} E_{it} &\in \{E_{it}^{(1)}, ..., E_{it}^{(d_E)}\} \\ &\in \{100, 300, 550, 850, 1200, 1600, 2000\}. \end{aligned}$$

The range of the support ensures that all possible range of experience states is covered under the specification of funding from R01 and Non-R01 grants. While the experience process in Section 4.1 is deterministic, following Benkard (2004), I make the state transition for experience stochastic. Specifically, to determine the experience level of i at period t, I first calculate $\hat{E}_{it}$ according to the deterministic process (7). Then I find the smallest upper bound, E_{upper}, to E_{it} and the largest lower bound, E_{lower}, to E_{it}, and calculate

$$E_{it} = \begin{cases} E_{\text{lower}} & \text{with probability} \quad \frac{\hat{E}_{it} - E_{\text{lower}}}{E_{\text{upper}} - E_{\text{lower}}} \\ E_{\text{upper}} & \text{with probability} \quad 1 - \frac{\hat{E}_{it} - E_{\text{lower}}}{E_{\text{upper}} - E_{\text{lower}}}. \end{cases}$$

[46] The dollar value is in thousands.

Total Factor Productivity

For TFP, I limit the number of productivity states to be $d_\omega = 8$, and use the following discrete support

$$\begin{aligned} \omega_{it} &\in \{\omega_{it}^{(1)}, ..., \omega_{it}^{(d_\omega)}\} \\ &\in \{-0.6, -0.3, 0, 0.2, 0.4, 0.6, 0.8, 1\}. \end{aligned}$$

To construct the transition probability, I first obtain the variance of ξ_{it} from data, and calculate the corresponding probability transition based on the TFP structure (10) using $\xi_{it} \sim Normal(0, \sigma_\xi^2)$.[47] I allow for more points on the higher end of the support and fewer points on the lower end of the support of TFP because PIs with TFP at the lower end of the support will not be funded. Therefore, having a coarser support of the lower end will eliminate unnecessary state variables but retain the results of the DP problem.

Table 2: State Variables

Variable	Explanation	Dimension	State Transition
F_{it}^{NonR01}	non-R01 Funding	3	conditional expectation
F_{it}^{R01}	proposed R01 Funding	1	deterministic
C_{it}	Grant Cycle: Four Years	8	deterministic
E_{it}	Experience	7	Stochastic: $E_{i,t+1} = \delta E_{it} + F_{it}$
ω_{it}	Productivity	8	Stochastic: $\omega_{i,t+1} = \rho\omega_{it} + \xi_{i,t+1}$

4.2.2 The Set of Applicants

The set of R01 applicants consists of the union of entrants and incumbents:

$$\mathcal{N}_t = \mathcal{N}_t^{\text{entrant}} \cup \mathcal{N}_t^{\text{incumbent}}.$$

The funding allocation depends crucially on the composition of the applicants. Since I do not observe the set of applications, I have to make a few assumptions on the sets $\mathcal{N}_t^{\text{entrant}}$ and $\mathcal{N}_t^{\text{incumbent}}$. These assumptions are based on annual R01 grant report published by the NIH.[48]

[47] To obtain the variance of ξ_{it}, I first back out the residual $\epsilon_{it} = \omega_{it} + \varepsilon_{it}$ from the production function (6), this allows me to obtain the sample variance $\sigma_\epsilon^2 = \sigma_\omega^2 + \sigma_\varepsilon^2$. I then use the AR1 structure (10) together with the i.i.d. assumption of measurement error ε_{it} to obtain the variance of ξ_{it}. Specifically, I run the project $\epsilon_{it} = \alpha\epsilon_{i,t-1} + \eta_{it}$. The autocorrelation coefficient of this projection α contains information of both the variance of the TFP σ_ω^2 and the variance of the residual σ_ϵ^2. Next, I use the fact that the variance of TFP is $\sigma_\xi^2/(1-\rho^2)$ to solve for σ_ξ^2.

[48] https://gsspubssl.nci.nih.gov/blog/articles?funding_patterns/2011

There are a total of 4477 R01 applications, and 651 of these are funded in 2011. The success rate is about 15%. Of the 4477 applicants, 3005 applications came from experienced investigators and 1472 applications came from new investigators. An experienced investigator is a PI who has previously received a R01 grant, and a new investigator is a Postdoc or PI who has not previously received an R01 grant.[49] For example, PI who has received an R01 grant is not qualified to be a new investigator. On the other hand, a PI who has not received an R01 grant but has received other types of NIH grants, such as R21, is qualified to be a new investigator. Furthermore, the NIH determines whether new investigators are early stage investigators. An early stage investigator is defined to be a new investigator within 10 years of receiving their highest degree. Of the 1472 applications from new investigators, 564 came from early stage investigators.

The baseline model of the paper focuses on a budget allowing for 600 R01 awards each year. I allow for 1000 applicants to apply as new investigators. Of the 1000 new investigators, I specify 500 of them to enter with experience level 1, and the other 500 of them to enter with experience level 2. The initial distribution of TFP is drawn from $Normal(0,(1+\rho^2)\sigma_\xi^2)$. Furthermore, I assume both early and non-early investigators to have received zero Non-R01 grants.

The incumbent PIs are defined as PIs that have active R01 grants and/or PIs that do not currently have active R01 grants but had them previously. Specifically,

$$\mathcal{N}_t^{\text{incumbent}} = \mathcal{N}_t^{\text{Active R01}} \cup \mathcal{N}_t^{\text{Expired R01}}$$

where $\mathcal{N}_t^{\text{Active R01}}$ is the set of PIs with active R01 grants, and $\mathcal{N}_t^{\text{Expired R01}}$ is the set of PIs currently had their R01 grants expired. Note that $\mathcal{N}_t^{\text{expired R01}}$ contains experienced PIs with Non-R01 grants.

5 Approximate Dynamic Programming

5.1 Dynamic Resource Allocation Formulation

The DP problem in (4) suffers from three curses of dimensionality. First, the action is high-dimensional and combinatoric. Second, given the large dimensionality of PI characteristics, the state space of the DP problem is enormous. Third, the conditional expectation of the value function depends on the action, and it is difficult to evaluate when the state space is large. Therefore, standard approaches to solving (4) in the DP literature, such as brute-

[49] https://www.cancer.gov/grants-training/policies-process/overview/grants-process.pdf

force value function iteration, is infeasible. In order to compute the solution of (4), I first show that the problem (4), under proper transformation, can be represented as a dynamic resource allocation (DRA) problem that is well known in the operations research literature. Second, due to the concave structure of the DRA, I can utilize techniques in approximate dynamic programming to construct a tractable approximation of the value function, and thereby evaluate the solution to the DP problem to under such an approximation.

5.1.1 State Vector

To proceed, I formulate (4) into a more manageable representation. First, I discretize characteristics F, E, and ω to d_F, d_E, and d_ω levels respectively as discussed in Section 4.2. As a consequence, PIs are equivalent up to the characteristics; hence, we can associate each PI with a particular PI type. I represent PI type X as a tensor product of F, E, ω, and C, and denote the finite type space by $\mathcal{X}$. $\mathcal{X}$ has d_X elements where $|\mathcal{X}| = d_F \times d_E \times d_\omega \times d_C = d_X < \infty$.[50] Since each PI takes a particular value from $\mathcal{X}$, I can represent the set of available PIs as a d_x-dimensional resource vector $\mathbf{N}_t = [N_{1t}, ..., N_{d_x t}]' = (N_{xt})_{x \in \mathcal{X}}$, where $N_{xt} = \sum_{i \in \mathcal{N}_t} \mathbf{1}\{X_i = x\}$, is the number of type k PIs available at year t.[51] The vector $\mathbf{N}_t$ is the state vector of the PI-population distribution by the types. Consequently, we can represent the state of (4) at period t as $\mathbf{S}_t = \{\mathbf{N}_t, B_t\}$. We assume $B_t = B$ to simplify the presentation, so that we can represent the state of the problem as the vector $\mathbf{N}_t$.[52] d_x is more than 1000 in the baseline model in this paper, implying the number of state variable is more than 1000. Given the number of applicants is about 3000, the approximate size of the state space of the funding allocation problem, $|\mathbf{S}|$, is about $1000^{3000} = 10^{9000}$.

5.1.2 Decision

Another consequence of discretizing the observed characteristics (or aggregation) of PIs is that, I can simplify the planner's decision of choosing from a subset with combinatorics to choosing the number of PIs of different types to fund, subject to budget and resource constraints. Define $\mathcal{D}$ to be the set of decisions the planner can make. In our model, a planner can choose whether to fund a PI. Hence, I represent the decision by $\mathcal{D} = \{0, 1\}$ where

$$\mathcal{D} = \begin{cases} 0 & \text{don't fund} \\ 1 & \text{fund.} \end{cases}$$

[50] Note that the grant cycle state variable only takes discrete values.

[51] $\mathbf{1}\{.\}$ is an indicator function.

[52] For a more general model, we can model the budget variable as the number of existing PIs, and the same analysis presented in the present paper would still go through.

Therefore, the planner's decision can be represented as a $2d_X$-dimensional vector

$$\begin{aligned}\mathbf{q}_t &= \{q_{xt}^d\}_{x\in\mathcal{X},d\in\mathcal{D}}. \\ &= [q_{1t}^0, ..., q_{d_X t}^0, q_{1t}^1, ..., q_{d_X}^1]'\end{aligned}$$

where q_{xt}^d represents the number of type x PI under NIH's decision d at year t. Note that the decision for each type x PI needs to satisfy the following balance constraints

$$q_{xt}^0 + q_{xt}^1 = N_{xt} \ \forall x \in \mathcal{X}.$$

5.1.3 Output Per Period

Note that PIs that are not funded but have either active R01 grants will continue to produce. As a result, we define output associated with a type x lab as

$$Y_x = \begin{cases} Y_x^0 & \text{if not funded} \\ Y_x^1 & \text{if funded} \end{cases}$$

where Y_x^0 is the output a type x PI would produce if it is not funded, and Y_x^1 is the output a type x would produce if it is funded. Notice that PIs of types with no active R01 grant and Non-R01 grants will produce 0 output, otherwise $Y_x^0 > 0$. Therefore, we can represent the payoff per period as

$$\begin{aligned}Y(\mathbf{q}_t) &= \sum_{x\in\mathcal{X}} Y_x^0 q_{xt}^0 + \sum_{x\in\mathcal{X}} Y_x^1 q_{xt}^1 \\ &= \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} Y_x^d q_{xt}^d\end{aligned}$$

5.2 Dynamic Programming Problem

With this discretization, the DP problem (4) can be simplified to the following representation

$$\begin{aligned}V_t(\mathbf{N}_t) &= \max_{\mathbf{q}_t} \left\{ \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} Y_x^d q_{xt}^d + \beta\mathbb{E}\left[V_{t+1}(\mathbf{N}_{t+1})|\mathbf{N}_t, \mathbf{q}_t\right] \right\} & (12) \\ \text{such that } \sum_{x\in\mathcal{X}} F_x q_{xt}^1 &\leq B_t & (13) \\ q_{xt}^0 + q_{xt}^1 &= N_{xt} \quad \forall x \in \mathcal{X} & (14) \\ q_{xt}^d &\geq 0 \quad \forall x \in \mathcal{X}, d \in \mathcal{X} & (15) \\ q_{xt}^d &\in \mathbb{Z}^+ \quad \forall x \in \mathcal{X}, d \in \mathcal{D}. & (16)\end{aligned}$$

Equation (12) is the value function where the planner chooses a decision vector $\mathbf{q}_t$ to maximize the discounted sum of research output at every period. (13) is the budget constraint, (14) is the feasibility constraint, and (16) restricts the decision to be positive integer. To put it succinctly, I define the set of feasible choices under constraints (13), (14), and (16) as $\mathcal{Q}(\mathbf{N}_t)$:

$$\begin{aligned}\mathcal{Q}(\mathbf{N}_t) \;=\; & \{\mathbf{q}_t : \mathbf{F}'\mathbf{q}_t \leq B, \\ & q^0_{xt} + q^1_{xt}, = N_{xt} \quad \forall x \in \mathcal{X} \\ & q^d_{xt} \in \mathbb{Z}^+ \quad \forall x \in \mathcal{X}, d \in \mathcal{D}\Big\}.\end{aligned}$$

In the case where B_t varies with t, the feasibility set can be written as $\mathcal{Q}(\mathbf{N}_t, B_t)$. In sum, the DP problem can be expressed as following

$$V_t(\mathbf{N}_t) \;=\; \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t)} \left\{Y(\mathbf{q}_t) + \beta \mathbb{E}\left[V_{t+1}(\mathbf{N}_{t+1}) | \mathbf{N}_t, \mathbf{q}_t\right]\right\}, \tag{17}$$

where $Y(\mathbf{q}_t) = \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} Y^d_x q^d_{xt}$.

This representation (17) is much easier to handle than (4); however, it still suffers from the curses of dimensionality because d_x can be large.

