Productivity Beliefs and Efficiency in Science

Fabio Bertolotti
Bank of Italy & CEPR

Kyle R. Myers
Harvard University & NBER

Wei Yang Tham
University of Toronto

September 10, 2026

Investment in science is key to growth. But how can we evaluate the efficiency of these investments when inputs are not allocated via prices and output is not quantified? Here, we develop an approach to estimating productivity and misallocation that overcomes these challenges. In our model of science, heterogeneous researchers derive utility from income and scientific output. Thus, researchers' willingness-to-pay for scientific inputs can reveal their productivity beliefs. We estimate the model using a representative survey of US academic researchers. We find productivity beliefs are highly skewed, and efficient reallocations can yield gains equivalent to billions of dollars in funding.

Bertolotti: fabiobertolotti@hotmail.it. Myers: kmyers@hbs.edu. 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 concerns 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 indirect evidence of allocative inefficiency in science. Information asymmetries, adjustment costs, sociological biases, and the lack of explicit price mechanisms all suggest that many researchers may have significantly more or fewer resources than would be optimal. But 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 better support the most productive researchers? Answering this requires answering the following first: what is the distribution of productivity 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 not allocated via prices easily expressed in monetary units. These features prevent the use of most conventional methods for productivity estimation.

To overcome these challenges, we develop a new methodology 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. Our approach adds a novel component: the direct solicitation of producers' willingness to pay (WTP) for inputs. Because these WTP decisions reflect future production expectations, we refer to the productivity estimates they reveal as producers' *productivity beliefs*.

We assume that each individual researcher derives utility from consumption, leisure, and scientific output, which they produce using inputs they can adjust subject to frictions.

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

In our WTP experiments, we offer these individuals the possibility to use their income to purchase inputs from us (the experimenters) without incurring any adjustment costs (e.g., a linear price schedule). The key is that, conditional on technologies, incomes, and preferences, differences in researchers' WTP for inputs reflect (only) their productivity differences. Thus, variation in this WTP can be used to recover the distribution of productivity across researchers and, through counterfactuals, quantify the value from more efficiently allocating inputs.

In order to apply this methodology, we develop a model of researchers' production and labor supply decisions. Researchers choose how much time to spend conducting research and raising funds, given their salaries, administrative duties, and guaranteed funding. They differ in their productivity, the funding intensity of their research, and their ability to raise funds. The model also allows for differences in preferences over income, scientific output, and time. Together, researchers' observed choices and their WTP for alternative contracts identify their technologies and preferences. Importantly, the model distinguishes the ability to produce research from the ability to obtain resources for it. 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.

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).[3] 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. These tradeoffs resemble decisions researchers face when considering job offers, and this provides the necessary data on their WTP for inputs.

The WTP responses are economically substantial and behave in plausible ways. The median researcher values an additional dollar of research funding at 10 cents of income and an hour free of administrative duties at $68. Researchers with higher implied hourly wages place greater value on their time. Conditional on those wages, researchers who spend more time fundraising are willing to pay more for guaranteed funding. The components of the

[3]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. We reproduce several of these comparisons in the Appendix.

model explain a large fraction of the variation in researchers' responses (approximately 90%), and we find little evidence that responses are significantly influenced by sample selection or systematic noise.

The estimates reveal substantial heterogeneity in researchers' productivity beliefs. After normalizing productivity within each market, the 90–10 percentile ratio is 9.[4] These differences are large, although comparisons with firm-level productivity estimates are necessarily imperfect because the underlying measures differ (e.g., Syverson 2011; De Loecker and Syverson 2021), and this scale of dispersion in productivity across individual workers has been observed in some high-skilled settings (e.g., Sackman, Erikson, and Grant 1968). As an external check, model-implied output is positively correlated with field-normalized citations even after controlling for input levels, which indicates that our estimates contain information about researchers' production functions beyond inputs alone.

With our productivity estimates in hand, we explore correlations of the technological parameters and find patterns suggestive of allocative inefficiencies (e.g., more productive researchers tend to raise fewer dollars per hour devoted to fundraising). Furthermore, we use the productivity estimates to learn more about the planner's implicit weight put on different scientific markets. We use a revealed preference approach to infer the multipliers that rationalize observed funding across markets (i.e., since we can observe the total amount of funding allocated to each field, and we can estimate the marginal returns to more funding in each field, we can infer the weights that rationalize the observed allocation of funding). This exercise provides us with parameters to scale private valuations into social valuations, as well as a novel view on the revealed preferences that scientific institutions, as a whole, have for different fields of science.

To study allocative efficiency, we compare observed allocations with an optimal benchmark where a perfectly informed planner directly assigns inputs to maximize the social value of output. This benchmark preserves aggregate inputs within each market and restricts individual reallocations using bounds informed by the observed data and survey experiments. Using this benchmark, we estimate that the observed allocative efficiency of science is roughly 48%; science produces only about half of the social value that is theoretically possible using observed input levels. Importantly, almost all of the misallocation is with respect to funding, not time. Reallocating research time alone yields little improvement.

[4] 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 (cf., Hall and Mairesse 2024).

To provide a more tangible valuation of the potential gains from reallocation, we estimate equivalent variation in budgetary terms: matching the output of the optimal benchmark would require increasing research budgets by roughly 195%. Extrapolating to the population level, this amounts to approximately $32 billion in funding per year. The scope for improvement varies across markets, but substantial inefficiencies are present in every field and career-stage. Even just bringing each market to the efficiency of the most efficient *observed* market, rather than the theoretical optimum, yields gains equivalent to a 75% increase in science-wide research funding.

These benchmarks establish the potential for large gains, but they do not explain how scientific institutions could achieve them. So our last exercises use the model as a laboratory to evaluate alternative funding policies (while allowing for behavioral responses). We explore a range of experimental policies motivated by ongoing debates, and we find they differ considerably in their performance. Our most exploratory experiments allow researchers to exchange income for funding (e.g., directly purchase research grants). Their practical feasibility is far from established, but they prove to be very efficient, which suggests that the information revealed by researchers' demand for inputs could also be useful for designing allocation mechanisms.

Our estimates come with some important caveats. First, we face the challenges common to all existing productivity estimation techniques.[5] Second, our estimates concern ex-ante productivity beliefs rather than realized productivity, and our WTP data comes 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 an important step forward for theoretical and empirical studies in the economics of science and innovation.

After a brief discussion of related literature, Section 2 presents the model and identification strategy. Section 3 describes the survey, and Section 4 reports the WTP experiments and their validation. Section 5 presents the estimates and Section 6 uses them to quantify allocative efficiency and evaluate alternative policies.

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

Related Literature

Our paper sits at an (almost empty) intersection of two large bodies of literature: social studies of basic science and its effect on downstream innovation (i.e., Merton 1973; Mansfield 1980; Stephan 1996; Fleming and Sorenson 2004; Branstetter 2005; Iaria, Schwarz, and Waldinger 2018; Bikard and Marx 2020; Marx and Fuegi 2020; Arora et al. 2024; Krieger, Schnitzer, and Watzinger 2024); and economic studies of productivity and factor misallocation (i.e., Syverson 2011; Restuccia and Rogerson 2017; De Loecker and Syverson 2021). More generally, we engage with the classic challenge in labor economics of estimating worker-level productivity metrics (i.e., Hoffman and Stanton 2024).

Broadly speaking, 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 examined the allocation of research funding across disciplines and its impact on bibliometric proxies of scientific output and innovation (Ahmadpoor and Jones 2017; Howell 2017; Fleming et al. 2019; Moretti, Steinwender, and Van Reenen 2019; Myers 2020; Myers and Lanahan 2022; Azoulay and Li 2022; Yin et al. 2022). Moreover they have quantified some status-based frictions (Azoulay, Fons-Rosen, and Graff Zivin 2019) and have also studied potential sources of frictions such as: political lobbying (Hegde and Sampat 2015), information asymmetries (Hager, Schwarz, and Waldinger 2024), reward systems (Azoulay, Graff Zivin, and Manso 2011; Clauset, Arbesman, and Larremore 2015; Ma, Mondragón, and Latora 2015; Ma and Uzzi 2018), gender (Hofstra et al. 2020; Huang et al. 2020), career stages (Sinatra et al. 2016; Way et al. 2017; Liu et al. 2018), 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) find 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. Complementary to our work focusing on misallocation within scientific markets, Chen, Liu, and Ma (2026) quantify the optimal allocation of scientific funding across fields with a general-equilibrium growth model, and Zacchia (2020) looks at knowledge spillovers across scientists.

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. 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 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 (cf., 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 model researchers' production and consumption and show how their WTP for inputs can reveal their productivity beliefs. We use the term "beliefs" because researchers choose inputs and report their WTP before producing the associated output; these decisions therefore reveal their perceived productivity, rather than their subsequently realized productivity. After an illustrative model designed to deliver intuition of our method, we describe the full model of researchers' labor supply in Section 2.2, and then we discuss identification and estimation in Section 2.3. Our approaches to handling measurement error and engaging with the externalities of science are described in Sections 2.5 and 2.6, respectively. Additional details regarding the model and estimation are contained in Appendix A.

2.1. Illustrative Example

The following simple example shows how soliciting researchers' WTP for an input can be used to estimate their productivity beliefs. Throughout the example, and in the full model below, production functions describe researchers' perceived relationships between inputs and output; we do not separately model uncertainty about those relationships.[6]

Utility maximizing researchers are indexed by i and are endowed with income, M (e.g., a guaranteed salary). Researchers produce scientific 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 $c_i > 1$. These convex costs could be due to adjustment costs or other frictions. Researchers produce output according to a linear research production function: $Q_i = \alpha_i \times X_i$, where α_i is the focal productivity parameter. Researchers' utility function is additively separable in income, research output, and costs. They choose X_i to maximize their expected utility per:

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

where b_i is the marginal utility of scientific output. We cannot observe the productivity, preference, or cost parameters (α_i, b_i, c_i), nor can we observe output (Q_i). Our focal goal is to estimate productivity (α_i).

[6]This logic is familiar from production-function estimation based on input choices such as the "factor share" approach (De Loecker and Syverson 2021): those choices reflect producers' information about productivity when the inputs are selected.

Our solution is to solicit researchers' willingness to pay (WTP) out of their income (M) to obtain a fixed quantity of input *without* any frictions. Specifically, researchers are asked to report their WTP to obtain Δ units of X without incurring the convex costs. By definition, their WTP is the price that makes them indifferent between purchasing Δ units of X and not purchasing them. Formally, their *WTP* solves:

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

Solving for i's optimal choices in the no-purchase and purchase scenarios and rearranging yields a simple expression for researchers' productivity beliefs:

$$\alpha_i = \frac{\mathrm{WTP}_i/\Delta}{b_i} .$$

In this simple example, estimating productivity does not require output quantities nor input costs. However, it does require observing, estimating, 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. Full Model of Researchers' Utility and Scientific Production

There are N researchers, indexed by i, that operate in scientific markets indexed by k; each operates in a single market. Researchers derive utility from their income and the scientific knowledge they produce, and disutility from work. Specifically, they supply their labor to maximize utility conditional on: (*i*) 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);[7] (*ii*) a vector of individual-specific attributes θ_i related to scientific production, further specified below; and (*iii*) a vector of heterogeneous preference parameters μ_i.[8]

Given their contract and parameters, researchers choose how much time to spend on research (R_i) and fundraising (F_i), which, together with their duties (D_i), sum to yield their

[7] 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.

[8] This source of heterogeneity helps us limit the degree to which variation in WTP that is actually driven by variation in preferences might cause us to, for example, overstate productivity dispersion.

total work hours (H_i).

The quantity of scientific output (e.g., knowledge), Q_i, that researchers produce is a Cobb-Douglas function of total research funding B_i and research time R_i, which includes researcher-specific productivities (α_i) and output elasticities (per γ_i). Total research funding B_i is the sum of three components: a minimum funding amount, $B_{\min}$; guaranteed funding G_i; and the product of their fundraising ability and fundraising time ($\phi_i \times F_i$). All work hours are non-negative and have a total upper bound of $H_{\max}$.

