# Productivity Beliefs and Efficiency in Science

Fabio Bertolotti
Bank of Italy
& CEPR

Kyle R. Myers
Harvard University
& NBER

Wei Yang Tham
University of
Toronto

June 12, 2026

What is the distribution of productivity across researchers? How efficiently are resources allocated in science? We answer these questions with a new method that overcomes challenges that have long plagued economic analyses of science. Our model of researchers' labor supply shows how their willingness to pay for inputs reveals their subjective productivity beliefs. We estimate the model using a nationally representative survey of academic researchers, and we find productivity beliefs to be highly skewed. Counterfactuals show that reallocating the observed research budget more efficiently would, at the population level, be equivalent to billions of dollars in additional funding.

Fabio Bertolotti: fabiobertolotti@hotmail.it. Kyle Myers: kmyers@hbs.edu. Wei Yang Tham: weiyang.tham@rotman.utoronto.ca. This project received financial support from The Alfred P. Sloan Foundation and the Harvard Business School. There are no financial relationships or other potential conflicts of interest that apply to the authors. This project would not have been possible without the excellent work of Rachel Mural, Nina Cohodes, and Yilun Xu and input from Marie Thursby, Jerry Thursby, and Karim Lakhani. We received helpful comments from Stefano Baruffaldi, Nick Bloom, Judy Chevalier, Matt Clancy, Zoë Cullen, Shane Greenstein, Matt Grennan, Daniel Gross, Bronwyn Hall, Sabrina Howell, Bruce Kogut, Petra Moser, David Rivers, Jesse Shapiro, Tim Simcoe, Chad Syverson, Joel Waldfogel, as well as seminar participants at Columbia, JOMT, MIT, the Workshop on the Organisation Economics and Policy of Scientific Research, the Banff Productivity Innovation and Economic Growth Conference, the NBER Productivity Lunch, Boston University TPRI, IIOC, Washington University in St. Louis, the Conference on the Economics of Innovation in Memory of Zvi Griliches, and Catholic University of Milan.

# 1. Introduction

Science has continuously proven to be the engine of economic growth.[1] Academic science, in particular, is now a $100 billion engine for the United States (e.g., Azoulay et al. 2019; Fleming et al. 2019; Gibbons and NCSES 2024). However, there is growing evidence that science is getting harder (Jones 2009; Bloom et al. 2020), and there are a growing number of debates about the functioning of scientific institutions.[2] Understanding productivity and efficiency in science has become increasingly important.

A long line of work, which we review below, has documented instances of allocative inefficiencies in science; researchers are denied resources for reasons unrelated to their productivity. While it is clear we can do better at allocating resources, it is unclear *how much* better we can do. How much more knowledge could be produced if we improved the allocation of resources to best support the most talented scientists? Answering this question first requires answering another question: what is the distribution of marginal returns across researchers? In this paper, we formalize and answer both of these questions.

Progress on these questions has been slow because estimating productivity in science poses some major challenges. First, the output of science is difficult to observe; how does one quantify a unit of knowledge? Bibliometric data is popular; however, as noted by Adams and Griliches (1998), "*what constitutes a scientific paper makes for an elastic yardstick of scientific achievement*." Second, scientific inputs are rarely allocated via price mechanisms and researchers face large adjustment costs (Myers 2020; Baruffaldi and Gaessler 2025). These facts violate key assumptions at the core of most conventional methods for productivity estimation.

To overcome these challenges, we develop a method for estimating productivity without data on output quantities or input prices. As in conventional methods, we leverage the fact that producers' input choices reveal information about their productivity. However, our approach involves a new step: the direct solicitation of producers' willingness to pay (WTP) for inputs.

We assume that each individual researcher derives indirect utility from their output and other sources of consumption, and we allow their productivities, technologies, and preferences to be idiosyncratic. Individuals have earnings (i.e., salaries), which they can use for consumption or, in our experiments, to purchase inputs from us, the experimenters.

[1] For motivating theory, see: Jones (2005). For a recent review, see: Bryan and Williams (2021).
[2] For coverage in popular press, see: Rohn (2017); Ioannidis (2018); Piper (2019); Mills (2021).

The key is that, conditional on individuals' earnings, technologies, and preferences, one producer will only have a higher WTP compared to another if they believe they are more productive. Thus, the variation in WTP for inputs can then be used to recover the distribution of individual-level productivity and, through counterfactuals, quantify the value from more efficiently allocating inputs.

In order to apply this method to the market for science, we first present a new model of researchers' labor supply and consumption decisions. Researchers have heterogeneous preferences and derive utility from three sources: their salary, their scientific output (i.e., the quantity of knowledge they produce), and their leisure time. An initial contract specifies their salary, administrative duties, and guaranteed funding. Given their contract, researchers choose how to allocate their time across fundraising (e.g., writing grants)[3], research, and leisure given their beliefs, budgets, and constraints. Each researcher has a unique production function, defined by two researcher-specific parameters: (*i*) a funding-intensity parameter that determines the relative weight of the two key inputs, funding and time, and (*ii*) a productivity parameter that describes the efficiency with which researchers produce scientific output using their inputs.

To identify researchers' WTP for inputs, we use the logic of compensating variation to solve for the salary change that would leave the researcher indifferent to alternative contracts, where these alternatives specify counterfactual amounts of guaranteed research funding or administrative duties. This variation in salary — our measure of researchers' WTP for inputs — depends on productivity because the expected change in scientific output, given the counterfactual funding or duties, depends on the researcher's productivity beliefs.

In order to generate the data necessary to estimate the model, we make use of a nationally representative survey of research-active professors across all major fields of science at roughly 150 major institutions of higher education in the US; for details, see Myers et al. (2023).[4] The survey solicits researchers' salaries, time allocations, and access to inputs. Importantly, the survey also includes a series of hypothetical experiments that solicit researchers' willingness to trade off their salary for more research funding or fewer administrative duties. Throughout the paper, we define scientific markets as the interaction of fields of study and career-stage. We focus on variation and counterfactuals *within* these markets to avoid heterogeneity that is beyond the scope of our model.

[3] We allow for heterogeneity in researchers' ability to fundraise.

[4] Evidence is provided in Myers et al. (2023) suggesting the presence of non-response bias in the sample is very low on observable dimensions such as institutional rank, grant funding, and publication rates; some of this is reprinted in the Appendix.

Researchers' willingness-to-pay responses are of plausible magnitudes and behave as expected. The median researcher is willing to pay 10 cents for $1 of research funding and $68 for one less hour of administrative duties. The components of the model explain a large fraction of the variation in researchers' responses (approximately 90%), and we find little evidence that the responses are driven by sample selection or systematic noise. Furthermore, researchers' willingness to pay for free time is strongly correlated with their implied hourly wages as expected.

As a sign of face validity, our model-based estimates of scientific output are positively correlated with popular bibliometric proxies for output (e.g., citations) even after conditioning on input levels. Interestingly, our estimates indicate that approximately 38.2% of the variance in scientific output across individual researchers is due to variance in their productivity.[5]

Our estimates of productivity beliefs vary widely across researchers, even after accounting for outliers. Within markets, defined by field of study by career-stage intersections, the ratio of the $90^{\text{th}}$ and $10^{\text{th}}$ percentiles of productivity is approximately 8.1. This represents a high degree of dispersion compared to what is observed in commercial markets at the firm level (e.g., Syverson 2011; De Loecker and Syverson 2021); however, this scale of dispersion in productivity across individual workers has been observed in some high-skilled settings (e.g., Sackman, Erikson, and Grant 1968). The high degree of dispersion suggests that it may be hard for the market to facilitate positive selection on productivity.

To evaluate allocative efficiency, we compare observed allocations to two alternatives. The first is a first-best benchmark, where a planner directly assigns funding and research time within each market to maximize the social value of scientific output. The second is a uniform funding policy simulation, where researchers within the same market face the same guaranteed funding and fundraising ability and then re-optimize their choices. While this allocation is likely suboptimal, it requires no knowledge of any individual researcher's productivity, and it is useful for assessing the importance of heterogeneous "fundraising skill" or "grantsmanship" in determining current (in-)efficiencies.

We find that the observed allocation generates roughly half of the social output that the same aggregate resources could produce: moving to the first-best raises social output by 97%. The uniform funding policy captures a substantial share of these gains, closing

[5] We are unable to distinguish the degree to which our productivity estimates reflect an individual's latent capabilities versus any cumulative advantage they have acquired (c.f., Hall and Mairesse 2024).

roughly half of the gap between the observed allocation and the first-best. To provide a more tangible valuation of the gains from reallocation, we ask how much funding would need to increase under the observed allocation structure to match the output gains from these counterfactuals. Matching the first-best would require increasing research budgets by roughly +168%, which, on the scale of our focal population, amounts to more than $32.30B per year. Across a range of robustness tests and alternative specifications, we persistently find large implied gains from reallocation on the scale of billions of dollars per year, including counterfactuals where we only improve allocative efficiency to be as good as we observe in-sample.

Finally, in our more speculative analyses, we relax the within-market focus and compare markets of science to one another. Allocative efficiency ranges from roughly 39% in the least efficient fields to roughly 60% in the most efficient, and the markets that allocate resources most efficiently tend to be those with unequal funding guarantees but homogeneous costs of fundraising. We also use the model to engage with across-market allocations. Unlike most prevailing debates about funding that cannot separate the extent to which aggregate input level variation across fields is due to differences in productivity (e.g., one field is more productive than another) versus demand (e.g., one field has larger positive externalities than another), we can. Our approach holds aggregate productivity fixed and allows us to recover the welfare weights that would rationalize the observed allocation of resources across fields. These estimates provide a new starting point for debates about the allocation of funding across science more broadly.

Our estimates come with some important caveats. First, we face a set of challenges common to all existing productivity estimation techniques.[6] Second, we face some unique limitations: (*i*) our method provides only ex-ante productivity beliefs and has no way of identifying ex-post differences in production; and (*ii*) the identifying variation in our data is based on stated preferences from hypothetical experiments. Throughout the paper we engage with these limitations by providing robustness tests of our assumptions and face validity tests of the data. Furthermore, we implement a novel approach to measurement error that leverages Bayesian shrinkage logic. Overall, the results from these tests give us confidence in our conclusions. Still, we interpret our results as plausible upper bounds on the productivity dispersion and gains from reallocation in science. Given the dearth of quantitative evidence in this area, we view our estimates and methodology as valu-

[6] Namely, (*i*) noise due to mismeasurement or misspecification; and (*ii*) the presence of unmodeled heterogeneity in output prices (or preferences over payoffs).

able inputs for future theoretical and empirical studies in the economics of science and innovation.

After a brief review of our connection to the literature, the rest of the paper proceeds as follows: Section 2 details the model of researchers' production and consumption as well as the framework and key assumptions underpinning our methodology for identifying the parameters of the model; Section 3 details the survey data used in estimation; Section 4 describes and reports the results of the survey experiments at the core of our identification strategy; Section 5 provides the model estimates; Section 6 contains our counterfactual allocation exercises; and Section 7 concludes with a discussion of our results and the usefulness of our methodology more generally.

**Related Literature**

Our paper sits primarily at the intersection of two bodies of literature: empirical studies of science and markets for innovation (i.e., Merton 1973; Stephan 1996; Bryan and Williams 2021); and economic studies of producers' productivity and factor misallocation (i.e., Syverson 2011; Restuccia and Rogerson 2017; De Loecker and Syverson 2021). More generally, we are focused on a classic challenge in labor economics of estimating worker-level productivity metrics (i.e., Hoffman and Stanton 2024).

The meta-science literature has long been interested in misallocation. Following a long line of sociological work (Merton 1973; Zuckerman 1988; Shapin 1995) economists have quantified some status-based frictions (Azoulay, Fons-Rosen, and Graff Zivin 2019) and have also studied potential frictions such as: political lobbying (Hegde and Sampat 2015), information asymmetries regarding researchers' output (Hager, Schwarz, and Waldinger 2024), and competitive pressures from priority-based credit mechanisms (Hill and Stein 2025). However, the scientific production function is rarely formalized in a way that yields the structural parameters necessary for quantifying productivity dispersion or allocative efficiency. Rosenbloom et al. (2015) finds the marginal returns to scientific funding (per bibliometrics) are relatively similar across US universities, which is suggestive of allocative efficiency. Taking a more dynamic view, Qiu (2023) estimates the discount rate that rationalizes funding allocations at the US National Institutes of Health and finds a relatively high discount rate, which suggests young scientists may be under-funded.

Within the economics of science literature, our work pairs particularly well with Adda and Ottaviani (2024), who show both theoretically and empirically how uncertainty in the fundraising process (i.e., noise in peer review for grants) combined with heterogeneous

productivity can influence researchers' decisions about fundraising and, in turn, influence the allocation of resources across scientific fields. And while our model is deterministic, the funding guarantees and fundraising abilities that researchers take as given in our framework can be read as the reduced form of exactly these equilibrium forces: each researcher's expected return to an hour of fundraising, as shaped by the evaluation mechanisms they face. Adda and Ottaviani (2024) characterize how the design of evaluation generates heterogeneous returns to fundraising; we estimate the researcher-level heterogeneity directly and quantify the allocative costs. See Carnehl, Ottaviani, and Preusser (2025) for an excellent, more general review of the economic theories relevant to resource allocation mechanisms in science.

Looking towards the productivity literature, our methodology is centered on understanding producers' factor demand, which follows the logic of identification based on input cost shares (e.g., Hsieh and Klenow 2009) or the inversion of input demand into control functions (e.g., Olley and Pakes 1996; Levinsohn and Petrin 2003; Ackerberg, Caves, and Frazer 2015). Much work has been done to extend these methods to account for features such as unobservable prices (De Loecker et al. 2016), adjustment costs (Petrin and Sivadasan 2013; Asker, Collard-Wexler, and De Loecker 2014), as well as measurement error and heterogeneity (Kim, Petrin, and Song 2016; Bils, Klenow, and Ruane 2021; Gollin and Udry 2021). To our knowledge, we are the first to formalize an approach for estimating production functions without data on output quantities and input prices.

We also follow a growing body of work using surveys to study productivity in settings where inputs and outputs are subjective or difficult to measure (e.g., Bloom, Sadun, and Van Reenen 2012; Atkin, Khandelwal, and Osman 2019). We are not the first to solicit producers' WTP for inputs (c.f., Cole et al. 2013; Wossen et al. 2024); however, prior work typically has an inherent interest in producers' demand for a specific input. Our work also runs parallel to the development of methods that use subjective expectations for estimating production functions (e.g., Keiller, de Paula, and Van Reenen 2024) and (quasi-)experimental variation to estimate misallocation (e.g., Sraer and Thesmar 2023; Carrillo et al. 2023).

## 2. Methodological Framework and Model of Science

In this section, we show how a researcher's expectations about their productivity can be estimated by using information about their WTP for a fixed amount of inputs. After an illustrative model designed to deliver intuition of our methodology, we describe our

specific model of researchers' labor supply, including functional form assumptions, in Section 2.2, and then we discuss the identification and estimation in Section 2.3. Additional details regarding the model and estimation are contained in Appendix A.

### 2.1. Illustrative Example

To illustrate our approach, the following simple example shows how soliciting researchers' WTP for an input can be used to estimate their productivity.

Utility maximizing researchers are indexed by $i$ and are endowed with expected income, $M$ (e.g., a guaranteed salary). Researchers produce research output, $Q$, using a single variable input, $X$ (e.g., research funding). Researchers obtain the input by incurring heterogeneous, convex costs given by: $X_i^{c_i}$, where $c_i > 1$. Such convex costs summarize potential frictions identified in the meta-science literature. Researchers produce research output according to a linear research production function: $Q_i = \alpha_i \times X_i$, where $\alpha_i$ is the focal productivity parameter. Researchers' utility function is linear in income and research output less costs. They maximize utility by choosing $X_i$:

$$\max_{X_i} \; M_i + b_i \times Q_i - X_i^{c_i} \quad \text{subject to} \quad Q_i = \alpha_i \times X_i \,,$$

where $b$ is the marginal utility of research output relative to income, the numeraire. We cannot observe the productivity or cost parameters $(\alpha_i, c_i)$, and we cannot observe output $(Q_i)$. Our goal is to estimate productivity $(\alpha_i)$.[7]

Our solution is to solicit researchers' willingness to pay $(WTP_i)$ for a fixed quantity of input using their income $(M_i)$. Consider an experiment where researchers are asked to report their WTP to obtain $\Delta$ units of $X$ without incurring the convex costs. By definition, they will choose $WTP_i$ that makes them indifferent between purchasing and not purchasing the additional units of $X$. Formally, $WTP_i$ solves:

$$\underbrace{\max_{X_i} \left[ M_i + b_i \times \alpha_i \times X_i - X_i^{c_i} \right]}_{\substack{\text{expected utility in} \\ \text{no-purchase scenario}}} = \underbrace{\max_{\widetilde{X}_i} \left[ (M_i - WTP_i) + b_i \times \alpha_i \times (\widetilde{X}_i + 1) - \widetilde{X}_i^{c_i} \right]}_{\substack{\text{expected utility in} \\ \text{scenario paying } WTP \text{ for } \Delta}} .$$

[7] Note that, even if we could also observe or estimate producers' expected input choice, $X_i^*$, then the first order condition, $X_i^* = (b \times \alpha_i / c_i)^{1/(c_i - 1)}$, still does not separately identify the productivity and cost parameters $(\alpha_i, c_i)$.

Solving for $i$'s optimal choice of $X$ in both scenarios, plugging them in, and rearranging yields a simple expression for identifying researchers' productivity beliefs:

$$\alpha_i = \frac{WTP_i}{b_i} \, .$$

In this simple example, estimating productivity does not require output quantities nor input costs. However, it does require estimating, observing or making assumptions about researchers' marginal utility from output, $b_i$ (e.g., output prices). We emphasize this point below as it is a key component of our methodology.

### 2.2. Researchers' Labor Supply and Functional Forms

Researchers supply their labor to maximize their utility conditional on a contract from their primary institution, which is a triplet of state variables $\mathbf{S}_i = (M_i, G_i, D_i)$: salary $M_i$ (\$), guaranteed funding $G_i$ (\$), and administrative and teaching duties $D_i$ (hours).[8] Given their contract, researchers choose how much time to spend on research ($R_i$), fundraising ($F_i$), which together with duties ($D_i$) sum up to yield their total work hours ($H_i$).

The quantity of knowledge $Q_i$ that researchers produce is a Cobb-Douglas function of total funding $B_i$ and research time $R_i$, which includes researcher-specific productivities ($\alpha_i$) and output elasticities (per $\gamma_i$). Additionally, we assume a minimum funding amount $B_{\min}$, and that all work hours are non-negative and have an upper bound $H_{\max}$.

Researchers' indirect utility is given by:

$$\mathcal{U}(\mathbf{S}_i\,,\,\theta_i\,,\,\mu_i) = \max_{R_i\,,F_i\,,H_i} m_i(M_i) + b_i(Q_i) - c_i(R_i\,,\,F_i\,,\,D_i) \tag{1a}$$

subject to

$$Q_i = \alpha_i \times B_i^{\gamma_i} \times R_i^{1-\gamma_i} \tag{1b}$$

$$R_i + F_i + D_i = H_i \leq H_{\max} \tag{1c}$$

$$B_i = B_{\min} + G_i + \phi_i \times F_i \, , \tag{1d}$$

where $\theta_i = (\alpha_i\,,\,\gamma_i\,,\,\phi_i)$ is the vector of individual-specific attributes related to scientific activity and $\mu_i$ is a vector of parameters that govern the shape of $m_i(...)$, $b_i(...)$, and $c_i(...)$.

[8] We use the term "*contract*" loosely, since, for example, a researcher's guaranteed funding may come in part from future flows of funding from grants obtained outside the institution. In the context of the methodological framework presented in Section 2, the function $m_i(...) + b_i(...)$ serves the role of the "benefit" function, while labor disutility $c_i(...)$ represents the cost function, which is parameterized to be convex.

The functional forms of these benefit and cost functions are assumed to be:

$$m_i(\ldots) = \omega \times \frac{M_i^{1-\sigma_i}}{1-\sigma_i} \tag{2a}$$

$$b_i(\ldots) = \frac{Q_i^{1-\eta_i}}{1-\eta_i} \tag{2b}$$

$$c_i(\ldots) = \psi \times \frac{(R_i + F_i + D_i^{\xi_i})^{1+\zeta_i}}{1+\zeta_i} \,. \tag{2c}$$

The $\xi_i$ parameter allows for additional disutility from duty-related work (e.g., administration or teaching), which improves the model's fit to the data. We also assume the following: $\eta_i \in (0,1)$, $\psi > 0$, $\xi_i > 0$, and $\zeta_i > 0$.

Solving the researcher's maximization problem, Equation (1), yields optimal choices of research time, fundraising time, and total working hours as a function of states, attributes, and parameters. We represent these optimal choices as policy functions $\mathcal{R}(\mathbf{S}_i, \theta_i, \boldsymbol{\mu}_i)$, $\mathcal{F}(\ldots)$, and $\mathcal{H}(\ldots)$. The derivations of these policy functions are described in Appendix A.1.

### 2.3. Identification, Key Assumptions, and Estimation

In what follows, we describe how our WTP solicitation identifies productivity beliefs using a generalized version of the models above. Despite revolving around forecasts, we present the following framework as static; there is no dynamic optimization.[9]

For clarity here, we focus on a compact version of the researchers' objective function per:

$$\max_{X_i} \quad u_i\big(M_i\,,\, \alpha_i \times f_i(X_i)\big) - c(X_i) \quad \text{subject to} \quad g_i(X_i) = 0\,, \tag{3}$$

which is a generalized version of Equation (1). Notably, we allow for non-separable utility of output $f_i(X_i)$ and income $M_i$ to emphasize that assumption is not necessary for identification.

Solving the first order conditions of this problem, letting $X_i^*$ be the argument that maximizes Equation (3), and ignoring constraints in what follows for simplicity, yields:

$$\alpha_i \times \frac{\partial u_i\big(M_i\,,\, \alpha_i \times f_i(X_i)\big)}{\partial X_i} \times \frac{\partial f_i(X_i)}{\partial X_i} - \frac{\partial c_i(X_i)}{\partial X_i} = 0\,, \tag{4}$$

[9] In so doing, we implicitly load dynamic considerations related to career progression onto other parameters in the model. We also omit the expectation operator despite variables being forecasts.

which shows that productivity ($\alpha_i$) is an implicit function of observables and the unknown parameters that govern the functions $u_i(...)$,$f_i(...)$, and $c_i(...)$. At this point, productivity is not separately identified from these unknown parameters.