5.3 Post-decision state variable

Recall that PIs in $\hat{\mathcal{N}}_t$ at period t proceed to the subsequent period $t+1$, and their old types (at period t) transition to new types (at period $t+1$) with known probability. Hence, the realization of the state variable at period $t+1$ depends on the state variable $\mathbf{N}_t$ and the action $\mathbf{q}_t$. I define the post-decision state variable as

$$\mathbf{N}^q_t = (N^q_{xt})_{x\in\mathcal{X}} \tag{18}$$

which represents the number of PIs of each type available after decision $\mathbf{q}$ but right before period $t+1$. The pre-decision state variable at period $t+1$ depends on $\mathbf{N}^q_t$, which can be written as

$$\mathbf{N}_{t+1} = \mathbb{P}(\cdot|\mathbf{N}^q_t), \tag{19}$$

where the pre-decision state vector $\mathbf{N}_{t+1}$ is drawn from a known distribution conditional on $\mathbf{N}^q_t$. In the special case with no grant cycle state variables, action $\mathbf{q}_t$ determines the available PIs in the subsequent period. Therefore, $N^q_{xt} = q^1_{xt}\ \forall x \in \mathcal{X}$. In this case, we can interpret $\mathbf{N}^q_t$ as the PIs that are funded at t just before transitioning to $t+1$. In general, $N^q_{xt} \neq q^1_{xt}$

since the transition of characteristics such as grant cycle and experience dynamics is mostly deterministic.

Using the post-decision state variable representation in (18) and (19), I can rewrite the DP problem (17) as

$$V_t(\mathbf{N}_t) \quad = \quad \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t)} \left\{ Y(\mathbf{q}_t) + \beta \mathbb{E}\left[V_{t+1}(\mathbf{N}_{t+1}) | \mathbf{N}_t^q \right] \right\}. \tag{20}$$

Observe that, when $\beta = 0$, the problem (20) is a simple knapsack problem, which can be solved with commercial software with large d_x values. The key idea of the ADP solution to (20) is to leverage this unique structure of the problem by approximating the conditional expectation by a form in which the solution to the approximated $V_t(\mathbf{N})$ is easy to compute.

To proceed, first denote the conditional expectation by

$$V_{t-1}^q(\mathbf{N}_{t-1}^q) := \mathbb{E}\left[V_t(\mathbf{N}_t) | \mathbf{N}_{t-1}^q \right]. \tag{21}$$

Simple algebra yields

$$\begin{aligned} V_{t-1}^q(\mathbf{N}_{t-1}^q) &= \mathbb{E}\left[V_t(\mathbf{N}_t) | \mathbf{N}_{t-1}^q \right] \\ &= \mathbb{E}\left[\max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t)} \left\{ Y(\mathbf{q}_t) + \beta \mathbb{E}\left[V_{t+1}(\mathbf{N}_{t+1}) | \mathbf{N}_t^q \right] \right\} | \mathbf{N}_{t-1}^q \right] \qquad (22) \\ &= \mathbb{E}\left[\max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t)} \left\{ Y(\mathbf{q}_t) + \beta V_t^q(\mathbf{N}_t^q) \right\} | \mathbf{N}_{t-1}^q \right]. \qquad (23) \end{aligned}$$

Equality (22) is obtained by substituting (20) into (21), and equality (23) is obtained by substituting (21) into (22). The conditional expectation is taken over the realization of $\mathbf{N}_t$ conditional on the post decision state variable $\mathbf{N}_{t-1}^q$. Let $\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t)$ denote a particular realization r_t of $\mathbf{N}_t$ conditional on $\mathbf{N}_{t-1}^q$, where r_t denotes a particular random number generation at t. Therefore, a particular realization of the expectation (22) is

$$V_{t-1}^q(\mathbf{N}_{t-1}^q, r_t) = \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} Y(\mathbf{q}_t) + \beta V_t^q(\mathbf{N}_t^q). \tag{24}$$

This particular representation is called the certainty equivalent representation of the expectation, or a Monte-Carlo draw of the expectation.[53] I call equation (24) the Subproblem to the dynamic resource allocation problem.

[53]This type of method is used in Rust (1997).

5.4 Value Function Approximation

If we know the function $V_t^q(\mathbf{N}_t^q)$, we can obtain the solution to the DP problem (20). The idea to solving (12) is to find a suitable approximation for $V_t^q(\mathbf{N}_t^q)$, and obtain a solution based on such approximation. It is well known in the literature that $V_t^q(\mathbf{N}_t^q)$ is concave in $\mathbf{N}_t^q$.[54] The approximate dynamic programming method proceeds in two steps. The first step is to approximate the expectation function V_t^q in such a way that it captures the structure of problem and the solution to the Subproblem (24) is easy to compute. The second step is to use an iterative procedure similar to policy iteration to obtain the parameters of the approximated expectation function. In general, the expectation function V_t^q can be well approximated by

$$\hat{V}_t^q(\mathbf{N}_t^q;\theta_t) = \sum_{b\in\mathcal{B}} \phi_b(\mathbf{N}_t^q;\theta_{bt}), \tag{25}$$

where $\phi_b(\cdot)$ is a basis function, and θ_{bt} is a parameter associated with the basis $\phi_b(\cdot)$.

The present chapter adopts the particular approximation bellow,

$$\hat{V}_t^q(\mathbf{N}_t^q;\theta_t) = \sum_{x\in\mathcal{X}} \phi_x(\mathbf{N}_t^q;\theta_{xt}) \tag{26}$$

where $\phi_x(\cdot)$ is a function approximating the value of N_{xt}^q, the number of type x PIs remaining post-decision, and $\theta_t = (\theta)_{x\in\mathcal{X},t\in\mathcal{T}}$ is a vector of parameters of the value function. This approximation assumes that the value function is separable in types.

A convenient functional form is linear, and so the linearly approximated expectation function can be written as

$$\hat{V}_t^q(\mathbf{N}_t^q;\theta_t) = \sum_{x\in\mathcal{X}} \theta_{xt} N_{xt}^q. \tag{27}$$

This convenient functional form is a good starting to solving (20).

Although easy to implement, the linear approximation (27) imposes a strong functional form assumption on the value function. One simple nonlinear approximation is to use a piecewise-linear concave form instead. Let $\bar{S}_x$ denote the maximum number of type x PIs allowed in the analysis, and define $\mathcal{S}_x = \{0, 1, ..., \bar{S}_x\}$ as the support of N_{xt}^q. Then the piecewise-linear approximation can be written as

$$\begin{aligned} \phi_x(N_{xt}^q, \theta_{xt}) &= \sum_{s\in\mathcal{S}_x} \theta_{xt}(s) z_{xt}(s) && (28)\\ z_{xt}(s) &= \mathbb{I}\{N_{xt}^q \le s\} && (29)\\ \theta_{xt}(s) &\ge \theta_{xt}(s+1)\ \forall s < \bar{s}_x, && (30) \end{aligned}$$

[54] Appendix B provides a simple sketch of the proof.

with normalization $\theta_{xt}(0) = 0$ for all $x \in \mathcal{X}$. Equations (28), (29), and (30) defines the concave piecewise-linear function of type x PI, where parameters $(\theta_{xt}(s))_{s\in\mathcal{S}_x}$ determine the shape of the piecewise-linear approximation of the value of a type x PI. More precisely, an increment of type x PI from s to $s+1$ increases the value of the value function of a type x PI by $\theta_{xt}(s+1)$. It is clear that the linear approximation (27) is a special case of (28), with additional restrictions $\theta_{xt}(s) = \theta_{xt}(s')$ for $s \neq s'$ and for all $x \in \mathcal{X}$.

The piecewise-linear approximation can be interpreted as a nonparametric approximation to the value of the number of type x PIs. However, the additional restrictions (30) impose a concave shape on the nonparametric function ϕ_x and restrict the diminishing returns to value of the PIs.

5.5 An ADP Algorithm

Algorithm 1 ADP Algorithm Pseudo Code

Step0 Initialization:

1. Set M, the maximum number of iteration
2. initialize θ_t^0 for $t \in \mathcal{T} = \{1, ..., T\}$.
3. Set iteration counter $m = 1$.

Step1 Set $t = 1$ and initialize $\mathbf{N}_1^m$ (sample from a possible state). For $t = 1, 2, ..., T-1$, do

1. draw r_{t+1}^m and determine the state $\mathbf{N}_{t+1}(\mathbf{N}_t^{q,m}, r_{t+1}^m)$
2. Solve

$$\begin{aligned}
\tilde{V}_t^q(\mathbf{N}_t^{q,m}, r_{t+1}^m) &= \max_{\mathbf{q}_{t+1}\in\mathcal{Q}(\mathbf{N}_{t+1}(\mathbf{N}_t^{q,m}, r_{t+1}^m))} Y(\mathbf{q}_{t+1}) + \beta V_{t+1}^q(\mathbf{N}_{t+1}^q; \theta_{t+1}^{m-1}).\\
\mathbf{q}_{t+1}^m &= \arg\max_{\mathbf{q}_{t+1}\in\mathcal{Q}(\mathbf{N}_{t+1}(\mathbf{N}_t^{q,m}, r_{t+1}^m))} Y(\mathbf{q}_{t+1}) + \beta V_{t+1}^q(\mathbf{N}_{t+1}^q; \theta_{t+1}^{m-1})\\
\mathbf{N}_{t+1}^{q,m} &= \Delta\mathbf{q}_{t+1}^m
\end{aligned}$$

3. Update parameter
   (a) compute slope
   $$\tilde{v}_{xt}^m = \tilde{V}_t(\mathbf{N}_t^{q,m} + e_x, r_{t+1}^m) - \tilde{V}_t(\mathbf{N}_t^{q,m}, r_{t+1}^m) \quad \forall x \in \mathcal{X}.$$
   (b) sup-norm project or L_2 projection
   $$\theta_t^m = UPDATE(\theta_t^{m-1}, \tilde{V}_t^q(\mathbf{N}_t^m, r_{t+1}^m), \tilde{\mathbf{v}}_t^m)$$

Step2 If $m < M$, $m = m + 1$. Go back to Step1.

Under linear approximation (27) or concave piecewise-linear approximation (28), the Subproblem (24) is an integer programming problem conditional on parameters β and θ. This Subproblem can be solved quickly even when d_x is large. See Appendix A for more details. Provided that (24) can be solved quickly, the algorithm to compute the allocation of (12) is to obtain the parameters $(\theta_t)_{t\in\mathcal{T}}$ of the value function, and this algorithm resembles the "policy-iteration" type methods.

This section presents a sketch to obtain the parameter $(\theta_t)_{t\in\mathcal{T}}$, and I demonstrate the algorithm for obtaining the parameters when the value function is linearly approximated. See Appendix B for the algorithm when the value function is piecewise-linearly approximated.

When the value function is linearly approximated, e.g. (27), we can interpret the coefficient θ_{xt} as the marginal value of one additional type x PI at period t. We update the parameter iteratively using policy iteration.

Algorithm 1 demonstrates the pseudo-algorithm for the ADP procedure. Specifically, for iteration m, the additional value of type x PI at period t is

$$\tilde{v}^m_{xt}(\theta^{m-1}_{t+1}) = \tilde{V}_t(\mathbf{N}^{q,m}_t + e_x, r^m_{t+1}; \theta^{m-1}_{t+1}) - \tilde{V}_t(\mathbf{N}^{q,m}_t, r^m_{t+1}; \theta^{m-1}_{t+1}), \tag{31}$$

where e_x is a d_x dimensional vector with 1 in the x-th element and 0 everywhere else. Note that the evaluation of $\tilde{V}_t$ is conditional on parameter θ^{m-1}_{t+1}, which is the parameter of iteration $m-1$ at period $t+1$. This approach resembles backward dynamic programming with policy iteration.

The parameter θ^m_{xt} can then be updated using the following formula:

$$\theta^m_{x,t} = (1-\alpha_m)\theta^{m-1}_{xt} + \alpha_m \tilde{v}^m_{xt} \quad \forall x \in \mathcal{X}, t \in \mathcal{T}, \tag{32}$$

where $\tilde{v}^m_{xt}$ is obtained from (31). The formula (31) can be thought of derivative base updating rule, such as the steepest descent type algorithm. Note that calculating $\tilde{v}^m_{xt}$ involves parameters θ^{m-1}_{t+1} to update parameter θ^m_t. In particular, we notice that $\tilde{V}_t(\mathbf{N}^{q,m}_t, r^m_{t+1})$ is a function of θ^{m-1}_{t+1}, and we are using the parameters of the approximated value function at period $t+1$ to update parameters of the approximated value function at period t.[55] The convergence property of this algorithm follows from Powell (2011).

[55]The parameter α_m is the step size, and goes to 0 as $m \to \infty$. Following Topaloglu and Powell (2006) I set $\alpha_m = 20/(40+m)$.