Researchers' indirect utility is given by:

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

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 production, and μ_i is the vector of parameters that govern the shape of $m_i(\cdot)$, $b_i(\cdot)$, and $c_i(\cdot)$. The functional forms of these benefit and cost functions are assumed to be:

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

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

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

The ξ_i parameter allows for additional disutility from duty-related work (e.g., administration or teaching).[9]

It is helpful to clarify "productivity" per this model. As written, the α_i parameter reflects quantity-based productivity (i.e., TFPQ); it determines output quantity for given levels of funding and research time. However, it is plausible that our Cobb-Douglas specification,

[9] We also assume the following: $\eta_i \in (0, 1)$, $\psi > 0$, $\xi_i > 0$, and $\zeta_i > 0$. Therefore, the vector μ_i collects the parameters $(\omega, \psi, \sigma_i, \eta_i, \xi_i, \zeta_i)$.

which imposes constant returns-to-scale, is misspecified. If there are truly nonlinear (i.e., decreasing) returns to scale, this will be reflected primarily in our estimates of the η_i parameter governing $b_i(\cdot)$. That is, we cannot separate researchers' decreasing returns to scale from their decreasing marginal (indirect) utility from scientific output. Thus, as a conservative approach to reporting productivity, we report a welfare-relevant productivity metric using the following:

$$A_i \equiv \frac{\alpha_i^{1-\eta_i}}{1-\eta_i}\,, \tag{3}$$

This metric reflects the private utility generated by productivity level α_i at a unit Cobb–Douglas input index. Thus, in the language of traditional productivity studies, A can be viewed as a revenue-based productivity metric (i.e., TFPR).

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 $\mathcal{R}(\mathbf{S}_i, \theta_i, \mu_i)$, $\mathcal{F}(\cdot)$, and $\mathcal{H}(\cdot)$. The derivations of these policy functions are described in Appendix A.1.

2.3. Identification and Estimation

Individual Attributes. We identify fundraising abilities (ϕ_i) by dividing the expected amount of funds raised ($\phi_i \times F_i$) by fundraising time (F_i). As we will discuss in Section 3 below, we observe those variables in our data. Conditional on utility parameters and observed allocations, we identify funding intensity (γ_i) and productivity beliefs (α_i) from the optimality conditions of Equation (1) at an interior solution:

$$(1-\gamma_i) \times \alpha_i^{1-\eta_i} \times B_i^{(1-\eta_i)\times\gamma_i} \times R_i^{(1-\eta_i)\times(1-\gamma_i)-1} = \psi \times (H_i - D_i + D_i^{\xi_i})^{\zeta_i}\,. \tag{4}$$

Equation (4) identifies the productivity parameters because the marginal benefit of research time is equated with its marginal utility cost. This identification strategy requires that the observed utility-maximizing allocation features a positive fundraising time. For observations that do not meet this condition in our sample, we extend the approach and estimate fundraising ability and input shares based on observable characteristics. We then identify the productivity beliefs according to Equation (4). Appendix A.2 provides further details about this extension and the derivation of optimality conditions.

WTP for inputs. To identify the parameters of the utility function (μ_i), we solicit researchers' WTP for research inputs using their own income M_i. In a series of four experiments, described in detail below in Section 4, we ask researchers to report their income level that would make them indifferent between their current expectations and a counterfactual allocation of inputs (e.g., more research funding, fewer administrative duties).

To see the value of these experiments, consider researcher i's optimization problem in the scenario where they are willing to forgo WTP_i dollars of income to receive Δ units of additional guaranteed funding:

$$\begin{aligned}\tilde{U}(\mathbf{S}_i, \theta_i, \mu_i, \Delta, \mathrm{WTP}_i) = \max_{\tilde{R}_i, \tilde{F}_i, \tilde{H}_i} \; & m_i(M_i - \mathrm{WTP}_i) + b_i(\tilde{Q}_i) - c_i(\tilde{R}_i, \tilde{F}_i, D_i) \\ & \text{subject to} \\ & \tilde{Q}_i = \alpha_i \times \tilde{B}_i^{\gamma_i} \times \tilde{R}_i^{1-\gamma_i} \\ & \tilde{R}_i + \tilde{F}_i + D_i = \tilde{H}_i \leq H_{\max} \\ & \tilde{B}_i = B_{\min} + G_i + \Delta + \phi_i \times \tilde{F}_i \,.\end{aligned}$$

In this scenario, the individual's WTP_i is subtracted from their income since they are both in monetary units and treated as perfect substitutes in the utility function. The funding obtained, Δ, is added to the budget, but the researcher's cost function, $c_i(\cdot)$, still only depends on the quantity of inputs they choose to obtain after the experiment.

Here, the researcher's WTP for Δ dollars of funding equates the following:

$$U(\mathbf{S}_i, \theta_i, \mu_i) = \tilde{U}(\mathbf{S}_i, \theta_i, \mu_i, \Delta, \mathrm{WTP}_i)\,, \tag{5}$$

which is the generic version of the indifference equation in our illustrative example above. An individual's WTP makes them indifferent to the purchase of the inputs.

Estimation. Let Δ_j be the amount of input offered to the researcher in experiment j and WTP_{ij} the associated researcher's willingness to pay implied by the model. Under invertibility conditions that we specify in the next subsection, it is possible to express WTP_{ij} as an implicit function of observed input allocations ($\boldsymbol{X}_i$), states, attributes, and parameters based on Equation (5). Because attributes (θ_i) can be expressed as functions of parameters, observable allocations, and states through structural conditions, WTP_{ij} is an implicit function of utility parameters (μ_i) and observables only, which we denote by $\widehat{WTP}(\mathbf{S}_i, \mu_i, \boldsymbol{X}_i, \Delta_j)$. We can then estimate them through GMM by minimizing the distance between $\widehat{WTP}(\mathbf{S}_i, \mu_i, \boldsymbol{X}_i, \Delta_j)$ and WTP_{ij}, which is observed in the data. This approach requires parameterizing the heterogeneity in utility parameters such that $\boldsymbol{\mu}_i$ is a function of researcher-specific observable characteristics.[10] Therefore, the GMM estimator for $\boldsymbol{\mu}$ solves the problem:

$$\min_{\boldsymbol{\mu}} \quad \sum_{j=1}^{J}\sum_{i=1}^{N}\left(\widehat{WTP}(\mathbf{S}_i, \boldsymbol{\mu}, T_i, \boldsymbol{X}_i, \Delta_j) - WTP_{ij}\right)^2 , \tag{6}$$

2.4. Key Assumptions

The identification and estimation strategy outlined above requires some important assumptions on the environment and on the properties of the utility function.

Assumption 1 — Rank Condition. The matrix $\mathbb{E}\left[\mathbf{J}_i(\boldsymbol{\mu})'\mathbf{J}_i(\boldsymbol{\mu})\right]$ *is positive definite at the true parameters,* where $\mathbf{J}_i(\boldsymbol{\mu})$ denotes the Jacobian of researcher *i*'s vector of model-implied WTP responses with respect to the common parameters $\boldsymbol{\mu}$. Under our model, this condition ensures local identification: no non-zero change in the parameters leaves model-implied WTP unchanged across all researchers and experiments. In Appendix B, we apply our methodology to alternative economic settings and illustrate how convex costs of acquiring inputs (e.g., due to frictions or adjustment costs) make WTP variation informative about productivity.

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

Assumption 3 — Homogeneous or Observable Output Benefits. The benefits due to output, given by $\partial b_i/\partial Q_i$*, 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).

[10] Let T_i be an observable characteristic of researcher i. We assume that σ_i, ζ_i, and ξ_i are exponential functions of T_i, e.g., $\sigma_i = \exp\left(\delta_{\sigma,0} + \delta_{\sigma,1}T_i\right)$, with constants $(\delta_{\sigma,0}, \delta_{\zeta,0}, \delta_{\xi,0})$ and slopes $(\delta_{\sigma,1}, \delta_{\zeta,1}, \delta_{\xi,1})$, respectively, while η_i is a logistic function of T_i, i.e., $\eta_i = \left(1 + \exp(\delta_{\eta,0} + \delta_{\eta,1}T_i)\right)^{-1}$. Defining $\boldsymbol{\delta} \equiv (\delta_{\sigma,0}, \delta_{\zeta,0}, \delta_{\xi,0}, \delta_{\eta,0}, \delta_{\sigma,1}, \delta_{\zeta,1}, \delta_{\xi,1}, \delta_{\eta,1})$, then $\boldsymbol{\mu} = (\omega, \psi, \boldsymbol{\delta})$.

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.

Assumption 4 — Observable States and Current Allocations. Individuals' budget constraints, per M_i, and optimal input plans in the absence of the WTP experiment are observable. These variables are necessary for identification via the first-order conditions.[11]

2.5. 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 has 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 ρ, 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. Given the homogeneous parameter estimates for μ we reconstruct the individual-specific parameters and production function attributes 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 with medians.[12]

This procedure is not a silver bullet. The parameter ρ is not identified from any of our data and must be calibrated ad hoc, and the simplicity of this method assumes away plausible scenarios of non-classical measurement error that may be relevant (e.g., a researcher's measurement error may be correlated with their productivity). Still, this procedure provides a transparent way to discipline the amount of economically meaningful

[11]There are some knife-edge cases and simple models where these variables need not be observed and productivity is identified by the experiment alone, such as in the illustrative example above.

[12]The posterior median minimizes posterior expected absolute-error loss (Gelman et al. 2013).

misallocation in the data and may serve as a useful starting point for future work on measurement error in misallocation studies.

Our preferred calibration is to set $\rho = 0.4$. This value is motivated by the correlation between researchers' self-reported income and their publicly-reported income, which we find to be approximately 0.75.[13] Still, 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.

2.6. Social Value of Output

The output of science (i.e., knowledge), almost by definition, is a product with externalities. Thus, a full assessment of the social value of current or counterfactual allocations of inputs requires taking a stance on the nature of these externalities. The WTP experiment is not helpful here, as it only identifies the private value that researchers attach to their own scientific output.[14]

We engage with these externalities in two steps. First, we assume that externalities are uniform within scientific markets, and a simple functional form can be used to aggregate researchers' output into social value by way of a linear externality term. This motivates an approach of only conducting counterfactual exercises *within* markets. Second, we use a revealed preference argument to back out the market-specific externality terms that rationalize the observed allocation of funding across markets (i.e., the funding system's implicit beliefs about the size of externalities in each market). We summarize our approach to recovering these externality terms below and have more detail in Appendix A.5.

Social Value Function. Let $\mathcal{B}_k \equiv \sum_{i \in k} B_i$ denote the observed, aggregate research funding assigned to market k, and let $B_i(\mathcal{B}_k)$ denote the funding that reaches researcher i. We define the social value of output produced in k as:

$$\mathcal{V}_k(\mathcal{B}_k; \chi_k) \equiv \chi_k \times \sum_{i \in k} v_i\left(Q_i\left(B_i(\mathcal{B}_k)\right)\right) . \tag{7}$$

[13] If all of the discrepancy between these two values was due to classical measurement error, it would imply a ρ of 0.44. However, the publicly-reported incomes are not ground truth and so we are comfortable shading downward to our preferred $\rho = 0.4$.

[14] Even an experiment that elicited every scientist's WTP for resources given to every other scientist would omit benefits accruing outside the scientific community, which is arguably the largest source of positive externalities.

Here, $Q_i(B_i(\mathcal{B}_k))$ denotes equilibrium output, including the researcher's behavioral response to funding; we suppress its other arguments for ease of notation. The function $v_i(\cdot)$ is a dollar-valued analogue of the private output-value function $b_i(\cdot)$.

The key parameter is χ_k, which is a dimensionless multiplier that scales private value into social value. Thus, each researcher appropriates $1/\chi_k$ of the social value they produce. Because χ_k multiplies the value of every feasible allocation by the same positive constant, its level does not affect which within-market allocation maximizes social value. Thus, our within-market counterfactuals do not require us to take a stance on χ_k to solve for optima or to report relative (i.e., percentage) gains in social value. However, valuing the aggregate, across-market gains from any reallocation in absolute terms does require taking a stance on χ_k.

Planner's Problem and Implied Social Multipliers. To calibrate χ_k, we use a revealed-preference benchmark. We assume the collection of institutions and policies that allocate research funding across markets behaves as if it solves:

$$\max_{\{\mathcal{B}_k \geq 0\}_k} \quad \sum_k \mathcal{V}_k(\mathcal{B}_k; \chi_k) - \mathcal{C} \times \sum_k \mathcal{B}_k \,, \tag{8}$$

where $\mathcal{C}$ is the marginal social cost of a dollar of research funding. Appendix A.5 states the assumptions behind this exercise and shows that an interior observed allocation implies:

$$\chi_k = \frac{\mathcal{C}}{\sum_{i \in k} \left(\frac{dB_i(\mathcal{B}_k)}{d\mathcal{B}_k} \times \frac{\partial U_i^* / \partial B_i}{\partial U_i^* / \partial M_i} \right)} \,, \tag{9}$$

where all terms on the right-hand side are evaluated at the observed allocation.