In order to identify the parameters in Equation 4, we can experimentally solicit individuals' WTP for some amount of the input using their own money $M_i$ (e.g., directly from us, the experimenters). For example, we offer the producer an opportunity to purchase $\Delta$ units of $X$ and solicit their WTP (e.g., via stated preferences or incentive-compatible revealed preferences), which would be deducted from their $M_i$.

To see the value of this experiment, consider the individual's optimization problem in the scenario where they choose to pay their WTP to purchase $\Delta$ inputs from us:

$$\max_{\widetilde{X}_i} \quad u_i\big(M_i - WTP_i\,,\, \alpha_i \times f_i(\widetilde{X}_i + \Delta)\big) - c_i(\widetilde{X}_i)\,, \tag{5}$$

where we define $\widetilde{X}_i^*$ as the argument that maximizes Equation (5). In this scenario, the individual's $WTP_i$ is subtracted from their $M_i$ since they are both in monetary units and treated as perfect substitutes in the benefit function. The inputs purchased, $\Delta$, are added to the production function, but the producer's cost function, $c_i(...)$, still only depends on the quantity of inputs they choose to obtain after the experiment.

Therefore, the producer's WTP for $\Delta$ units of the input equates the following:

$$u_i\big(M_i, \alpha_i \times f_i(X_i^*)\big) - c_i(X_i^*) = u_i\big(M_i - WTP_i, \alpha_i \times f_i(\widetilde{X}_i^* + \Delta)\big) - c_i(\widetilde{X}_i^*)\,, \tag{6}$$

which is the generic version of the indifference equation in our illustrative example above. The individuals' WTP make them indifferent between (*i*) the scenario where they don't purchase the inputs from us (left-hand side of Eq. 6) and (*ii*) the scenario where they do purchase from us (right-hand side of Eq. 6). Next, we walk through the key assumptions that facilitate identification of $\alpha_i$ through the inversion of Equation (6).

**Key Assumptions**

*Assumption 1 — Convex Input Costs.* *The direct costs of inputs are convex:* $\partial c_i(...)/\partial X_i > 0$ *and* $\partial^2 c_i(...)/\partial X_i^2 > 0$. These convexities can be due to any sort of friction or adjustment cost. If this assumption does not hold, then the producer's WTP in the experiment may

not depend on their productivity, as we illustrate below.[10]

*Assumption 2 — Monotonic Benefits. The benefit function $u_i(...)$ is strictly monotonically increasing in $Q_i$ and $M_i$.* This implies that more money or more output always increases the individuals' payoff and, importantly, that $u_i(...)$ is invertible.

*Assumption 3 — Homogeneous or Observable Output Benefits. The benefits due to output, given by $\partial u_i/\partial Q_i$, are either homogeneous or depend on observable covariates.* Unobservable heterogeneity in the returns to producing more output would not allow us to separately identify individual-specific returns to output from their productivity. Thus, these returns must either be observable (and assumed to capture all quality differences), a function of other observables (in which case they can be estimated), or unobservable and homogeneous (in which case individual productivity beliefs are identified up to a common level shifter). In the language of traditional producers, this assumption implies that, conditional on observables, the output is a commodity and all producers face common output prices.[11]

*Assumption 4 — Observable Money and Plans. Individuals' monetary constraints, per $M_i$, and optimal input plans $X_i^*$ are observable*. There are some knife-edge cases where these variables need not be observed and productivity is identified by the experiment alone, such as in the illustrative example above. However, those cases are likely rare in practice and current allocations must be observed for estimation.

**Estimation**

Let $\Delta$ be the amount of input offered to the researcher in the experiment and $WTP_i^{\text{obs}}$ be researchers' reported willingness to pay for those $\Delta$ units. Furthermore, let $\boldsymbol{\mu}_i$ be the vector of parameters that govern the production, benefit, and cost functions. Under Assumption 2

[10] This is perhaps the least intuitive of our assumptions. Formally, this assumption ensures the necessary rank condition for estimation. Intuitively, the WTP experiment operates via an implicit linear cost schedule, so if the individual already faces a linear cost schedule when obtaining inputs, our WTP experiment may simply reflect those linear costs as a corner solution (i.e., the producer would never pay us, the experimenters, more for an input than what they pay in their own input market).

[11] Recall our simple, illustrative example at the beginning of this section — we showed that $\alpha_i = WTP_i/b$; however, if in fact the benefits of output are given by $u_i$, which is unobservable, then we cannot separately identify $\alpha_i$ from $u_i$.

— the benefit function ($b$) is invertible — we can write a producer's $WTP_i$ as a function:

$$WTP_i = \Pi(\Delta, M_i, X_i^* ; \boldsymbol{\mu}_i) \equiv \widehat{WTP}(\Delta, M_i, X_i^* ; \boldsymbol{\mu}_i) \,, \tag{7}$$

which leaves only the unknown parameters $\boldsymbol{\mu}_i$ to be estimated. This approach makes use of the fact that productivity ($\alpha_i$) and optimal allocations in the purchase scenario of the WTP experiment ($\widetilde{X}_i^*$) are implicit functions of observables and unknown parameters.[12] Now, we have a theoretical prediction of a producer's $WTP_i$ for $\Delta$.

In general, the GMM estimator for $\boldsymbol{\mu}_i$ solves the minimization problem:

$$\min_{\boldsymbol{\mu}_i} \quad \sum_{i=1}^{N} \left( \widehat{WTP}(\Delta, M_i, X_i^* ; \boldsymbol{\mu}_i) - WTP_i^{\text{obs}} \right)^2 , \tag{8}$$

which identifies $\boldsymbol{\mu}_i$ given standard GMM identification conditions being met. Particularly relevant for identification is the rank condition, which requires that the matrix of derivatives of the stacked moment conditions with respect to the vector of parameters has full rank.[13] It can be shown that Assumption 1 (Convex Input Costs) is a sufficient condition for the rank condition being met.

Note that instead of soliciting individuals' WTP directly, the experimenter could solicit individuals' current income, $M_i$, as well as their income that would make them indifferent to receiving $\delta$, and this income-at-indifference could be used for identification. Also, if there are constraints on individuals' input levels, they could be offered to pay to relax those constraints. Furthermore, depending on the variation in the data, it may not be possible to allow for full heterogeneity in the parameters that govern the production, benefit, and cost functions. In this case, the experimenter could assume $\boldsymbol{\mu}_i$ is given by a vector of common unobservables, say $\boldsymbol{\mu}$, and idiosyncratic observable variables, say $Z_i$, and estimate $\boldsymbol{\mu}$ via the GMM estimator above.[14]

[12] Specifically, Assumptions 1 and 2 generally guarantee that $WTP_i$ is a direct function of $\alpha_i$, which, by Equation (A13) in the Appendix, depends on $\boldsymbol{\mu}$.

[13] This condition fails, for example, if the WTP does not directly depend on a specific parameter, or if a parameter is redundant.

[14] After recovering the parameters, we also shrink the dispersion in researchers' productivity within each market before running the counterfactuals.

### 2.4. Measurement Error and Bayesian Shrinkage

A challenge for any study of misallocation is distinguishing real inefficiencies from measurement errors. If researchers misreport their input levels, then our estimates of misallocation will be biased. One solution is to use panel data, together with assumptions about persistence, to separate transitory measurement error from persistent heterogeneity (e.g., Bils, Klenow, and Ruane 2021). However, our data is cross-sectional.

To engage with this challenge, we use a Bayesian shrinkage procedure. Appendix A.4 gives the full description. In brief, for each state and choice variable entering the model, we assume that the observed value is the sum of a latent truth and classical measurement error. This latent variable model motivates a posterior expression governed by a parameter, $\rho \in [0, 1]$, that is defined as the share of observed variance attributable to measurement error. Likewise, we can express the correlation between reported and true values as $\sqrt{1-\rho}$. With an assumed value of $\rho$, the procedure uses Monte Carlo draws to shrink researcher-level variables towards their market-level means. Then, within each draw, we replace the observed primitives with a draw-specific posterior, reconstruct the parameters implied by the structural model, and solve the observed and counterfactual economies using the posteriors. We then compute all nonlinear objects (e.g., aggregate output gains) within each draw before summarizing across draws.

This procedure is not a silver bullet. The parameter $\rho$ is not identified from any of our data and must be set exogenously, and the simplicity of this method assumes away plausible scenarios of non-classical measurement error that may be relevant (e.g., a researchers' measurement error may be correlated with their productivity). Our main estimates are all based on a calibration of $\rho = 0.2$, but we explore a range of values and show that an exceptionally large amount of measurement error would be necessary to render our qualitative conclusions invalid. Furthermore, this procedure provides a transparent way to discipline the amount of economically meaningful misallocation in the data and may prove as a useful point for future work on measurement error in misallocation studies.

### 2.5. Social Value of Output

The model above identifies productivity through the lens of researchers' private returns. Per the WTP experiment: a researcher is willing to give up salary for additional inputs because those inputs increase the private value they expect to receive from producing more scientific output. But the externalities of knowledge production are such that much of the social value of scientific output is not captured by the researcher who produces

it. These benefits may accrue to other scientists, firms, workers, patients, consumers, or future generations. Thus, the private-return object identified by the model is not, by itself, the object a social planner would use to allocate resources.

Ideally, we would estimate a researcher-specific appropriation parameter, say $\kappa_i$, that converts private output value into social output value. If researcher $i$ captures share $1/\kappa_i$ of the total surplus generated by their work, then $\kappa_i \times b_i(Q_i)$ would be the relevant social-value object. Estimating $\kappa_i$ is beyond what our research design can support.[15] This limitation is especially important because two researchers with the same private return to scientific output may generate very different social spillovers.

Our approach for still incorporating social value into our analyses is as follows. We define markets $m$ as groups of scientists in the same field and career-stage cells and let social output be:

$$\mathcal{V}_i \equiv \mathcal{V}_i(Q_i; \kappa) = \kappa \times \nu_{m(i)}(Q_i), \tag{9}$$

where $\nu_m(\cdot)$ is governed by the market-level mean of researchers' estimated output-curvature parameters.[16] Replacing researcher-specific output valuation functions, $b_i(\cdot)$, with a market-level valuation function, $\nu_m(\cdot)$, removes idiosyncratic features of a researcher's private circumstances that may affect their WTP for output, while preserving systematic differences across scientific markets. The scalar $\kappa \geq 1$ then scales this market-level output value to reflect average external benefits. Under this interpretation, researchers capture share $1/\kappa$ of the social value generated by their output.

This formulation is intentionally ad hoc. It does not identify heterogeneous scientific spillovers, and it does not imply that all fields or researchers create the same external value. Rather, it gives us a transparent way to convert model-implied output into the social-output units needed for counterfactual allocation exercises. In the empirical analysis, we report sensitivity to $\kappa$ and use $\kappa = 5$ as our baseline, implying that researchers capture 20 percent of the social value of their scientific output. This value lies in the range of estimates from studies of innovation and producer surplus, which find private capture rates from roughly 2% to 50% (Nordhaus 2004; Jena and Philipson 2008; Lakdawalla et al. 2010). We return to the role of this assumption in the efficiency exercises in Section 6.

[15] Note that even an extremely rich WTP design that elicited the full $N \times N$ matrix of each scientist's willingness to pay for every other scientist to receive additional resources would miss much of the surplus from science, since a large share of that surplus is realized outside the scientific community.

[16] Specifically, $\nu_m(Q) = \frac{Q^{1-\bar{\eta}_m}}{1-\bar{\eta}_m}$, where $\bar{\eta}_m$ is the mean of $\eta_i$ within market $m$.

### 2.6. Additional Notes on the Model and Estimation

*Heterogeneous Science and Preferences.* Ideally, we would have enough variation to estimate researcher-specific production parameters $(\alpha_i, \gamma_i, \phi_i)$ as well as all of the idiosyncratic utility-related parameters that govern researchers' utility from salary, output, and effort $(\sigma_i, \eta_i, \xi_i, \zeta_i)$. However, as shown above, we only have enough structural conditions to uniquely identify the production-related parameters. Thus, we specify each of the utility-related parameters as a parametric function of researchers' observable features (i.e., socio-demographics, subjective descriptions of their research, etc.) that we also have data on. This allows us to estimate a set of common parameters defining heterogeneity. We elaborate on this approach in Section 3.4 below.

*"Productivity".* In productivity studies, the distinction between quantity- and revenue-based productivity metrics is important: is the estimate of "productivity" TFPQ or TFPR? One way to interpret our model of scientific production and researchers' utility is that the $\alpha$ parameter reflects TFPQ — it describes the growth in output from a proportional increase in inputs — and $b(\alpha)$ reflects TFPR — it describes the marginal return to the researcher from a proportional increase in inputs. But recall, we are assuming constant returns-to-scale in production. Thus, any true (un-modeled) non-linear returns-to-scale will load on our estimates of the $b(\cdot)$ function (which is governed by the $\eta$ parameter). Thus, we prefer to report $b(\alpha)$ as our focal productivity metric and interpret it as a sort of TFPR metric since it reflects the value to the researcher (in utility terms) of being more efficient at converting inputs into outputs.

## 3. Survey and Data Overview

### 3.1. Survey Design

We estimate the parameters of our model using the National Survey of Academic Researchers (Myers et al. 2023) and provide a brief overview of the survey methodology here. The population target is U.S. professors who conduct research at major institutions of higher education. To construct the sampling frame, information on professors was collected from the 158 largest institutions in the US by total R&D funding using the National Science Foundation's 2019 Higher Education R&D survey (HERD; National Science Foundation 2023).

The population consisted of 264,035 unique e-mails. A total of 131,672 individuals were

e-mailed and 4,389 (3.33%) completed the survey.[17] We then restrict the sample to the 4,004 individuals (91.2% of respondents) who reported being a professor, spending a non-zero amount of time on research, and having a non-zero salary from their primary institution.

During recruitment, incentives and reminders were randomly assigned. The four incentive arms were: (*i*) no incentive, (*ii*) entry into a lottery to win a gift card, (*iii*) the ability to vote for a set of charities to receive a donation, and (*iv*) both the second and third incentives. The reminder arms were zero, one, or two follow-up emails. Each email was randomly assigned to one incentive arm and one reminder arm with equal probability, resulting in twelve possible combinations.

The randomized incentives and reminders provide us with instruments that we can use to implement a sample selection correction (i.e., Heckman 1979). The validity of this approach relies on having variables that cause entry into the sample (i.e., completing the survey) but do not affect the outcomes of interest. This allows us to adjust for unobservable differences between the population and our sample. Appendix Table B1 reports the results from a regression of an indicator for survey completion on the different incentive and reminder arms, showing that all arms had a statistically significant positive effect on researchers' propensity to complete the survey.

In addition to adjusting for unobservable differences, Myers et al. (2023) also checks the representativeness of the respondent sample by comparing it to the invited sample on observable characteristics. First, the authors explore a series of observable characteristics at the researcher level by comparing the grant and publication histories of respondents and non-respondents using the Dimensions database (Digital Science 2018), which collects and disambiguates scientific metrics for researchers worldwide.[18] Appendix Figure B1 (replicated from Myers et al. 2023) shows invite-respondent overlap on various measures of scientific inputs and outputs. Overall, there is little difference between the respondents and non-respondents both economically and statistically speaking.

Looking at the institutional level, Appendix Figure B2 (replicated from Myers et al. 2023) shows invite-respondent overlap on various measures of institution funding derived from

[17] The IRB approval permitted e-mailing only 50% of the population. The response rate is more than twice what has been obtained from sourcing academic researcher contacts from the corresponding author data contained within the publication record (e.g., Myers et al. 2020).

[18] Using a fuzzy name matching process, Myers et al. (2023) are able to confidently match 87,000 (66%) of the researchers to their records in Dimensions.

the HERD survey. As in the case of the researcher-level comparison, the distributions overlap substantially. In this case, there are some statistically significant differences; on average, respondents come from institutions with slightly less research funding (4–6%).

### 3.2. Summary Statistics: Researchers and their Inputs

Table 1 reports the summary statistics of the key covariates in our analyses. See Myers et al. (2023) for a more detailed investigation of these summary statistics.

*Markets.* We define scientific markets according to two dimensions: field and career stage. Using the name of the professor's department and their self-reported scientific discipline, we assign them to a set of twenty fields of study.[19] We define career stage as a binary distinction based on whether a researcher is younger than the within-field median age ("early career"), or equal to or older than the within-field median age ("late career"). Thus, we obtain forty scientific markets. Importantly for our later analyses, all counterfactuals we estimate constrain the reallocation of inputs to only occur *within* these field-by-career-stage markets.

*Salaries and Research Inputs.* The survey solicits a range of details regarding researchers' salary, their guaranteed funding (e.g., from prior grants or institutional guarantees), expected funds they will raise over the coming five years, and their time allocations. Importantly, most variables are elicited as expectations over the coming five years to ensure the responses span the same time horizon as the thought experiments described below. In the Appendix, we replicate a test of respondents' self-reporting by comparing their self-reported salaries to the publicly-reported salaries we are able to locate for a subset of researchers. Overall, there is a high degree of alignment (see Appendix Figure B3).

*Position and Socio-demographics.* Appendix Tables B2-B3 provides a full summary of all other major features collected regarding researchers' positions (e.g., rank and tenure status) and their socio-demographics (e.g., gender, race/ethnicity, citizenship).

[19] We also aggregate those fields into five broader "aggregate" fields for simplicity in reporting some results: Humanities and related; Engineering, Math, and related; Medicine and Health Sciences; Natural Sciences; and Social Sciences.

TABLE 1. Summary Statistics—Salary, Funding, Time, and Fields

| | *mean* | *s.d.* | *p*50 |
|---|---|---|---|
| *Salary and research funds, $/year* | | | |
| Total salary | 159,028.23 | 74,516.13 | 140,000.00 |
| Guaranteed & existing funding | 66,374.97 | 121,936.46 | 5,000.00 |
| Fundraising expectations | 93,909.19 | 144,013.69 | 20,000.00 |
| *Work time, hrs./week* | | | |
| Total work | 49.05 | 11.05 | 48.00 |
| Research | 18.69 | 9.79 | 17.40 |
| Fundraising | 4.53 | 5.18 | 2.65 |
| Administration | 7.55 | 6.24 | 5.80 |
| Teaching and other work | 18.28 | 9.97 | 17.20 |
| *Major field*, {0,1} | | | |
| Engineering, math, & related | 0.17 | 0.37 | |
| Humanities & related | 0.19 | 0.39 | |
| Medical & health sciences | 0.28 | 0.45 | |
| Natural sciences | 0.16 | 0.37 | |
| Social sciences | 0.21 | 0.40 | |

*Note*: Summary statistics for the 4,003 researchers in the estimation sample. Salary and funding are in dollars per year; time allocations are in hours per week. Funding and time variables are self-reported expectations over the coming five years. $\{0, 1\}$ denotes binary variables.

### 3.3. Summary Statistics: Subjective Output Measures

Our methodology for estimating productivity is explicitly designed to avoid the need to quantify the output produced by researchers' efforts. Still, it is useful to have a more qualitative understanding of what researchers are producing. Fortunately, the survey includes a number of subjective questions regarding researchers' output. Table 2 summarizes answers to these questions, which asked the frequency with which researchers intended to produce outputs of certain types (e.g., articles, books, materials, products) for certain types of audiences (e.g., peers, policymakers, businesses, general public) as well as other measures of riskiness and the degree to which the researcher focuses on asking new questions or answering existing questions.

The traditional proxy for scientific output is peer-reviewed journal articles, and the data indicate that this is the most common output type that researchers intend to produce. However, there is a considerable amount of variation and a significant amount of attention focused on producing output of types and for audiences that may never be codified in a journal article. More importantly, these measures are often negatively correlated in a way that suggests strategic substitution (see Appendix Table B4). For instance, researchers

TABLE 2. Summary Statistics—Subjective Output

| | *mean* | *s.d.* |
|---|---|---|
| *Intended research outputs,* {0,1,2} | | |
| Journal articles | 1.87 | 0.37 |
| Books | 0.52 | 0.68 |
| Materials or methods | 0.70 | 0.68 |
| Products | 0.47 | 0.64 |
| *Intended research audience,* {0,1,2} | | |
| Academic peers | 1.88 | 0.36 |
| Policymakers | 0.84 | 0.69 |
| Businesses | 0.54 | 0.62 |
| General public | 0.83 | 0.62 |
| *Riskiness of own research,* [0,10] | | |
| Own belief | 4.63 | 2.34 |
| Belief of peers' beliefs | 4.57 | 2.37 |
| *Theoretical vs. empirical,* [0,10] | | |
| Ask [0] or answer [10] questions | 4.89 | 2.54 |

*Note*: Summary statistics for the 4,003 researchers in the estimation sample. Output and audience variables code responses to questions of the form "How often is the following the intended output (audience) of your research?": rarely = 0, sometimes = 1, very often = 2. Risk is measured on a 0–10 scale, where 10 indicates very high risk. The final variable records whether the researcher focuses on asking new questions (0) or answering existing questions (10).

who focus more on publishing articles for their academic peers are less likely to focus on publishing books or developing products intended for non-academic audiences. This highlights the limitation of observable output proxies.

### 3.4. Heterogeneous Science and Preferences

As discussed in Section 2.6, we lack the variation necessary for allowing fully flexible heterogeneity in researchers' preferences. Thus, we choose to make the functions that describe researchers utility from income, scientific output, and effort ($m_i(\cdot)$, $b_i(\cdot)$, and $c_i(\cdot)$, respectively) depend on the large set of observable variables just described. Thus, instead of assuming that all economists have the same marginal utility of income, we can allow this preference to depend on the researchers' household structure. Likewise, instead of assuming that all economists have the same marginal utility from producing more knowledge, we can allow this to depend on researchers career stage and the subjective output measures described above.This is useful because it allows us to incorporate more heterogeneity into the consumption components of the model and limit the degree to

which variation in the data might otherwise cause us to overstate the heterogeneity in the production parameters.