6 Evaluating the Optimal Funding Allocation

This section uses the dynamic funding allocation model together with the production function estimates to analyze the short-term and long-term differences in research output generated by policies under different discount factors. The goal of this exercise is to quantify the magnitude of the differences in research output and examine the factors that drive these discrepancies by looking at the distribution of PIs over time. Lastly, I compare the predicted share of funding awarded to the young PIs with the observed data to calibrate the NIH's discount factor.

6.1 Results

I solve the model with different discount factors using the approximation dynamic programming techniques introduced in Section 5. I run 500 simulations of the funding allocation for different discount factors with initial conditions drawn from the stationary distribution of PIs under the myopic funding policy.

Figure 4 shows the aggregate research output per period for different discount factors, $\beta \in \{0, 0.1, 0.3, 0.5, 0.7, 0.8, 0.9\}$, over 30 years. The analysis demonstrates that there are substantial tradeoffs between forward-looking policies and the myopic policy. Despite having minor reductions in research output in the first four periods, forward looking policies produce more research output than the myopic policy as time progresses. The magnitude of this discrepancy increases with the discount factor. For instance, when the discount factor is 0.1, the allocation and the research output trajectories are similar to those of the myopic policy. What is striking, however, is that with a relatively small discount factor, for example, $\beta = 0.3$, the forward looking policy produces drastically higher research output than the myopic policy over time. For a discount factor of 0.9, the forward looking policy produces 1% less research output for the first four years but about 5% more research output per year thereafter.

Figures 5 and 6 show the plot of the cumulative differences in research output under different discount factors as compared to the myopic policy. These figures highlight the significant tradeoffs over time. The cumulative difference in research output between the forward-looking policy under a discount factor of 0.9 and the myopic policy, for example, exceeds the total annual research output at the end of the time horizon. This indicates that we could have missed a year of research output at the end of the thirty-year planning horizon if we do not account for the impact of funding on experience when deciding whom to fund.

There are two factors that drive these differences. First, forward looking policies put more weight on TFP than on experience as compared to the myopic policy. As a result, forward looking policies fund more productive young PIs. Second, since more high TFP PIs are selected at the expense of high experience and low TFP PIs, and TFP is persistent, forward-looking policies lead to a more productive distribution of PIs in the long run.

To illustrate this, Figure 7 plots the short-run distribution of winning PIs as characterized by experience and TFP. The distribution is calculated using the first five years of allocations starting from the stationary distribution of PIs under the myopic funding policy. This figure illustrates that the distribution of winning PIs gravitates toward the upper-left quadrant as the discount factor increases. This indicates that younger and more productive PIs receive a growing share of total funding as the discount factor increases.

As a result, long-run distributions of PIs under forward-looking policies are more productive as compared to the myopic policy. Figure 8 demonstrates the changes in the long-run distribution of PIs as the discount factor increases. We can see that as the discount factor increases, the long-run distribution gravitates upward (or higher TFP) and slightly toward the left (or slightly younger). Figure 9 plots the long-run distribution of winning PIs, showing the same pattern persist as compared to the short-run distribution of winning PIs.

In summary, forward looking policies invest more in productive young PIs at the expense of more experienced but less productive PIs. Since productivity is persistent, this leads to a more productive distribution of PIs in the long-run.

6.2 Optimality versus Reality

> "One of my personal priorities is developing the next generation of talented biomedical researchers." —Francis Collins

In the previous section, we learn that the tradeoffs in research output between forward-looking policies and the myopic policy are significant. The key driver of these differences hinge on the willingness to invest in young and productive PIs. Figure 10 plots the optimal share of funds to the young PIs under different discount factors. I define young PIs as the entrants, or those who have not yet received an R01 grants. Based on my model, the optimal share of funds to young PIs is 20.2% under the myopic policy, and this share increases as the discount factor increases. A forward-looking policy with a discount factor of 0.9 funds 28.9% more young PIs as compared to the myopic policy, and the optimal share of funds to young PIs is 25.9%.[56]

Determining the optimal distribution of funds between young and veteran PIs has been an issue policy makers have publicly grappled with.[57] How well is the NIH allocation doing

[56] This is calculated based on the long-run distribution of winning PIs over 500 simulation.

[57] See, e.g. Freeman and Van Reenen (2008), and Collins (2010; 2017; 2018).

in terms of investing in the future generation of researchers to maximize aggregate research output? The statistics published by the NIH show that 24.6% of R01 grants were awarded to young researchers (or the entrants) over the past 8 years, see Figure 10.[58] Under model predictions, the NIH would fund 18% less of R01 grants to young PIs if its policies were myopic. Using the share of grants allocated to young PIs, the discount factor that rationalizes the NIH's funding behavior is about 0.75.

These results indicate that the NIH appears to account for some intertemporal tradeoffs. However, they are still underfunding some young researchers as compared to policies with higher discount factors. Nevertheless, the results show that their actual funding behavior is consistent with their mission to fund more young researchers.

6.3 Effects of a Temporary Budget Cut

In March 2017, the current administration proposed a 20% cut on the NIH annual budget. Given the widespread perception that funding was already difficult to obtain, the scientific community as well as lawmakers strongly opposed the idea, worrying that such a proposal would jeopardize scientific innovation and delay medical cures. Six months later, the Congress rejected the proposal. In light of this episode, it is important to evaluate the impact of budget cuts on research output.

In this section, I employ the empirical framework developed in this paper to analyze the impact of a stylized version of the proposed budget cut on research output and distribution of PIs.[59] I use the same setup as the one in section 7, but focus on the scenario in which the R01 grant budget is cut by 20% for the first four years. Furthermore, I assume that PIs would not respond to such a budget cut but would continue to apply for the same type of R01 funding as before, so that the quality of the entrants (or the young PIs) would not change. As a result, the NIH can award 480 R01 grants during the budget cut period as opposed to 600 R01 grants in a normal period.

Figure 11 plots the cumulative differences in research output between polices with and without the budget cut over time.[60] The cumulative output reduction in the first four years is about 36% of the annual research output. Moreover, the total loss of research output continues to accumulate even when the budget returns to normal at the fifth year and thereafter. The differences eventually plateau out after 25 years, and the permanent reduction in research output from this policy is about 80% of the annual research output.

[58] See https://gsspubssl.nci.nih.gov/blog/articles?funding_patterns/2011

[59] The actual budget cut proposal was to cut payments to universities for overheads rather than budget cuts on certain types of research grants.

[60] I use the stationary distribution of PIs of the corresponding funding policy as the initial condition for this simulation.

Figure 12 illustrates the effect of the budget cut in terms of percentage cumulative differences in research output. It shows that the cumulative output is decreased by 10% during the budget cut period, and this reduction continues to be significant even ten years after the budget cut. Overall, this policy would result in a 3% reduction in aggregate research output over thirty years.

This policy has direct and indirect effects on research output, which leads to a large permanent reduction in aggregate output. The direct effect is a reduction in the scientific workforce during the period of budget cut. Figure 13 shows a plot of the number of winning PIs per year decomposed by experience levels during the budget cut period. This plot demonstrates that the policy would hurt mostly young and mid-career PIs; 30% of young PIs would not be funded, as compared to 16% of mid-career PIs and 16% of veteran PIs.[61] Of these 120 missing R01 awards, 53 of them would have been given to young PIs, 54 would have been given to mid-career PIs, and 12 would have been given to the veteran PIs. Moreover, this funding reduction leads to a decrease in PI experience accumulation, thus impacting output dynamically and indirectly. The missing PIs during the period of budget cut never get to accumulate the experience needed to become veterans. Hence, the absence of these productive experienced PIs that the NIH would have funded during a normal period explain the long-lasting reduction in aggregate output. In conclusion, a budget cut policy, even if temporary, would lead to a significant long-lasting reduction in aggregate research output.

7 Conclusion

I provide an empirical framework to compute the optimal share of funds between young and veteran investigators. I show that the dynamic impact of funding on research output is of first-order importance, and ignoring this dynamic impact when deciding whom to fund can be detrimental to the future of scientific advancement. The main result highlights the importance of time preferences on funding allocation and research output over time. I find that a forward-looking policy under a discount factor of 0.9 would fund 30% more young PIs than the myopic policy, and this translate to 5% more research output per year in the future.

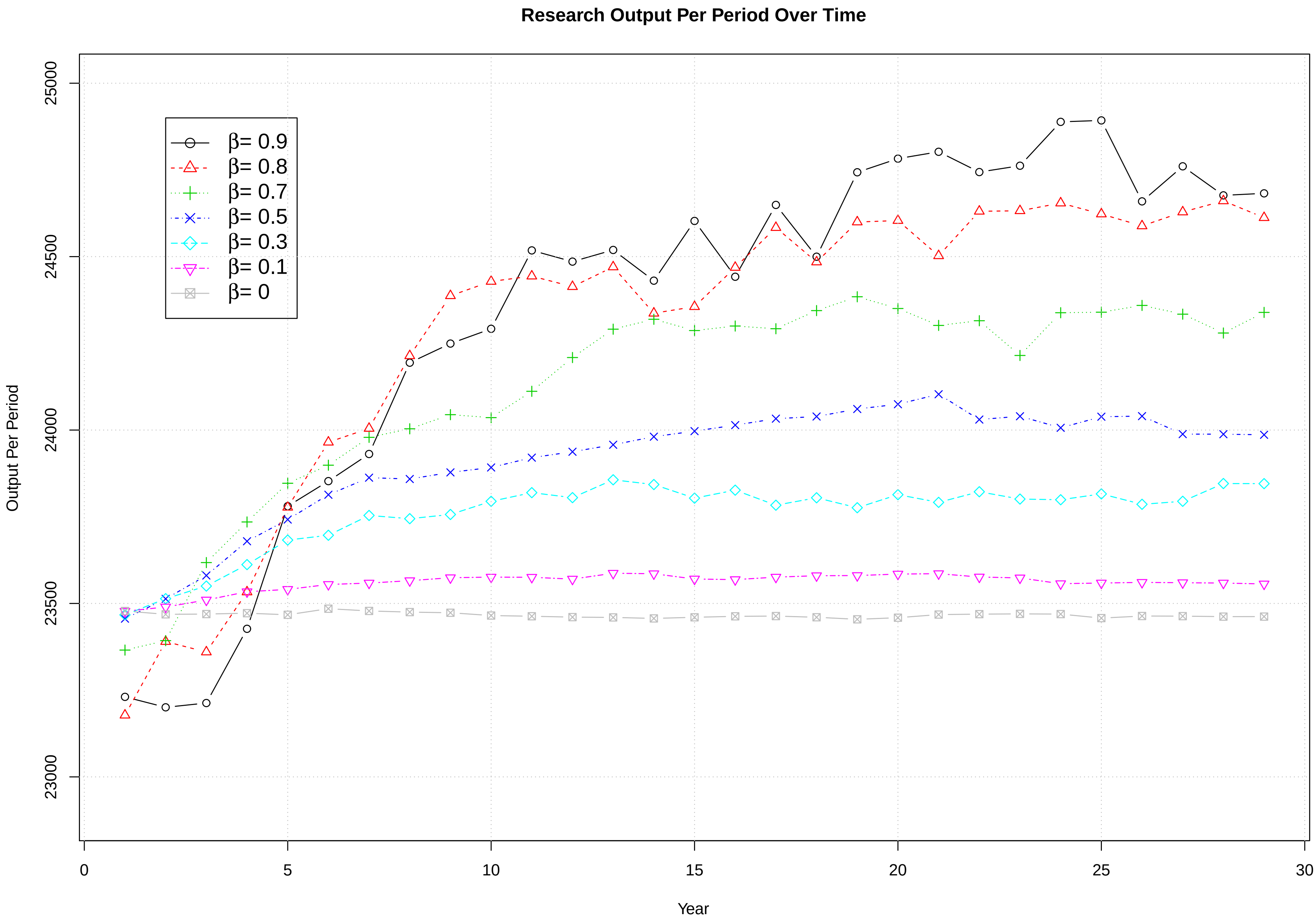
While the scientific community is concerned with the NIH for not funding enough young researchers, I show that the NIH seems to have accounted for some intertemporal tradeoffs. The NIH funds 18% more young researcher as compared to the myopic policy predicted by the model. Thus, the observed behavior is in line with the NIH's mission to fund the

[61] I define young PIs as PIs at the two lowest experience levels, mid-career PIs as PIs at middle three experience levels, and veteran PIs as PIs at the highest two experience levels.

younger generation. However, there is still room to fund more young PIs, as the forward-looking policy with a discount factor of 0.9 would fund 8% more young PIs as compared to the observed funding behavior. Lastly, I show that a temporary 20% funding cut would not only result in an immediate decline in research output, but would also lead to long-lasting negative impacts on research output. This finding resonates with the concerns of the scientific community that the budget cut policy proposed by the current administration can be disastrous for science, medical cure, and public health.