We normalize $\mathcal{C} = 1$. The implied social multipliers scale proportionally with this assumption: for any alternative marginal cost of funds $\mathcal{C}$, simply multiply our reported estimates by $\mathcal{C}$. The first term in the denominator of Equation (9) describes the incidence of an additional market funding dollar: how much of that dollar reaches researcher i. We must calibrate this component and, for simplicity, assume proportional marginal incidence (i.e., $dB_i(\mathcal{B}_k)/d\mathcal{B}_k = B_i/\mathcal{B}_k$), which assumes the average allocation also describes the marginal dollar's allocation. The second term in the denominator of Equation (9) is researcher i's marginal private return to funding, expressed in salary dollars by dividing by their marginal utility of income. This component is entirely identified by data and estimates.

To be clear, this is a revealed-preference accounting exercise, not a direct experimental estimate of externalities. It recovers the market-specific multipliers that rationalize observed funding levels given the problem we've defined. If observed allocations also reflect political considerations, historical commitments, institutional constraints, or omitted objectives, those forces will be absorbed into the implied χ_k.

3. Survey and Data Overview

3.1. Survey Design

We estimate the parameters of the 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 US 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 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 (3.3%) completed the survey.[15] We then restrict the sample to the 4,003 individuals (91% 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.

Regarding sample selection, Myers et al. (2023) check 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. Appendix Figure C1 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. At the institutional level, Appendix Figure C2 shows invite-respondent overlap on various measures of institution funding derived from the HERD survey. As in the case of the researcher-level comparison, the distributions overlap substantially. In this

[15] 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).

case, there are some statistically significant differences; on average, respondents come from institutions with slightly less research funding (4–6%). See Myers et al. (2023) for more.

We can also engage with sample selection on unobservables using the randomized incentives and reminders as instruments (i.e., Heckman 1979). The validity of this selection correction 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 C1 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. As described below, we use the sample selection correction term in our analyses as a covariate.

3.2. Summary Statistics: Researchers and their Inputs

Table 1 reports the summary statistics of the key covariates in our analyses.

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.[16] 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, 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, guaranteed funding (e.g., from prior grants or institutional guarantees), fundraising expectations, and their time allocations. Importantly, these 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 performed by Myers et al. (2023), which compares respondents' self-reported salaries to the publicly-reported salaries located for a subset of researchers. Overall, there is a high degree of alignment (see Appendix Figure C3).

[16] 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 Aggregate 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: Reports summary statistics for the 4,003 researchers in the estimation sample.

Position and Socio-demographics. Appendix Tables C2-C3 provide 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).

3.3. Summary Statistics: Subjective Output Measures

The survey data 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. The survey also shows negative correlations between researchers' reported emphasis on journal articles and their emphasis on several other output types in a way that suggests strategic substitution (see Appendix Table C4). For instance, researchers who focus more on publishing articles for their academic peers are less likely to focus on

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: Reports summary statistics for researchers' descriptions of their work. Output and audience frequencies are coded 0 = rarely, 1 = sometimes, and 2 = very often; a risk score of 10 denotes very high risk.

publishing books or developing products intended for non-academic audiences.

Our methodology for estimating productivity is explicitly designed to avoid the quantification of scientific output. Still, we use these subjective descriptions of researchers' output in our approach for handling heterogeneity described next.

3.4. Heterogeneous Payoffs and Preferences

As discussed in Section 2.3, the model incorporates heterogeneity in researchers' preferences via individual-specific parameters that describe researchers' utility from income, scientific output, and effort (i.e., per the functions $m_i(\cdot)$, $b_i(\cdot)$, and $c_i(\cdot)$, respectively). This heterogeneity is useful because it allows us to limit the degree to which variation in the WTP data driven by heterogeneous preferences might 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, C2, and C3) into a one-dimensional index. We estimate two clusters of researcher types ($k = 2$) and use each researcher's standardized log Euclidean distance from one cluster centroid as a one-dimensional index. 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 C.2 reports the results of the k-means estimation showing the distribution of researcher type T in the sample.

4. Survey Experiment

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

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 additional $250,000 in guaranteed research funding, available for five years, in exchange for a lower salary, (*ii*) an additional $1,000,000 in guaranteed research funding, available for five years, in exchange for a lower salary, (*iii*) an elimination of all administrative duties for five years in exchange for a lower salary, and (*iv*) an increase in duties by 20 hours per month for five years in exchange for a higher salary.[17] 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.[18] We denote such an income level with M_{ij} (or, equivalently, $WTP_{ij} \equiv M_i - M_{ij}$) where M_i represents the actual income level and $j \in \{1, 2, 3, 4\}$ indexes the four experiments. Appendix Figure C5 displays examples of how these thought experiments appeared to the survey respondents.

These experiments are designed to reflect real-world examples of the tradeoffs that academic researchers face on the job market (e.g., a promotion that involves more salary and more administrative duties; a job offer that involves less salary but a larger laboratory funding package, etc.). In the notation of the model, the counterfactual guaranteed funding

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

[18] 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.

and duties in the four experiments give rise to the following counterfactual states:

$$
\begin{aligned}
(G_{i1}, D_{i1}) &= \big((G_i + \$250{,}000), D_i\big) \\
(G_{i2}, D_{i2}) &= \big((G_i + \$1{,}000{,}000), D_i\big) \\
(G_{i3}, D_{i3}) &= \big(G_i, 0\big) \\
(G_{i4}, D_{i4}) &= \big(G_i, (D_i + 20 \text{ hours/month})\big),
\end{aligned}
$$

where G_i and D_i represent actual states. As noted in Section 2, the values of M_{ij} reported in these four experiments make the researcher indifferent between the current state (M_i, G_i, D_i) and counterfactual (M_{ij}, G_{ij}, D_{ij}). Therefore, we can map the reported values to their model-based counterpart $\widehat{M}_{ij}$ with a known function of observable data and the unknown parameters to be estimated.[19]

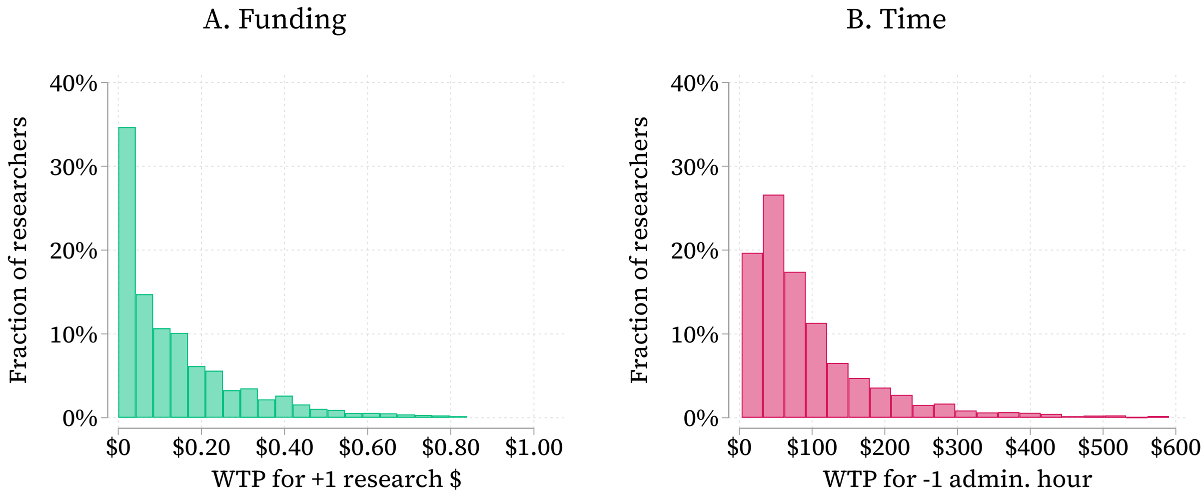
WTP for a Benchmark Good. When soliciting WTP measures, especially in an un-incentivized manner, it is possible that respondents under- or overstate their price sensitivity. In order to test for this, we follow Dizon-Ross and Jayachandran (2022) and explore respondents' WTP for a "benchmark good." Dizon-Ross and Jayachandran (2022) note that, if one solicits a subject's WTP 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 NSAR asks researchers 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. WTP Results

Figure 1 plots the distribution of WTP for guaranteed research funding and free time. The values shown are expressed per dollar of funding or per administrative hour removed. As evidenced by Figure 1, there is considerable variation in WTP responses across researchers. The mean, median, and standard deviation of the two WTP distributions are \$0.15, \$0.10, and \$0.18 for funding and \$102, \$68, and \$112 for time, respectively. Some of this variation reflects different preferences and constraints, but some variation reflects heterogeneous

[19] Note that 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.

FIGURE 1. Willingness to Pay Experiment Responses



Note: Shows histograms of WTP in salary dollars per additional funding dollar (Panel A) or administrative hour removed (Panel B), averaged across each input's two experiments. Each distribution excludes its bottom and top 1%.

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 to our approach, but they may be influenced by sample selection effects, behavioral biases, or noise. While the thought experiments do reflect real-world tradeoffs researchers make on the job market, these data are novel. So, it is not clear what plausible variation should look like here. Still, we can conduct some tests motivated by economic and statistical theory to explore the plausibility of our WTP distributions.

Motivated by the notion of opportunity costs, we first test how researchers' WTP varies with their implied hourly wage (i.e., 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 C6A, 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 C6B, we find this to be true as well.

As another test, Appendix Tables C5 and C6 describe analyses where we regress researchers' WTP responses on the variables in the model as well as (*i*) the large vector of observable features also solicited in the survey ($\mathbf{Z}_i$), (*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 82–96% of the variation in WTP responses (see Appendix Table C5). When we include the large vector of observable features as additional explanatory features, the R^2 statistics increase by only a few percentage points. Residual variation in WTP responses (conditional on the variables of the model) is not well explained by these heterogeneous observables. This evidence supports the assumption that the model's ingredients are the key determinants of researchers' answers and gives us confidence that we are not mis-specifying heterogeneity in the model by a significant degree.

We also find that the inverse Mills ratio is not a statistically significant predictor of responses (see Appendix Table C6). 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. 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 C6). 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; a one-standard-deviation increase in benchmark-good WTP is associated with changes of roughly 0.01–0.03 standard deviations in the experiment responses. Furthermore, the benchmark-good WTP is included in our vector of features that we use to make the researcher type index (which determines heterogeneity in the utility function parameters), which allows us to control for this to some degree. Overall, while the experiments are not incentivized, researchers' responses generally appear to behave as expected and are of plausible magnitudes.

5. Estimates for the Model of Science

In this section, we summarize the estimated preference parameters (Section 5.1), production and technology parameters (Sections 5.2–5.4), and the market-level social multipliers that rationalize the observed allocations across markets (Section 5.5). For more on the model's fit, Appendix Figure D1 reports absolute errors in predicted salaries at indifference. Across the four experiments, median absolute errors average roughly 9% (as a percentage of salaries).

5.1. A View of Researchers' Utility Functions

To illustrate our estimates of the utility function, Appendix Figure D2 shows how an average researcher's utility depends on the levels of key variables (i.e., 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 to the 90^{th} percentile, 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 53%. 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.[20] The fact that the disutility from spending more time on research is close to linear and practically small suggests that adjustments to research time (conditional on duties) may be relatively inexpensive compared to adjustments to research funding.

5.2. Funding Intensity

Figure 2 summarizes the funding intensity (output elasticity) parameter γ_i. Figure 2A shows the individual-level distribution, which exhibits considerable heterogeneity. The share of researchers that are very labor-intensive or very funding-intensive (i.e., $\gamma_i < 0.2$ or $0.6 < \gamma_i$) is 32%.