To balance the benefits of allowing for heterogeneity with the costs of estimation, we use $k$-means clustering to reduce the full set of observable, relevant features ($\mathbf{Z}_i$, which includes all of the variables in Tables 2, B2, and B3) into a one-dimensional index. We assume there are two clusters of researcher types ($k = 2$) and estimate each researcher's distance from these clusters with their Euclidean similarity score. This distance provides a one-dimensional description of all of the different ways researchers responded to questions that could plausibly be driven by heterogeneous preferences.

We refer to this resulting index as describing researchers' type, $T_i$. Appendix B.2 reports the results of the $k$-means estimation showing the distribution of researcher type $T$ in the sample.[20]

## 4. Survey Experiment

Here, we describe the thought experiments in the survey (Subsection 4.1) and how they connect to the model (Subsection 4.2). We then report the distribution of responses and conduct tests for face validity, sample selection, and other potential concerns (Subsection 4.3).

### 4.1. Soliciting Willingness to Pay

*WTP for Funding and Time.* Respondents are presented with four hypothetical scenarios, each offering different trade-offs between salary and inputs. They are asked to imagine that their primary institution has offered them: (*i*) an increase of $250,000 in guaranteed funding in exchange for a lower salary, (*ii*) an increase of $1,000,000 in guaranteed funding in exchange for a lower salary, (*iii*) an elimination of all administrative duties in exchange for a lower salary, and (*iv*) an increase in duties by 20 hours per month over a five-year

[20] Looking to estimation, we specify each of the consumption-related parameters $(\sigma_i, \eta_i, \xi_i, \zeta_i)$ to be simple, univariate functions of a researcher's type. For instance, the parameter that governs the utility from private consumption is parameterized as: $\sigma_i = \exp(\delta_{\sigma,0} + T_i\delta_{\sigma,1})$. We similarly parameterize $\zeta_i$ and $\xi_i$ with exponential functions with intercept and slope, respectively. The curvature of utility in scientific output, $\eta_i$, is modeled with a logistic function bounded in the interval $(0, 1)$. Therefore, we express individual-specific parameters $\boldsymbol{\mu}_i(T_i, \boldsymbol{\mu})$ as a function of a researcher's type $T_i$ and of the vector of common parameters to be estimated, denoted by $\boldsymbol{\mu} = (\omega, \psi, \boldsymbol{\delta})$, where $\boldsymbol{\delta}$ includes the deep parameters that govern the utility functions. Allowing the $\boldsymbol{\mu}_i$ parameters to be type-specific is not a panacea, but it helps reduce the degree to which variation in the experimental data that is truly driven by heterogeneous preferences or heterogeneous demand for scientific output contaminates our productivity estimates.

period in exchange for a higher salary.[21] All of these hypotheticals are posed over a five-year span to fix the time horizon for all respondents. In order to solicit the salary at which they are indifferent between the current position and each offer, respondents are asked to report the lowest offered salary at which they would be willing to accept the offer.[22] Importantly, the survey only solicits researchers' WTP for more funding. This matters because counterfactual reallocations will involve reducing some researchers' funding. In the case of time, the survey solicits both WTP (for more free time) and willingness to accept (WTA; for less free time) and we treat these as symmetric after being filtered through the model. Appendix Figure B5 displays examples of how these thought experiments appeared to the survey respondents.

*WTP for a Benchmark Good.* When soliciting willingness to pay, especially in an un-incentivized manner, it is always possible that respondents under- or overstate their price sensitivity. In order to test for this, we follow Dizon-Ross and Jayachandran (2022) and explore respondents' willingness to pay for a "benchmark good." Dizon-Ross and Jayachandran (2022) note that, if one solicits a subject's willingness to pay for a good whose value (to the subject) is plausibly uncorrelated with the value of the focal good, then any correlation between the two willingness-to-pay values can be attributed to systematic noise. The survey asks respondents to report the maximum amount they would be willing to pay per month for high-speed internet access at their primary residence. We use this as our benchmark good since researchers' scientific productivity should plausibly be uncorrelated with their demand for high-speed internet access at home.

### 4.2. Connection to Model: WTP and Compensating Variation

Four survey thought experiments elicit the compensating variation of individual researchers in relation to (*i*) an increase of \$250,000 in guaranteed funding $G_i$, (*ii*) an increase of \$1,000,000 in guaranteed funding $G_i$, (*iii*) a reduction of duties $D_i$ to 0, and (*iv*) an increase in duties by 20 hours per month over a 5-year period. In other words, we ask for income levels $\widehat{M}_{ij}$, where $j = \{1, 2, 3, 4\}$ indexes the four experiments, that would make the researcher's utility in the counterfactual scenario equal to their indirect utility $\mathcal{V}_i^*$ at current allocations. Formally, counterfactual guaranteed funding and duties in the four

[21] Researchers who report no administrative duties are not shown scenario (*iii*).

[22] Pilot tests indicated that researchers could more easily report the lowest salary that could make them take the offer as opposed to the literal salary at indifference, and since indifference is a vanishingly small amount less than this reported value, we treat their answer as the amount at indifference.

experiments are:

$$
\begin{aligned}
(\widetilde{G}_{i1}, \widetilde{D}_{i1}) &= \left((G_i + \$250{,}000), D_i\right) \\
(\widetilde{G}_{i2}, \widetilde{D}_{i2}) &= \left((G_i + \$1{,}000{,}000), D_i\right) \\
(\widetilde{G}_{i3}, \widetilde{D}_{i3}) &= \left(G_i, 0\right) \\
(\widetilde{G}_{i4}, \widetilde{D}_{i4}) &= \left(G_i, (D_i + 20 \text{ hours/month})\right),
\end{aligned}
$$

where $G_i$ and $D_i$ represent actual states.

As noted in Section 2, the values of $\widehat{M}_{ij}$ reported in these four experiments make the researcher indifferent between all possibilities, and therefore we can write these reported values to be a known function of observable data and the unknown parameters to be estimated.

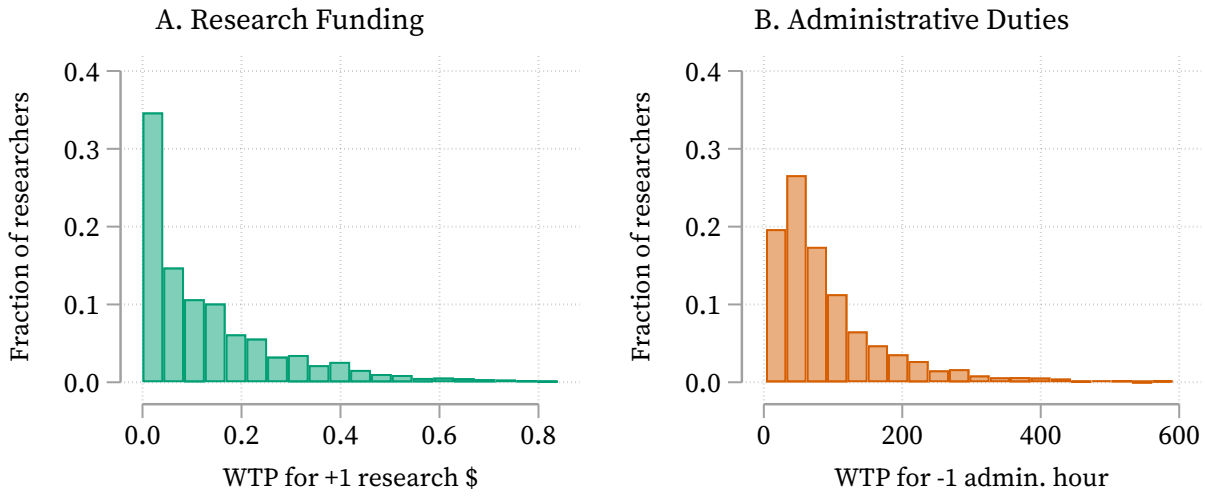
### 4.3. WTP Results

Figure 1 plots the distribution of WTP for guaranteed research funding and free time (in the form of WTP for less duties or WTA for more duties). The values shown are averaged over the two thought experiments for either factor and converted to a per-dollar basis. As evidenced by Figure 1, there is considerable variation in WTP responses across researchers. Some of this variation reflects different preferences and constraints, but another portion reflects heterogeneous productivity across researchers. The model described above allows us to separate these two forces.

Validating these WTP responses is a difficult but important task. They are the key ingredient of our entire exercise, but they may be driven in part by sample selection effects, behavioral biases, or noise. Although these are practical trade-offs that professors face throughout their careers, systematic data have not been collected in this way before to our knowledge. Thus, it is not clear what plausible variation would look like here. Still, we can conduct some tests motivated by economic and statistical theory to explore how reasonable these distributions of WTP are.

Motivated by the notion of opportunity costs, we test how WTP varies with researchers' implied hourly wage (per their annual salary divided by their total hours of work). Assuming this implied wage rate is a proxy for researchers' opportunity costs, it should be the case that researchers with higher hourly wages are willing to pay more for their free time. In Appendix Figure B6A, we see that this is indeed the case. Furthermore, it should be the case that, conditional on opportunity costs of time, researchers who expect to spend more time fundraising should have a higher WTP for guaranteed funding. In Appendix Figure

FIGURE 1. Willingness to Pay Experiment Responses




*Note*: Histograms of researchers' stated willingness to pay, per unit: dollars of salary per additional dollar of guaranteed research funding (Panel A) and per hour of administrative duties removed (Panel B), averaged over the two thought experiments for each input. The median researcher is willing to pay $0.10 per dollar of funding and $68 per administrative hour. Distributions are trimmed at the 1st and 99th percentiles.

B6B, we find this to be true.

As another test, Appendix Tables B5 and B6 describe analyses where we regress researchers' WTP responses on the variables in the model as well as (*i*) the large vector of observable features ($\mathbf{Z}_i$), (*ii*) an inverse Mills ratio constructed using the randomized survey participation incentives following Heckman (1979), and/or (*iii*) researchers' WTP for the benchmark good (high-speed internet) following Dizon-Ross and Jayachandran (2022). We find the variables in the model can explain 82–97% of the variation in WTP responses (see Appendix Table B5). This supports the model's ingredients as being key determinants of researchers' answers. When we include the large vector of observable features in addition as additional explanatory features, the $R^2$ statistics increase by only one to four percentage points. Residual variation in WTP responses (conditional on the variables of the model) are not well explained by these heterogeneous observables. This gives us confidence that we are not dramatically mis-specifying heterogeneity in the model.

We also find that the inverse Mills ratio is not a statistically significant predictor of responses (see Appendix Table B6). Under the assumptions outlined in Heckman (1979), this supports the notion that respondents did not differentially select into our sample as a function of their WTP for inputs and suggests some generalizability of our responses. Finally, we find that researchers' WTP for the benchmark good (high-speed internet at their house)

is correlated with their WTP for inputs (see Appendix Table B6). Under the assumption that researchers' demand for home internet is truly uncorrelated with their productivity, this provides some evidence of systematic noise in how respondents are reporting their WTP in these experiments. However, the economic magnitude of this relationship is quite small and on the scale of pennies on the dollar. Furthermore, the benchmark good WTP is included in our vector of features we use to make the researcher type index (that determines heterogeneity in the utility function parameters), which allows us to control for this to some degree. Overall, despite not being incentivized experiments, researchers' responses generally appear to behave as expected and are of plausible magnitudes.

# 5. Estimates for the Model of Science

## 5.1. A View of Researchers' Utility Functions

To illustrate our estimates of the utility function, Appendix Figure C1 shows how an average researcher's utility depends on the levels of key variables (salary, administrative duties, research budget, research time). The figure shows the percent change in a researcher's utility as the variable is increased from the $10^{th}$ percentile level to the $90^{th}$ percentile level while all other variables and parameters are held fixed at the sample averages.

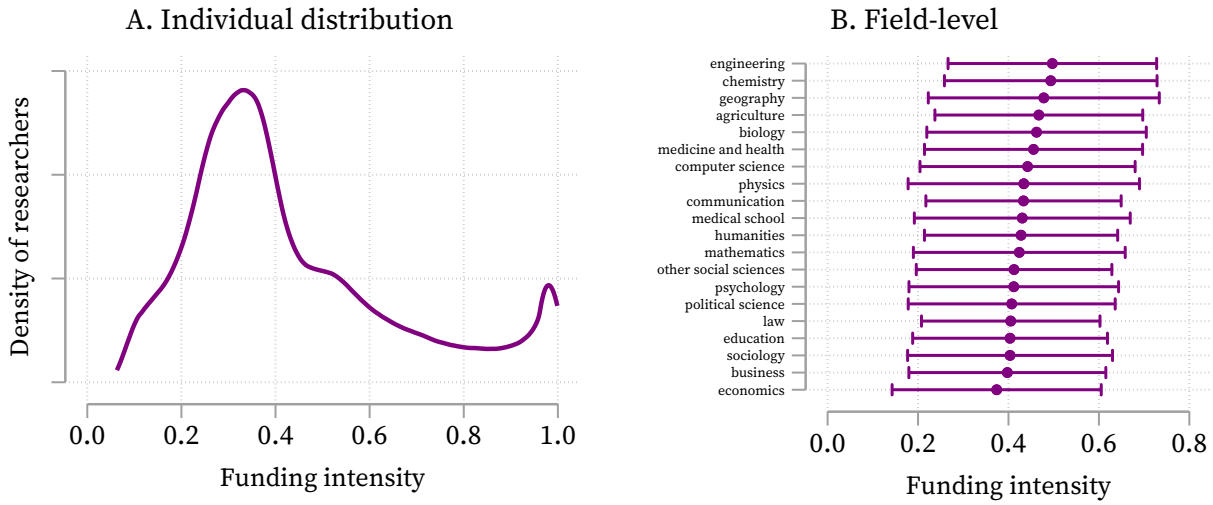
In terms of magnitude, salary and administrative duties are the most important for researchers' utility. Shifting the average researcher's salary from the $10^{th}$ percentile to the $90^{th}$ percentile increases their utility by roughly 55%. An equivalent relative increase in administrative duties reduces utility by about 38%. In contrast, similarly scaled increases in funding or research time (and likewise research output) raise utility by only a few percentage points.[23]

## 5.2. Funding Intensity

Figure 2 shows the unconditional distributions of the funding intensity parameter, which describes researchers' marginal returns to funding (per $\gamma_i$) relative to their own research time (per 1-$\gamma_i$). Figure 2A shows the individual-level distribution and highlights a significant degree of heterogeneity across researchers, with roughly 20% of our sample having a funding intensity either larger than 0.6 or smaller than 0.2.

[23] As evidence of fit, the median absolute difference between a researcher's stated WTP and the model's prediction ranges from 6–12 percentage points across the four questions. The numerical estimates of the common parameters and more detailed descriptions of the model's fit are available from the authors upon request.

FIGURE 2. Funding intensity, $\gamma$




*Note*: Panel A shows the distribution of the funding-intensity parameter $\gamma_i$, the elasticity of output with respect to funding, trimmed at the 1st and 99th percentiles. Panel B shows within-field means (markers) and ±1 standard deviation (whiskers).

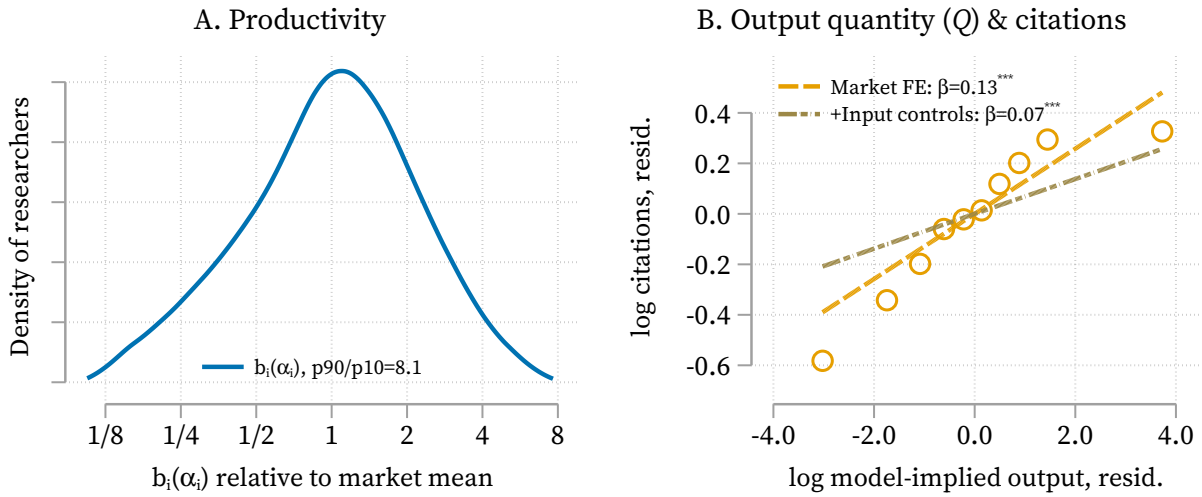
As both an interesting exercise and test of face validity, Figure 2B plots the average funding intensity parameters ($\gamma_i$) for the twenty fields represented in the sample. We find that $\gamma_i$ is highest in chemistry and engineering, where research is often capital intensive and involves the use of expensive lab equipment. In contrast, the social sciences (e.g., economics, business, sociology) display the lowest funding intensity on average, which is not surprising given the nature of research in those fields. Notably, the relatively large standard deviations reveal significant heterogeneity in funding intensity across researchers within each field of study.

### 5.3. Productivity and Output

Figure 3A shows the distribution of the productivity parameter ($\alpha$), scaled by the valuation function ($b(\cdot)$) and normalized to have a mean of 1 within each market. Our estimates reveal a large skew in researchers' beliefs about their productivity. Figure 3A also reports the ratio of the $90^{th}$ and $10^{th}$ percentiles, the "90–10 productivity ratio" measure of variance commonly reported in traditional productivity studies. The 90–10 productivity ratio is substantial: the $90^{th}$ percentile researcher believes they are roughly 8 times as productive as the $10^{th}$ percentile researcher.

It is difficult to benchmark this dispersion. Firm-level estimates are primarily from manufacturing sectors, which typically involve 90–10 productivity ratios on the scale of 1.5× to 5× (Syverson 2011; De Loecker and Syverson 2021). Most worker-level productivity estimates

FIGURE 3. Productivity & output




*Note*: Panel A shows the distribution of researcher productivity, $b_i(\alpha_i)$, relative to its market mean; the 90–10 ratio is computed after trimming at the 1st and 99th percentiles. Panel B shows a binned scatterplot of log citations on log model-implied output, residualized on market fixed effects; the second line additionally controls for input levels. $^{***}p < 0.01$.

also appear to exhibit this degree of dispersion, although the specifics of productivity metrics varies widely across studies (Hoffman and Stanton 2024). One notable exception, and the source of the "10× engineer" label is Sackman, Erikson, and Grant (1968), who used direct observation to estimate software engineers' productivity and found that some individuals were as much as an order of magnitude faster on coding tasks than others.

Another test of the validity of our model is whether our model-based output measure tracks observable proxies of research performance. To conduct this test, we use bibliometric data matched to approximately three-quarters of the sample and construct each researcher's cumulative citations over the five years preceding the survey. We lack post-survey bibliometric data; however, prior work documents that citation-based output is highly persistent at horizons of this length (e.g., Sinatra et al. 2016), so the timing mismatch between survey responses and bibliometric data is unlikely to attenuate the relationship substantially.

Figure 3B confirms that our model-based output measure and citation output are positively correlated within-markets. This is a reassuring test, but does not speak to the validity of our production function parameters per se: our model-based output depends on both productivity parameters and input levels, where the latter have long been known to correlate with citations. A sharper test conditions on input levels. If our productivity estimates are capturing genuine heterogeneity in $\alpha_i$ and $\gamma_i$, then the quantity-citation relationship should

survive after absorbing inputs. Figure 3B shows that it does: the conditional correlation is attenuated, as expected, but remains statistically significant.[24]

To understand how variation in productivity affects output, we decompose the variance in log output into components attributable to the parameter $\alpha$, input levels $B$ and $R$, and factor intensity $\gamma$. Note that the variance in log output due to the variance in log productivity is given by: $\text{Var}(\log(\alpha_i)) + 2 \times \text{Cov}(\log(\alpha_i)\,,\ \gamma_i \times \log(B_i)) + 2 \times \text{Cov}(\log(\alpha_i)\,,\ (1-\gamma_i) \times \log(R_i))$. Using this relationship, we estimate that 38.2% of the variance in output across researchers is due to the variance in productivity. Without adjusting for the covariances between productivity and the input levels and funding-intensity parameter, we obtain an estimate of 59.9%. This finding indicates that more productive researchers obtain more inputs that are well-suited to their production functions, and it motivates questions related to allocative efficiency that we turn to next.

## 6. Efficiency and Welfare

The preceding sections establish that individual researchers differ widely in their productivities and in inputs. We now ask what those differences imply at the level of markets. How much social value is lost because funding and research time are not allocated to the researchers with the highest marginal returns? And how much of that gap could be covered by policies that do not require knowing any individual researcher's productivity?

We answer these, and related questions, in three steps. First, Section 6.1 defines the counterfactual allocations we study, most notably (*i*) a first-best benchmark in which a planner who knows each researcher's productivity can optimally assign funding and research time directly, and (*ii*) a uniform funding policy that gives all researchers in a market identical funding guarantees and fundraising terms and then lets them re-optimize. Next, Section 6.2 quantifies the gains from moving between these (and other) allocations: how much misallocation exists, and how much of it simple policies could eliminate. Finally, Section 6.3, our most speculative analyses, look across markets to identify which fields allocate their resources most efficiently, what distinguishes them, and what market-specific welfare weights would rationalize the allocation of resources across markets we observe.

[24] Note, we view this output-citation test as validating the model's output measure *on average*: researchers whom the model ranks as highly productive do, in fact, produce more cited work. But this test does not imply that citations are a sufficient statistic for research value at the individual level. The bibliometric correlation provides a check on the average quality of our output measure, not an endorsement of citations as the object of interest.