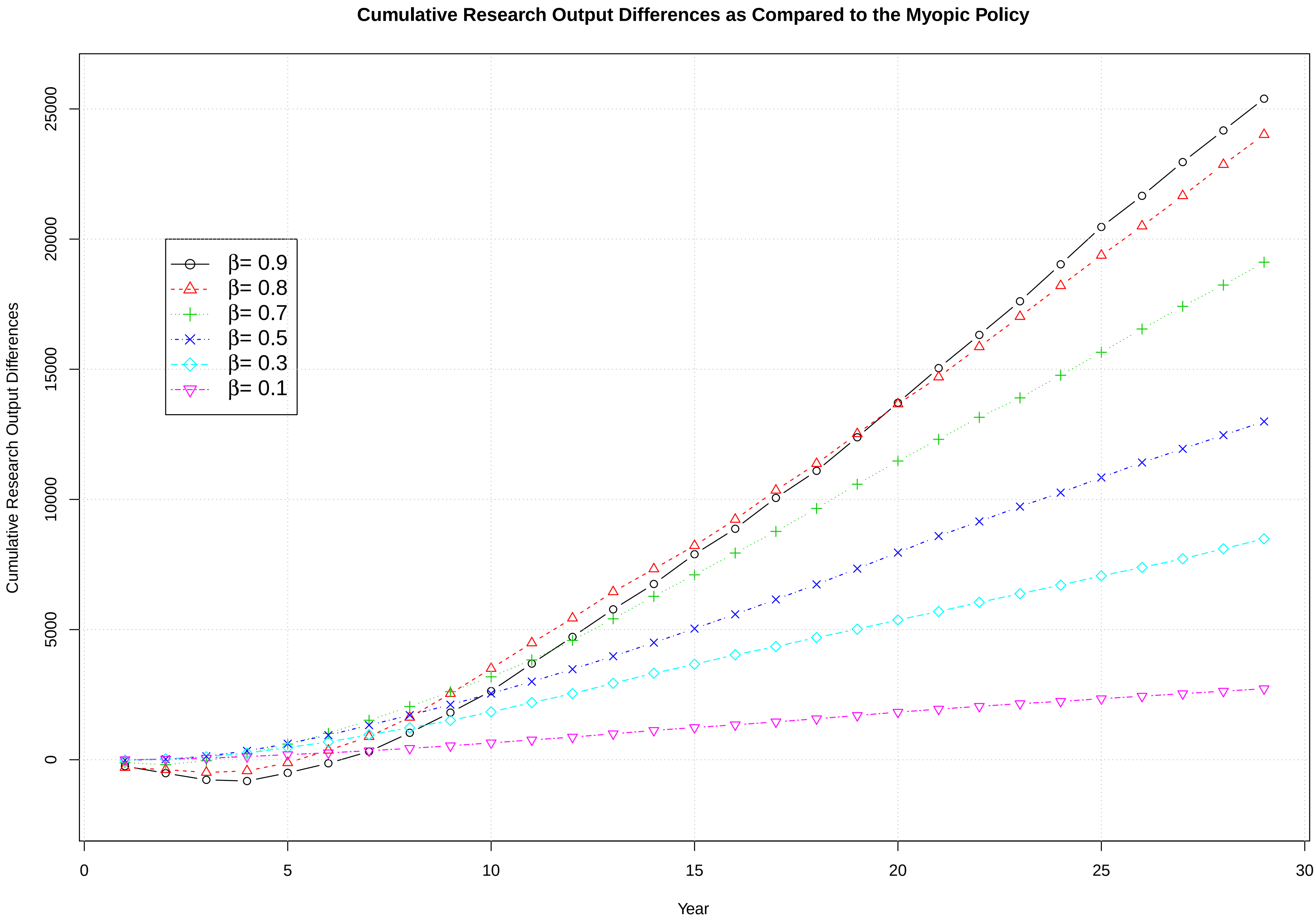
8 Appendix A: Graphs

Figure 4:



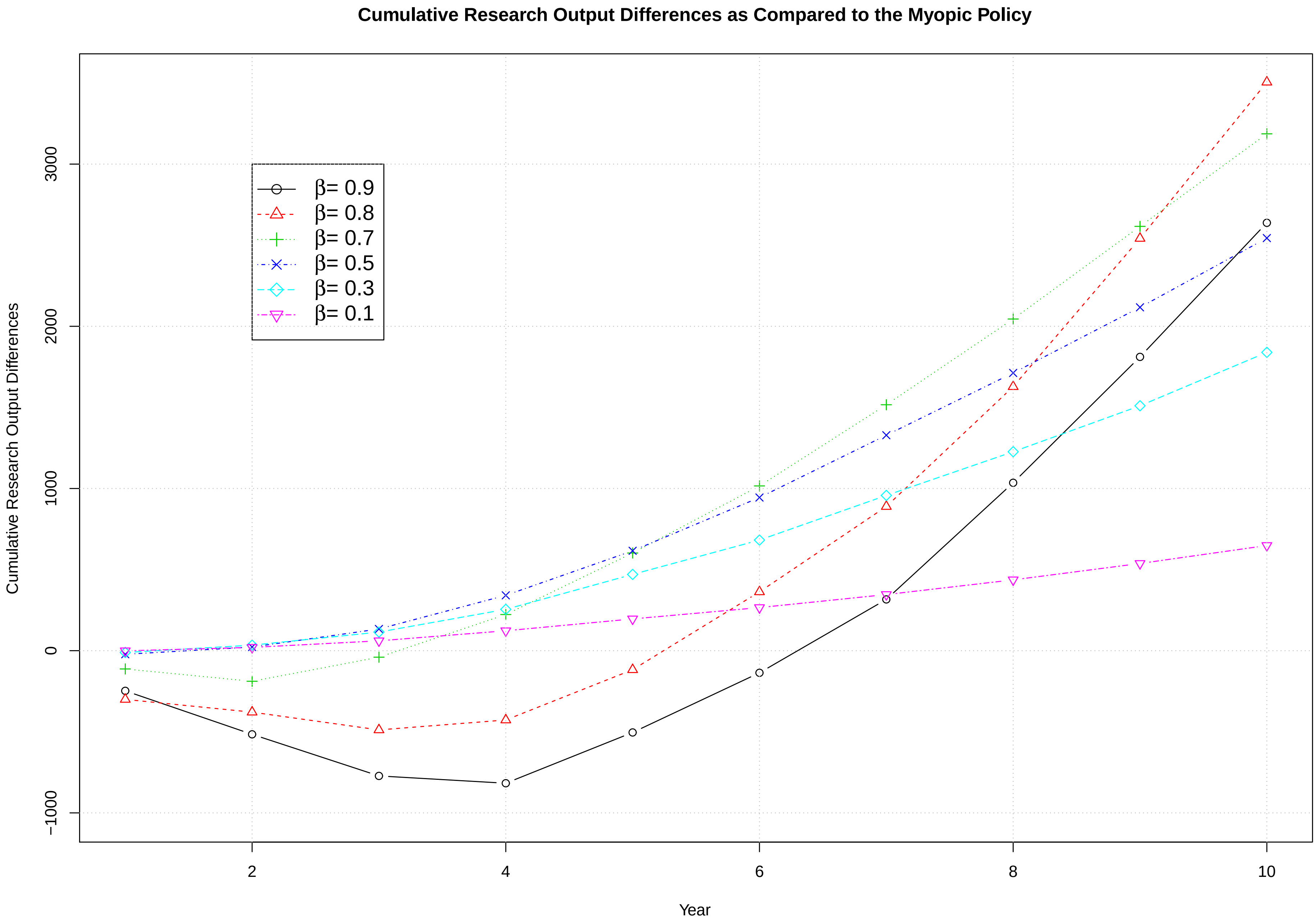
Notes: The mean research output per period by funding policies under different discount factors β across 500 simulations. The initial condition uses the stationary distribution of PIs under the myopic policy.

Figure 5:



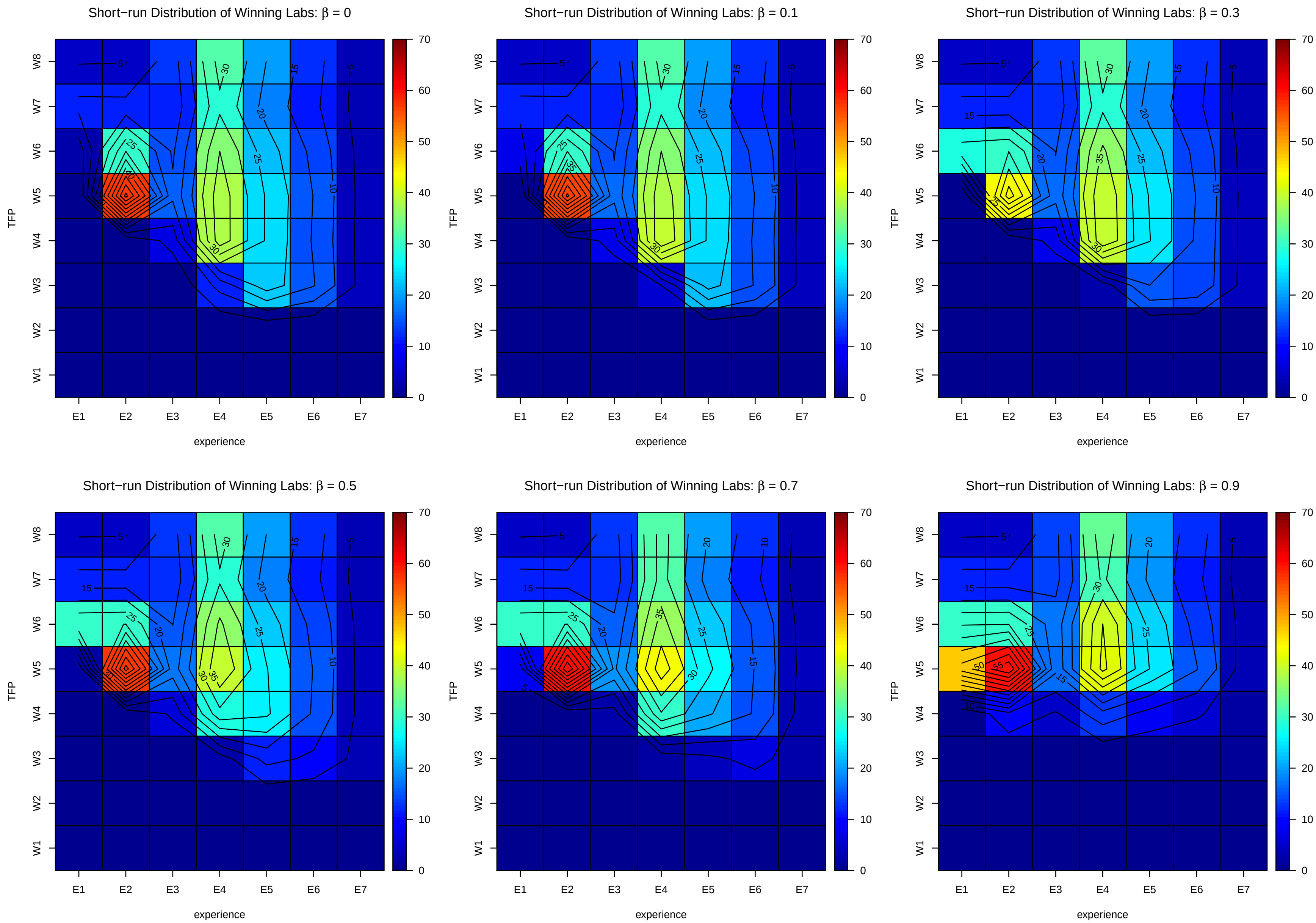
Notes: The mean cumulative output differences between forward-looking policies under different discount factors β as compared to the myopic policy across 500 simulations. The initial condition uses the stationary distribution of PIs under the myopic policy.