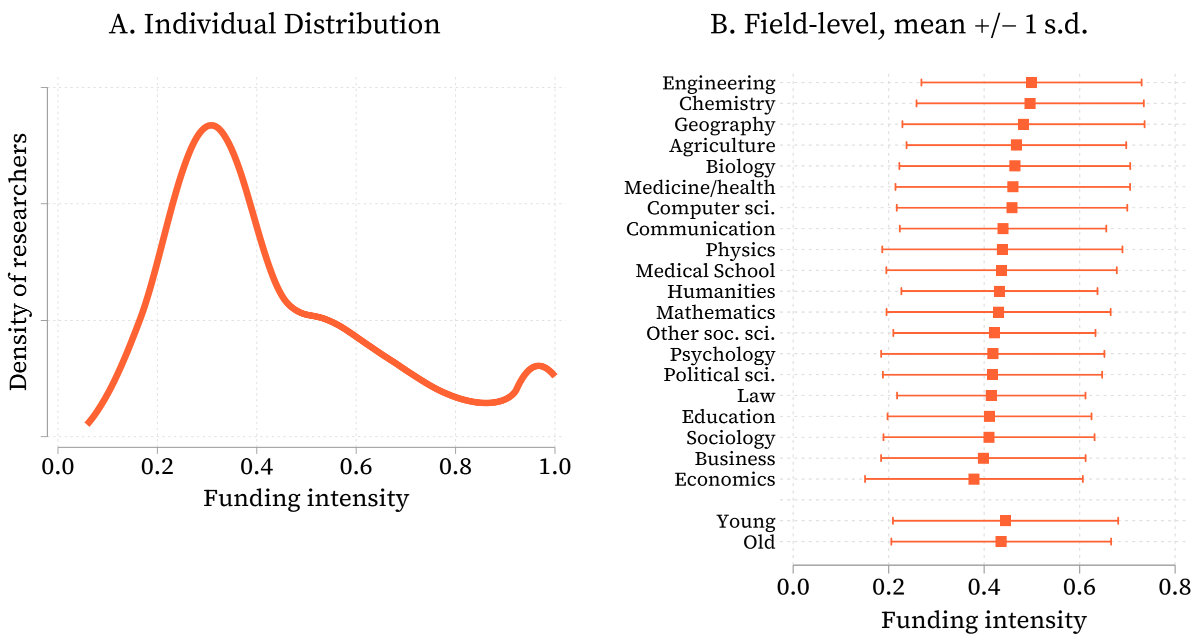
As both an interesting exercise and test of face validity, Figure 2B reports the means and standard deviations for the 20 scientific fields and two age groupings that define our scientific markets. Intuitively, traditionally capital-intensive fields (e.g., engineering, chemistry) tend to be more funding intensive, while theoretical and social sciences (e.g., economics, law, sociology) are less so. Still, the within-field standard deviations reveal substantial heterogeneity in funding intensity across researchers within fields, highlighting the importance of allowing for individual-level heterogeneity in this technological attribute.

5.3. Productivity and Output

Figure 3A shows the distribution of the productivity parameter A_i, normalized to have a mean of 1 within each market. Our estimates reveal a large skew in researchers' beliefs

[20]The numerical estimates of the common parameters are available upon request.

FIGURE 2. Funding Intensity, γ



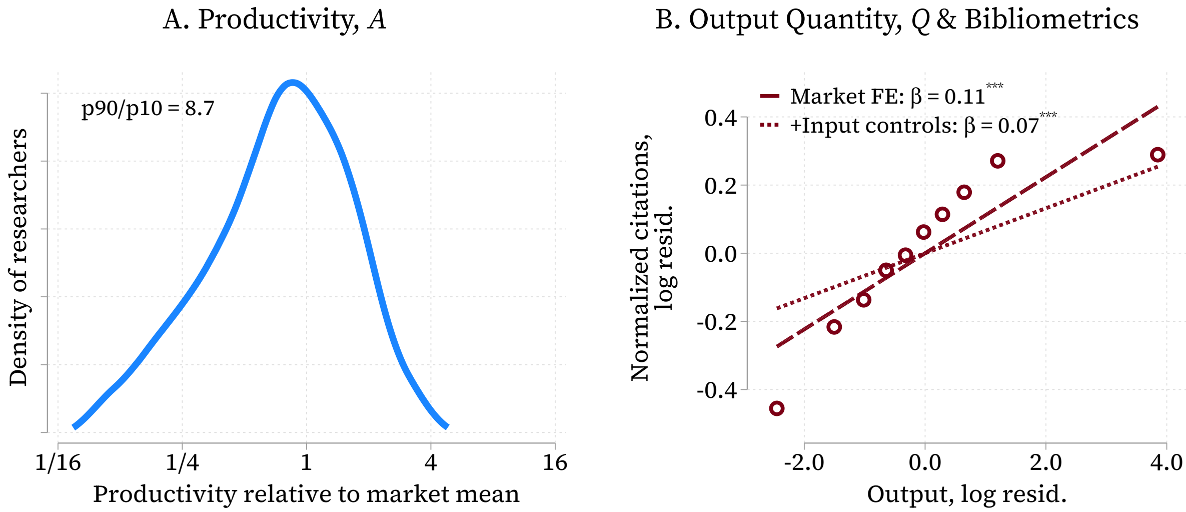
Note: Panel A shows the distribution of researcher-level funding intensity, γ_i; values below the 1st percentile and above the 99th percentile are omitted for display. Panel B shows field and career-stage means; whiskers extend one within-group standard deviation in each direction.

about their productivity. Figure 3A also reports the ratio of the 90^{th} and 10^{th} percentiles, the "90–10 productivity ratio" measure of dispersion commonly reported in traditional productivity studies. The 90–10 productivity ratio here is substantial: the 90^{th} percentile researcher believes they are roughly 9 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× (e.g., Syverson 2011; De Loecker and Syverson 2021). Most worker-level productivity estimates also appear to exhibit this degree of dispersion, although the specifics of productivity metrics vary widely across studies (Hoffman and Stanton 2024). One notable exception, and the source of the oft-quoted "10× engineer" moniker, 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.

Motivated by Gabaix (2009) and related work studying the upper tail of talent, we examine the highest values of researchers' productivity. Appendix Figure D3 shows approximately linear relationships between log rank and log productivity among the top 5% and top 1% of researchers, which indicate Pareto tails. The implied Pareto tail exponents — the negatives

FIGURE 3. Productivity & Output



Note: Panel A shows A_i relative to its market mean; the density excludes the bottom and top 1%, but p90/p10 uses the full sample. Panel B shows ten binned means of log field-normalized citations and log output after removing market fixed effects; the short-dashed fit adds log funding and log research time. Citations, output, funding, and research time are trimmed at the 1st and 99th percentiles. Stars use median draw-specific robust p-values: $^{*}\, p < 0.10$, $^{**}\, p < 0.05$, and $^{***}\, p < 0.01$.

of the fitted slopes — are both roughly 4, considerably larger than Zipf's law (which is an exponent of one). The magnitudes indicate that doubling a productivity threshold leaves roughly 6% as many researchers above it. By comparison, Zipf's law, with a tail exponent of one, would leave half as many. Thus, despite substantial dispersion in productivity overall, the upper tails are considerably less concentrated than Zipf's benchmark.

As another check on the estimates, we examine whether model-implied output is positively related to observable research performance per traditional bibliometrics. We have noted that a strength of our approach is that it does not require observing scientific output; nevertheless, if citations are positively related to broader scientific output on average within markets, they provide a useful external check on the model's output measure.

We use bibliometric data matched to approximately 77% of the sample and measure research performance using field-normalized citations from researchers' pre-survey publication records since post-survey data is not yet available at scale. While these bibliometrics describe past performance (and the model captures researchers' expectations about future output), there is good evidence that these bibliometrics are persistent (within individuals) over time (Sinatra et al. 2016).

Figure 3B shows that model-implied output and citations are positively correlated within markets. This relationship is reassuring, but also could simply reflect differences in inputs:

researchers with more funding or research time may produce both more model-implied output and more cited work. A sharper check therefore controls for input levels. When we include log-transformed funding and research time as controls, the focal coefficient falls from 0.11 to 0.07, but remains statistically significant.[21] This finding suggests that the estimated production parameters, and thus the predicted output, contain information about productivity beyond that captured by input levels alone.

5.4. Technology Correlations

Dispersion in output depends both on productivity differences and on how inputs are distributed across researchers. To examine their respective contributions, let $x_i \equiv \gamma_i \times \log(B_i) + (1-\gamma_i) \times \log(R_i)$ denote the input component. The Cobb-Douglas production function implies that $\log(Q_i) = \log(\alpha_i) + x_i$. The variance of log output can be decomposed as follows:

$$\underbrace{\mathrm{Var}(\log(Q_i))}_{100\%} = \underbrace{\mathrm{Var}(\log(\alpha_i))}_{58\%} + \underbrace{\mathrm{Var}(x_i)}_{64\%} + \underbrace{2 \times \mathrm{Cov}\left(\log(\alpha_i),\, x_i\right)}_{-22\%}, \tag{10}$$

with the percentages below each term expressing its contribution relative to total variance in log-output. The variance of log productivity alone represents 58%; including its covariance with the input component reduces its contribution to 36%. This latter number is the percentage reduction in log-output variance that would result from eliminating productivity heterogeneity while holding each researcher's inputs and funding intensity fixed. The negative covariance indicates that differences in inputs and funding intensity partly offset productivity differences in determining output.

This decomposition describes how productivity and realized inputs combine to generate output dispersion across the sample. We next examine how researchers' production characteristics relate to the funding opportunities they face within markets. Table 3 reports pairwise correlations among productivity (α_i), funding intensity (γ_i), fundraising ability (ϕ_i), and guaranteed funding (G_i), after removing market fixed effects.

Within markets, productivity is positively correlated with funding intensity, but negatively correlated with fundraising ability. Fundraising ability is also negatively correlated with funding intensity. Thus, researchers with more productive or more funding-intensive technologies tend to raise fewer dollars per hour devoted to fundraising. Guaranteed

[21] The reported significance measures condition on the model-parameter draws; they do not incorporate all uncertainty in the estimated parameters.

TABLE 3. Production and Input Correlations

		γ_i	α_i	ϕ_i	G_i
funding intensity	γ_i	–			
research productivity	α_i	0.39^{***}	–		
fundraising ability	ϕ_i	-0.33^{***}	-0.20^{***}	–	
guaranteed funding	G_i	0.44^{***}	0.06^{**}	0.07^{***}	–

Note: Reports Pearson correlations after removing market fixed effects. Correlations and two-sided tests pool draws using Fisher transformations and Rubin's variance formula. $^{*}\,p < 0.10$, $^{**}\,p < 0.05$, and $^{***}\,p < 0.01$.

funding exhibits a different pattern: it is positively correlated with funding intensity, but only weakly correlated with productivity and fundraising ability.

Together, these results show that research productivity, fundraising ability, and access to guaranteed funding are distinct features of researchers' circumstances. They suggest a potential for misallocation, but cannot speak directly to any notion of allocative efficiency. Next, in Section 6, we use the model to directly quantify the gains available from reallocating inputs.

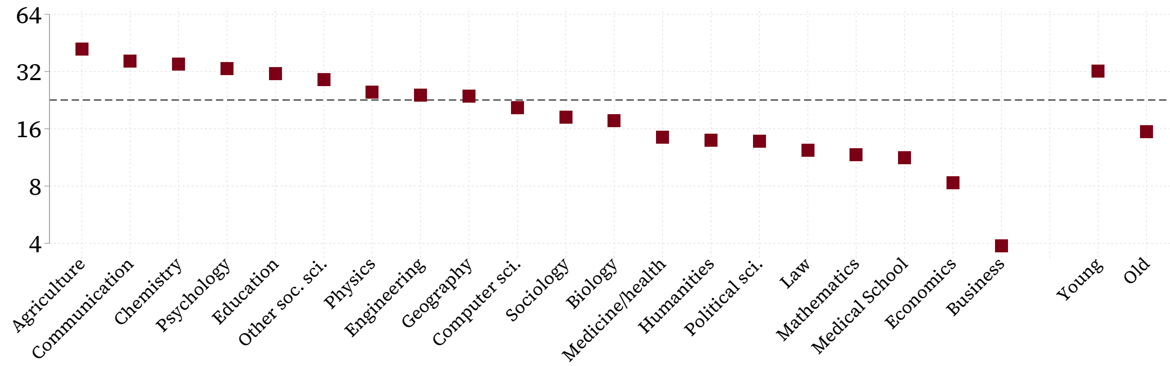
5.5. Implied Social Multipliers

Figure 4 reports the implied social multipliers that rationalize observed funding across markets. Following Section 2.6, we normalize the marginal cost of research funds to one and assume that an additional dollar reaches researchers in proportion to their observed funding. We estimate the multipliers for the 40 field-by-career-stage markets and report field and career-stage averages, weighting each market by its number of researchers. The dashed line shows the corresponding science-wide average, which is 23.

The estimates reveal substantial heterogeneity. Field-level multipliers range from roughly 4 to 42. There is also a sizable difference across career stages: the average multiplier for younger researchers is roughly 2 times larger than for older researchers. This is consistent with (among other things) life-cycle models and learning-by-doing dynamics that suggest larger externalities from investing in younger researchers (e.g., Levin and Stephan 1991; Qiu 2023).

A larger multiplier means that a larger gap between social and private value is needed to rationalize a market's observed funding. Under our benchmark, markets with lower marginal private returns to funding require larger social multipliers for the marginal social return to equal the marginal cost of funds. Thus, these estimates describe the funding system's implicit valuations, not independently measured differences in scientific

FIGURE 4. Implied Social Multipliers, χ



Note: Shows researcher-count-weighted field and career-stage averages of χ_k. The dashed line shows the corresponding science-wide average.

spillovers. They may also absorb historical commitments, institutional constraints, or other objectives that influence funding decisions. That is, a larger multiplier does not, by itself, indicate that a market should receive more funding: the multiplier is inferred precisely to rationalize the funding it already receives.