### 6.1. Counterfactual Allocations and Policies

All counterfactuals share three features designed to arrive at conservative results and keep the simulations inside the regions where the model is disciplined by data. First, reallocation occurs only within markets, where a market is a field-by-career-stage cell as defined in Section 3. We never, for example, take funding from a young biologist and give it to an old psychologist. Second, no researcher's input levels may move outside the range observed in their market. Third, no change may exceed the support of the variation used in the WTP solicitations (i.e., no researcher can receive more than an additional $1,000,000 of funding, the largest amount considered in the survey experiments).[25]

*First Best.* The central benchmark is a first-best allocation of observed aggregate resources to maximize the social value of scientific output:[26]

$$\max_{\{B_i,R_i\}_{i\in m}} \sum_{i\in m} \mathcal{V}_i(B_i,R_i) \quad \text{subject to} \quad F_i = 0,\ \sum_{i\in m} B_i = \sum_{i\in m} B_i^{\text{obs}},\ \sum_{i\in m} R_i = \sum_{i\in m} R_i^{\text{obs}}. \tag{10}$$

Here, $\mathcal{V}_i(B_i,R_i)$ denotes the social value of the output researcher $i$ produces with inputs $(B_i,R_i)$, per Equation 9. Because the planner assigns funding and research time directly, fundraising is unnecessary, hence $F_i = 0$.[27] The aggregate constraints ensure that any gains come from reallocating existing resources, not from expanding them.

*Wedges.* Let $\lambda_m^B$ and $\lambda_m^R$ denote the shadow prices on funding and research time from the first-best problem in market $m$. For input $X \in \{B,R\}$ and allocation $a$, define researcher $i$'s wedge as

$$w_{iX}^a \equiv \frac{\partial \mathcal{V}_i^a / \partial X_i}{\lambda_m^X}. \tag{11}$$

In an interior allocation with no misallocation, $w_{iX}^a = 1$ for every researcher in the market. A wedge above one indicates that researcher $i$'s marginal social product exceeds the market shadow price — the researcher has too little of the input — and a wedge below one indicates the opposite. The dispersion of these wedges is a standard diagnostic for misallocation

[25] All counterfactuals are solved within each Monte Carlo draw of the measurement-error procedure described in Section 2 and Appendix A.4, and we report medians across draws. The baseline specification sets the externality scale to $\kappa = 5$ and the measurement-error share to $\rho = 0.2$. Appendix Table C1 varies these and other implementation choices.

[26] Recall that we define the social value of output as $\mathcal{V}_i \equiv \kappa \times \nu_{m(i)}(Q_i)$.

[27] We also consider an alternative benchmark that holds each researcher's fundraising time at its observed level. Researchers' duties (e.g., administration, teaching) are held at observed levels. We treat these as producing social value outside the model.

(e.g., Hsieh and Klenow 2009).[28]

*Allocative Efficiency Metric.* We summarize misallocation with a scalar metric of allocative efficiency. For any allocation $a$, let $a^\star$ denote the first-best allocation that uses the same aggregate funding and research time as $a$. Then

$$\text{Allocative Efficiency}_m(a) \equiv 100 \times \frac{\mathcal{V}_m^a}{\mathcal{V}_m^{a^\star}},$$

which equals 100% when allocation $a$ is on its own frontier. For the observed allocation, this metric compares observed social output to the first-best (i.e., the $\mathcal{V}_m$ achieved with the solution to Eq. 10). For policy simulations in which behavioral responses change aggregate inputs, the benchmark $a^\star$ is re-solved at the policy's own realized aggregates, so that we properly adjust for the change in aggregate resources.

*Uniform Funding Policy.* Our main policy experiment removes a feature of scientific markets that attracts persistent debate: researchers in the same market access funding on very different terms. The "uniform" policy replaces those heterogeneous terms with common ones. Every researcher in market $m$ receives the market-average funding guarantee ($G_i \to \overline{G}_m$) and the market-average fundraising ability ($\phi_i \to \overline{\phi}_m$). Researchers then choose their time allocations maximizing their private utility, exactly as in the observed economy. Aggregate funding in each market is still constrained to its observed level, while aggregate research time adjusts endogenously through researchers' labor supply. Unlike the omniscient planner that must exist to implement the first-best allocation, implementing this policy requires no knowledge of any individual's productivity.

*Social-Objective Alignment.* Researchers' private payoffs from producing more scientific output, per $b_i(\cdot)$, need not coincide with the social value function we have assumed, per $v_m(\cdot)$. To assess this distortion, we simulate a counterfactual in which each researcher's private payoff from output is replaced with $v_{m(i)}(\cdot)$, holding the funding environment fixed. We interpret the results from this particular policy simulation very cautiously. Our definition of the social value output function is both ad hoc in functional form, and also entirely devoid of concerns over heterogeneous "directions" in which researchers may

[28] As in that literature, dispersion in measured wedges is an upper bound on inefficiency to the extent that it reflects unmodeled technologies or constraints (e.g., Bergquist, Lashkari, and Verhoogen 2026). Our within-market focus, support restrictions, and measurement-error adjustments are all aimed at limiting this concern.

produce knowledge. Thus, this simulation can speak only to potential distortions in the *level* of private payoffs that researchers receive *relative to our assumed definition of social value*.

### 6.2. Counterfactual Results

To begin, we focus on the comparison of observed allocations to the theoretical first-best benchmark achieved by solving Equation 10. Appendix Figure C2 shows how the distributions of inputs and outputs compare in these two allocations; Panels A–C all show a clear, positive relationship between observed funding, research time, and output versus the optimal levels thereof. Researchers whom the model identify as being more productive do, in observed allocations, have more funding, spend more time on research, and produce more output. Scientific markets select on productivity *on average*.

Panels D–F of Appendix Figure C2 highlight the distributional differences between observed and optimal inputs and outputs. For example, the optimal funding distribution is *more compressed* than the observed one (the standard deviation falls by roughly 25%), while the optimal output distribution is *more dispersed* (the standard deviation of social output rises by roughly 56%). A substantial part of the observed variance in funding does not reflect productivity differences, and observed output varies too little, because the most productive researchers are under-resourced

Appendix Figure C3 plots the distributions of the wedges $w_{iX}^{\text{obs}}$ separately for each input, highlighting a major difference. Funding wedges are widely dispersed: for a large share of researchers, the marginal social product of funding is more than four times above or below their market's shadow price, and the tails of the distribution extend well beyond a factor of sixteen. On the other hand, research-time wedges concentrate within a factor of about four. The implication is that misallocation in science is mostly about money, not time.

Table 3 (Col. 1) summarizes the transition from the observed allocation to the first-best. First considering researchers' private utility, the first-best allocation yields a 3.7% increase in utility, which is equivalent to roughly an additional $9,000 (6%) in salary.

Looking to output and efficiency: the observed allocation attains an allocative efficiency of 50.7%: with the same aggregate funding and research time, optimally assigned, social output would be 97% higher.[29] Interpreting this change is difficult given the units of output,

[29] Interestingly, and as a very loose benchmark, this is on the same magnitude of allocative efficiency that

TABLE 3. Counterfactual Allocations

| | | Observed to first-best (1) | Observed to uniform funding (2) |
|---|---|---|---|
| *Change in inputs & outputs* | | | |
| funding, $B$ | mean | ~0% | ~0% |
| | s.d. | -24.6% | -35.5% |
| research time, $R$ | mean | ~0% | +6.1% |
| | s.d. | +5.9% | -4.6% |
| fundraising time, $F$ | mean | -100% | +168% |
| | s.d. | -100% | +322% |
| utility, $\mathcal{U}$ | mean | +3.7% | +2.1% |
| | s.d. | +4.8% | +1.4% |
| private output, $b(Q)$ | mean | +116% | +49.4% |
| | s.d. | +74.4% | +19.8% |
| social output, $v(Q)$ | mean | +97.3% | +55.4% |
| | s.d. | +56.3% | +16.1% |
| *Allocative efficiency* | | | |
| initial allocation | | 50.7% | 50.7% |
| counterfactual allocation | | 100% | 78.5% |
| *Budgetary equivalent of social reallocation gains* | | | |
| researcher-level | % | +168% | +88.8% |
| | $/year | +$266,286 | +$141,343 |
| population-level | $/year | +$32.3B | +$17.1B |
| *Salary equivalent of private reallocation gains* | | | |
| researcher-level | % | +5.9% | +3.2% |
| | $/year | +$9,235 | +$5,034 |

*Note*: Each column reports changes from moving the observed allocation to the counterfactual in the column header: the first-best benchmark (column 1) and the uniform funding policy (column 2), as defined in the main text. Allocative efficiency is aggregate social output relative to the first-best frontier with the same aggregate resources. Budgetary (salary) equivalents give the increase in funding (salary) under observed allocations that would generate the same gain in social output (utility) as the reallocation; population-level figures scale researcher-level dollars to the population represented by the sampling frame.

which are effectively in utils. Thus, we conduct an exercise that solves for the budgetary equivalent, under observed allocations, of a gain in social output; this metric answers the question: how much must the aggregate budget be increased under observed allocations to achieve the same growth in social output achieved through the counterfactual allocation? Table 3 (Col. 1) reports that the gain in social output is equivalent to a more-than-doubling of researchers' funding budgets (+168%), which, on the scale of the population, is equivalent

Bils, Klenow, and Ruane (2021) report for manufacturing in India.

to more than $30 billion per year. In other words, our estimates imply that it would take a tremendous amount of additional research funding to be allocated under the current regime to achieve the same social output that could be achieved with an optimal allocation of the current budget.

Two clarifications about what drives these gains. First, they do not come from eliminating fundraising. Appendix Table C2 shows that a planner who optimally reallocates funding and research time *while leaving every researcher's observed fundraising time in place* achieves nearly identical gains, and moving from that allocation to the full first-best adds almost nothing except for minor increases in researchers' utility. The time cost of the fundraising-based system is a second-order loss; the first-order loss is that funding does not flow to the most productive. Second, the gains do not appear to be an artifact of assumed parameters or measurement error (at least in the way we engage with measurement error). Appendix Table C1 re-computes the transition to the first-best while varying the externality scale $\kappa$, the share of variance attributed to measurement error ($\rho$, from 0 to 0.8), the correlation structure of the measurement-error shocks, winsorization of input values, and the number of Monte Carlo draws. Across these specifications, the observed allocative efficiency ranges from roughly 27% to 51%, and the population-level budget equivalent never falls below $18 billion per year — a figure attained only when we attribute 80% of the observed within-market variance to reporting noise. Under several plausible alternatives the implied gains are substantially *larger* than our baseline.

Still, the first-best benchmark clearly does not appreciate the realities of social planning. Perhaps a more practical question is: what if every market matched the allocative efficiency of the *best in-sample market* —an efficiency level some segment of science already attains? Appendix Figure C4 reports the results from this exercise. Bringing every market up to the best in-sample level would raise aggregate social output by roughly a third, which corresponds to a budget equivalent of about 55% of current funding. Even this in-sample benchmark indicates efficiency gains could be worth billions of dollars per year. The gains vary across fields, with budget equivalents ranging from about 25% in chemistry (which contains the most efficient markets) to nearly 140% in communication.

How much of the gap to the first-best can a policy close if the planner cannot observe productivity? Table 3 (Col. 2) reports the effects of the uniform funding policy. Allocative efficiency—measured against the policy's own resource envelope—rises from 50.7% to 78.5%, closing roughly half of the efficiency gap. Mean social output rises by 55%, with a budget equivalent of 88% of current funding, or roughly $17 billion per year at the

population level. The behavioral responses are instructive: with funding terms equalized, fundraising time more than quadruples on average. Many researchers who previously raised nothing begin to fundraise. Despite this, research time also rises by 11%, because access to funding raises the marginal product of time spent on research. The dispersion of funding falls by 39%, and researchers' mean utility rises slightly (+1.0%). A policy that simply puts everyone in a market on the same funding terms thus recovers a meaningful fraction of what an omniscient planner could achieve, while leaving researchers slightly better off on average.

Appendix Table C2 reports additional counterfactuals and policy simulations. Aligning researchers' private payoffs with the market-level social value function alone does almost nothing: social output rises by 1% and allocative efficiency is effectively unchanged (Col. 6). This suggests that there may not be large distortions in the *level* of incentives that researchers receive for their work; however, per our discussion above about the limitations of this particular simulation, we still cannot speak to directional distortions that may loom large (e.g., Carnehl and Schneider 2025).

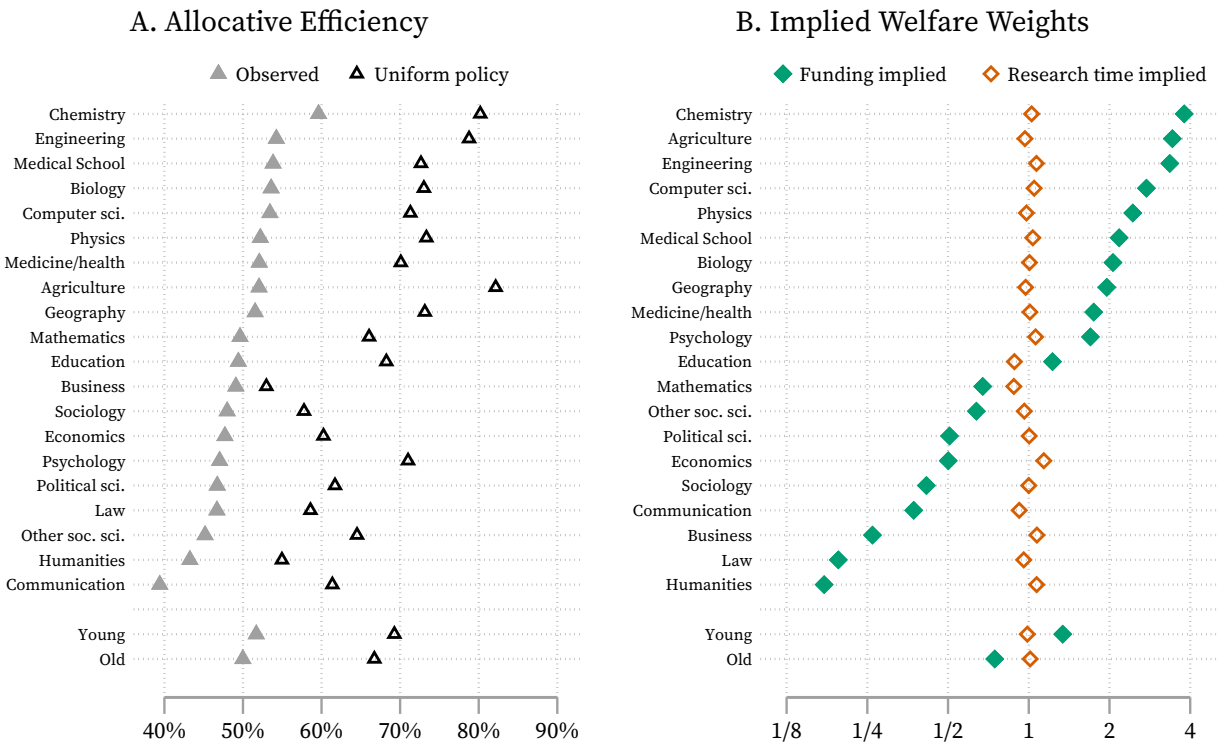
### 6.3. Efficiency and Welfare across Markets of Science

The proceeding analyses focus entirely on within-market efficiency, so as not to conflate inefficiencies with truly efficient idiosyncrasies unique to a market of science. Here, we relax that concern and take differences across markets more seriously as being informative of policies and/or preferences and ask: how efficient is each market, what distinguishes the efficient ones, and what do observed cross-market allocations reveal about the social values implicitly assigned to different fields of science?

*Efficiency by Field and Career Stage.* Figure 4A reports allocative efficiency by (*i*) field of study (aggregating the two career stage markets within each field) and (*ii*) career stage (aggregating the twenty fields within each career stage). Observed efficiency ranges from about 40% in communication to about 60% in chemistry. The ordering appears far from random: the most funding-intensive, laboratory-based, "hard-science" fields (e.g., chemistry, engineering, medical school disciplines) sit at the top, while less funding-intensive, "social-science" fields (e.g., humanities, law, communication) sit at the bottom. Early- and late-career markets are about equally efficient, with perhaps a slight edge to markets of young researchers. The uniform funding policy raises efficiency in every group, with gains ranging from a few percentage points in business to roughly thirty percentage points in agriculture, indicating that no field's observed funding heterogeneity is so well-targeted

that equalizing terms would decrease aggregate productivity.

FIGURE 4. Efficiency and Externalities




*Note*: Panel A shows allocative efficiency–aggregate social output relative to each group's own first-best frontier–by field and career stage, under the observed allocation and the uniform funding policy. Panel B shows the input-specific welfare weights that rationalize the observed cross-market allocation of funding (solid) and research time (hollow), normalized to geometric mean one; a weight above one indicates that the observed allocation implicitly values that field's output more than the average field's.

Why do some markets allocate resources better than others? As an exploratory exercise, we regress (log) market-level allocative efficiency on the (log) within-market variances of the model's states, choices, and parameters, using an adaptive lasso to select among them (see Appendix Table C3). We focus on the variance of market features because this moment of the distribution is plausibly less likely to be influenced by the "nature of science" within each field and more likely to reflect policies and norms in place. One predictor dominates: the variance of guaranteed funding. Its standardized coefficient is 0.83–0.92 depending on the predictor set, and the post-lasso regressions explain 75–80% of the cross-market variation; Appendix Figure C5 shows the raw relationship, a cross-market correlation of 0.88. Conditional on the variance of guarantees, the variance of fundraising ability ($\phi$) enters negatively.

Taken at face value, the pattern says that efficient markets allocate guaranteed funding unequally—as the first-best does, given very unequal productivities—while keeping the *cost* of raising additional funds more uniform. Inefficient markets exhibit the reverse: relatively flat guarantees combined with dispersed fundraising ability, which generates funding variation untethered from productivity. We emphasize the caveats: these are descriptive correlations and guaranteed funding is likely partly due past fundraising success, so high variance in guarantees may reflect another unobservable feature. Still, the result motivates the use of policies that lead to more homogeneous costs of fundraising while allowing for very heterogeneous funding guarantees.

*Implied Welfare Weights across Fields.* Finally, we zoom out from allocations within markets to the allocation across them. Debates about whether particular fields are over- or under-funded are difficult because they often conflate two distinct objects: how productive a field's researchers are, and how much society values a unit of that field's output. Our estimates let us hold the first object fixed and back out the second in a revealed preference sort of exercise. Specifically, we ask: given the observed allocation of inputs across markets and our estimates of marginal social products, what market-specific welfare weights $g_{mX}$ would a planner need for the observed cross-market allocation of input $X \in \{B, R\}$ to be efficient? The observed allocation is rationalized when:

$$g_{mX} \times \partial \mathcal{V}_m / \partial X_m = \text{a constant } \forall m \, ,$$

with these weights $g$ reflecting, for instance, field-specific externalities. Note that if a single planner allocated both inputs with one set of weights, the funding-implied and time-implied weights would coincide for each market (i.e., $g_{mB} = g_{mR}$). But clearly the two inputs here are allocated by different actors: research time largely by researchers themselves, funding largely by institutions and agencies. Thus, we compute the weights separately for each market-input pair $(g_{mB}, g_{mR})$.

Figure 4B reports the results, normalized so each input's weights have geometric mean one since we have no way to calibrate the constant term in the expression above . The research-time-implied weights are nearly uniform: every field, and both career stages, has a weight close to one. This does not certify the time allocation as socially efficient. Research time is chosen privately, so its returns are equated by researchers' own labor supply decisions; the uniform weights tell us the cross-field allocation of time looks like a private equilibrium, not that the equilibrium is optimal. Thus, these weights are not useful for comparing whether one market is over- or under-resourced relative to the social

optimum.

The funding-implied weights tell a very different story: they span a more-than-twenty-fold range, from about 4 in chemistry and 3.5 in agriculture and engineering down to one-quarter in business, one-fifth in law, and one-sixth in the humanities. The funding weight implied for early-career markets (about 1.3) is nearly twice that for late-career markets (about 0.75). Read literally, the U.S. science ecosystem allocates funding as if a unit of output from the laboratory natural sciences and engineering generates several times the external value of a unit from the social sciences and humanities (after accounting for differences in productivity and funding intensity). We interpret these weights cautiously, since they inherit all of the model's assumptions and identify only relative valuations. But they provide one of the first quantitative estimates of the field-level social valuations implicit in how the U.S. funds science, and a benchmark against which explicit valuations (e.g., from spillover studies, or from policymakers' stated priorities) can be compared.

## 7. Discussion

By some measures, the market for science believes that the benefits from a more efficient allocation are low and declining (or, the costs are high and increasing). Consider the share of science committed to *meta-science* — efforts to improve the efficiency of science writ large: in recent years, approximately 0.1-1.0% of all publications relate to meta-science; furthermore, while US investments in science grew roughly 3-fold over the past fifty years, the number of meta-science studies per dollar declined by nearly 80% in that same period.[30] This could reflect a rational, fully-informed allocation of resources. Or, it may reflect significant uncertainty about the degree to which the aggregate productivity of the scientific enterprise could be improved. In this paper, we use a new approach to productivity estimation in order to shed new light on the benefits from improving allocative efficiency in science.

Our approach allows us to estimate researchers' productivity beliefs without observing output quantities or input prices. We find observed input allocations to be positively correlated with those that maximize a plausible measure of the social value of science, but we also find very large gains to be had from more efficient allocations. In the counterfactuals we explore, aggregate productivity could be nearly doubled. Notably, simply equalizing the terms on which researchers in the same market access funding, our uniform funding

[30] See Appendix D for the methods used to produce these estimates.

policy simulation, could close roughly half of the gap between observed and first-best allocative efficiency.

Our results connect closely to work on the design of research funding mechanisms, most notably Adda and Ottaviani (2024). Much of the policy debate around grantmaking focuses on the average level of time researchers spend pursuing funding. Our counterfactuals suggest that, from the perspective of aggregate output, this concern may be second order. The first-order problem is selection — whether the mechanism directs funding to the researchers with the highest marginal returns — and this is precisely the margin highlighted by Adda and Ottaviani (2024). Our finding that the social sciences have lower allocative efficiency is consistent with Adda and Ottaviani (2024), who find these fields to have noisier evaluation processes (which generate the inefficiencies in their model). Combining researcher-level estimates of productivity and fundraising ability like ours with equilibrium models of grant competition like theirs strikes us as a promising path toward the optimal design of scientific funding mechanisms.