Figure 6:



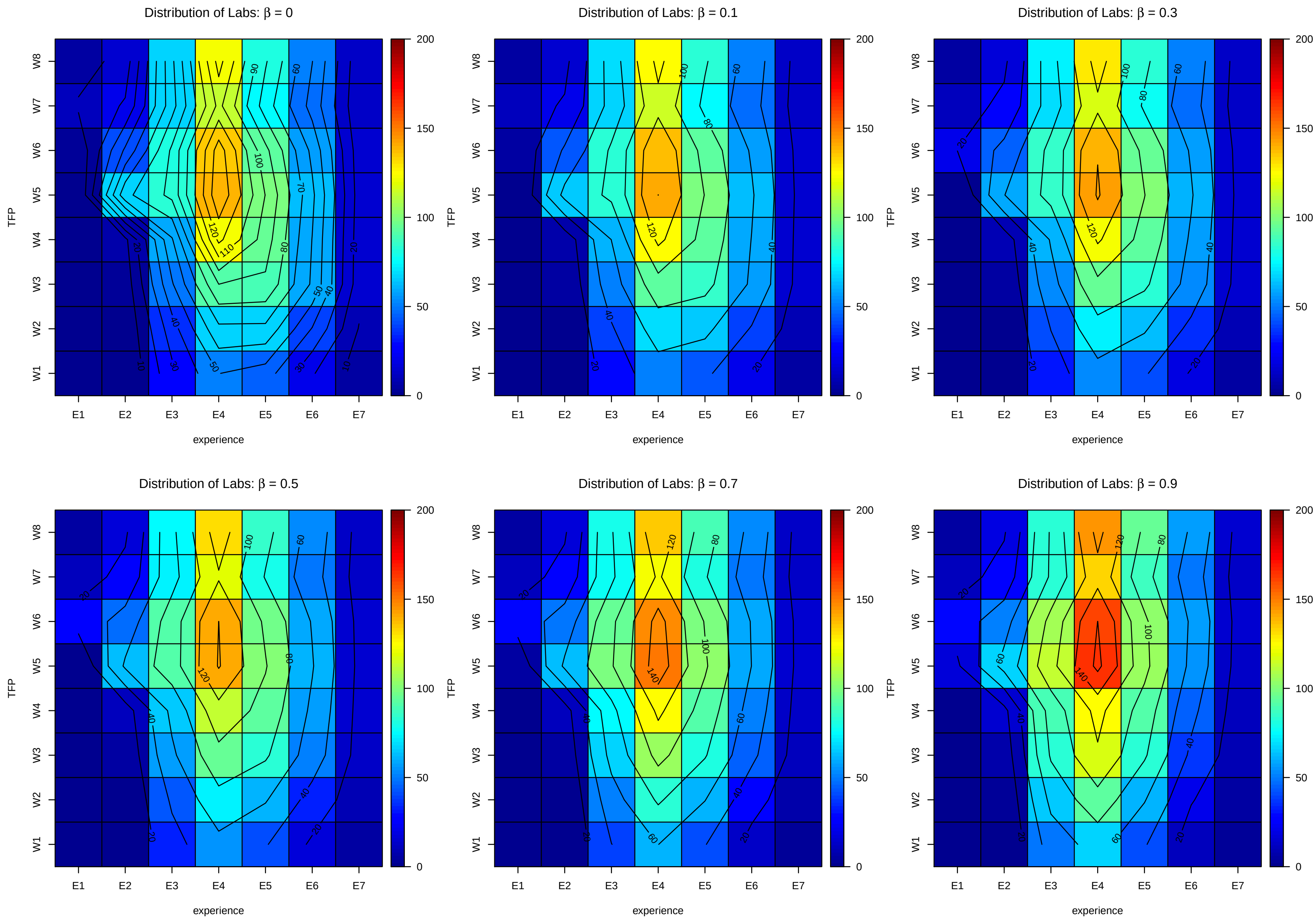
Notes: The mean cumulative output differences between forward-looking policies during the first ten years under different discount factors β as compared to the myopic policy across 500 simulations. The initial condition uses the stationary distribution of PIs under the myopic policy.

Figure 7:



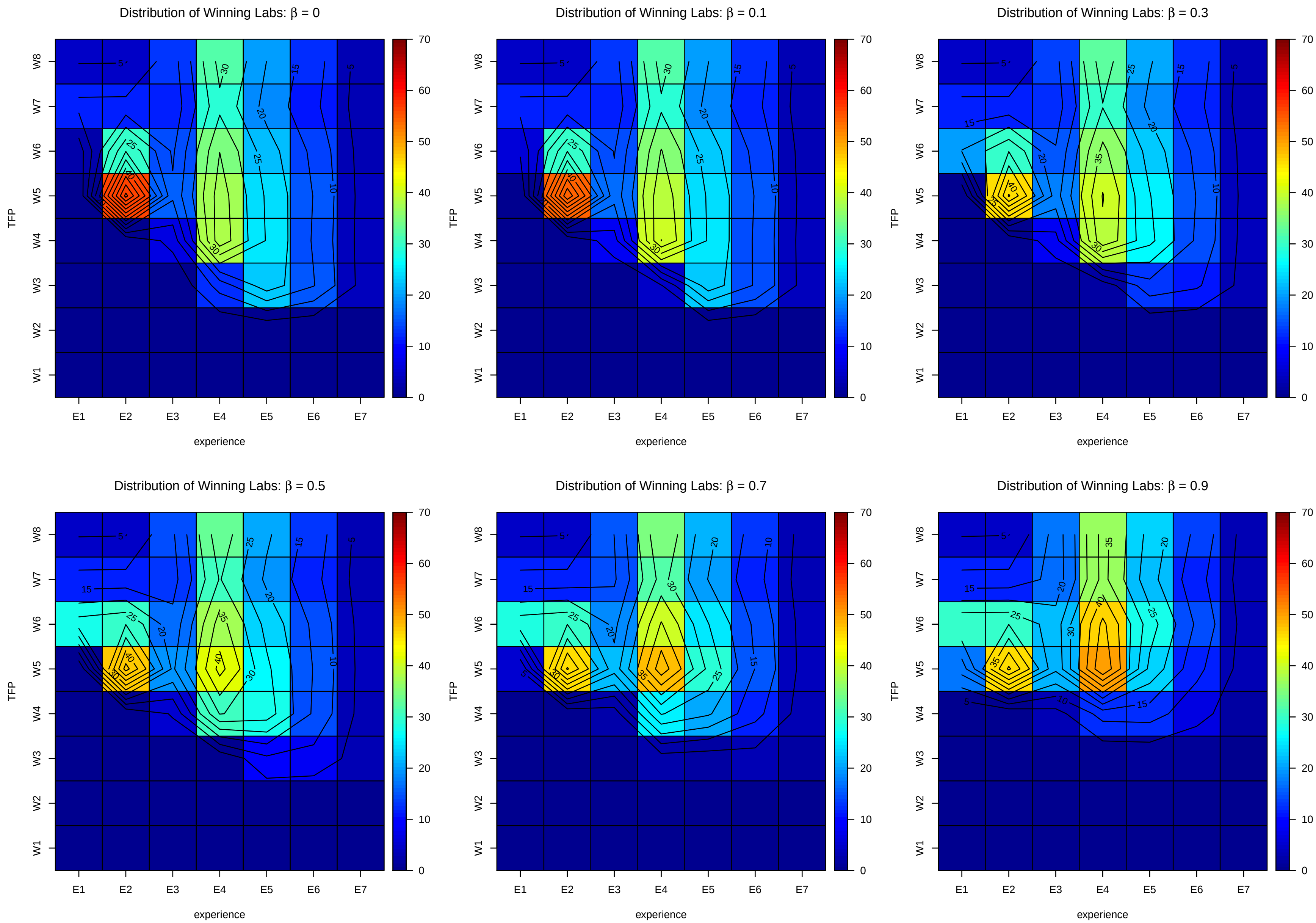
Notes: $W1$ is the lowest TFP level, and $W8$ is the highest TFP level; $E1$ is the lowest experience level, and $E7$ is the highest experience level. Each panel plots the short-run distribution of winning PIs under the corresponding discount factor, showing that the distribution gravitates toward younger and more productive PIs.

Figure 8:



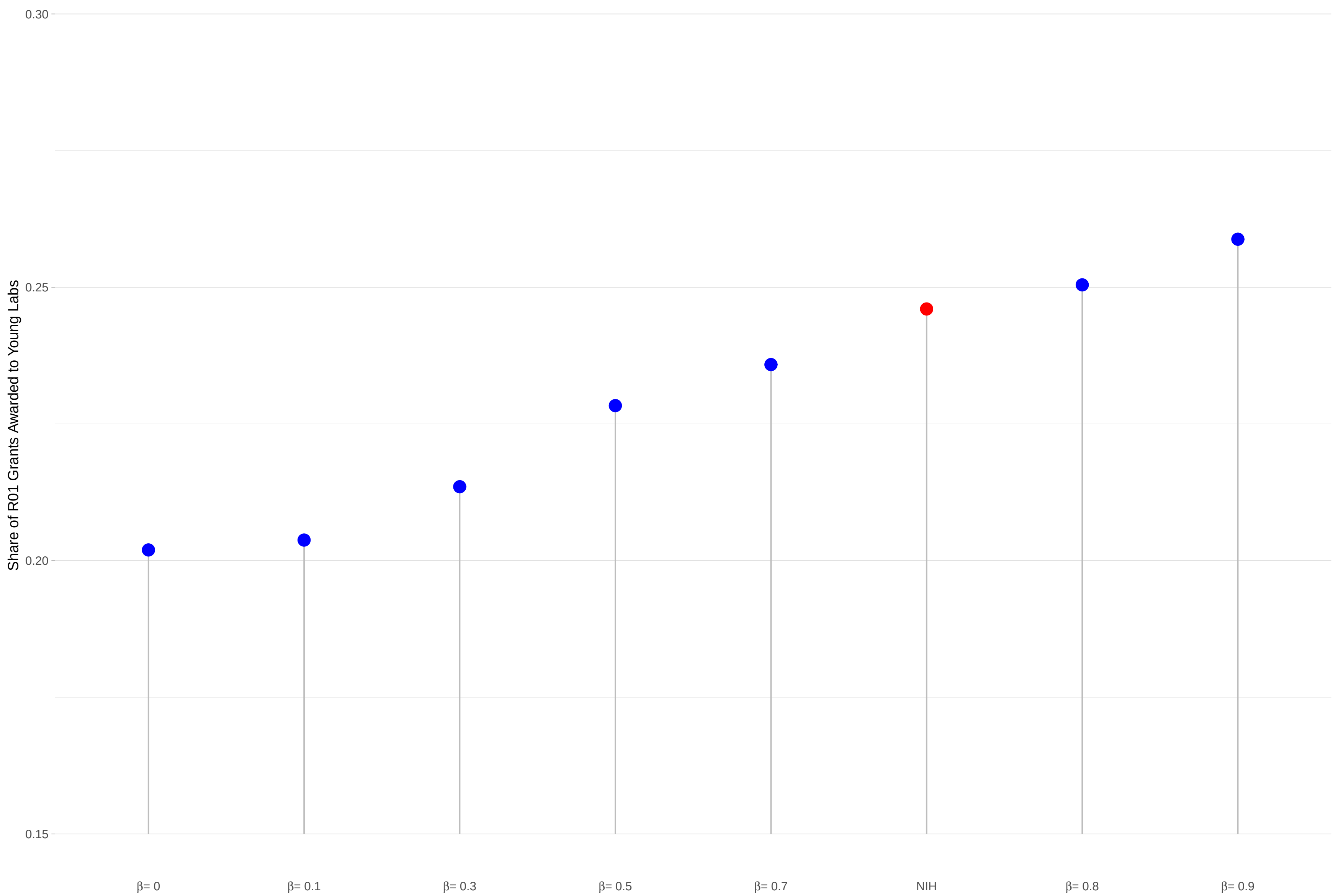
Notes: $W1$ is the lowest TFP level, and $W8$ is the highest TFP level; $E1$ is the lowest experience level, and $E7$ is the highest experience level. Each panel plots the short-run distribution of winning PIs under the corresponding discount factor, showing that the distribution gravitates toward more productive and slightly younger PIs.

Figure 9:



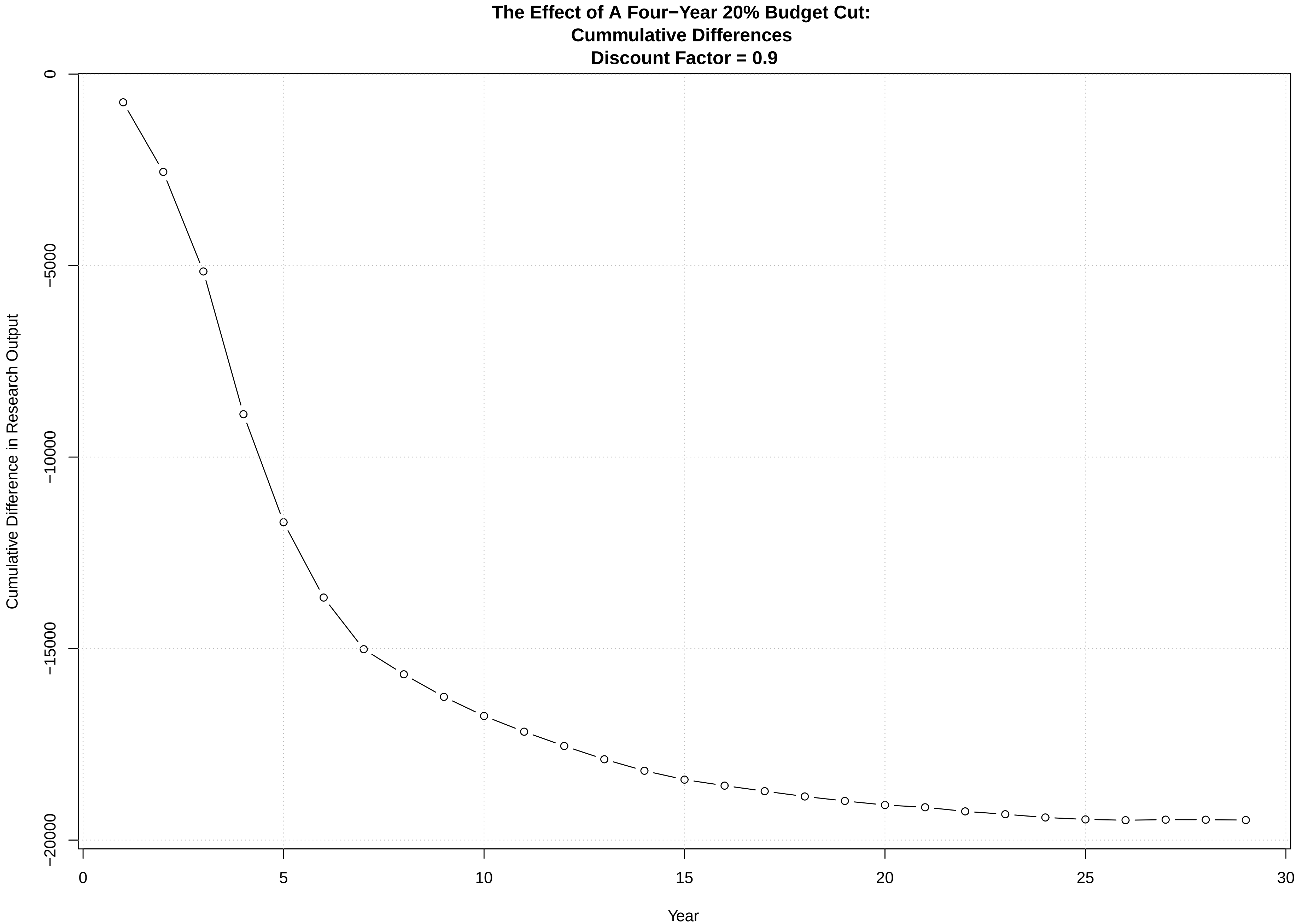
Notes: $W1$ is the lowest TFP level, and $W8$ is the highest TFP level; $E1$ is the lowest experience level, and $E7$ is the highest experience level. Each panel plots the short-run distribution of winning PIs under the corresponding discount factor, showing that the distribution gravitates toward younger and more productive PIs.

Figure 10:



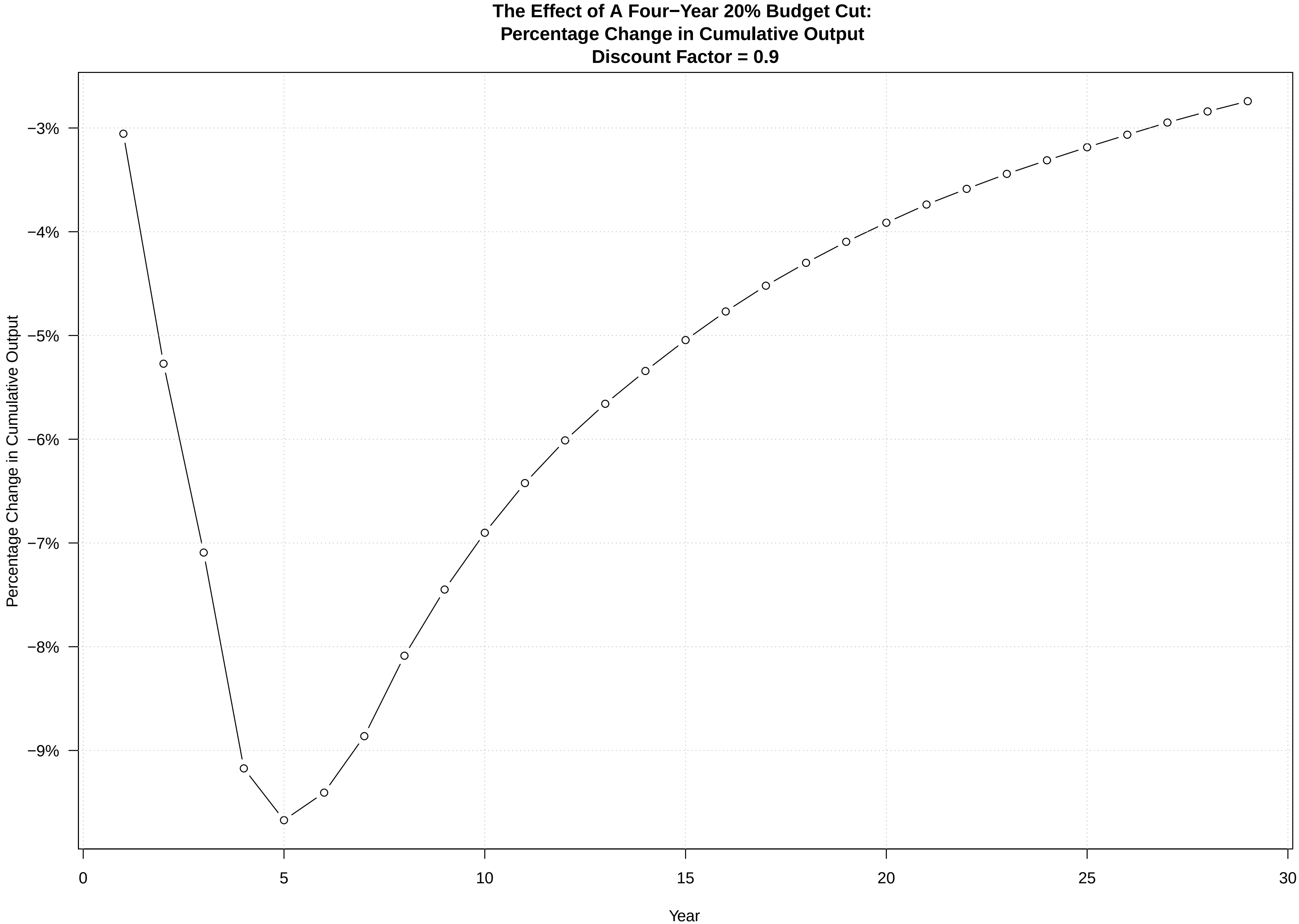
Notes: The share of R01 grants award to young PIs. Young is defined as the entrants, that is, PIs at the first two experience levels. The NIH's share is calculated using statistics published by the NIH's NCI R01 grant archive from 2011 to 2017. The discount factor that rationalizes the NIH's R01 funding behavior is about 0.75.

Figure 11:



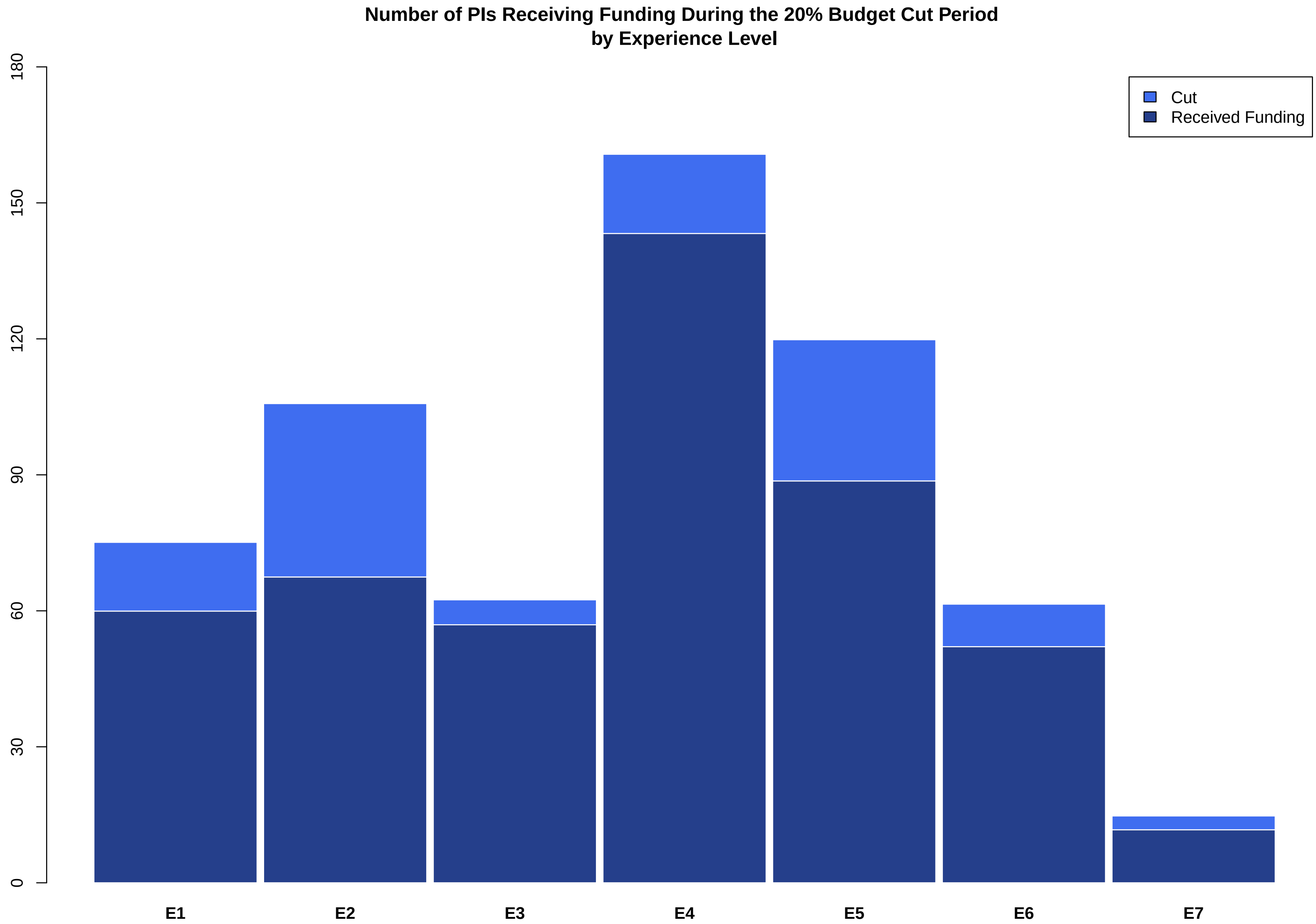
Notes: The effect of a four-year 20% budget cut measured by mean cumulative differences across 500 simulations.

Figure 12:



Notes: The effect of a four-year 20% budget cut as measured by the percentage change in mean cumulative difference across 500 simulations.

Figure 13:



Notes: The number of affected PIs from the four-year 20% budget cut is colored in light blue. $E1$ is the lowest experience level, and $E7$ is the highest experience level.

9 Appendix B: Computation

9.1 Structure of the Subproblem

The approximations (27) and (28) make the Subproblem (24) a simple integer programming problem given unknown parameter θ. To see this, the structure of (24) relies on the post-decision state variables. In the case without the grant cycle, the post-decision state variable is just the action variable $\mathbf{q}_t$. In the case with the grant cycle, the post decision state variable is a function of $\mathbf{N}_t$ and $\mathbf{q}_t$. Furthermore, it is important to notice that the grant cycle and experience dynamics consist of deterministic state transitions.

We can model the post-decision state variable with two steps to accommodate for deterministic state transitions. First, define the transition function

$$\Delta_{x'}(x,d) = \begin{cases} 1 & \text{if applying decision } d \text{ on a type } x \text{ lab transforms the lab into} \\ & \text{type } x' \text{ in the subsequent period just before the realization of the next period,} \\ 0 & \text{otherwise.} \end{cases}$$

Second, write the post-decision state variable as

$$N^q_{x't} = \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} \Delta_{x'}(x,d) q^d_{xt}. \tag{33}$$

In matrix notation,

$$\mathbf{N}^q_t = \Delta \mathbf{q}_t.$$

where Δ is a $d_x \times 2d_x$ deterministic state transition matrix.

Linear Approximation

Under linear approximation of V_t^q, $\tilde{V}_{t-1}$ can be formulated as

$$\begin{aligned}
\tilde{V}_{t-1}(\mathbf{N}_{t-1}^q, r_t) &= \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} \left\{ \sum_{x\in\mathcal{X}, d\in\mathcal{D}} Y_x^d q_{xt}^d + \beta \tilde{V}_t(\mathbf{N}_t^q) \right\} \\
&= \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} \left\{ \sum_{x\in\mathcal{X}, d\in\mathcal{D}} Y_x^d q_{xt}^d + \beta \sum_{x\in\mathcal{X}} \theta_{xt} N_{xt}^q \right\} \\
&= \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} \left\{ \sum_{x\in\mathcal{X}, d\in\mathcal{D}} Y_x^d q_{xt}^d + \beta \sum_{x'\in\mathcal{X}} \theta_{x't} \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} \Delta_{x'}(x,d) q_{xt}^d \right\} \\
&= \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} \left\{ \sum_{x\in\mathcal{X}, d\in\mathcal{D}} Y_x^d q_{xt}^d + \beta \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} \sum_{x'\in\mathcal{X}} \theta_{x't} \Delta_{x'}(x,d) q_{xt}^d \right\} \\
&= \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} \left\{ \sum_{x\in\mathcal{X}, d\in\mathcal{D}} \left(Y_x^d + \beta \sum_{x'\in\mathcal{X}} \theta_{x't} \Delta_{x'}(x,d) \right) q_{xt}^d \right\}.
\end{aligned}$$

In matrix notation, we have

$$\begin{aligned}
\tilde{V}_{t-1}(\mathbf{N}_{t-1}^q, r_t) &= \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(r_t))} \{ \mathbf{Y}' \mathbf{q}_t + \beta \theta' \Delta \mathbf{q}_t \} \\
\text{such that } \mathbf{F}' \mathbf{q}_t &\leq B_t \\
\Upsilon \mathbf{q}_t &= \mathbf{N}_t \\
q_{xt}^d &\in \mathbb{Z}^+ \quad \forall x \in \mathcal{X}, d \in \mathcal{D}
\end{aligned} \tag{34}$$

where $\Upsilon = [diag(d_x), diag(d_x)]$ with $diag(d_x)$ denoting the $d_x \times d_x$ identity matrix. Notice that (34) is an integer programming problem given parameters θ and β.