These estimates provide the weights we use to aggregate social value across markets in the counterfactual exercises that follow. Because the multiplier is common to researchers within a market, it does not affect which within-market allocation maximizes social value. Nor does rationalizing the total funding assigned to a market imply that its funding is efficiently distributed among researchers. The next section examines this separate question by comparing observed allocations with feasible reallocations of the same aggregate inputs.

6. Allocative Efficiency and Counterfactuals

The preceding section documents substantial differences in researchers' technologies and funding opportunities. We now ask what these differences imply for the allocation of resources. How much more social value could be produced with improved allocation of funding and research time? How does the scope for improvement differ across markets? And how well could alternative and more practical allocation policies perform?

Our central measure is allocative efficiency. For any allocation a, let V_k^a denote the social value of output in market k, and let $a^\star$ denote the allocation that maximizes this value using the same aggregate input levels (subject to any relevant constraints). We define:

$$\text{Allocative Efficiency}_k(a) \equiv 100 \times \frac{\mathcal{V}_k^a}{\mathcal{V}_k^{a^\star}} \,. \tag{11}$$

For example, an observed allocative efficiency of 50% means that a market produces half the social value attainable with its observed aggregate inputs. This measure distinguishes gains from reallocating resources from gains from expanding them.[22]

All counterfactuals involve reallocations within the field-by-career-stage markets defined above. We do not move resources across fields or career stages, nor do we change researchers' technologies. Social value is measured using the functions and implied multipliers from Section 2.6.[23]

6.1. Optimal Benchmark

To establish our theoretically optimal benchmark allocation (i.e., the allocation that determines $\mathcal{V}_k^{a^\star}$), we consider a planner who knows each researcher's technology and can assign research funding and research time directly. The planner uses observed aggregate input levels and solves:

$$\begin{aligned} \max_{\{B_i^{\text{opt}}, R_i^{\text{opt}}\}_{i\in k}} \quad & \sum_{i\in k} \mathcal{V}_i(B_i^{\text{opt}}, R_i^{\text{opt}}) \\ & \text{subject to} \\ & \sum_{i\in k} B_i^{\text{opt}} = \sum_{i\in k} B_i \,, \quad \sum_{i\in k} R_i^{\text{opt}} = \sum_{i\in k} R_i \,, \quad F_i^{\text{opt}} = 0 \;\forall\; i \in k \,. \end{aligned} \tag{12}$$

Direct funding assignment removes the need to fundraise, hence the last constraint. However, the hours released from fundraising are not added to aggregate research time: the constraint on total R remains binding. Researchers' administrative and teaching duties are held fixed. We also impose researcher-level constraints to limit extrapolation beyond

[22] In our policy experiments, if the aggregate input level changes due to researchers' behavioral responses, we resolve for the benchmark ($\mathcal{V}_k^{a^\star}$).

[23] Results shown in the Appendix D show qualitatively similar findings when we assume the social multipliers, χ_k, equal 1 for all markets. This is because the common multiplier χ_k does not influence the optimal within-market allocations; it only affects the aggregation of metrics across markets.

observed allocations and the variation in the WTP experiments.[24]

Beyond our metric of allocative efficiency, we can investigate misallocation with input wedges. Let λ_k^B and λ_k^R denote the shadow prices on funding and research time implied by the optimal benchmark in market k. For input $X \in \{B, R\}$ and an allocation a, define researcher i's wedge as:

$$w_{iX}^a \equiv \frac{\partial \mathcal{V}_i^a / \partial X_i}{\lambda_k^X} . \tag{13}$$

In an interior solution 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 (e.g., Hsieh and Klenow 2009).[25]

Optimal Benchmark Allocation Results. Table 4 reports how inputs, outputs, and productivity metrics change as we move from the observed allocation to the optimal benchmark. Column 1 optimizes both inputs, while Columns 2 and 3 optimize research time and funding separately to investigate factor-specific misallocation. The single-input exercises hold both the other input and fundraising time at their observed levels.

The observed allocation achieves an allocative efficiency of 48% . Put another way, the theoretically optimal benchmark allocation of observed resources can increase the social value of output by roughly 108% . Almost all of this gain can be obtained by reallocating funding alone: this exercise, shown in Column 3 of Table 4, yields an allocative efficiency of 98% . This indicates that research funding is much more misallocated than research time and is clearly illustrated in Appendix Figure D5, which plots the distribution of input

[24] These constraints are:

$$\max\left(B_k^{\min}, (B_i - \$1{,}000{,}000)\right) \leq B_i^{\text{opt}} \leq \min\left(B_k^{\max}, (B_i + \$1{,}000{,}000)\right)$$
$$\max\left(R_k^{\min}, (R_i - 1{,}200\text{hrs})\right) \leq R_i^{\text{opt}} \leq \min\left(R_k^{\max}, (R_i + 1{,}200\text{hrs})\right)$$
$$D_i + R_i^{\text{opt}} \leq H_k^{\max} .$$

Here, the market minima and maxima are calculated from the observed data. Technically, these researcher-level constraints imply that our benchmark may best be described as a constrained second-best. Practically, these restrictions allow substantial reallocation while ruling out arbitrarily large changes and ensuring that we operate inside the support where the model is disciplined by the variation in the WTP experiments.

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

wedges for funding and time; the variance of log wedges for funding is roughly 7 , while it is only 0.52 for research time.

In terms of the distribution of inputs, the optimal benchmark involves a more uniform distribution of funding (i.e., the standard deviation of funding falls by 24%) and a slightly more concentrated distribution of research time (i.e., the standard deviation of research time increases by 10%). This reallocation leads to a concentration of output both in private and social value terms (i.e., the standard deviation of researcher-level social value increases by 65%). Appendix Figure D4 illustrates these changes, including binned scatterplots of observed versus optimal input and output levels, as well as Lorenz curves to help visualize the two allocations.

TABLE 4. Observed versus Optimal Benchmark Comparison

		Input(s) optimized		
		B, R	R only	B only
		(1)	(2)	(3)
Change in inputs & outcomes				
research budget: B_i	mean	~0%	~0%	~0%
	s.d.	−24.3%	~0%	−24.4%
research time: R_i	mean	~0%	~0%	~0%
	s.d.	+9.6%	+8.6%	~0%
fundraising time: F_i	mean	−100%	~0%	~0%
	s.d.	−100%	~0%	~0%
utility: $\mathcal{U}_i$	mean	+4.5%	~0%	+4.3%
	s.d.	+8.5%	~0%	+8.5%
output, private value: $b_i(Q_i)$	mean	+93.4%	−0.1%	+91.4%
	s.d.	+52.1%	~0%	+51.9%
output, social value: $\mathcal{V}_i(Q_i)$	mean	+108%	+0.9%	+104%
	s.d.	+64.7%	+0.1%	+63.8%
Private utility, compensating variation: Income				
% of own income		+5.4%	~0%	+5.1%
$/year		+$9,318	−$13	+$8,892
Social value, equivalent variation: Funding				
researcher-level, $/year		+$266,253	+$1,671	+$254,209
population-level, $/year		+$32.3B	+$0.20B	+$30.8B
Allocative efficiency				
initial allocation = 48.1%				
benchmark allocation		100%	48.5%	98.0%

Note: Reports changes from observed allocations under the benchmark reallocations in Section 6.1. Mean compensating variation excludes nonfinite values. Population funding equivalents divide the sample total by the 3.3% response rate.

How does scientists' private utility change under the optimal benchmark? It increases by roughly 4% . To convert this magnitude to a dollar-based valuation, we solve for researchers' compensating variation, which describes the dollars (out of their income) they'd be willing to pay to have their input allocation solution in the optimal benchmark. On average, researchers' compensating variation is about $9,300 per year, or 5 % of their annual income; this is a non-trivial amount of private value.

To value the gains in social value through the lens of the people and institutions funding this science, we also solve for the equivalent variation in research funding, which describes the additional research funding that would need to be invested using current allocations to achieve the same growth in social value that the optimal benchmark achieves with observed input levels. This exercise indicates that the planner would need to increase aggregate research funding budgets by 195 % per year. On the scale of the full population of researchers, this amounts to more than $32 billion per year; this is certainly a non-trivial amount of social value.

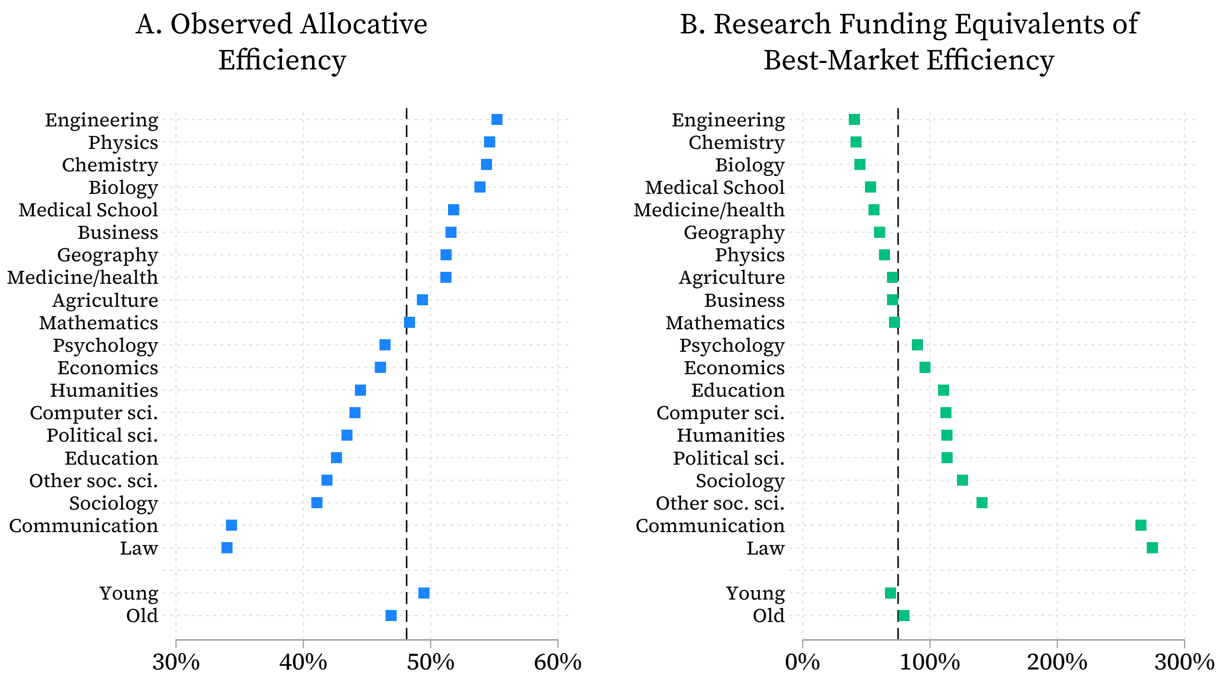
Sensitivity Tests. Appendix Table D1 shows that these aggregate, science-wide conclusions do not hinge on our use of the revealed preference estimates for the social multipliers (χ_k), with similar results obtained using $\chi_k = 1$ for all markets. In Appendix Table D2, we report a wide range of results varying the measurement-error share assumption (ρ), the (in-)dependence of errors across inputs, the treatment of preference heterogeneity, and different choices about winsorizing outliers. Across these specifications, the estimates of observed efficiency range from 37 % to 54 %, and the population-level budgetary equivalents span $27 billion to $93 billion per year. Thus, our preferred approach reported here in the main text is on the conservative side of the spectrum, in the sense that it is one of the most optimistic views of observed allocations.

6.2. Allocative Efficiency Across Markets

The aggregate results conceal considerable variation across scientific markets. Figure 5A reports observed allocative efficiency by field and career stage. There is considerable variation in efficiency across markets, with the low end of the spectrum in the range of 35–40%, while others have efficiencies near 55%.

Recall, the theoretically optimal benchmark assumes the planner has perfect information on all researchers' technologies and can costlessly assign inputs. This is clearly not realistic. So as an initial attempt to appreciate practical feasibility, we resolve for the gains from reallocation using *not* the theoretical benchmark (which is 100% efficient), but the best

FIGURE 5. Across-market Comparisons



Note: Panel A shows observed allocative efficiency by field and career stage. Panel B shows funding increases equivalent to the social-value gains from matching the highest observed market efficiency, allowing research time to respond. Dashed lines show science-wide values.

observed allocative efficiency across markets in our sample (which is roughly 68%).[26] Figure 5B reports the equivalent variation in budgets from this exercise. We find that if each market was as efficient as the most-efficient market, the gains in social value would still be economically large and equivalent to a science-wide research budget increase of 75 %.