We also shed new light on debates surrounding the allocation of funding across science. To summarize the challenge, consider this quote from Congressman Sherwood Boehlert, the chairman of the U.S. House Science Committee in 2001, who summarized the central difficulty of allocating research budgets:

> "*The science and engineering community tends to feel, with some reason, that every field and sub-discipline is underfunded . . . but money is not likely to become available to 'fully fund' every field even if we could agree on what that would mean.*"

Our across-market analysis is a step towards a more disciplined discussion about "*what that would mean*". Per our model, the funding-implied welfare weights shown in Figure 4B are key objects to debate. For example, if you think the positive social externalities of knowledge produced by chemists is roughly eight times larger than it is for political scientists, then the model would suggest that the *relative* amount of funding in those two fields is currently efficient; if you think knowledge in chemistry is more or less than eight times valuable than knowledge in political science, then the model reveals an inefficiency.

The model abstracts away from many important considerations that likely drive actual allocations (e.g., dynamic returns), and our data relies on researchers' stated preferences. These and other limitations of our approach motivate us to interpret these magnitudes as plausible upper bounds on the gains from more efficient use of scientific inputs. Still, these large potential gains provide a glimmer of hope amid declining R&D productivity across

most sectors of the economy (Bloom et al. 2020) and the persistently growing burden of knowledge that raises the cost of conducting frontier science (Jones 2009). Furthermore, our approach to identifying productive scientists may prove useful in talent selection processes more generally (e.g., Agarwal and Gaule 2020).

Our approach has deep roots in the economics of labor (i.e., surveys of time use and work-leisure trade-offs), industrial organization (i.e., production function estimation), marketing (i.e., using surveys to identify demand functions), and macroeconomics (i.e., models of factor misallocation). By drawing on insights from these fields, our methodology allows us to overcome many of the challenges that have long plagued our understanding of productivity and efficiency in science.

Our analyses reveal researchers' *beliefs* about their productivity and therefore their *beliefs* about how well inputs are allocated. This is a crucial limitation, since there are many potential biases affecting these beliefs. Still, science is inherently about forecasting uncertain outcomes; therefore, the optimal mechanisms for identifying productive researchers and allocating them more inputs will need to tackle this challenge of engaging with researchers' forecasts of their productivity. Specifically, one important next step in this line of work will be to ensure that the producers being studied report their willingness-to-pay for inputs truthfully. Of course, the theoretical underpinnings of how to elicit true willingness-to-pay estimates have long been established (e.g., Becker, DeGroot, and Marschak 1964). However, the magnitude of costs in our setting is on the scale of tens to hundreds of thousands of dollars. Thus, developing the practical details of incentivizing truthful responses from producers at this scale would be a fruitful endeavor.

More broadly, our approach may prove useful in other settings. There are many markets populated by a large number of producers acquiring inputs in a highly decentralized way to produce outputs that are not easy to observe. In Appendix E, we provide a range of alternative formulations of our model they may motivate its usefulness elsewhere. For example, in developing economies, accurate producer-level data on outputs can be difficult to obtain (e.g., Tybout 2000); in entrepreneurship, many organizations never produce observable output before exiting (e.g., Decker et al. 2014); and in nonprofit sectors, the output may be so high-dimensional that reaching consensus on a suitable proxy is challenging (e.g., Philipson and Lakdawalla 2001). In each of these examples, our methodology could provide a way forward to better understand the distribution and determinants of productivity and efficiency.

# References


Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica* 83 (6): 2411–2451.

Adams, James D, and Zvi Griliches. 1998. "Research Productivity in a System of Universities." *Annals of Economics and Statistics* 49-50: 127–162.

Adda, Jérôme, and Marco Ottaviani. 2024. "Grantmaking, grading on a curve, and the paradox of relative evaluation in nonmarkets." *The Quarterly Journal of Economics* 139 (2): 1255–1319.

Agarwal, Ruchir, and Patrick Gaule. 2020. "Invisible geniuses: Could the knowledge frontier advance faster?" *American Economic Review: Insights* 2 (4): 409–424.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic inputs and resource (mis) allocation." *Journal of Political Economy* 122 (5): 1013–1063.

Atkin, David, Amit K. Khandelwal, and Adam Osman. 2019. "Measuring Productivity: Lessons from Tailored Surveys and Productivity Benchmarking." *AEA Papers and Proceedings* 109: 444–49.

Azoulay, Pierre, Christian Fons-Rosen, and Joshua S Graff Zivin. 2019. "Does science advance one funeral at a time?" *American Economic Review* 109 (8): 2889–2920.

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven Sampat. 2019. "Public R&D investments and private-sector patenting: Evidence from NIH funding rules." *Review of Economic Studies* 86 (1): 117–152.

Baruffaldi, Stefano, and Fabian Gaessler. 2025. "The returns to physical capital in knowledge production: Evidence from lab disasters." *American Economic Journal: Applied Economics* Forthcoming.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak. 1964. "Measuring utility by a single-response sequential method." *Behavioral Science* 9 (3): 226–232.

Bergquist, Lauren F, Danial Lashkari, and Eric Verhoogen. 2026. "Wedges: A Microeconomic Perspective on Misallocation." *Mimeo*.

Bils, Mark, Peter J Klenow, and Cian Ruane. 2021. "Misallocation or mismeasurement?" *Journal of Monetary Economics* 124: S39–S56.

Bloom, Nicholas, Charles I Jones, John Van Reenen, and Michael Webb. 2020. "Are ideas getting harder to find?" *American Economic Review* 110 (4): 1104–1144.

Bloom, Nicholas, Raffaella Sadun, and John Van Reenen. 2012. "The organization of firms across countries." *The Quarterly Journal of Economics* 127 (4): 1663–1705.

Bryan, Kevin A, and Heidi L Williams. 2021. "Innovation: Market failures and public policies." In *Handbook of Industrial Organization*, vol. 5, 281–388: Elsevier.

Carnehl, Christoph, Marco Ottaviani, and Justus Preusser. 2025. "Designing scientific grants." *Entrepreneurship and Innovation Policy and the Economy* 4 (1): 139–178.

Carnehl, Christoph, and Johannes Schneider. 2025. "A quest for knowledge." *Econometrica* 93 (2): 623–659.

Carrillo, Paul, Dave Donaldson, Dina Pomeranz, and Monica Singhal. 2023. "Misallocation in firm production: A nonparametric analysis using procurement lotteries." *Mimeo*.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery. 2013. "Barriers to household risk management: Evidence from India." *American Economic*

*Journal: Applied Economics* 5 (1): 104–135.

De Loecker, Jan, Pinelopi K Goldberg, Amit K Khandelwal, and Nina Pavcnik. 2016. "Prices, markups, and trade reform." *Econometrica* 84 (2): 445–510.

De Loecker, Jan, and Chad Syverson. 2021. "An industrial organization perspective on productivity." In *Handbook of Industrial Organization*, vol. 4, 141–223: Elsevier.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The role of entrepreneurship in US job creation and economic dynamism." *Journal of Economic Perspectives* 28 (3): 3–24.

Digital Science. 2018. "Dimensions [Software]." Available from https://app.dimensions.ai. Accessed in 2023, under licence agreement.

Dizon-Ross, Rebecca, and Seema Jayachandran. 2022. "Improving Willingness-to-Pay Elicitation by Including a Benchmark Good." *AEA Papers and Proceedings* 112: 551–555.

Fleming, Lee, Hillary Greene, G Li, Matt Marx, and Dennis Yao. 2019. "Government-funded research increasingly fuels innovation." *Science* 364 (6446): 1139–1141.

Gibbons, Michael T, and NCSES. 2024. "Higher Education R&D Expenditures Increased 11.2%, Exceeded $108 Billion in FY 2023." National Science Foundation 25-313. https://ncses.nsf.gov/pubs/nsf25313.

Gollin, Douglas, and Christopher Udry. 2021. "Heterogeneity, measurement error, and misallocation: Evidence from African agriculture." *Journal of Political Economy* 129 (1): 1–80.

Hager, Sebastian, Carlo Schwarz, and Fabian Waldinger. 2024. "Measuring science: Performance metrics and the allocation of talent." *American Economic Review* 114 (12): 4052–4090.

Hall, Bronwyn H, and Jacques Mairesse. 2024. "Explorations of Cumulative Advantage Using Data on French Physicists."

Heckman, James J. 1979. "Sample selection bias as a specification error." *Econometrica* 47 (1): 153–161.

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

Hill, Ryan, and Carolyn Stein. 2025. "Race to the bottom: Competition and quality in science." *The Quarterly Journal of Economics* 140 (2): 1111–1185.

Hoffman, Mitchell, and Christopher T. Stanton. 2024. "People, Practices, and Productivity: A Review of New Advances in Personnel Economics." In *Handbook of Labor Economics*,: Elsevier.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4): 1403–1448.

Ioannidis, John P.A. 2018. "Science Funding is Broken."

Jena, Anupam B, and Tomas J Philipson. 2008. "Cost-effectiveness analysis and innovation." *Journal of Health Economics* 27 (5): 1224–1236.

Jones, Benjamin F. 2009. "The burden of knowledge and the "Death of the renaissance man": Is innovation getting harder?" *Review of Economic Studies* 76 (1): 283–317.

Jones, Charles I. 2005. "Growth and ideas." In *Handbook of economic growth*, vol. 1, 1063–1111: Elsevier.

Keiller, Agnes Norris, Áureo de Paula, and John Van Reenen. 2024. "Production function estimation using subjective expectations data.", National Bureau of Economic Research.

Kim, Kyoo, Amil Petrin, and Suyong Song. 2016. "Estimating production functions with control functions when capital is measured with error." *Journal of Econometrics* 190 (2): 267–279.

Lakdawalla, Darius N, Eric C Sun, Anupam B Jena, Carolina M Reyes, Dana P Goldman, and Tomas J Philipson. 2010. "An economic evaluation of the war on cancer." *Journal of Health Economics* 29 (3): 333–346.

Levinsohn, James, and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *Review of Economic Studies* 70 (2): 317–341.

Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: The University of Chicago Press.

Mills, M. Anthony. 2021. "Fix Science, Don't Just Fund it."

Myers, Kyle R. 2020. "The Elasticity of Science." *American Economic Journal: Applied Economics* 12 (4): 103–134.

Myers, Kyle R, Wei Yang Tham, Nina Cohodes, Karim Lakhani, Rachel Mural, Jerry G Thursby, Marie C Thursby, and Yilun Xu. 2023. "New facts and data about US research professors."

Myers, Kyle R, Wei Yang Tham, Yian Yin, Nina Cohodes, Jerry G Thursby, Marie C Thursby, Peter Schiffer, Joseph T Walsh, Karim R Lakhani, and Dashun Wang. 2020. "Unequal effects of the COVID-19 pandemic on scientists." *Nature Human Behaviour* 4 (9): 880–883.

National Science Foundation. 2023. "Higher Education Research and Development Survey (HERD)." Available from https://www.nsf.gov/statistics/srvyherd/. Accessed in 2023.

Nordhaus, William D. 2004. "Schumpeterian profits in the American economy: Theory and measurement."

Olley, G Steven, and Ariel Pakes. 1996. "The dynamics of productivity in the telecommunications equipment industry." *Econometrica* 64 (6): 1263.

Petrin, Amil, and Jagadeesh Sivadasan. 2013. "Estimating lost output from allocative inefficiency, with an application to Chile and firing costs." *Review of Economics and Statistics* 95 (1): 286–301.

Philipson, Tomas, and Darius Lakdawalla. 2001. "Medical care output and productivity in the nonprofit sector." In *Medical Care Output and Productivity*, edited by David M Cutler and Ernst R. Berndt, 119–140: University of Chicago Press.

Piper, Kelsey. 2019. "Science funding is a mess. Could grant lotteries make it better?".

Qiu, YJ Jeff. 2023. "The Matthew effect, research productivity, and the dynamic allocation of NIH grants." *RAND Journal of Economics* 54 (1): 135–164.

Restuccia, Diego, and Richard Rogerson. 2017. "The causes and costs of misallocation." *Journal of Economic Perspectives* 31 (3): 151–174.

Rohn, Jenny. 2017. "The way research is funded is harmful to science — it's time for a change."

Rosenbloom, Joshua L, Donna K Ginther, Ted Juhl, and Joseph A Heppert. 2015. "The effects of Research & Development funding on scientific productivity: Academic chemistry, 1990-2009." *PLoS ONE* 10 (9): e0138176.

Sackman, Harold, Warren J Erikson, and E Eugene Grant. 1968. "Exploratory experimental studies comparing online and offline programming performance." *Communications of the ACM* 11 (1): 3–11.

Shapin, Steven. 1995. "Here and Everywhere: Sociology of Scientific Knowledge." *Annual Review of Sociology* 21 (1): 289–321.

Sinatra, Roberta, Dashun Wang, Pierre Deville, Chaoming Song, and Albert-László Barabási. 2016.

"Quantifying the evolution of individual scientific impact." *Science* 354 (6312): aaf5239.

Sraer, David, and David Thesmar. 2023. "How to use natural experiments to estimate misallocation." *American Economic Review* 113 (4): 906–938.

Stephan, Paula E. 1996. "The economics of science." *Journal of Economic Literature* 34 (3): 1199–1235.

Syverson, Chad. 2011. "What determines productivity?" *Journal of Economic Literature* 49 (2): 326–365.

Tybout, James R. 2000. "Manufacturing firms in developing countries: How well do they do, and why?" *Journal of Economic Literature* 38 (1): 11–44.

Wossen, Tesfamicheal, David J Spielman, Arega D Alene, and Tahirou Abdoulaye. 2024. "Estimating seed demand in the presence of market frictions: Evidence from an auction experiment in Nigeria." *Journal of Development Economics* 167: 103242.

Zuckerman, H. 1988. "The Sociology of Science." In *Handbook of Sociology*, edited by NJ Smelser, 511–574: Sage Publications.

# Supplemental Appendix

## A. Additional Model Details

### A.1. Derivation of Policy Functions

The policy functions $\mathcal{R}(\mathbf{S}_i, \theta_i, \mu_i)$, $\mathcal{F}(\mathbf{S}_i, \theta_i, \mu_i)$ and $\mathcal{H}(\mathbf{S}_i, \theta_i, \mu_i)$ characterize the solution $(R_i^*, F_i^*, H_i^*)$ to problem (1) for each individual scientist $i$ as a function of states $\mathbf{S}_i$, attributes $\theta_i$, and parameters $\mu_i$. They determine indirect utility $\mathcal{U}_i^* = \mathcal{U}(\mathbf{S}_i, \theta_i, \mu_i)$ at current allocations. Substituting the constraints, the utility maximization problem is:

$$\mathcal{U}(\cdot) = \max_{F_i, H_i} m_i(M_i) + b_i(\alpha_i \left(B_{\min} + G_i + \phi_i F_i\right)^{\gamma_i} \left(H_i - F_i - D_i\right)^{1-\gamma_i}) - c_i(H_i, D_i) \,, \tag{A1}$$

with the policy functions $\mathcal{F}(\cdot,\cdot,\cdot)$ and $\mathcal{H}(\cdot,\cdot,\cdot)$ solving the optimality conditions:

$$\frac{\partial b_i(Q_i)}{\partial Q_i}\frac{\partial Q_i}{\partial F_i} + \lambda_{F,i} = 0 \tag{A2}$$

$$\frac{\partial b_i(Q_i)}{\partial Q_i}\frac{\partial Q_i}{\partial H_i} - \frac{\partial c_i(H_i, D_i)}{\partial H_i} - \lambda_{H,i} = 0 \,, \tag{A3}$$

and $\mathcal{R}(\cdot,\cdot,\cdot)$ being residually determined based on the time-constraint (1c). To derive the policy functions, we start from model's optimality conditions (A2) and (A3) and evaluate them using the functional forms described in the main text. We obtain the conditions:

$$Q_i^{1-\eta_i}\Big[\gamma_i\phi_i(B_{\min} + G_i + \phi_i F_i)^{-1} - (1-\gamma_i)(H_i - D_i - F_i)^{-1}\Big] + \lambda_{F,i} = 0 \tag{A4}$$

$$Q_i^{1-\eta_i}(H_i - D_i - F_i)^{-1}(1-\gamma_i) - \psi(H_i - D_i + D_i^{\xi_i})^{\zeta_i} - \lambda_{H,i} = 0 \,. \tag{A5}$$

We characterize the policy functions in two intervals of $H_i$. The first is $H_i \in (D_i, H_{i,F>0}]$ and the second is $H_i \in (H_{i,F>0}, H_{\max}]$, depending on whether the optimal fundraising time allocation is a corner solution ($F_i = 0$) or not ($F_i > 0$). We identify the threshold $H_{i,F>0}$ below which optimal fundraising is zero by solving (A5) for $F_i$ as a function of $H_i$ assuming that $F_i$ is strictly positive, i.e., $\lambda_{F,i} = 0$:

$$F_i = \gamma_i(H_i - D_i) - \frac{1-\gamma_i}{\phi_i}(B_{\min} + G_i) \,. \tag{A6}$$

Therefore, (A6) implies that fundraising time is strictly positive as long as:

$$H_i > D_i + \frac{1-\gamma_i}{\gamma_i \phi_i}(B_{\min} + G_i)\,, \tag{A7}$$

which identifies the threshold $H_{i,F>0}$ that is strictly larger than $D_i$ as long as $\gamma_i < 1$. As a consequence, we can solve for optimal hours $H_i$ using equation (A5), which takes the piece-wise functional form:

$$\begin{cases} (1-\gamma_i)[\alpha_i(B_{\min}+G_i)^{\gamma_i}]^{1-\eta_i}(H_i - D_i)^{(1-\gamma_i)(1-\eta_i)-1} - \psi(H_i - D_i + D_i^{\xi_i})^{\zeta_i} - \lambda_{H,i} = 0 & \text{if } H_i \le H_{i,F>0} \\ \left[\alpha_i(\phi_i\gamma_i)^{\gamma_i}(1-\gamma_i)^{1-\gamma_i}\right]^{1-\eta_i}\left(H_i - D_i\frac{B_{\min}+G_i}{\phi_i}\right)^{-\eta_i} - \psi(H_i - D_i + D_i^{\xi_i})^{\zeta_i} - \lambda_{H,i} = 0 & \text{if } H_i > H_{i,F>0} \end{cases} \tag{A8}$$

and defines an implicit solution $\mathcal{H}_i$ for $H_i$ as a function of all parameters and state variables. (A8) is continuous but not differentiable at $H_{i,F>0}$. Moreover, the Lagrange multiplier $\lambda_{H,i}$ equals zero for $H_i < H_{\max}$ and $H_i$ is always strictly larger than $D_i$ because $\lim_{H_i \to D_i}$ (A8) $= +\infty$. Hence, the non-negativity constraint on research time is always met. Given the policy function $\mathcal{H}_i$ for hours, the functions defining optimal allocation to fundraising time and research time are:

$$\mathcal{F}_i = \begin{cases} 0 & \text{if } \mathcal{H}_i \le H_{i,F>0} \\ \gamma_i(\mathcal{H}_i - D_i) - \frac{1-\gamma_i}{\phi_i}(B_{\min} + G_i) & \text{if } \mathcal{H}_i > H_{i,F>0} \end{cases} \tag{A9}$$

and

$$\mathcal{R}_i = \mathcal{H}_i - D_i - \mathcal{F}_i\,. \tag{A10}$$

### A.2. Estimation Outline

We infer individual attributes and parameters using a mix of calibration and estimation. First, we set minimum funding $B_{min} = \$5,000$ and fix total hours endowment $H_{max}$ to 62 work hours per week (approximately the 90th percentile of the observed work hours). Second, we estimate individual attributes and common parameters of the utility function, including deep parameters that determine individual-specific parameters in $\mu_i$ as functions of researcher's type $T_i$. We use survey information on hours worked, research time, fundraising time, guaranteed funding, expected additional funds raised, duties, and the salaries reported in the alternative offer experiments, which we map to model counterparts $(H_i, R_i, F_i, G_i, \phi_i F_i, D_i, \{M_{ij}\}_{j=1}^4)$, respectively.

Conditional on some common parameters, we can infer individual-specific attributes

$\theta_i$ from researchers' optimality conditions and model's structure. One important note is that we have separate estimation routines for inferring individual attributes $\theta_i$ among researchers with non-zero fundraising time ($F_i > 0$) and those at a corner solution ($F_i = 0$). For the former group, we first infer fundraising ability $\phi_i$ by exploiting the identity equating the observed, expected additional funding ($EG_i$) to $\phi_i F_i$. Therefore, $\forall i = 1, ..., N_F$:

$$\widehat{\phi}_i = \frac{EG_i}{F_i} \,. \tag{A11}$$

Expression (A11) is well-defined if and only if observed $F_i > 0$, which is why the inference strategy must differ for researchers with $F_i = 0$ at observed allocations.[31] As a second step, conditional on $\widehat{\phi}_i$, we infer factor shares $\widehat{\gamma}_i$ from researchers' first order condition in fundraising time, which we derive in Appendix A.1. For the chosen functional forms, this takes the expression:

$$\widehat{\gamma}_i = \frac{B_i}{B_i + \phi_i R_i} = \frac{B_{\min} + G_i + \widehat{\phi}_i F_i}{B_{\min} + G_i + \widehat{\phi}_i F_i + \widehat{\phi}_i R_i} \,. \tag{A12}$$

Equation (A12) states that $\gamma_i$ constitutes the weight of total funding over the total dollar-value of inputs used in scientific activity, with the last term of the denominator being the dollar-valued opportunity cost of research time relative to fundraising. Finally, the optimality condition for total hours worked determines productivity $\widehat{\alpha}_i$ as a function of parameters, observed allocations, and $\widehat{\phi}_i$ and $\widehat{\gamma}_i$:

$$(1-\widehat{\gamma}_i)\widehat{\alpha}_i^{1-\eta_i}(B_{\min} + G_i + \widehat{\phi}_i F_i)^{(1-\eta_i)\widehat{\gamma}_i} R_i^{(1-\eta_i)(1-\widehat{\gamma}_i)-1} = \psi(H_i - D_i + D_i^{\xi_i})^{\zeta_i} \,, \tag{A13}$$

This equation holds exactly if $H_i < H_{\max}$. Therefore, Equations (A11), (A12), and (A13) determine individual attributes as functions of parameters and observed allocations for researchers with positive fundraising time. Unfortunately, for the group of researchers reporting zero fundraising time, we can neither infer $\widehat{\phi}_i$ from Equation (A11), nor can we compute $\widehat{\gamma}_i$.[32] Therefore, we assume that $\phi_i$ and $\gamma_i$ are parametric polynomial functions of state variables, and we estimate these functions using the other sub-sample of researchers with $F_i > 0$.[33]

[31] For numerical stability, we constrain $\widehat{\phi}_i \geq 1$.