Piecewise-linear Approximation

Under piecewise-linear approximation of the value function, the Subproblem (24) can be written as

$$\tilde{V}_{t-1}^q(\mathbf{N}_{t-1}^q, r_t) := \max_{\mathbf{q}_t \in \mathcal{Q}(\mathbf{N}_t(\mathbf{N}_{t-1}^q, r_t))} Y(\mathbf{q}_t) + \beta \sum_{x\in\mathcal{X}} \sum_{s=1}^{\bar{S}_x} \theta_{xt}(s) z_{xt}(s) \tag{35}$$

$$\begin{aligned}
\text{such that } N_{x't}^q &= \sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} \Delta_{x'}(x,d) q_{xt}^d \; \forall s \in \mathcal{S}_x, x' \in \mathcal{X}, \\
z_{xt}(s) &= \mathbb{I}\{N_{xt}^q \leq s\} \; \forall s \in \mathcal{S}_x, x \in \mathcal{X}, \\
\theta_{xt}(s) &\geq \theta_{xt}(s+1) \; \forall s < \bar{S}_x.
\end{aligned} \tag{36}$$

In general, the Subproblem as formulated in (35) is difficult to evaluate. Instead, we rewrite the problem as

$$
\begin{aligned}
\tilde{V}^q_{t-1}(\mathbf{N}^q_{t-1}, r_t) &:= \max_{\mathbf{q}_t, \mathbf{z}_t} \sum_{x\in\mathcal{X}, d\in\mathcal{D}} Y^d_x q^d_{xt} + \beta \sum_{x\in\mathcal{X}} \sum_{s=1}^{\bar{S}_x} \theta_{xt}(s) z_{xt}(s) && (37)\\
\text{such that } \sum_{x\in\mathcal{X}} F_x q^1_{xt} &\leq B_t && (38)\\
q^0_{xt} + q^1_{xt} &= N_{xt} \quad \forall x \in \mathcal{X} && (39)\\
\sum_{d\in\mathcal{D}} \sum_{x\in\mathcal{X}} \Delta_{x'}(x,d) q^d_{xt} - N^q_{x't} &= 0 \quad \forall x' \in \mathcal{X} && (40)\\
\sum_{s=0}^{\bar{S}_x} z_{xt}(s) - N^q_{xt} &= 0 \quad \forall x \in \mathcal{X}. && (41)\\
z_{xt}(s) &\in \{0,1\} \quad \forall x \in \mathcal{X}, s \in \mathcal{S} && (42)\\
\theta_{xt}(s) &\geq \theta_{xt}(s+1) \quad \forall x \in \mathcal{X}, s \in \mathcal{S} && (43)
\end{aligned}
$$

This alternative representation is equivalent to the Subproblem as represented in (35). Note that (37) includes additional action variables $(z_{xt}(s))_{x,s}$, and the collection of variables $(z_{xt}(s))_{x,s}$ has to satisfy the constraints (41) and (42). The constraint (41) is a balance constraint, and the constraint (42) allows $z_{xt}(s)$ to take values from 0 to 1. Due to the concavity constraints (43), it is easy to see that $z_{xt}(s) \geq z_{xt}(s+1)$ for all $s \in \mathcal{S}_x$ and $x \in \mathcal{X}$. The Subproblem in this representation is a linear programming problem, which can be solved using standard LP packages.

The problem (35) can be represented in matrix form. Without loss of generality, we assume $\bar{S}_x = \bar{S}$, and let $\mathbf{z}_t$ be a $\bar{S} \cdot d_x$-by-1 zero-one decision vector for all $t \in \mathcal{T}$,

$$
\mathbf{z}_t = [z_{1t}(1), ..., z_{1t}(\bar{S}), ..., z_{xt}(1), ..., z_{xt}(\bar{S}), ..., z_{d_x t}(1), ..., z_{d_x t}(\bar{S})]',
$$

and θ_t be a $\bar{S} \cdot d_x$-by-1 vector for all $t \in \mathcal{T}$,

$$
\theta_t = [\theta_{1t}(1), ..., \theta_{1t}(\bar{S}), ..., \theta_{xt}(1), ..., \theta_{xt}(\bar{S}), ..., \theta_{d_x t}(1), ..., \theta_{d_x t}(\bar{S})]'.
$$

Furthermore, let ι be an $\bar{S}$ dimensional vector of ones, and define

$$
\Xi = -\begin{bmatrix} 1 & 0 & 0 \\ & \ddots & \\ & & 1 \end{bmatrix} \otimes \iota'
$$

where $\otimes$ is the Kronecker product. Finally, let

$$\begin{aligned}
\mathbf{q}_t &= [q^0_{1t}, ..., q^0_{d_x t}, q^1_{1t}, ..., q^1_{d_x t}]' \\
\mathbf{F} &= [F_1, ..., F_{d_x}, 0, ..., 0]' \\
\mathbf{z}_t &= [z_{1t}(1), ..., z_{1t}(\bar{S}), ..., z_{xt}(1), ..., z_{xt}(\bar{S}), ..., z_{d_x t}(1), ..., z_{d_x t}(\bar{S})]' \\
\theta_t &= [\theta_{1t}(1), ..., \theta_{1t}(\bar{S}), ..., \theta_{xt}(1), ..., \theta_{xt}(\bar{S}), ..., \theta_{d_x t}(1), ..., \theta_{d_x t}(\bar{S})]' \\
\mathbf{w}_t &= [\mathbf{q}'_t : \mathbf{z}'_t]' \\
\mathit{\Gamma} &= \begin{bmatrix} \Delta : & \Xi \\ \Upsilon : & \mathbf{0} \end{bmatrix}.
\end{aligned}$$

The matrix representation of a DP resource allocation problem with grant cycle under the piecewise-linear approximation is

$$\begin{aligned}
\max_{\mathbf{w}_t} & \quad \mathbf{Y}'_t\mathbf{q}_t + \beta\theta'_t\mathbf{z}_t \\
\text{such that } \Gamma\mathbf{w}_t &= [\mathbf{0}; \mathbf{N}_t]' \\
\mathbf{F}'\mathbf{q}_t &\leq B_t
\end{aligned}$$

It takes less than a second to solve this problem for more than 1000 types of PIs.

[Note] To restrict when can one apply for a grant, we can simply impose restrictions on constraints $q^0_{xt} + q^1_{xt} = N_{xt} \quad \forall x \in \mathcal{X}$. More precisely, we restrict the number of PI types with non-actionable grant cycles by setting $q^0_{xt} = N_{xt}$. Notice that, in matrix notation, the constraints $q^0_{xt} + q^1_{xt} = N_{xt}$ for all x are indicated in $\Upsilon\mathbf{q}_t = \mathbf{N}_t$. To incorporate the restriction $q^0_{xt} = N_{xt}$ for PI types with non-actionable grant cycles, we can simply set some of the elements in Υ to zeros. This, however, doesn't seem to speed up computation of the subproblem. An alternative is to rewrite the problem including only active PI types.

Define $d_{x^{na}}$ as the number of PI types that are non-actionable, and d_{x^a} as the number of PI types that are actionable. We use the superscripts na and a to indicate non-actionable PIs and actionable PIs respectively. To simplify the action space, let Δ^a be a $d_x \times (2d_{x^a})$ state transition matrix focused only on the actionable types. Let Δ^{na} be the $d_x \times (d_{x^{na}})$ state transition matrix for the non-actionable types. Let $\Upsilon^a = [diag(d_{x^a}) : diag(d_{x^a})]$. Then the restricted DP problem can be written as

$$\begin{aligned}
\max_{\mathbf{w}_t} & \quad \mathbf{Y}'_t\mathbf{q}^a_t + \beta\theta'_t\mathbf{z}_t \\
\text{such that } \begin{bmatrix} \Delta^a : & \Xi \\ \Upsilon^a : & \mathbf{0}^a \end{bmatrix} \mathbf{w}^a_t &= [\Delta^{na}\mathbf{N}^{na}_t; \mathbf{N}^a_t]' \\
\mathbf{F}^a\mathbf{q}^a_t &\leq B_t
\end{aligned}$$

This reduces the number of constraints as well as the dimension of the decision vector. Note also that linear approximation refers to $\bar{S} = 1$.

9.2 Updating Parameters

Piecewise-linear Approximation

The piecewise linear approximation updating scheme is similar to the linear approximation updating scheme, and proceeds by two steps. Let $\hat{v}_{xt}^m$ be defined as (32), and the first step sets

$$u_{xt}^m(s) = \begin{cases} (1-\alpha_m)\theta_{xt}^{m-1}(s) + \alpha_m \hat{v}_{xt}^m & \text{if} \quad s = N_{tx}^{q,m} + 1 \\ \theta_{xt}^{m-1}(s) & \text{if} \quad \text{o.w.} \end{cases}$$

for all x. The second step projects the vector $\mathbf{u}_{xt}^m = [u_{xt}^m(1), ..., u_{xt}^m(\bar{s}_x)]$ onto the set $\mathcal{Z} = \{z \in \mathbb{R}^{\bar{q}} : z_1 \geq z_2 \geq ... \geq z_{\bar{q}}\}$. That is, we set

$$\hat{\theta}_{xt}^m = \arg\min_{\mathbf{z}\in\mathcal{Z}} \|\mathbf{z} - \mathbf{u}_{xt}^m\|_2 .$$

Alternatively, we can make use of sup-norm, i.e.

$$\hat{\theta}_{xt}^m = \arg\min_{\mathbf{z}\in\mathcal{Z}} \|\mathbf{z} - \mathbf{u}_{xt}^m\|_\infty .$$

These projections have closed-form solutions, see Topaloglu and Powell (2006) for more details.

9.3 Concavity

The ADP strategy relies on the concavity of the value function. In this subsection, I provide a sketch of the proof. The planner solves the following linear programming problem in the terminal period T.

$$\begin{aligned} V_T(\mathbf{N}_T) &= \max_{\mathbf{q}_T} \mathbf{q}_T'\mathbf{Y} \\ \text{such that } \mathbf{F}'\mathbf{q} &\leq B \\ \mathbf{q} &\leq \mathbf{N}_T. \end{aligned}$$

In this simple problem, the value function $V_T(\mathbf{N}_T)$ is concave in $\mathbf{N}_T$. To see this, consider resource vector $\mathbf{N}^1$ and $\mathbf{N}^2$ and $\mathbf{N} = \lambda\mathbf{N}^1 + (1-\lambda)\mathbf{N}^2$ for any $\lambda \in (0,1)$. And let $\mathbf{q}^1$ be the optimal solution to $V_T(\mathbf{N}^1)$ and $\mathbf{q}^2$ be the optimal solution to $V_T(\mathbf{N}^2)$. Since $\lambda\mathbf{q}^1 + (1-\lambda)\mathbf{q}^2 \leq \lambda\mathbf{N}^1 + (1-\lambda)\mathbf{N}^2 = \mathbf{N}$, $\mathbf{q} = \lambda\mathbf{q}^1 + (1-\lambda)\mathbf{q}^2$ is a feasible solution to $V_T(\mathbf{N})$. Note that

$V_T(\mathbf{N}^1) = \mathbf{q}^{1\prime}Y$ and $V_T(\mathbf{N}^2) = \mathbf{q}^{2\prime}Y$, so $\lambda V_T(\mathbf{N}^1)+(1-\lambda)V_T(\mathbf{N}^2) = \mathbf{q}'Y$ where $\mathbf{q}$ is a feasible solution to $V_T(\lambda\mathbf{N}^1+(1-\lambda)\mathbf{N}^2)$. Therefore, $V_T(\lambda\mathbf{N}^1+(1-\lambda)\mathbf{N}^2) \geq \lambda V_T(\mathbf{N}^1)+(1-\lambda)V_T(\mathbf{N}^2)$.