Correlates of Market-level Efficiency. Identifying the causal determinants of allocative efficiency (e.g., specific policy effects) is clearly important but well beyond the scope of this paper. To provide a preliminary view of suggestive correlations, we conceptualize the variances of variables and parameters in the model as being reduced-form proxies for different policies across markets. For instance, if the variance of fundraising ability is larger in one market than another, that difference may reflect differences in fundraising policies. As an explorative exercise, Appendix Table D3 reports the results from a Lasso-based stability selection process to identify the most important predictors of market-level allocative efficiency, using market-level variances as policy proxies. The clear standout is

[26]This is analogous to the exercise in Hsieh and Klenow (2009) that assigns the allocative efficiency of the US to India and China.

the variance of guaranteed research funding ($\text{Var}_k(G_{i\in k})$), which is the most commonly selected feature across the selection runs. Appendix Figure D6 plots the unconditional relationship between $\text{Var}_k(G)$ and allocative efficiency for all forty markets, which reveals a stark, positive relationship. Markets that have higher variance funding guarantees also tend to have high allocative efficiency. This and other ongoing metascience debates motivate our policy experiments in the next section.

6.3. Policy Experiments

The benchmark results show the potential gains to reallocation are large, but do not explain how a funding institution could achieve them. Here, we use our model as a laboratory and consider a range of policy experiments that change how researchers obtain inputs. Given the result that virtually all of the loss from misallocation is driven by funding policies, we focus on that single input. Importantly, in all of these experiments we allow for researchers' behavioral response (e.g., endogenous changes to their choices of R and/or F). Furthermore, if either the policy or behavioral responses change aggregate input levels, we resolve the optimal benchmark at the new aggregate input levels for calculating allocative efficiency.[27] Table 5 reports the results from these experiments, which we describe below.

Uniform Policies. Concerns about the burden of grant applications and unequal access to competitive funding have motivated proposals to provide researchers with more uniform baseline support (e.g., Ioannidis 2011; Vaesen and Katzav 2017). We experiment with two versions of uniformity. Column 1 assigns researchers in each market a common level of guaranteed funding, G, while Column 2 assigns them a common fundraising ability, ϕ. The two policies yield similar reductions in funding dispersion (i.e., the standard deviation of funding falls by roughly 40 % in both cases), but their efficiency gains differ substantially. Allocative efficiency with uniform G rises only a few percentage points, while a uniform ϕ (i.e., having all researchers face the same shadow price of fundraising) increases efficiency to 67 %. Notably, this is not a fundraising-time-saving result: fundraising time actually increases under the uniform ϕ policy while efficiency still improves.

[27] For the uniform, bibliometric, and block-or-cap policies, we implement the policy and then proportionally adjust funding, subject to the individual bounds, to preserve the observed market budget. Researchers can adjust research time after this final funding assignment, but fundraising already undertaken remains a sunk cost. Researchers do not anticipate this final funding adjustment when choosing fundraising effort. These experiments therefore evaluate the policy rules together with a budget-balancing adjustment, rather than fully anticipated equilibria under the allocation rules alone.

Bibliometric Output-based Policies. Citation information already influences grant allocations, but its value as a proxy for scientific productivity remains unclear (e.g., Stephan, Veugelers, and Wang 2017; Hager, Schwarz, and Waldinger 2024). We therefore consider two policies that allocate funding based on bibliometric output. Both use the field-normalized citation measure introduced above.[28] Column 3 uses this measure directly, while Column 4 uses a naïve input adjustment and divides the bibliometric by the researcher's observed total funding. The raw citation policy raises allocative efficiency to 60 %, compared with only 51 % under the funding-scaled policy. The contrast between them illustrates how an intuitive productivity proxy (i.e., an output-to-input proxy) need not be a better allocation rule than an output proxy in isolation when both proxies have measurement error.

Block or Cap Policies. There continue to be calls for greater scrutiny of awards to already well-funded researchers and hypotheses that spreading resources more broadly would improve efficiency (e.g., Berg 2012; Fortin and Currie 2013). We consider three policies that move in this direction in ways that are less dramatic than the uniform policies described above. The partial block grant in Column 5 converts 25% of observed fundraising-based funding into an equal addition to guaranteed funding and reduces each researcher's return to fundraising by 25%. This policy produces the largest efficiency gain among the non-market rules considered here: allocative efficiency reaches 72 %. The broader base funding policy in Column 6 moves each researcher's guaranteed funding 25% toward the market mean, leaving their fundraising return unchanged. Column 7 caps fundraising-based funding at the market's observed 90^{th} percentile and distributes the amount above that threshold in the observed allocation equally as additional guaranteed funding. These two policies produce smaller gains: allocative efficiency reaches 50 % and 53 %, respectively. These results do not imply that broader support or caps are generally ineffective. Rather, compared to the partial block grant exercise of Column 5, they show that redistributing guarantees or limiting the upper tail alone yields relatively modest improvements.

Market Economies. Our most exploratory experiments consider scenarios closer to traditional capital markets and treat research funding as a tradable good that researchers can purchase using their own income. First, in Column 8, we allow researchers to competitively trade their (observed) funding allocations at a common, market-clearing price. Here, buyers pay sellers, so trading preserves both aggregate market funding and aggregate researcher income. This decentralized approach can almost match the theoretical bench-

[28] For researchers we cannot match to a bibliographic record, we use the predicted relationship between bibliometrics and the model-implied output estimated using researchers matched to their bibliographies.

mark, with the discrepancy arising because researchers trade per their private valuations, not the social value of their inputs. Columns 9–11 consider a simpler mechanism in which the planner sells funding at a posted price. These policies retain guaranteed funding at observed levels, and place all fundraising-based funds into a pool. Researchers can purchase $50,000 tranches from this pool at prices of 2, 10, or 20 cents of income per dollar of research funding. These payments go to the planner and are recycled into research funding, so aggregate research budgets can increase without additional external funding. All three posted-price policies substantially improve efficiency relative to what is observed. The intermediate price of 10 cents of income per dollar of research funding (near the sample median) yields the largest gains in efficiency, and appears to balance the selection effects of higher prices (which screen on productivity beliefs) with the budget constraint effects of those higher prices (which screen on incomes and preferences).

Two conclusions emerge from these experiments. First, policies that improve allocative efficiency need not make researchers better off on average. The uniform and bibliometric policies generate positive social gains but negative mean private compensating variation. Thus, researchers' willingness to support a reform may differ from its value to the institutions funding science. The partial block grant illustrates that this tradeoff is not inevitable: it produces the largest efficiency gain among the non-market policies, and researchers' compensating variation is positive and non-trivial on average. These results suggest that the practical design of reform matters for both the allocation of resources and researchers' private welfare.

Second, the market-based mechanisms produce the largest gains in allocative efficiency while also being valued positively (on average) by researchers. Certainly, the practical feasibility of such policies is far from clear, especially in the case of the competitive trading arrangement. And of course, these results abstract from implementation costs and a range of important forces beyond the scope of our model (e.g., the extensive margin of participating in a given market; the *directional* changes researchers may undertake given these policies). Still, these experiments provide some of the first evidence that researchers' willingness to exchange income for funding, if managed properly, could provide a mechanism to help scientific institutions improve resource allocation.

7. Discussion

This paper develops an approach to estimating productivity beliefs without observing output quantities or input prices. Applying it to academic science reveals substantial differ-

TABLE 5. Policy Experiments

		Uniform		Bibliometric based		Blocks or caps			Market economy			
		G only	ϕ only	raw	funding scaled	partial block grant	broader base funding	fundraiser funds cap	competitive trading	posted sale		
										p = .02	p = .10	p = .20
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Change in inputs & outcomes												
research budget: B_i	mean	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	+0.6%	+1.9%	+2.4%
	s.d.	–38.4%	–41.5%	–33.7%	–38.3%	–18.1%	–2.3%	–2.0%	–24.6%	–25.8%	–19.7%	–19.8%
research time: R_i	mean	+12.0%	+10.9%	+9.2%	+5.0%	+14.3%	+12.7%	+7.6%	+5.6%	+7.4%	+6.6%	+9.2%
	s.d.	–1.2%	–0.9%	+1.7%	+2.4%	–0.4%	–1.8%	+2.2%	+5.4%	+4.0%	+3.2%	+1.3%
fundraising time: F_i	mean	–57.4%	+124%	–100%	–100%	–80.0%	–27.3%	+1.6%	–100%	–100%	–100%	–100%
	s.d.	–2.5%	–18.3%	–100%	–100%	–52.5%	+15.1%	+21.8%	–100%	–100%	–100%	–100%
income: M_i	mean, $/year	$0	$0	$0	$0	$0	$0	$0	$0	–$847	–$2,627	–$3,215
Private utility, compensating variation: Income												
% of own income		–16.0%	–2.5%	–10.6%	–23.4%	+2.9%	–0.3%	+0.7%	+5.6%	+4.1%	+3.2%	+2.6%
$/year		–$23,758	–$4,634	–$18,247	–$36,249	+$4,962	–$836	+$1,219	+$9,356	+$7,137	+$5,793	+$4,683
Social value, equivalent variation: Funding												
researcher-level, $/year		+$17,488	+$88,445	+$55,507	+$13,138	+$110,254	+$8,436	+$21,145	+$266,362	+$217,020	+$233,364	+$213,148
population-level, $/year		+$2.12B	+$10.7B	+$6.73B	+$1.59B	+$13.4B	+$1.02B	+$2.56B	+$32.3B	+$26.3B	+$28.3B	+$25.9B
Allocative efficiency												
initial allocation = 48.1%												
policy allocation		51.9%	66.8%	60.5%	50.7%	71.6%	49.9%	53.0%	99.8%	91.1%	94.2%	90.3%

Note: Reports changes from the observed allocation under the policies in Section 6.3. Policies share the benchmark bounds on individual B_i and R_i, with market caps on F_i and H_i. Columns 1–8 preserve market funding; Columns 9–11 recycle researchers' payments into funding. Compensating variation includes income transfers and averages finite individual values. Population funding equivalents divide sample totals by the 3.3% response rate. Allocative efficiency uses a newly solved benchmark at each policy's realized aggregate inputs.

ences in researchers' technologies and large potential gains from reallocating resources, in particular, funding. In theory, a more efficient allocation could roughly double the social value of scientific output using the same amount of resources. Explorations of (relatively) more practical policies indicate that substantial gains could be had from, for instance, making the costs of fundraising more uniform.

Our results connect closely to work on the design of scientific funding mechanisms, most notably Adda and Ottaviani (2024). Much of the policy debate around grantmaking focuses on the *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 — which 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 forward.

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

> "*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 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 4 are key objects to debate. For example, if you think the positive social externality of knowledge produced by chemists is roughly 2 times larger than it is for political scientists, then the model indicates that the *relative* amount of funding in those two fields is currently efficient; if you think knowledge in chemistry is more or less than 2 times as 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 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 estimates describe researchers' beliefs about their own productivity. Accordingly, our counterfactual gains are evaluated using those beliefs, rather than independently measured production outcomes. These beliefs may be mistaken even when researchers report their WTP truthfully. 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 with a large number of producers acquiring inputs in a highly decentralized way to produce outputs that are hard to observe. In developing economies, accurate producer-level data on outputs is costly to obtain (e.g., Tybout 2000); in entrepreneurship, many organizations never produce observable output before exiting (e.g., Decker et al. 2014); in the nonprofit sector, output is often so high-dimensional that reaching consensus on suitable proxies is challenging (e.g., Philipson and Lakdawalla 2001). In Appendix B, we provide alternative formulations of our model that may provide a way forward to better understanding the distribution and determinants of productivity and efficiency in these and other important settings.

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.

Ahmadpoor, Mohammad, and Benjamin F Jones. 2017. "The dual frontier: Patented inventions and prior scientific advance." *Science* 357 (6351): 583–587.

Arora, Ashish, Sharon Belenzon, Konstantin Kosenko, Jungkyu Suh, and Yishay Yafeh. 2024. "The Rise of Scientific Research in Corporate America." *Organization Science* 36 (4): 1466–1488.

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.

Azoulay, Pierre, Joshua S Graff Zivin, and Gustavo Manso. 2011. "Incentives and creativity: Evidence from the academic life sciences." *RAND Journal of Economics* 42 (3): 527–554.

Azoulay, Pierre, and Danielle Li. 2022. "Scientific Grant Funding." *Innovation and Public Policy*: 117.

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

Berg, Jeremy M. 2012. "Well-funded investigators should receive extra scrutiny." *Nature* 489 (7415): 203.