[32] The Lagrange multiplier $\lambda_{F,i}$ is strictly positive and unknown.

[33] We use the following specifications: $\phi_i = \exp\left\{\sum_{p=1}^{3} \beta_{G,p} G_i^p + \sum_{p=1}^{3} \beta_{D,p} D_i^p + \sum_{p=1}^{3} \beta_{M,p} M_i^p\right\}$, and $\gamma_i = \left(1 + \exp\left\{\sum_{p=1}^{3} \iota_{G,p} G_i^p + \sum_{p=1}^{3} \iota_{D,p} D_i^p + \sum_{p=1}^{3} \iota_{M,p} M_i^p\right\}\right)^{-1}$, which we estimate using poisson and logistic regressions. Whenever the fitted $(\widehat{\phi}_i, \widehat{\gamma}_i)$ combination would imply, given the observed state variables and for

Finally, given the estimates attributes $\widehat{\phi}_i$ and $\widehat{\gamma}_i$, the state variables, and the vector of common parameters $\widehat{\mu}$, we infer productivity beliefs $\widehat{\alpha}_i$ through Equation (A13). Given estimates of individual attributes $\widehat{\theta}_i = (\widehat{\alpha}_i, \widehat{\gamma}_i, \widehat{\phi}_i)$ as functions of parameters $\mu = (\omega, \psi, \delta)$, calibrated values $(B_{\min}, H_{\max})$, and observed allocations, we estimate the vector $\mu$ by generalized method of moments as:

$$\widehat{\mu} = \arg\min_{\mu} \mathcal{L}(\mu) = \arg\min_{\mu} \sum_{i=1}^{N} \sum_{j=1}^{4} \left( M_{ij}^{\text{obs}} - \widehat{M}_{ij}(\mathbf{S}_{ij}, \widehat{\theta}_i(\mu_i(\mu, T_i)), \mu_i(\mu, T_i)) \right)^2 , \qquad \text{(A14)}$$

where $M_{ij}$ is researcher $i$'s answer to thought experiment $j$ in the data and $\widehat{M}_{ij}$ is its model-based counterpart, which is itself an implicit function of the vector of counterfactual states $\mathbf{S}_i$, researcher's type $T_i$, and of the estimand $\mu$ which determines individual attributes $\widehat{\theta}_i$ and individual-specific parameters $\mu_i$.[34] Parameter estimates $\widehat{\mu}$ solve (A14) conditional on parameter restrictions specified in Section 2.2. Appendix A.3 provides additional details on the search algorithm.

### A.3. Estimation Algorithm

Here, we describe the estimation of the common parameters $\mu = (\omega, \psi, \delta_{\sigma,0}, \delta_{\sigma,1}, \delta_{\eta,0}, \delta_{\eta,1}, \delta_{\xi,0}, \delta_{\xi,1}, \delta_{\zeta,0}, \delta_{\zeta,1})$. First, we define a grid of parameter values at which we perform a preliminary evaluation of the loss function in (A14). The grid is defined by the Cartesian product of the following:

$$\begin{aligned}
&\omega^{(0)} = [0.1, 1, 10] , \quad \psi^{(0)} = [0.00001, 1, 10] , \\
&\delta_{\sigma,0}^{(0)} = [-100, 0] , \quad \delta_{\sigma,1}^{(0)} = [0] , \\
&\delta_{\eta,0}^{(0)} = [\ln(0.8) - \ln(0.2), 0, \ln(0.2) - \ln(0.8)] , \quad \delta_{\eta,1}^{(0)} = [0] \\
&\delta_{\xi,0}^{(0)} = [-11.5, \ln(2)] , \quad \delta_{\xi,1}^{(0)} = [0] , \\
&\delta_{\zeta,0}^{(0)} = [-11.5, 0] , \quad \delta_{\zeta,1}^{(0)} = [0] .
\end{aligned}$$

Second, we select the grid-points where the value of the loss function is within 0.5% of the minimum, and we use them as initial points for the numerical solution. We perform a preliminary "Nelder-Mead simplex direct search" with a high tolerance on the loss function

a specific vector of parameters $\mu_i$, that the optimal (observed) time allocation to fundraising would be strictly positive, we re-scale $\widehat{\phi}_i$ below the individual-specific lower bound below which the optimal fundraising time at observed states is indeed null.

[34] To save notation, we omit that $\widehat{M}_{ij}$ are also a function of calibrated $(B_{\min}, H_{\max})$.

and standard tolerance on parameter values. We then select the parameters vectors where the value of the loss function is within 1% of the minimum, and we use them as initial points with the same search algorithm and with lower loss function tolerance. We repeat this step twice until we are left with a single candidate optimal parameter vector.

### A.4. Measurement Error and Bayesian Shrinkage

We account for classical measurement error in the researcher-level data with a latent variable model and auxiliary assumption about the correlation between reports and the truth, which facilitates a Bayesian shrinkage procedure. To simplify the complexity of this approach, we treat the common parameters estimated from the WTP moments as fixed, and then integrate over posterior draws of the latent researcher-level primitives. Thus, the uncertainty captured here is measurement error in reported data, not sampling uncertainty in the common parameters estimated in the GMM routine.

Let $m(i)$ denote researcher $i$'s market. Let $Y_{ij}^{obs}$ denote an observed researcher-level datapoint, where $j$ indexes variables such as salary, guaranteed funding, research time, etc. For each market $m$ and variable $j$, define the mean and variance as $\mu_{mj}$ and $s_{mj}^2$, respectively.

We assume that the observed datapoint is the sum of a latent true value and classical reporting error:

$$Y_{ij}^{obs} = Y_{ij}^{*} + \nu_{ij},$$

with

$$Y_{ij}^{*} \mid m(i) = m \sim N\Big(\mu_{mj}, (1-\rho) \times s_{mj}^2\Big), \qquad \nu_{ij} \mid m(i) = m \sim N\Big(0, \rho \times s_{mj}^2\Big)\ ,$$

and $Y_{ij}^{*} \perp \nu_{ij}$. This model preserves the observed within-market variance as the sum of true heterogeneity and reporting noise.[35]

The scalar $\rho \in [0, 1]$ is an exogenous parameter that describes the share of observed within-market variance attributed to reporting error. Thus, $\rho = 0$ implies no measurement error and no shrinkage, while larger values of $\rho$ imply that more of the observed cross-sectional dispersion is noise, and the expression $\sqrt{1-\rho}$ describes the correlation between $Y_{ij}^{obs}$ and $Y_{ij}^{*}$.

[35] We use log transformations when shrinking positive-valued variables and exponentiate the posterior.

The Gaussian structure gives a closed-form posterior. Since

$$\mathrm{Cov}(Y_{ij}^{*}, Y_{ij}^{obs} \mid m) = (1-\rho) \times s_{mj}^{2}, \qquad \mathrm{Var}(Y_{ij}^{obs} \mid m) = s_{mj}^{2},$$

the posterior mean is

$$\mathbb{E}\left[Y_{ij}^{*} \mid Y_{ij}^{obs}, m(i) = m\right] = \mu_{mj} + (1-\rho) \times (Y_{ij}^{obs} - \mu_{mj}),$$

and the posterior variance is

$$\mathrm{Cov}\left(Y_{ij}^{*} \mid Y_{ij}^{obs}, m(i) = m\right) = \rho \times (1-\rho) \times s_{mj}^{2}.$$

Therefore, for a given random draw $b$ in a Monte Carlo simulation, the posterior is

$$Y_{ijb}^{*} = \mu_{m(i)j} + (1-\rho) \times \left(Y_{ij}^{obs} - \mu_{m(i)j}\right) + \sqrt{\rho \times (1-\rho)} \times s_{m(i)j} \times \varepsilon_{ijb}, \qquad \varepsilon_{ijb} \sim N(0,1).$$

The baseline specification draws $\varepsilon_{ijb}$ independently across researchers, variables, and Monte Carlo draws. We also consider alternative covariance structures. In one specification, all variables for a researcher receive the same shock, so $\varepsilon_{ijb} = \varepsilon_{ib}$. In another, the vector of shocks for researcher $i$ is drawn from a multivariate normal distribution whose within-market correlation matrix equals the observed cross-variable correlation matrix of the transformed primitives. These alternatives change the dependence structure of measurement error across variables, but not the marginal posterior mean or variance.

For each draw $b$, the latent primitives $Y_{ib}^{*}$ are constructed and treated as the data. Given these primitives, we reconstruct the researcher-specific model objects $\phi_{ib}$, $\gamma_{ib}$, $\alpha_{ib}$, $\sigma_{ib}$, $\eta_{ib}$, $\xi_{ib}$, and $\zeta_{ib}$ using the same structural relationships as in the main estimator, while holding the common parameters fixed. We restrict posterior primitives to the observed within-market support and winsorize reconstructed parameters at pre-specified percentiles; these restrictions are imposed within draw before any counterfactual is solved.

All nonlinear objects (i.e., parameters, counterfactuals, gains from reallocation) are calculated within each posterior draw before aggregation to obey Jensen's inequality. We use 50 Monte Carlo draws by default, since tests using more draws (one of which we report below) show virtually no changes in any of our key statistics beyond this number of draws despite computation time increasing substantially (since counterfactual allocations must be solved within each draw).

# B. Additional Survey Results

## B.1. Additional Survey Results and Descriptive Statistics

Table B1 shows the effects of the randomized participation incentives and reminders on survey completion. Figures B1–B2 illustrate sample representativeness and Figure B3 documents the alignment between self- and publicly-reported salaries (see Myers et al. (2023) for more on the sample). Table B2 describes the summary statistics for the sample on numerous dimensions. Table B4 shows the pairwise correlations of the subjective output measures.

TABLE B1. Survey Completion per Randomized Treatments

| | (1) |
|---|---|
| Either incentive | 0.00297** |
| | (0.00117) |
| Both incentives | 0.00797*** |
| | (0.00141) |
| 1 reminder | 0.0112*** |
| | (0.00114) |
| 2 reminders | 0.0188*** |
| | (0.00119) |
| Constant | 0.0197*** |
| | (0.00107) |
| *N* obs. | 131,672 |

*Note*: Estimates from a linear regression of an indicator for survey completion on randomly assigned incentive and reminder arms for the 131,672 researchers e-mailed. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## B.2. Reducing Dimensions of Heterogeneity: $k$-means Clustering Results

Ideally, our model would allow for rich heterogeneity in researchers' preferences (i.e., the $u_{1,i}(\cdot)$, $u_{2,i}(\cdot)$, and $u_{3,i}(\cdot)$ functions). This would help ensure that (true) variation in preferences would not mistakenly be attributed as (estimated) variation in productivity. We lack enough data to estimate researcher-specific preference functions, but we do have the large vector of observables from the survey $\mathbf{Z}_i$.[36] This motivates the following approach. We use $k$-means clustering to collapse the multi-dimensional heterogeneity

[36]We could allow the parameters that govern the preference functions to depend on deep parameters and these observables. For example, for the preference function parameter $\sigma_i$, we could assume that

TABLE B2. Summary Statistics of Other Variables

| | *mean* | *s.d.* |
|---|---|---|
| *From HERD: Institution-level R&D* | | |
| Total R&D, $M | 611.32 | 451.18 |
| R&D per researcher, $M | 0.61 | 0.86 |
| Share federal gov.'t R&D, [0,1] | 0.52 | 0.12 |
| Share basic R&D, [0,1] | 0.63 | 0.20 |
| *From Dimensions: Research output* | | |
| Publications per year | 5.28 | 7.09 |
| Citations per year | 23.07 | 49.10 |
| Co-authors per publication per year | 9.09 | 68.40 |
| *Position details* | | |
| Assistant professor, {0,1} | 0.26 | 0.44 |
| Associate professor, {0,1} | 0.26 | 0.44 |
| Full professor, {0,1} | 0.41 | 0.49 |
| Other rank, {0,1} | 0.07 | 0.25 |
| Not on tenure track, {0,1} | 0.19 | 0.39 |
| Pre-tenure, {0,1} | 0.22 | 0.42 |
| Tenured, {0,1} | 0.58 | 0.49 |
| Years until next contract eval. | 3.68 | 1.82 |
| Duration of contract | 2.45 | 1.93 |

*Note*: Summary statistics for the 4,003 researchers in the estimation sample. "From HERD" variables describe the researcher's institution (National Science Foundation 2023); "From Dimensions" variables are from matched publication records (Digital Science 2018).

TABLE B3. Summary Statistics of Other Variables (cont.'d)

| | *mean* | *s.d.* |
|---|---|---|
| *Gender identity*, {0,1} | | |
| Female | 0.40 | 0.49 |
| Male | 0.55 | 0.50 |
| Other or N.R. | 0.05 | 0.22 |
| *Racial/ethnic identity*, {0,1} | | |
| Asian | 0.13 | 0.33 |
| Black | 0.03 | 0.18 |
| Hispanic | 0.06 | 0.23 |
| White | 0.77 | 0.42 |
| Other or N.R. | 0.05 | 0.21 |
| *Citizenship*, {0,1} | | |
| Citizen, domestic-born | 0.71 | 0.45 |
| Citizen or perm. resident, foreign-born | 0.24 | 0.43 |
| Other or N.R. citizenship | 0.05 | 0.21 |
| 1st–3rd generation in U.S. | 0.30 | 0.46 |
| Other or N.R. generation in U.S. | 0.70 | 0.46 |
| *Other covariates* | | |
| Age | 48.78 | 12.05 |
| Household total income | 260,336.00 | 215,388.31 |
| Married or domestic partnership, {0,1} | 0.82 | 0.39 |
| Single, {0,1} | 0.13 | 0.34 |
| Other or N.R. relationship, {0,1} | 0.05 | 0.22 |
| Dependents in household | 0.98 | 1.13 |
| Risk-taking in personal life, [0,10] | 5.26 | 2.13 |

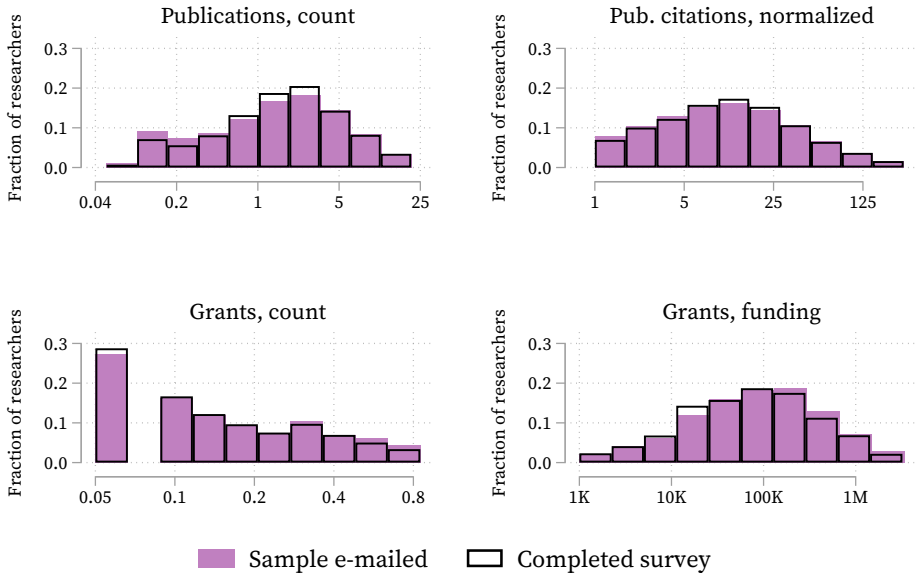
*Note*: Summary statistics for the 4,003 researchers in the estimation sample. N.R. denotes not reported.

TABLE B4. Pairwise Correlations of Subjective Output Measures

| | Articles | Books | Methods | Products | Academic | Policy | Business | Public |
|---|---|---|---|---|---|---|---|---|
| Articles | | | | | | | | |
| Books | −0.13$^{s}$ | | | | | | | |
| Methods | 0.08$^{s}$ | −0.20$^{s}$ | | | | | | |
| Products | −0.13$^{s}$ | −0.12$^{s}$ | 0.23$^{s}$ | | | | | |
| Academic | 0.48$^{s}$ | −0.01 | 0.02 | −0.24$^{s}$ | | | | |
| Policy | 0.01 | −0.00 | 0.10$^{s}$ | 0.25$^{s}$ | −0.07$^{s}$ | | | |
| Business | −0.03 | −0.07$^{s}$ | 0.20$^{s}$ | 0.38$^{s}$ | −0.12$^{s}$ | 0.24$^{s}$ | | |
| Public | −0.12$^{s}$ | 0.19$^{s}$ | 0.04 | 0.21$^{s}$ | −0.20$^{s}$ | 0.28$^{s}$ | 0.13$^{s}$ | |

*Note*: Pairwise correlations of the intended output-type and audience measures summarized in Table 2. $^{s}$ $p < 0.01$.

FIGURE B1. Sample Representativeness per Publication and Grant Funding Measures



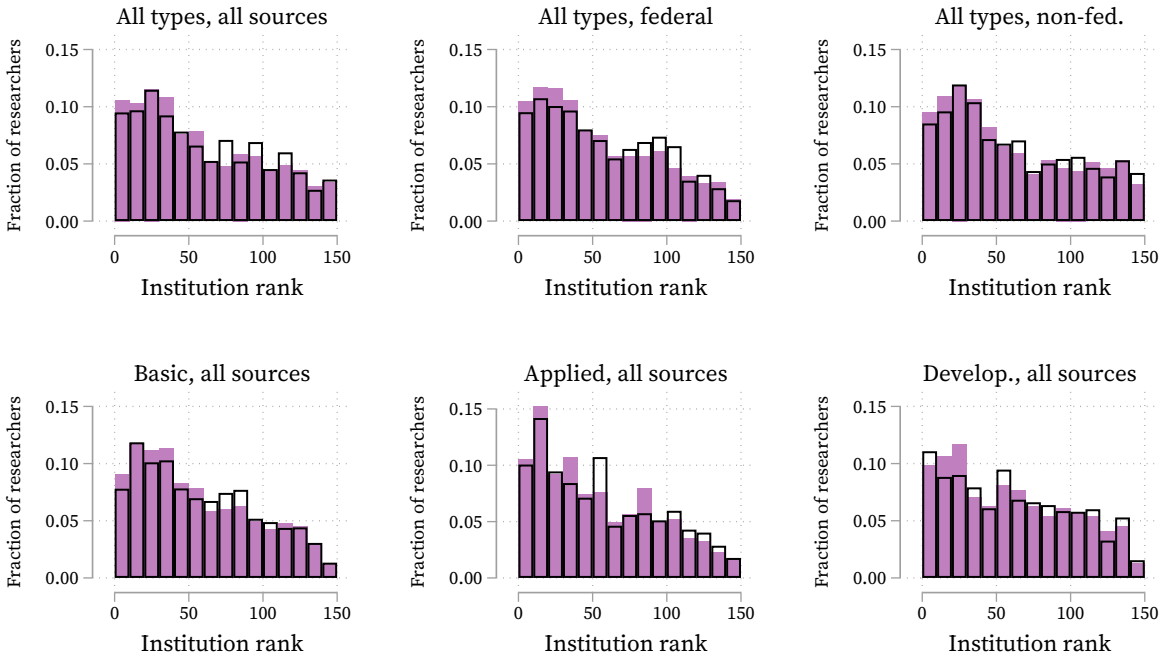

*Note*: Shows the distribution of publication and grant outcomes split by whether the researcher was emailed (i.e., the full sample) versus those who completed the survey; note the log $x$ axes. See Myers et al. (2023) for regression-based estimates of the differences.

from the observables $\mathbf{Z}_i$ into a single-dimensional index, $T_i$. Specifically, we assume there are two clusters of researcher types ($k$=2), and we use $k$-means clustering to estimate each researcher's euclidean distance from one of those groups; we use that distance as $T_i$, which provides a smooth, continuous measure of heterogeneity. Thus, for the preference function parameter $\sigma_i$, we assume that $\sigma_i = \exp(\delta_{\sigma,0} + T_i \times \delta_\sigma)$. Figure B4 shows the distribution of the one-dimensional euclidean similarity scores. Researchers with larger values of $T_i$ tend to be, among other things, older, white, males that are full professors in the humanities and natural sciences. A table reporting the mean differences between the two types of researchers is available upon request.

## B.3. Additional Experimental Details and Results

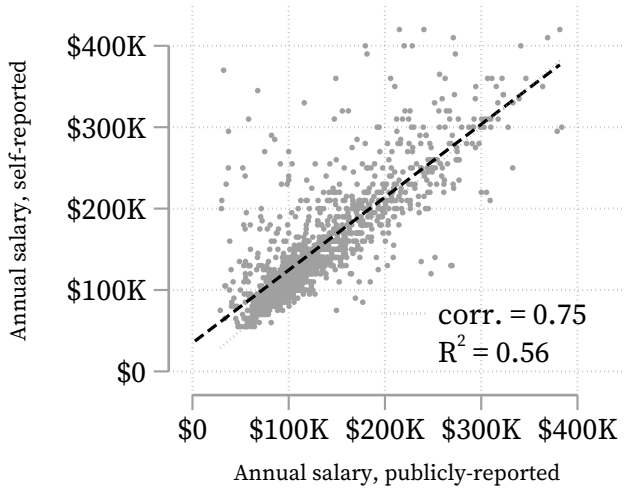
$\sigma_i = \exp(\delta_{\sigma,0} + \sum_z Z_i \delta_{\sigma,Z})$. However, this presents some practical estimation challenges (i.e., estimating approximately (35×3=) 105 additional parameters; the $\delta_{\sigma,z}$ parameters in the previous example).