Now consider period $T-1$, the planner solves the following problem

$$\begin{aligned} V_{T-1}(\mathbf{N}_{T-1}) &= \max_{\mathbf{q}_{T-1}} \mathbf{q}'_{T-1}Y + \beta\mathbb{E}[V_T(\mathbf{N}_T)|\mathbf{q}_{T-1}] \\ \text{such that } \mathbf{F}'\mathbf{q}_{T-1} &\leq B \\ \mathbf{q}_{T-1} &\leq \mathbf{N}_{T-1}. \end{aligned}$$

In the above formulation, I implicitly assume $\mathbf{q}_{T-1} = \mathbf{N}^q_{T-1}$ for simplicity. Let $\mathbf{q}^1$ be the solution to $V_{T-1}(\mathbf{N}^1)$ and $\mathbf{q}^2$ be the solution to $V_{T-1}(\mathbf{N}^2)$. We want to show $V_{T-1}(\lambda\mathbf{N}^1 + (1-\lambda)\mathbf{N}^2) \geq \lambda V_{T-1}(\mathbf{N}^1) + (1-\lambda)V_{T-1}(\mathbf{N}^2)$.

First, let $V^q(\mathbf{q}) = \mathbb{E}[V_T(\mathbf{N}_T)|\mathbf{q}]$, and I claim that $V^q(\mathbf{q})$ is concave in $\mathbf{q}$. Without loss of generality, let $\mathbf{q}$ and $\mathbf{N}$ to be two-dimensional vectors. Then we can write

$$\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}] = \int_{\varepsilon_1}\int_{\varepsilon_2} V_T([\varepsilon_1 q_1 + \varepsilon_2 q_2, (1-\varepsilon_1)q_1 + (1-\varepsilon_2)q_2])dF_{\varepsilon_1}(\varepsilon_1)dF_{\varepsilon_2}(\varepsilon_2).$$

Let $\mathbf{q} = \lambda\mathbf{q}^1 + (1-\lambda)\mathbf{q}^2$. Then

$$\begin{aligned} \mathbb{E}[V_T(\mathbf{N})|\mathbf{q}] &= \int_{\varepsilon_1}\int_{\varepsilon_2} V_T([\varepsilon_1 q_1 + \varepsilon_2 q_2, (1-\varepsilon_1)q_1 + (1-\varepsilon_2)q_2])dF_{\varepsilon_1}(\varepsilon_1)dF_{\varepsilon_2}(\varepsilon_2) \\ &= \int_{\varepsilon_1}\int_{\varepsilon_2} V_T([\varepsilon_1(\lambda q_1^1 + (1-\lambda)q_1^2) + \varepsilon_2(\lambda q_2^1 + (1-\lambda)q_2^2), \\ &\quad (1-\varepsilon_1)(\lambda q_1^1 + (1-\lambda)q_1^2) + (1-\varepsilon_2)(\lambda q_2^1 + (1-\lambda)q_2^2)])dF_{\varepsilon_1}(\varepsilon_1)dF_{\varepsilon_2}(\varepsilon_2) \\ &= \int_{\varepsilon_1}\int_{\varepsilon_2} V_T([\lambda(\varepsilon_1 q_1^1 + \varepsilon_2 q_2^1) + (1-\lambda)(\varepsilon_1 q_1^2 + \varepsilon_2 q_2^2), \\ &\quad \lambda((1-\varepsilon_1)q_1^1 + (1-\varepsilon_2)q_2^1) + (1-\lambda)((1-\varepsilon_1)q_1^2 + (1-\varepsilon_2)q_2^2))dF_{\varepsilon_1}(\varepsilon_1)dF_{\varepsilon_2}(\varepsilon_2) \\ &\geq \lambda\int_{\varepsilon_1}\int_{\varepsilon_2} V_T([\varepsilon_1 q_1^1 + \varepsilon_2 q_2^1, (1-\varepsilon_1)q_1^1 + (1-\varepsilon_2)q_2^1])dF_{\varepsilon_1}(\varepsilon_1)dF_{\varepsilon_2}(\varepsilon_2) \\ &+ (1-\lambda)\int_{\varepsilon_1}\int_{\varepsilon_2} V_T[\varepsilon_1 q_1^2 + \varepsilon_2 q_2^2, (1-\varepsilon_1)q_1^2 + (1-\varepsilon_2)q_2^2))dF_{\varepsilon_1}(\varepsilon_1)dF_{\varepsilon_2}(\varepsilon_2) \\ &= \lambda\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}^1] + (1-\lambda)\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}^2]. \end{aligned}$$

The inequality goes through because V_T is concave in the vector $\mathbf{N}$.

Second, observe that $\mathbf{q} = \lambda\mathbf{q}^1 + (1-\lambda)\mathbf{q}^2$ is a feasible solution to $\lambda\mathbf{N}^1 + (1-\lambda)\mathbf{N}^2 = \mathbf{N}$.

Then

$$
\begin{aligned}
V_{T-1}(\mathbf{N}) &\geq \mathbf{q}'Y + \beta\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}] \\
&= [\lambda\mathbf{q}^1 + (1-\lambda)\mathbf{q}^2]'Y + \beta\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}] \\
&\geq [\lambda\mathbf{q}^1 + (1-\lambda)\mathbf{q}^2]'Y + \beta\lambda\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}^1] + \beta(1-\lambda)\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}^2] \\
&= \lambda\left(\mathbf{q}^{1\prime}Y + \beta\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}^1]\right) + (1-\lambda)\left(\mathbf{q}^{2\prime}Y + \beta\mathbb{E}[V_T(\mathbf{N})|\mathbf{q}^2]\right) \\
&= \lambda V_{T-1}(\mathbf{N}^1) + (1-\lambda)V_{T-1}(\mathbf{N}^2)
\end{aligned}
$$

References

Ackerberg, D. (2016): "Timing Assumptions and Efficiency: Empirical Evidence in a Production Function Context," *working paper*.

Ackerberg, D. and J. Hahn (2015): "Some Non-Parametric Identification Results using Timing and Information Set Assumptions," *working paper*.

Ackerberg, D., C. Lanier Benkard, S. Berry, and A. Pakes (2007): "Chapter 63 Econometric Tools for Analyzing Market Outcomes," .

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83, 2411–2451.

Adams, J. D. and Z. Griliches (1998): "Research productivity in a system of universities," *Annales d'Economie et de Statistique*, 49/50, 127–162.

Azoulay, P., J. Graff Zivin, D. Li, and B. Sampat (2017): "Public R&D Investments and Private-sector Patenting: Evidence from NIH Funding Rules," *working paper*.

Azoulay, P., J. Graff Zivin, and G. Manso (2012): "NIH Peer Review: Challenges and Avenues for Reform," .

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

Benitez-Silva, H., G. Hall, G. Hitsch, G. Pauletto, and J. Rust (2000): "A Comparison of Discrete and Parametric Approximation Methods for Continuous-State Dynamic Programming Problems," *working paper*.

Benkard, C. L. (2000): "Learning and forgetting: The dynamics of aircraft production," *American Economic Review*, 90, 1034–1054.

——— (2004): "A dynamic analysis of the market for wide-bodied commercial aircraft," *Review of Economic Studies*, 71, 581–611.

Bertsekas, D. P. and J. N. Tsitsiklis (1996): "Neuro-Dynamic Programming (Optimization and Neural Computation Series, 3)," *Athena Scientific*.

Blau, D. M. and B. A. Weinberg (2017): "Why the US science and engineering workforce is aging rapidly," *Proceedings of the National Academy of Sciences*, 114, 3879–3884.

Bloom, N., M. Schankerman, and J. Van Reenen (2013): "Identifying Technology Spillovers and Product Market Rivalry," *Econometrica*, 81, 1347–1393.

Blundell, R. and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 115–143.

——— (2000): "Gmm Estimation With Persistent Panel Data: an Application To Production Functions," *Econometric Reviews*, 19, 1–340.

Brumm, J. and S. Scheidegger (2017): "Using Adaptive Sparse Grids to Solve High Dimensional Dynamic Models," *Econometrica*, 85, 1575–1612.

Budish, E., B. N. Roin, and H. Williams (2015): "Do firms underinvest in long-term research? Evidence from cancer clinical trials." *The American economic review*, 105, 2044–2085.

Collins, F. (2010): "Scientists need a shorter path to research freedom," .

——— (2017): "New NIH Approach to Grant Funding Aimed at Optimizing Stewardship of Taxpyaer Dollars," .

——— (2018): "Testimony on Prioritizing Cures: Science and Stewardship at the National Institutes of Health," .

Collins, F. S. (2015): "Exceptional opportunities in medical science: A view from the National Institutes of Health," .

Cook, I., S. Grange, and A. Eyre-Walker (2015): "Research groups: How big should they be?" *PeerJ*, 3, 1–13.

Crawford, G. S. and M. Shum (2005): "Uncertainty and learning in pharmaceutical demand," *Econometrica*.

Doraszelski, U. and J. Jaumandreu (2013): "R and D and productivity: Estimating endogenous productivity," *Review of Economic Studies*, 80, 1338–1383.

Fang, F. C. and A. Casadevall (2016): "Research funding: The case for a modified lottery," *mBio*, 7.

Fortin, J. M. and D. J. Currie (2013): "Big Science vs. Little Science: How Scientific Impact Scales with Funding," *PLoS ONE*, 8.

Freeman, R. B. and J. Van Reenen (2008): "Be careful what you wish for: A cautionary tale about budget doubling," .

Griliches, Z. (1979): "Issues in assessing the contribution of research and development to productivity growth," *The Bell Journal of Economics*, 10, 92–116.

Griliches, Z. and J. A. Hausman (1986): "Errors in variables in panel data," *Journal of Econometrics*.

Igami, M. (2018): "Artificial Intelligence as Structural Estimation: Economic Interpretations of Deep Blue, Bonanza, and AlphaGo," *Yale University Working Paper*.

Jacob, B. A. and L. Lefgren (2011a): "The impact of NIH postdoctoral training grants on scientific productivity," *Research Policy*.

——— (2011b): "The impact of research grant funding on scientific productivity," *Journal of Public Economics*, 95, 1168–1177.

Johnson, J. A. (2018): "NIH Funding: FY1994-FY2019," *Congressional Research Service Report*.

Jones, B. F. (2010): "Age and Great Invention," *Review of Economics and Statistics*, 92, 1–14.

Jones, B. F. and B. A. Weinberg (2011): "Age dynamics in scientific creativity," *Proceedings of the National Academy of Sciences*, 108, 18910–18914.

Kaiser, J. (2002): "NIH Budget Researches Doubling Point," *Science*.

Keane, M. P. and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*.

Lee, R. S. (2013): "Vertical integration and exclusivity in platform and two-sided markets," *American Economic Review*.

Levin, S. G. and P. E. Stephan (1991): "Research Productivity Over the Life Cycle: Evidence for Academic Scientists," *The American Economic Review*, 81, 114–132.

Levinsohn, J. and A. Petrin (2003): "Estimating production functions using inputs to control for unobservables," *Review of Economic Studies*, 70, 317–341.

Li, D. (2017): "Expertise versus Bias in evaluation: Evidence from the NIH," *American Economic Journal: Applied Economics*, 9, 60–92.

Li, D. and L. Agha (2015): "Big names or big ideas: Do peer-review panels select the best science proposals?" *Science*, 348, 434–438.

Macilwain, C. (1998): "Biomedical Research Shines in Clinton Budget Bonanza," *Nature*, 521–522.

March, J. G. (1991): "Exploration and Exploitation in Organizational Learning," *Organization Science*.

Merton, R. K. (1968): "The Matthew Effect in Science: The reward and communication systems of science are considered," *Science*, 159, 56–63.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263.

Pakes, A. and P. McGuire (2001): "Stochastic algorithms, symmetric markov perfect equilibrium, and the 'curse' of dimensionality," *Econometrica*.

Powell, W. B. (2010): "Merging AI and OR to solve high-dimensional stochastic optimization problems using approximate dynamic programming," *INFORMS Journal on Computing*.

——— (2011): *Approximate Dynamic Programming: Solving the Curses of Dimensionality: Second Edition*.

Preston, A. E. (2004): *Leaving Science: Occupational Exit from Scientific Careers*.

Rust, J. (1997): "Using Randomization to Break the Curse of Dimensionality," *Econometrica*, 65, 487.

Stephan, P. (1996): "The Economics of Science," *Journal of Economic Literature*, XXXIV, 1199–1235.

——— (2012): *How Economics Shape Science*, Harvard University Press.

Tabakovic, H. and T. Wollmann (2016): "The Impact of Money On Science: Evidence from Unexpected NCAA Football Outcomes," *Working Paper*.

Topaloglu, H. and W. B. Powell (2006): "Dynamic-Programming Approximations for Stochastic Time-Staged Integer Multicommodity-Flow Problems," *INFORMS Journal on Computing*, 18, 31–42.

Zerhouni, E. (2006): "NIH in the Post-Doubling Era: Realities and Strategies," *Science*, 314.