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

Bikard, Michaël, and Matt Marx. 2020. "Bridging Academia and Industry: How Geographic Hubs Connect University Science and Corporate Technology." *Management Science* 66 (8): 3425–3443.

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.

Branstetter, Lee. 2005. "Exploring the Link Between Academic Science and Industrial Innovation."

Annales d'Économie et de Statistique: 119–142.
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.
Carrillo, Paul, Dave Donaldson, Dina Pomeranz, and Monica Singhal. 2023. "Misallocation in firm production: A nonparametric analysis using procurement lotteries." *Mimeo*.
Chen, Xugan, Ernest Liu, and Song Ma. 2026. "The Economic Value of Science.", Working Paper.
Clauset, Aaron, Samuel Arbesman, and Daniel B. Larremore. 2015. "Systematic Inequality and Hierarchy in Faculty Hiring Networks." *Science Advances* 1 (1): e1400005.
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.
Fleming, Lee, and Olav Sorenson. 2004. "Science as a Map in Technological Search." *Strategic Management Journal* 25 (8–9): 909–928.
Fortin, Jean-Michel, and David J Currie. 2013. "Big science vs. little science: how scientific impact scales with funding." *PloS one* 8 (6): e65263.
Gabaix, Xavier. 2009. "Power laws in economics and finance." *Annual Review of Economics* 1 (1): 255–294.
Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis*. 3rd ed.: Chapman & Hall/CRC.
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.

Hofstra, Bas, Vivek V. Kulkarni, Sebastian Munoz-Najar Galvez, Bryan He, Dan Jurafsky, and Daniel A. McFarland. 2020. "The Diversity–Innovation Paradox in Science." *Proceedings of the National Academy of Sciences* 117 (17): 9284–9291.

Howell, Sabrina. 2017. "Financing innovation: Evidence from R&D grants." *American Economic Review* 107 (4): 1136–64.

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.

Huang, Junming, Alexander J. Gates, Roberta Sinatra, and Albert-László Barabási. 2020. "Historical Comparison of Gender Inequality in Scientific Careers Across Countries and Disciplines." *Proceedings of the National Academy of Sciences* 117 (9): 4609–4616.

Iaria, Alessandro, Carlo Schwarz, and Fabian Waldinger. 2018. "Frontier Knowledge and Scientific Production: Evidence from the Collapse of International Science." *The Quarterly Journal of Economics* 133 (2): 927–991.

Ioannidis, John PA. 2011. "Fund people not projects." *Nature* 477 (7366): 529.

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

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.

Krieger, Joshua Lev, Monika Schnitzer, and Martin Watzinger. 2024. "Standing on the Shoulders of Science." *Strategic Management Journal* 45 (9): 1670–1695.

Levin, Sharon G, and Paula E Stephan. 1991. "Research productivity over the life cycle: Evidence for academic scientists." *American Economic Review* 81 (1): 114–132.

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

Liu, Lu, Yang Wang, Roberta Sinatra, C. Lee Giles, Chaoming Song, and Dashun Wang. 2018. "Hot Streaks in Artistic, Cultural, and Scientific Careers." *Nature* 559 (7714): 396–399.

Ma, Athen, Raúl J. Mondragón, and Vito Latora. 2015. "Anatomy of Funded Research in Science." *Proceedings of the National Academy of Sciences* 112 (48): 14760–14765.

Ma, Yifang, and Brian Uzzi. 2018. "Scientific Prize Network Predicts Who Pushes the Boundaries of

Science." *Proceedings of the National Academy of Sciences* 115 (50): 12608–12615.
Mansfield, Edwin. 1980. "Basic Research and Productivity Increase in Manufacturing." *The American Economic Review* 70 (5): 863–873.
Marx, Matt, and Aaron Fuegi. 2020. "Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles." *Strategic Management Journal* 41 (9): 1572–1594.
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."
Moretti, Enrico, Claudia Steinwender, and John Van Reenen. 2019. "The Intellectual Spoils of War? Defense R&D, Productivity, and International Spillovers." *The Review of Economics and Statistics*: 1–46.
Myers, Kyle. 2020. "The Elasticity of Science." *American Economic Journal: Applied Economics* 12 (4): 103–134.
Myers, Kyle R., and Lauren Lanahan. 2022. "Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy." *American Economic Review* 112 (7): 2393–2423.
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.
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.

Stephan, Paula, Reinhilde Veugelers, and Jian Wang. 2017. "Reviewers are blinkered by bibliometrics." *Nature* 544 (7651): 411–412.

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.

Vaesen, Krist, and Joel Katzav. 2017. "How much would each researcher receive if competitive government research funding were distributed equally among researchers?" *PLoS One* 12 (9): e0183967.

Way, Samuel F., Allison C. Morgan, Aaron Clauset, and Daniel B. Larremore. 2017. "The Misleading Narrative of the Canonical Faculty Productivity Trajectory." *Proceedings of the National Academy of Sciences* 114 (44): E9216–E9223.

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.

Yin, Yian, Yuxiao Dong, Kuansan Wang, Dashun Wang, and Benjamin F. Jones. 2022. "Public Use and Public Funding of Science." *Nature Human Behaviour* 6 (10): 1344–1350.

Zacchia, Paolo. 2020. "Knowledge spillovers through networks of scientists." *Review of Economic Studies* 87 (4): 1989–2018.

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

Supplemental Appendices

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 θ_i, and parameters μ_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:

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

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 the 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 (A4) 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:

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

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:

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

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 $\boldsymbol{\mu}_i$ as functions of a 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 θ_i

from researchers' optimality conditions and the model's structure. One important note is that we have separate estimation routines for inferring individual attributes θ_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 ϕ_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.[29] 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 γ_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 $\widehat{\gamma}_i$.[30] Therefore, we assume that ϕ_i and γ_i are parametric polynomial functions of state variables, and we estimate these functions using the other sub-sample of researchers with $F_i > 0$.[31]

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

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

[31] 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 estimated 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 $\boldsymbol{\mu} = (\omega, \psi, \boldsymbol{\delta})$, calibrated values $(B_{\min}, H_{\max})$, and observed allocations, we estimate the vector μ by generalized method of moments as:

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

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 $\mathbf{S}_i$, the researcher's type T_i, and of the estimand μ which determines individual attributes $\widehat{\theta}_i$ and individual-specific parameters $\boldsymbol{\mu}_i$.[32] 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 $\boldsymbol{\mu} = (\omega, \psi, \delta_{\sigma,0}, \delta_{\sigma,1}, \delta_{\eta,0}, \delta_{\eta,1}, \delta_{\xi,0}, \delta_{\xi,1}, \delta_{\zeta,0}, \delta_{\zeta,1})$. 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:

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

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 $\boldsymbol{\mu}_i$, 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.

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

and standard tolerance on parameter values. We then select the parameter 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 an 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 μ_{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.[33]

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 ρ 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}^{*}.

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

The Gaussian structure gives a closed-form posterior. Since

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

the posterior mean is

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

and the posterior variance is

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

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

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

The baseline specification draws ε_{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 ϕ_{ib}, γ_{ib}, α_{ib}, σ_{ib}, η_{ib}, ξ_{ib}, and ζ_{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 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).

A.5. Social Value and Implicit Social Multipliers

Here, we derive the expression for the revealed preference-based estimates of the market-specific externality terms, χ_k, as shown in Equation 9. The exercise asks which values of χ_k would make the observed allocation of research funding across markets consistent with the planner's problem in Equation 8.

Dollar-valued output. The private output-value function $b_i(\cdot)$ is measured in researcher i's utility units. We express this value in salary dollars using the researcher's marginal utility of income at the observed allocation:

$$v_i(Q) \equiv \frac{b_i(Q)}{m_i'(M_i)} = \frac{Q^{1-\eta_i}}{(1-\eta_i) \times \omega\,(M_i)^{-\sigma_i}} \; . \tag{A15}$$

Thus, $v_i(Q)$ is a money-metric version of researcher i's private output value. This normalization preserves the researcher-specific curvature of output value and the researcher-specific marginal utility of income. It does not require utility itself to be comparable across researchers.

The social value of the output produced in market k is therefore:

$$\mathcal{V}_k(\mathcal{B}_k;\chi_k) \equiv \chi_k \times \sum_{i\in k} v_i\left(Q_i\left(B_i(\mathcal{B}_k)\right)\right) \; . \tag{A16}$$

The restriction imposed by Equation A16 is that the same proportional social-value multiplier χ_k applies to every researcher in market k. The functions $v_i(\cdot)$, the production functions $Q_i(\cdot)$, and all of the variables and parameters entering them remain researcher-specific. Thus, χ_k is the only common-within-market externality shifter. Because $v_i(\cdot)$ is measured in dollars, χ_k is dimensionless.

Marginal incidence of market funding. For any amount of funding assigned to market k, we assume that the within-market allocation rule exhausts the market budget: $\sum_{i\in k} B_i(\mathcal{B}_k) = \mathcal{B}_k$. We define researcher i's share of a marginal dollar assigned to market k as:

$$s_{ik} \equiv \frac{dB_i(\mathcal{B}_k)}{d\mathcal{B}_k} \; , \qquad s_{ik} \geq 0 \; , \qquad \sum_{i\in k} s_{ik} = 1 \; . \tag{A17}$$

Thus, s_{ik} describes the incidence of an additional market funding dollar.

Private-dollar returns to research funding. Let $U_i^*(B_i, M_i, \ldots)$ denote researcher i's indirect utility after choosing their time allocations optimally. Researcher i's marginal private return to a direct dollar of research funding, expressed in salary dollars, is:

$$\begin{aligned}\pi_i^B &\equiv \frac{\partial U_i^*/\partial B_i}{\partial U_i^*/\partial M_i} \\ &= \frac{b_i'(Q_i) \times \partial Q_i/\partial B_i}{m_i'(M_i)} = \frac{\gamma_i \times Q_i^{1-\eta_i}}{B_i \times \omega M_i^{-\sigma_i}} .\end{aligned} \tag{A18}$$

The last equality follows from $b_i'(Q_i) = Q_i^{-\eta_i}$, $\partial Q_i/\partial B_i = \gamma_i Q_i/B_i$, and $m_i'(M_i) = \omega M_i^{-\sigma_i}$. The envelope theorem allows us to omit derivatives of the researcher's optimal time choices from this local private-value calculation.

Equation A18 treats B_i as an additional dollar of research funding that enters total funding one-for-one and holds the researcher's time choices fixed when taking the derivative. This does not mean that behavior is held fixed in the counterfactual exercises. The equilibrium values $Q_i(B_i(\mathcal{B}_k))$ used to evaluate alternative allocations incorporate researchers' behavioral responses. For the revealed-preference exercise, we use the direct-funding return in Equation A18 as the funding system's local measure of the value of an additional dollar.

Given an incidence rule, the private-dollar return to one additional dollar assigned to market k is:

$$\pi_k^B \equiv \sum_{i \in k} s_{ik} \times \pi_i^B . \tag{A19}$$

The revealed-preference assumption is that the marginal social value of market funding equals this incidence-weighted private-dollar return multiplied by the market-specific social multiplier:

$$\frac{d\mathcal{V}_k(\mathcal{B}_k; \chi_k)}{d\mathcal{B}_k} = \chi_k \times \pi_k^B . \tag{A20}$$

Equation A20 is the additional assumption that connects researchers' model-implied private returns to the funding system's valuation of marginal research funding.

Planner's problem. We represent the collection of institutions and policies that allocates research funding across markets as solving:

$$\max_{\{\mathcal{B}_k \geq 0\}_k} \quad \sum_k \mathcal{V}_k(\mathcal{B}_k; \chi_k) - \mathcal{C} \times \sum_k \mathcal{B}_k , \tag{A21}$$

where $\mathcal{C}$ is the marginal social cost of supplying one additional dollar of research funding. The term "planner" refers to the funding system as a whole and does not require a single decision maker.

For any market receiving positive funding, the first-order condition evaluated at the observed allocation is:

$$\chi_k \times \pi_k^B = \mathcal{C} \,. \tag{A22}$$

Combining Equations A18–A22 yields the market-specific social multiplier that rationalizes the observed allocation:

$$\begin{aligned} \chi_k &= \frac{\mathcal{C}}{\pi_k^B} \\ &= \frac{\mathcal{C}}{\sum_{i\in k}\left(\frac{dB_i(\mathcal{B}_k)}{d\mathcal{B}_k} \times \frac{\partial U_i^*/\partial B_i}{\partial U_i^*/\partial M_i}\right)} \\ &= \frac{\mathcal{C}}{\sum_{i\in k}\left(s_{ik} \times \frac{\gamma_i \times Q_i^{1-\eta_i}}{B_i \times \omega \times M_i^{-\sigma_i}}\right)} \,. \end{aligned} \tag{A23}$$

All terms on the right-hand side of Equation A23 are evaluated at the observed allocation.