FIGURE B2. Sample Representativeness per Institutional Funding Measures



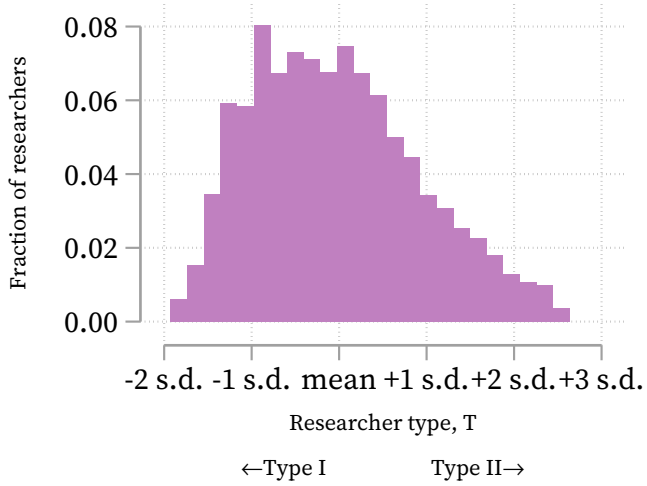

*Note*: Shows the distribution of institution-level funding split by whether the researcher was emailed (i.e., the full sample) versus those who completed the survey. See Myers et al. (2023) for regression-based estimates of the differences.

FIGURE B3. Correlation of Self- and Publicly-reported Annual Salaries



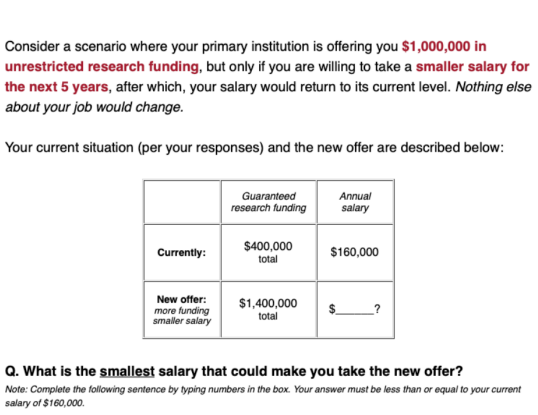

*Note*: Self-reported versus publicly reported annual salaries for the subset of researchers matched to public salary records.

FIGURE B4. Researcher Heterogeneity




*Note*: Distribution of the researcher type index $T_i$: the standardized log Euclidean distance from the $k$-means cluster centers, used to parameterize preference heterogeneity.

FIGURE B5. Example Images of Survey Experiment

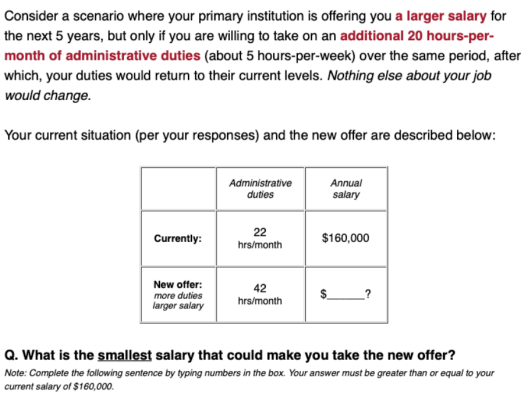

Consider a scenario where your primary institution is offering you **$1,000,000 in unrestricted research funding**, but only if you are willing to take a **smaller salary for the next 5 years**, after which, your salary would return to its current level. *Nothing else about your job would change.*

Your current situation (per your responses) and the new offer are described below:

| | *Guaranteed research funding* | *Annual salary* |
|---|---|---|
| **Currently:** | $400,000 total | $160,000 |
| **New offer:** *more funding smaller salary* | $1,400,000 total | $_____? |

**Q. What is the smallest salary that could make you take the new offer?**

*Note: Complete the following sentence by typing numbers in the box. Your answer must be less than or equal to your current salary of $160,000.*


Consider a scenario where your primary institution is offering you **a larger salary** for the next 5 years, but only if you are willing to take on an **additional 20 hours-per-month of administrative duties** (about 5 hours-per-week) over the same period, after which, your duties would return to their current levels. *Nothing else about your job would change.*

Your current situation (per your responses) and the new offer are described below:

| | *Administrative duties* | *Annual salary* |
|---|---|---|
| **Currently:** | 22 hrs/month | $160,000 |
| **New offer:** *more duties larger salary* | 42 hrs/month | $_____? |

**Q. What is the smallest salary that could make you take the new offer?**

*Note: Complete the following sentence by typing numbers in the box. Your answer must be greater than or equal to your current salary of $160,000.*

*Note*: Shows two screenshots of the survey experiment designed to solicit researchers' willingness to trade their salary for additional guaranteed funding or less research time (in the form of more administrative duties).

FIGURE B6. Correlations of Implied Valuations




*Note*: Binned scatterplots with fitted lines. Panel A: willingness to pay for one less administrative hour against the implied hourly wage (salary divided by annual work hours). Panel B: willingness to pay for an additional research dollar against weekly fundraising hours, conditional on the implied hourly wage. Variables are trimmed at the 1st and 99th percentiles.

TABLE B5. Proportion of Response Variation Correlated with Model and Non-model Variables

| | WTP +$250K | | WTP +$1M | | WTP +admin. | | WTP –admin. | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $R^2$ | 0.95 | 0.96 | 0.87 | 0.89 | 0.82 | 0.86 | 0.96 | 0.97 |
| Model vars. | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| **Z** | | ✓ | | ✓ | | ✓ | | ✓ |
| *N* obs. | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 |

*Note*: $R^2$ from regressions of researchers' responses to the four thought experiments on the model's state and choice variables ("Model vars."), with and without the full vector of observables used to construct the type index ($Z$).

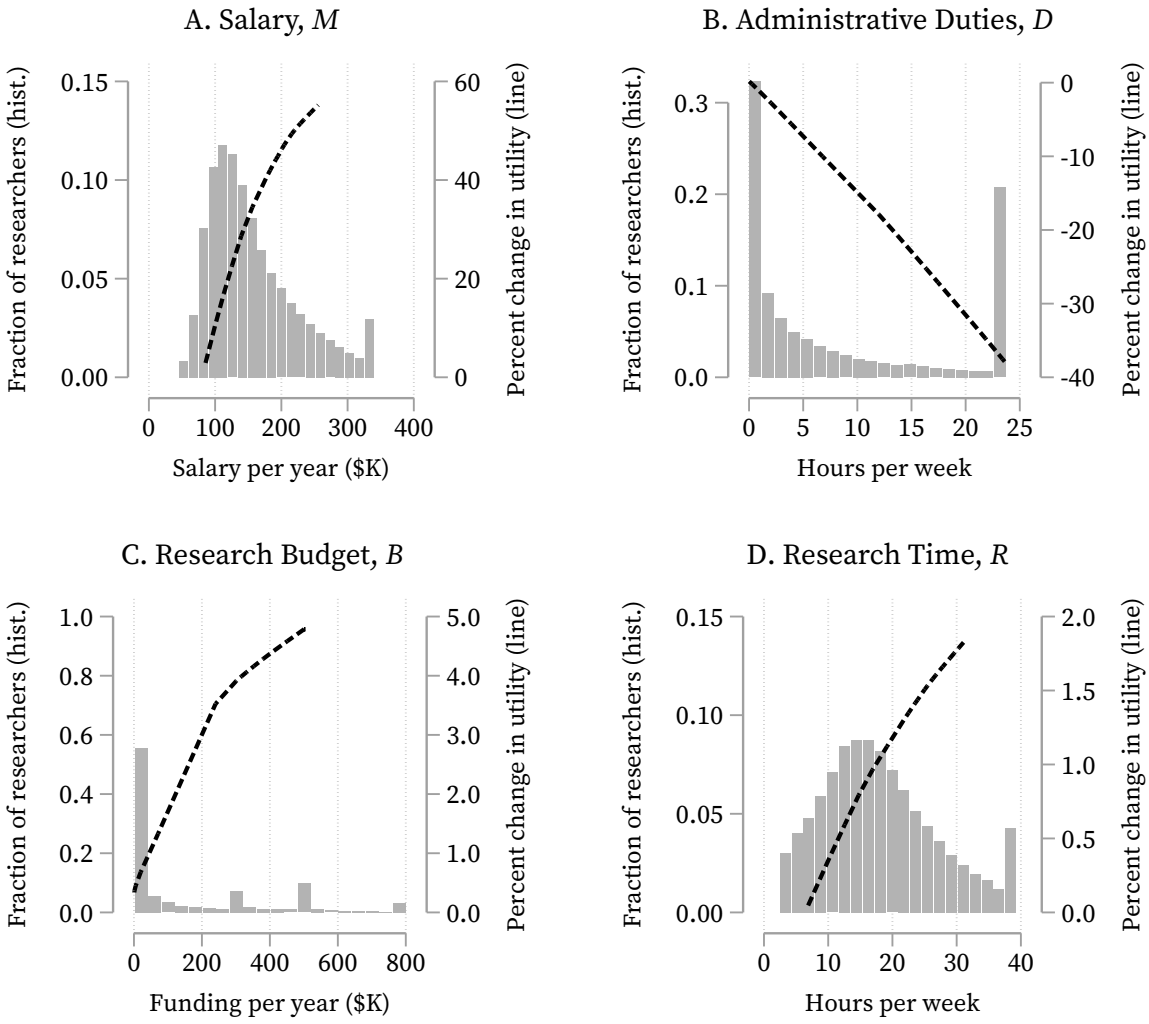
TABLE B6. Potential Role of Non-response Bias and Stated Preference WTP Bias

| | WTP +$250K | | WTP +$1M | | WTP +admin. | | WTP –admin. | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| IMR | | −0.00068 | | −0.00364 | | −0.00431 | | −0.00073 |
| | | (0.00379) | | (0.00593) | | (0.00683) | | (0.00327) |
| Benchmark | | −0.0123*** | | −0.0255*** | | 0.0135* | | −0.01000*** |
| | | (0.00413) | | (0.00645) | | (0.00743) | | (0.00356) |
| $R^2$ | 0.95 | 0.95 | 0.87 | 0.87 | 0.82 | 0.82 | 0.96 | 0.96 |
| Model vars. | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| *N* obs. | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 | 4,003 |

*Note*: Estimates from regressions of researchers' (standardized) responses to the four thought experiments on the model's state and choice variables, the inverse Mills ratio from the randomized participation incentives (Heckman 1979), and willingness to pay for a benchmark good–home high-speed internet (Dizon-Ross and Jayachandran 2022). Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

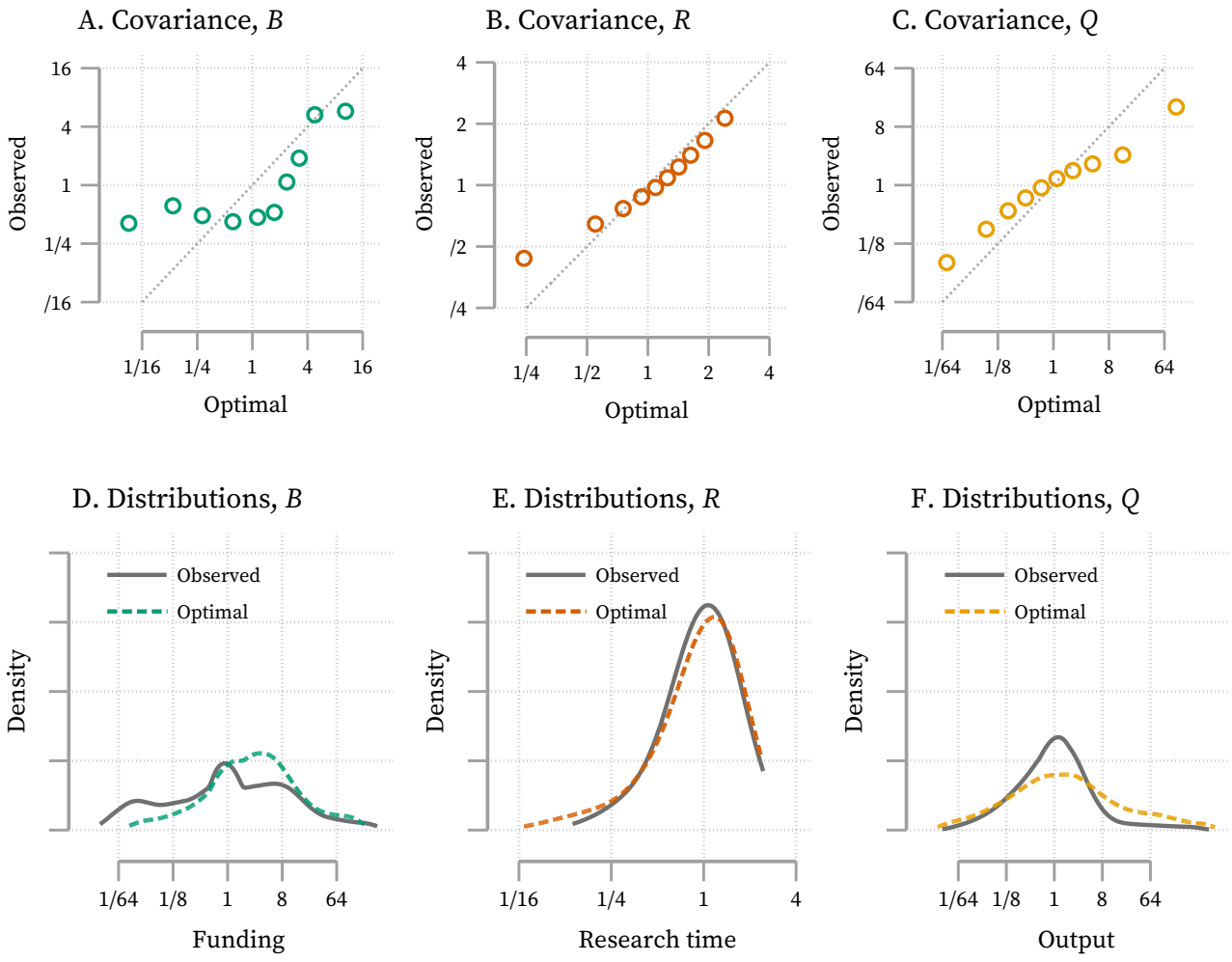
# C. Additional Model and Counterfactual Results

FIGURE C1. Utility Function Visualization



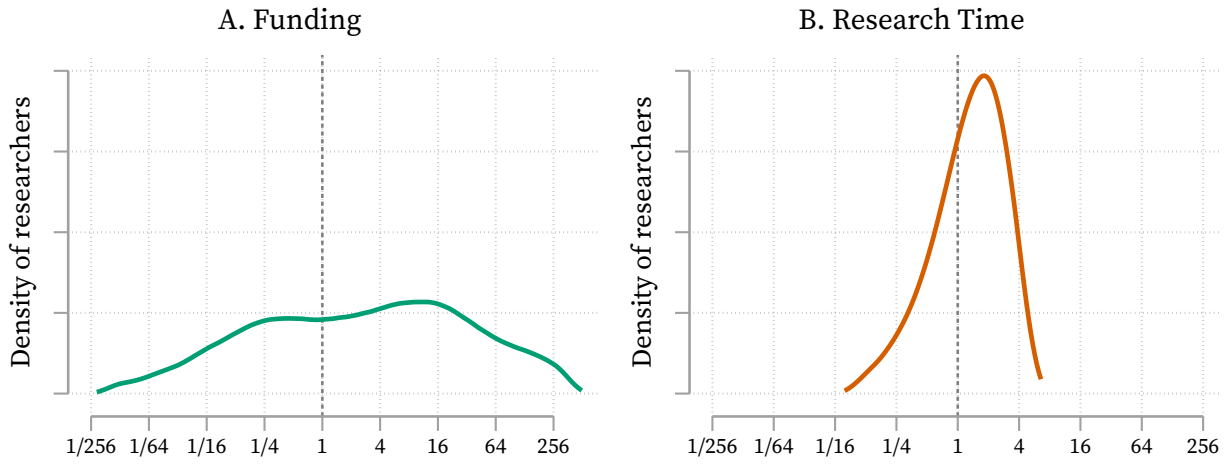

*Note*: Each panel shows how the average researcher's utility varies with one variable, holding all other variables and parameters at sample means and excluding behavioral responses. Lines show the percent change in utility relative to the variable's 10th-percentile value; histograms show the sample distribution of each variable, trimmed at the 1st and 99th percentiles. For simplicity, excludes behavioral responses.

FIGURE C2. Observed and optimal input and output levels




*Note*: Panels A–C show binned scatterplots of observed funding, research time, and output against their first-best counterparts, residualized on market fixed effects; dotted lines are 45-degree lines. Panels D–F show kernel densities of observed and first-best allocations, normalized by market medians; note the log scales. Variables are trimmed at the 1st and 99th percentiles.

FIGURE C3. Marginal Product Wedges



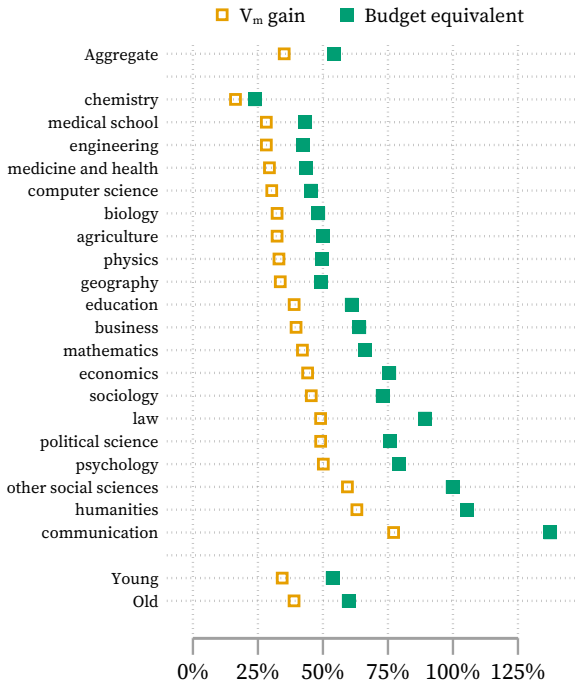

*Note*: Kernel densities of researcher-level wedges: the marginal social product of funding (Panel A) and of research time (Panel B), relative to the market's first-best shadow price (cf. Hsieh and Klenow 2009). A value of one indicates no misallocation; note the log scales, common to both panels. Distributions are trimmed at the 1st and 99th percentiles.

TABLE C1. Counterfactual Allocations: Alternative Specifications

| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Productivity dispersion* | | | | | | | | | | | | |
| $u_i(\alpha_i)$, p90/p10 | | 6.29 | 9.05 | 8.48 | 6.82 | 4.47 | 9.40 | 8.59 | 8.82 | 9.05 | 5.62 | 9.02 |
| *Change in inputs & outputs* | | | | | | | | | | | | |
| funding, $B$ | mean | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 |
| | s.d. | -19.5 | -24.6 | -24.2 | -23.2 | -22.3 | -24.4 | -26.9 | -24.6 | -24.6 | -30.5 | -24.6 |
| research time, $R$ | mean | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 |
| | s.d. | +2.8 | +5.9 | +8.4 | +12.6 | +22.9 | +6.3 | +4.7 | +6.0 | +5.9 | -3.7 | +6.0 |
| fundraising time, $F$ | mean | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 |
| | s.d. | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 | -100 |
| utility, $\mathcal{U}$ | mean | +0.9 | +3.7 | +4.5 | +4.2 | +2.8 | +3.6 | +3.1 | +3.2 | +3.7 | +1.8 | +3.7 |
| | s.d. | +0.3 | +4.8 | +7.5 | +8.5 | +6.9 | +4.4 | +3.1 | +3.2 | +4.8 | -0.6 | +4.8 |
| private output, $b(Q)$ | mean | +189 | +116 | +116 | +128 | +167 | +120 | +142 | +128 | +116 | +459 | +116 |
| | s.d. | +120 | +74.4 | +72.4 | +73.9 | +79.9 | +75.0 | +78.3 | +77.0 | +74.4 | +112 | +74.4 |
| social output, $v(Q)$ | mean | +100 | +97.3 | +101 | +112 | +135 | +100 | +118 | +103 | +97.3 | +277 | +97.6 |
| | s.d. | +103 | +56.3 | +54.4 | +56.4 | +64.2 | +57.1 | +61.7 | +59.5 | +56.3 | +134 | +56.3 |
| *Allocative efficiency* | | | | | | | | | | | | |
| initial allocation | | 50.0 | 50.7 | 49.8 | 47.1 | 42.5 | 50.0 | 45.8 | 49.3 | 50.7 | 26.5 | 50.6 |
| counterfactual allocation | | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| *Budgetary equivalent of social reallocation gains* | | | | | | | | | | | | |
| researcher-level | % | +309 | +168 | +171 | +197 | +277 | +173 | +226 | +190 | +168 | +1451 | +168 |
| population-level | $/year | +$56.9B | +$32.3B | +$28.2B | +$24.1B | +$18.4B | +$33.5B | +$43.9B | +$36.7B | +$32.3B | +$272.8B | +$32.6B |
| *Salary equivalent of private reallocation gains* | | | | | | | | | | | | |
| researcher-level | % | +1.4 | +5.9 | +7.1 | +6.5 | +4.2 | +5.7 | +4.9 | +5.1 | +5.9 | +2.8 | +5.9 |
| | $/year | +$2K | +$9K | +$11K | +$10K | +$6K | +$9K | +$8K | +$8K | +$9K | +$4K | +$9K |
| *Specification choices* | | | | | | | | | | | | |
| $\tilde{\kappa}$ | | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 2 | 5 | 5 |
| $\rho$ | | 0 | .2 | .4 | .6 | .8 | .2 | .2 | .2 | .2 | .2 | .2 |
| shock mode | | iz | iz | iz | iz | iz | iz | izcov | iz | iz | ii | iz |
| winsor pctl. | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 5 | 1 | 1 | 1 |
| MC draws | | 1 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 250 |
| fix $X$ | | | | | | | ✓ | | | | | |

*Note*: Each column reports the transition from the observed allocation to the first-best under the specification listed in the bottom panel: the social-output scale $\tilde{\kappa}$; the share of within-market variance attributed to measurement error $\rho$; the measurement-error shock structure (iz: independent across variables; ii: common within researcher; izcov: matched to the observed cross-variable correlations); the winsorization percentile; the number of Monte Carlo draws; and whether researcher attributes are held fixed across draws. Column 2 is the baseline specification. See Table 3 for row definitions. All entries are medians across draws.

FIGURE C4. Gains from Matching the Best Observed Market




*Note*: Gains from raising every market's allocative efficiency to the highest level observed in any market. For each group, hollow markers show the resulting gain in aggregate social output; solid markers show the budget equivalent–the proportional increase in funding under observed allocations that would generate the same gain. Disciplines are ordered by output gain.

TABLE C2. Counterfactual Allocations: Alternative Allocations

| | | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|---|
| *Change in inputs & outputs* | | | | | | | | | | |
| funding, $B$ | mean | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 | ~0 |
| | s.d. | -24.6 | -24.6 | ~0 | -39.1 | +23.3 | +7.5 | -30.2 | -35.5 | +16.9 |
| research time, $R$ | mean | ~0 | ~0 | ~0 | +11.2 | ~0 | +17.0 | ~0 | +6.1 | ~0 |
| | s.d. | +5.9 | +6.0 | -0.1 | -2.4 | +8.0 | -0.6 | +2.7 | -4.6 | +11.1 |
| fundraising time, $F$ | mean | -100 | ~0 | -100 | +346 | -100 | +149 | -100 | +168 | -100 |
| | s.d. | -100 | ~0 | -100 | +458 | -100 | +338 | -100 | +322 | -100 |
| utility, $\mathcal{U}$ | mean | +3.7 | +3.7 | +0.1 | +1.0 | +2.7 | -0.1 | +3.9 | +2.1 | +1.7 |
| | s.d. | +4.8 | +4.8 | ~0 | -0.7 | +5.6 | +0.0 | +4.8 | +1.4 | +3.4 |
| private output, $b(Q)$ | mean | +116 | +116 | ~0 | +16.4 | +86.3 | +0.4 | +116 | +49.4 | +45.3 |
| | s.d. | +74.4 | +74.4 | ~0 | -13.2 | +100 | +0.1 | +74.2 | +19.8 | +45.4 |
| social output, $v(Q)$ | mean | +97.3 | +97.3 | ~0 | +33.9 | +47.9 | +1.0 | +96.3 | +55.4 | +27.4 |
| | s.d. | +56.3 | +56.3 | ~0 | -7.6 | +69.3 | +0.1 | +56.1 | +16.1 | +34.6 |
| *Allocative efficiency* | | | | | | | | | | |
| initial allocation | | 50.7 | 50.7 | 100.0 | 50.7 | 67.6 | 50.7 | 50.9 | 50.7 | 78.5 |
| counterfactual allocation | | 100 | 100.0 | 100 | 67.6 | 100 | 50.9 | 100 | 78.5 | 100 |
| *Budgetary equivalent of social reallocation gains* | | | | | | | | | | |
| researcher-level | % | +168 | +168 | ~0 | +51.8 | +85.3 | +1.4 | +170 | +88.8 | +43.8 |
| population-level | $/year | +$32.3B | +$32.3B | +$0 | +$10.0B | +$16.4B | +$263.6M | +$32.8B | +$17.1B | +$8.4B |
| *Salary equivalent of private reallocation gains* | | | | | | | | | | |
| researcher-level | % | +5.9 | +5.8 | +0.1 | +1.6 | +4.2 | ~0 | +6.2 | +3.2 | +2.6 |
| | $/year | +$9K | +$9K | +$164 | +$2K | +$7K | +$0 | +$10K | +$5K | +$4K |
| *Allocation labels* | | | | | | | | | | |
| Baseline | | Observed | Observed | cfFobs | Observed | s3 | Observed | s2 | Observed | s1 |
| Counterfactual | | cf1 | cfFobs | cf1 | s3 | bs3 | s2 | bs2 | s1 | bs1 |

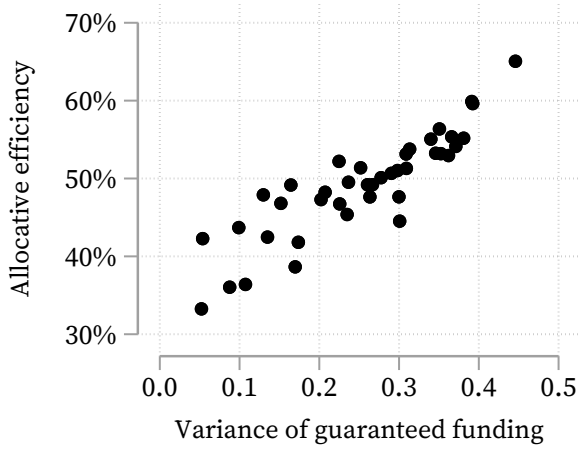
*Note*: Each column reports the change from the Baseline to the Counterfactual allocation named in the bottom rows: cf1 (first-best, no fundraising), cfFobs (first-best with fundraising time fixed at observed levels), s1 (uniform funding access), s2 (social-objective alignment), s3 (both), and bs1–bs3 (the first-best frontiers at each policy's realized aggregate resources). See Table 3 for row definitions and the main text for allocation definitions. All entries are medians across draws.

TABLE C3. Correlates of Market-Level Allocative Efficiency

| *selected predictor* | *coef.* | *std. error* | *std. coef.* |
|---|---|---|---|
| *Set 1: Variances: $M, G, F, R, A, O, H$* | | | |
| log variance guaranteed funding | 0.220*** | (0.032) | 0.825 |
| log variance admin duties | -0.041** | (0.017) | -0.141 |
| Lasso selection observations | 40 | | |
| Post-lasso observations | 40 | | |
| Post-lasso $R^2$ | 0.746 | | |
| Selected predictors | 2 | | |
| *Set 2: Variances: $M, G, F, R, A, O, H, \phi, \gamma, \alpha$* | | | |
| log variance guaranteed funding | 0.245*** | (0.029) | 0.921 |
| log variance fundraising productivity | -0.066** | (0.026) | -0.262 |
| log variance TFP alpha | -0.012 | (0.014) | -0.067 |
| Lasso selection observations | 40 | | |
| Post-lasso observations | 40 | | |
| Post-lasso $R^2$ | 0.801 | | |
| Selected predictors | 3 | | |

*Note*: Adaptive-lasso regressions of log market-level allocative efficiency on log within-market variances of the model's state and choice variables (Set 1), adding the estimated parameters (Set 2), across the 40 field-by-career-stage markets. Coefficients are from post-lasso OLS on the selected predictors, with heteroskedasticity-robust standard errors in parentheses; standardized coefficients rescale by the ratio of standard deviations. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, $^{***}\,p < 0.01$.

FIGURE C5. Allocative Efficiency and the Dispersion of Guaranteed Funding




*Note*: Each point is one of the 40 field-by-career-stage markets. The vertical axis is observed allocative efficiency; the horizontal axis is the within-market variance of guaranteed funding, in squared millions of dollars. The cross-market correlation is 0.88.

## D. Trends in Science and Metascience

Data on total US R&D spending is sourced from the National Center for Science and Engineering Statistics (NCSES).[37] Data on meta-science studies is sourced from (*i*) the *OpenAlex* publication database, which covers virtually all fields of science,[38] and (*ii*) the *PubMed* publication database, which focuses primarily on the life and biomedical sciences.[39]

OpenAlex is valuable in that its coverage spans a tremendous range of scientific fields; however, the categorization scheme used to index publications by topics is relatively coarse and lacks a clear set of topic codes for meta-science. In order to identify meta-science publications in OpenAlex, we restrict our focus to all journal-article publications flagged with the following topic codes: *scientometrics and bibliometrics research; scientific research and philosophical inquiry; philosophy and history of science; academic writing and publishing; science and science education; international science and diplomacy; innovation, technology, and society; research, science, and academia*.

PubMed is clearly a strict subset of scientific fields; however, publications are indexed with relatively fine-grain topic codes that facilitate the identification of meta-scientific studies. In order to identify meta-science publications in PubMed, we restrict our focus to all journal-article publications flagged with the following topic codes: *efficiency; research personnel; research support as topic; grants peer review; knowledge discovery; scholarly communication; allocative efficiency; grants; funding, capital*.

Clearly, in both cases, the set of topics we use to flag meta-science will include some false positives, and we surely have a non-trivial number of false negatives. However, we have no reason to suspect these error rates should be evolving significantly over time.

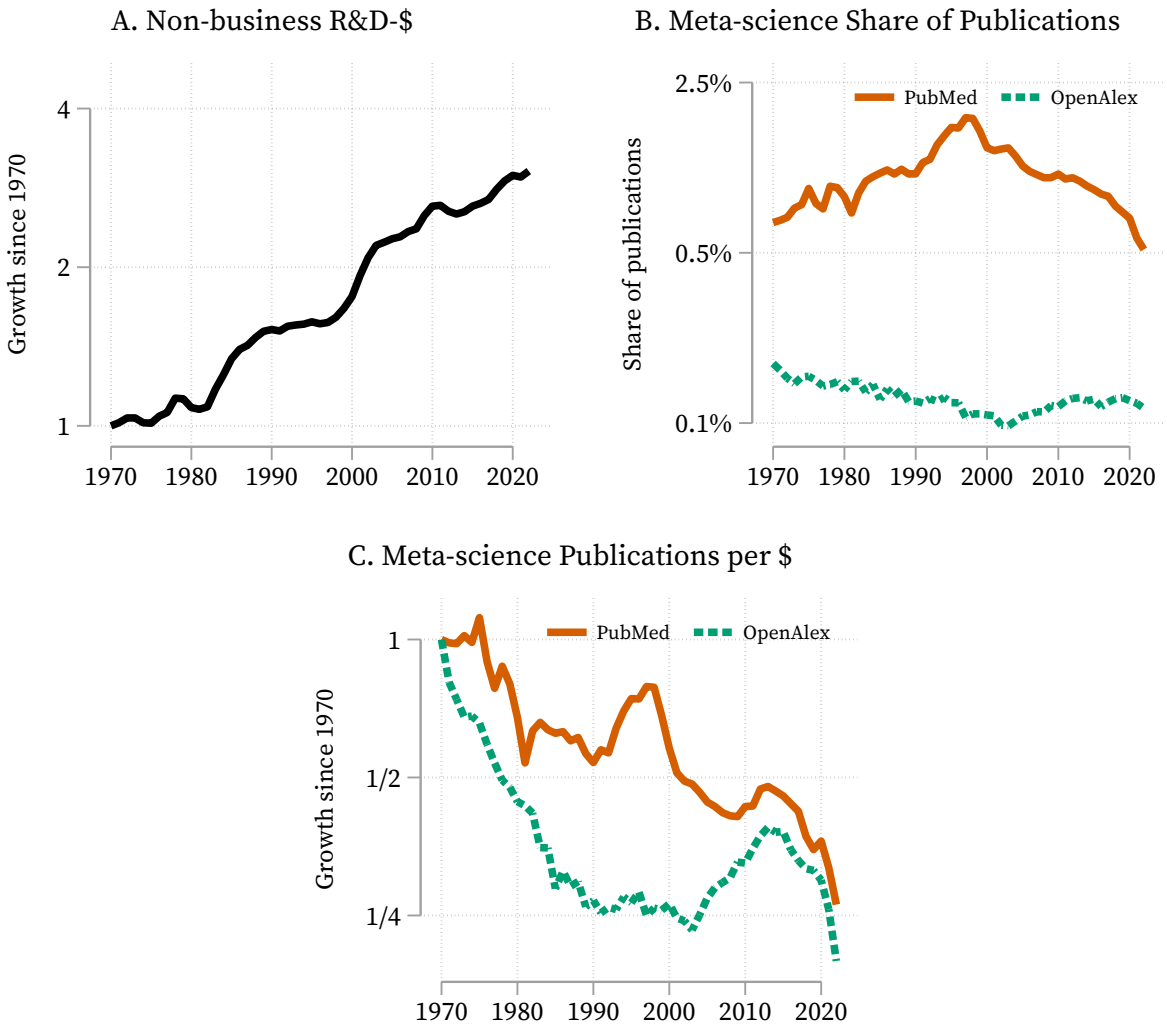
Figure D1 reports the trends for the past 50 years for total non-business R&D in real-$ (see Panel A), the share of all publications focused on meta-science (see Panel B), and the count of meta-science studies per $ shown in Panel A (see Panel C).

[37] See: National Center for Science and Engineering Statistics (NCSES). 2025. *U.S. R&D Totaled $892 Billion in 2022; Estimate for 2023 Indicates Further Increase to $940 Billion*. NSF 25-327. Alexandria, VA: U.S. National Science Foundation. Available at https://ncses.nsf.gov/pubs/nsf25327/.

[38] See: Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. ArXiv. https://arxiv.org/abs/2205.01833.

[39] See: PubMed [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – 2025. Available from: https://pubmed.ncbi.nlm.nih.gov.

FIGURE D1. Relative Growth in Science and Meta-science since 1970




*Note*: Panel A shows growth in real U.S. non-business R&D spending since 1970. Panel B shows the share of publications classified as meta-science in PubMed and OpenAlex (topic definitions in the text). Panel C shows meta-science publications per real non-business R&D dollar, normalized to 1970.

# E. Example Applications of Methodology

## E.1. Application 1: Constant Returns to Scale and Convex Costs

Consider first the case of a firm that operates a linear production function in a single input, e.g., labor, and faces convex adjustment costs because, e.g., there are hiring and firing frictions. The optimization problem is:

$$\max_{l_i} \ \alpha_i l_i - w l_i - c l_i^{\psi} \tag{E1}$$

with $b(\cdot) = \alpha_i l_i + m_i$ and $m_i = 0$, $c(\cdot) = w l_i + c l_i^{\psi}$ ($c > 0$ and $\psi > 1$), and $\mu_i = \mu = (w, c, \psi)$. The optimality condition of the problem is:

$$\alpha_i - w - c\psi l_i^{\psi-1} = 0 \tag{E2}$$

In the first step of the estimation procedure, we evaluate the optimality condition at observed allocations to characterize $\alpha_i$ as a function of parameters and observed allocations:

$$\alpha_i(\hat{l}_i, w, c, \psi) = w + c\psi \hat{l}_i^{\psi-1} \tag{E3}$$

where $\hat{l}_i$ denotes the observed input allocation. Moreover, conditional on attributes, parameters, and prices, the solution is determined by:

$$l_i^* = \left( \frac{\alpha_i - w}{\psi c} \right)^{\frac{1}{\psi-1}} \tag{E4}$$

In the second step, we implement the thought experiment and we offer to the firm $\Delta > 0$ extra units of labor, which the firm can hire by just paying $WTP_i$ dollars and without incurring the convex adjustment cost. Of course, the firm can re-optimize the quantity of labor that hires at market conditions. Therefore, the problem is:

$$\max_{\tilde{l}_i} \ \alpha_i(\tilde{l}_i + \Delta) - w\tilde{l}_i - c\tilde{l}_i^{\psi} - WTP_i \tag{E5}$$

and the solution:

$$\tilde{l}_i^* = \left( \frac{\alpha_i - w}{\psi c} \right)^{\frac{1}{\psi-1}} = l_i^* \tag{E6}$$

Replacing the solution in the profit function and equating profits at current allocations and in the thought experiment, we obtain:

$$WTP_i = \alpha_i(\hat{l}_i, w, c, \psi)(\tilde{l}_i^* + \Delta) - w\tilde{l}_i^* - c\left(\tilde{l}_i^*\right)^{\psi} - \alpha_i(\hat{l}_i, w, c, \psi)l^* + wl^* + c\left(l^*\right)^{\psi} \tag{E7}$$

Using the fact that $\tilde{l}_i^* = l_i^*$, we obtain:

$$WTP_i = \alpha_i(\hat{l}_i, w, c, \psi)\Delta = \left(w + c\psi\hat{l}_i^{\psi-1}\right)\Delta \tag{E8}$$

which allows the identification of $w, c, \psi$ and thus $\alpha_i(\hat{l}_i, w, c, \psi)$, given $\hat{l}_i$ and $\Delta$. The same logic applies to the case of decreasing returns to scale, namely:

$$\max_{l_i} \ \alpha_i l_i^{\beta} - wl_i - cl_i^{\psi}$$

with $v(\cdot) = \alpha_i l_i^{\beta} + m_i$ for $m_i = 0$ and $\beta \in (0,1)$, $c(\cdot) = wl_i + cl_i^{\psi}$ for $c > 0$ and $\psi > 1$, and $\mu_i = \mu = (\beta, w, c, \psi)$. In the general case, the optimal input choice does not admit a closed-form expression, but we can show that $\frac{\partial \tilde{l}_i^*}{\partial \Delta} \neq -1$, i.e., the additional input does not perfectly crowd out the quantity of input sourced in the market. This is a sufficient condition such that $WTP_i$ does depend on $\alpha_i$ and thus on all estimands $\mu$.

### E.2. Application 2: Decreasing Returns to Scale and Linear Costs

Here, we show an example of how a linear cost schedule renders our approach unable to recover productivity. The setting is:

$$\max_{l_i} \ \alpha_i l_i^{\beta} - wl_i$$

i.e., the problem coincides with the previous application for $\psi = 0$. The first order condition is $\beta\alpha_i l_i^{\beta-1} = w$, which identifies:

$$\alpha_i(\hat{l}_i, \beta, w) = \left(\frac{w}{\beta}\right)\hat{l}_i^{1-\beta} \tag{E9}$$

and determines the optimal allocation as:

$$l_i^* = \left(\frac{\beta\alpha_i}{w}\right)^{\frac{1}{1-\beta}} \tag{E10}$$

Under the thought experiment, the optimization problem becomes:

$$\max_{\tilde{l}_i} \ \alpha_i(\tilde{l}_i + \Delta)^{\beta} - w\tilde{l}_i$$

which implies:

$$\tilde{l}_i^* = \left(\frac{\beta\alpha_i}{w}\right)^{\frac{1}{1-\beta}} - \Delta = l_i^* - \Delta \tag{E11}$$

Replacing in the profit function and deriving $WTP_i$, we obtain:

$$\begin{aligned} WTP_i =& \alpha_i(\tilde{l}_i^* + \Delta)^{\beta} - w\tilde{l}_i^* - \alpha_i l_i^* + wl_i^* \\ =& \alpha_i(l_i^* - \Delta + \Delta)^{\beta} - w(l_i^* - \Delta) - \alpha_i l_i^* + wl_i^* \\ =& w\Delta \end{aligned}$$

where from the first to the second line we replace $\tilde{l}_i^* = l_i^* - \Delta$. Because the willingness to pay just depends on $w$, $\beta$ cannot be identified, which implies that also $\alpha_i$ is not pinned down.

### E.3. Application 3: Individuals as Producers with Utility Function

Consider the problem of an agent that gets utility from income $m_i$ and output $\alpha_i d_i^{\beta} l_i^{1-\beta}$—produced using a fixed input $d_i$ and labor $l_i$—and gets disutility from working. The problem is:

$$\max_{l_i} \ \ln m_i + \left(\alpha_i d_i^{\beta} l_i^{1-\beta}\right)^{\eta} - \phi l_i^{\psi} \tag{E12}$$

with $b(\cdot) = \ln mi_i + \left(\alpha_i d_i^{\beta} l_i^{1-\beta}\right)^{\eta}$ for $\beta \in (0,1)$ and $\eta \in (0,1)$, $c(\cdot) = \phi l_i^{\psi}$ for $\phi > 0, \psi > 1$, and $\boldsymbol{\mu}_i = \boldsymbol{\mu} = (\beta, \eta, \phi, \psi)$. The optimality condition is:

$$\eta(1-\beta)\alpha_i^{\eta} d_i^{\eta\beta} l_i^{(1-\beta)\eta-1} = \phi\psi l_i^{\psi-1} \tag{E13}$$

Therefore:

$$\alpha_i(\hat{l}_i, d_i, \beta, \eta, \phi, \psi) = \left(\frac{\phi\psi \hat{l}_i^{\psi-1}}{\eta(1-\beta) d_i^{\eta} \hat{l}_i^{\eta(1-\beta)-1}}\right)^{\frac{1}{\eta}} \tag{E14}$$

and

$$l_i^* = \left(\frac{\eta(1-\beta)\alpha_i^{\eta} d_i^{\eta\beta}}{\phi\psi}\right)^{\frac{1}{\psi-\eta(1-\beta)}} \tag{E15}$$

In the setting of a thought experiment where $\Delta$ additional units of the fixed input are offered to the agent, the problem becomes:

$$\max_{\tilde{l}_i} \ \ln\left(m_i - WTP_i\right) + \left(\alpha_i (d_i + \Delta)^{\beta} \tilde{l}_i^{1-\beta}\right)^{\eta} - \phi \tilde{l}_i^{\psi} \tag{E16}$$

with

$$\tilde{l}_i^{*} = \left(\frac{\eta(1-\beta)\alpha_i^{\eta}(d_i + \Delta)^{\eta\beta}}{\phi\psi}\right)^{\frac{1}{\psi - \eta(1-\beta)}}$$

Therefore, $WTP_i$ solves:

$$\ln\left(m_i - WTP_i\right) + \alpha_i^{\eta}\left(\frac{\eta(1-\beta)}{\phi\psi}\right)^{\frac{\eta(1-\beta)}{\psi-\eta(1-\beta)}}(d_i + \Delta)^{\frac{\beta\psi\eta}{\psi-\eta(1-\beta)}} - \phi\left(\frac{\eta(1-\beta)\alpha_i^{\eta}(d_i+\Delta)^{\eta\beta}}{\phi\psi}\right)^{\frac{\psi}{\psi-\eta(1-\beta)}} +$$

$$- \ln m_i - \alpha_i^{\eta}\left(\frac{\eta(1-\beta)}{\phi\psi}\right)^{\frac{\eta(1-\beta)}{\psi-\eta(1-\beta)}} d_i^{\frac{\beta\psi\eta}{\psi-\eta(1-\beta)}} - \phi\left(\frac{\eta(1-\beta)\alpha_i^{\eta} d_i^{\eta\beta}}{\phi\psi}\right)^{\frac{\psi}{\psi-\eta(1-\beta)}} = 0 \tag{E17}$$

and depends on all parameters.