Our preferred specification assumes that an additional dollar assigned to market k is divided in proportion to the observed allocation across the N_k researchers in that market: $s_{ik} = B_i/\mathcal{B}_k$ for all i in k.

In the empirical implementation, we calculate the researcher-level return in Equation A18 within each Monte Carlo draw, take the observed-budget-share-weighted mean within each market, and then invert that mean. This ordering ensures that all nonlinear researcher-level objects are constructed before aggregation.

Because χ_k is a positive constant common to every researcher in market k, it does not affect which allocation maximizes social value within that market. The researcher-specific functions $v_i(\cdot)$ continue to determine the ranking of within-market allocations. Thus, the within-market counterfactuals do not require an ex ante value of χ_k, while comparisons and aggregation of social value across markets do.

To be clear, this is a revealed-preference accounting exercise, not a direct experimental estimate of scientific externalities. It recovers the market-specific multipliers that rationalize observed funding under the maintained incidence rule, the marginal-return benchmark, and the planner's problem. If observed funding also reflects political considerations, historical commitments, institutional constraints, or omitted objectives, those forces will be absorbed into the implied χ_k. For this reason, we refer to these objects as implicit social multipliers.

B. Example Applications of Methodology

B.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:

$$\max_{l_i} \quad \alpha_i l_i - w l_i - c l_i^{\psi} \tag{B1}$$

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 $\boldsymbol{\mu}_i = \boldsymbol{\mu} = (w, c, \psi)$. The optimality condition of the problem is:

$$\alpha_i - w - c\psi l_i^{\psi-1} = 0 \tag{B2}$$

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

$$\alpha_i(\hat{l}_i, w, c, \psi) = w + c\psi \hat{l}_i^{\psi-1} \tag{B3}$$

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

$$l_i^* = \left(\frac{\alpha_i - w}{\psi c}\right)^{\frac{1}{\psi-1}} \tag{B4}$$

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 it hires at market conditions. Therefore, the problem is:

$$\max_{\tilde{l}_i} \quad \alpha_i(\tilde{l}_i + \Delta) - w\tilde{l}_i - c\tilde{l}_i^{\psi} - WTP_i \tag{B5}$$

and the solution:

$$\tilde{l}_i^* = \left(\frac{\alpha_i - w}{\psi c}\right)^{\frac{1}{\psi-1}} = l_i^* \tag{B6}$$

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

$$WTP_i = \alpha_i(\hat{l}_i, w, c, \psi)(\tilde{l}_i^* + \Delta) - w\tilde{l}_i^* - c\left(\tilde{l}_i^*\right)^{\psi} - \alpha_i(\hat{l}_i, w, c, \psi)l_i^* + wl_i^* + c\left(l_i^*\right)^{\psi} \tag{B7}$$

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

$$WTP_i = \alpha_i(\hat{l}_i, w, c, \psi)\Delta = \left(w + c\psi\hat{l}_i^{\psi-1}\right)\Delta \tag{B8}$$

which allows the identification of w, c, ψ and thus $\alpha_i(\hat{l}_i, w, c, \psi)$, given $\hat{l}_i$ and Δ. 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 α_i and thus on all estimands μ.

B.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:

$$\alpha_i(\hat{l}_i, \beta, w) = \left(\frac{w}{\beta}\right)\hat{l}_i^{1-\beta} \tag{B9}$$

and determines the optimal allocation as:

$$l_i^* = \left(\frac{\beta\alpha_i}{w}\right)^{\frac{1}{1-\beta}} \tag{B10}$$

Under the thought experiment, the optimization problem becomes:

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

which implies:

$$\tilde{l}_i^* = \left(\frac{\beta\alpha_i}{w}\right)^{\frac{1}{1-\beta}} - \Delta = l_i^* - \Delta \tag{B11}$$

Substituting this solution into the profit function and deriving WTP_i, we obtain:

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

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, β cannot be identified, which implies that also α_i is not pinned down.

B.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:

$$\max_{l_i} \quad \ln m_i + \left(\alpha_i d_i^\beta l_i^{1-\beta}\right)^\eta - \phi l_i^\psi \tag{B12}$$

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

$$\eta(1-\beta)\alpha_i^\eta d_i^{\eta\beta} l_i^{(1-\beta)\eta-1} = \phi\psi l_i^{\psi-1} \tag{B13}$$

Therefore:

$$\alpha_i(\hat{l}_i, d_i, \beta, \eta, \phi, \psi) = \left(\frac{\phi\psi\hat{l}_i^{\psi-1}}{\eta(1-\beta)d_i^{\eta\beta}\hat{l}_i^{\eta(1-\beta)-1}}\right)^{\frac{1}{\eta}} \tag{B14}$$

and

$$l_i^* = \left(\frac{\eta(1-\beta)\alpha_i^\eta d_i^{\eta\beta}}{\phi\psi}\right)^{\frac{1}{\psi-\eta(1-\beta)}} \tag{B15}$$

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

$$\max_{\tilde{l}_i} \ \ln(m_i - WTP_i) + \left(\alpha_i(d_i+\Delta)^{\beta}\tilde{l}_i^{1-\beta}\right)^{\eta} - \phi\tilde{l}_i^{\psi} \tag{B16}$$

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:

$$\begin{aligned}
&\ln(m_i - WTP_i) + \alpha_i^{\frac{\eta\psi}{\psi-\eta(1-\beta)}}\left(\frac{\eta(1-\beta)}{\phi\psi}\right)^{\frac{\eta(1-\beta)}{\psi-\eta(1-\beta)}}(d_i+\Delta)^{\frac{\beta\psi\eta}{\psi-\eta(1-\beta)}} - \phi\left(\frac{\eta(1-\beta)\alpha_i^{\eta}(d_i+\Delta)^{\eta\beta}}{\phi\psi}\right)^{\frac{\psi}{\psi-\eta(1-\beta)}} + \\
&-\ln m_i - \alpha_i^{\frac{\eta\psi}{\psi-\eta(1-\beta)}}\left(\frac{\eta(1-\beta)}{\phi\psi}\right)^{\frac{\eta(1-\beta)}{\psi-\eta(1-\beta)}} d_i^{\frac{\beta\psi\eta}{\psi-\eta(1-\beta)}} + \phi\left(\frac{\eta(1-\beta)\alpha_i^{\eta}d_i^{\eta\beta}}{\phi\psi}\right)^{\frac{\psi}{\psi-\eta(1-\beta)}} = 0
\end{aligned} \tag{B17}$$

and depends on all parameters.

C. Additional Survey Results

C.1. Additional Survey Results and Descriptive Statistics

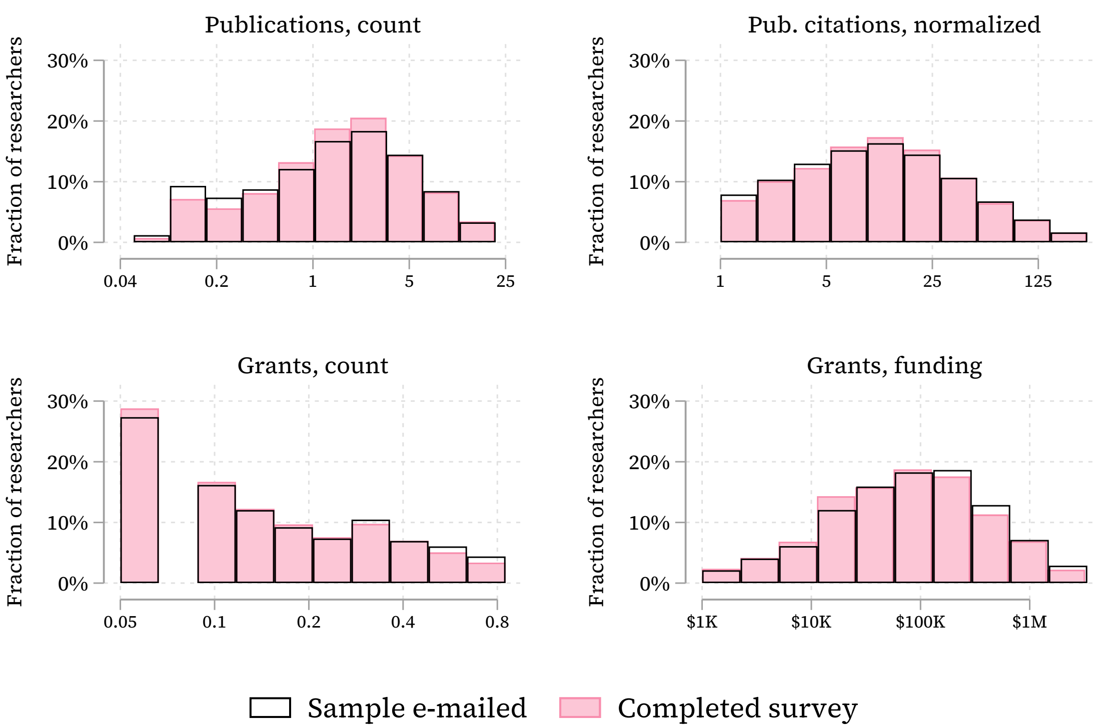
Table C1 reports regressions of survey completion on the randomized treatments. Figures C1 and C2 compare publication and funding measures for survey respondents and researchers e-mailed. Figure C3 compares self-reported and publicly reported salaries. Tables C2 and C3 report additional descriptive statistics, and Table C4 reports correlations among intended research outputs and audiences.

TABLE C1. Survey Completion per Randomized Treatments

	survey completion (1)
Either incentive	0.00297**
	(0.00117)
Both incentives	0.00797***
	(0.00141)
1 reminder	0.01122***
	(0.00114)
2 reminders	0.01879***
	(0.00119)
Constant	0.01971***
	(0.00107)
N obs.	131,672

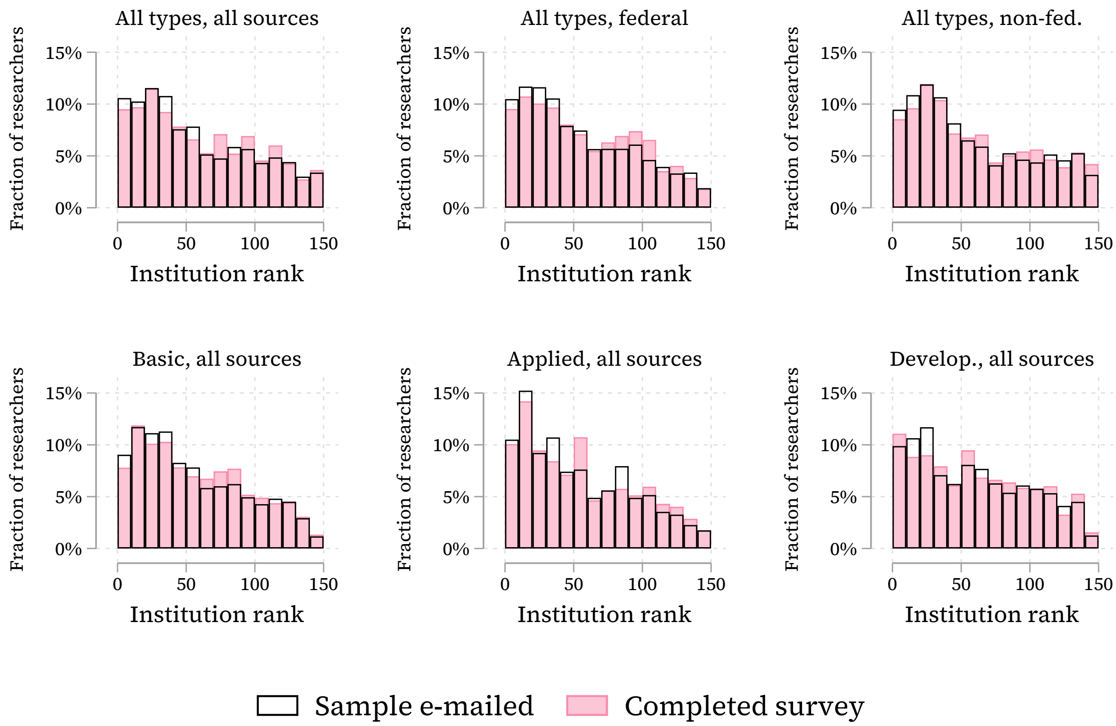
Note: Reports OLS estimates of survey completion on the randomized incentive and reminder assignments for all researchers e-mailed. Robust standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

FIGURE C1. Sample Representativeness per Publication and Grant Funding Measures



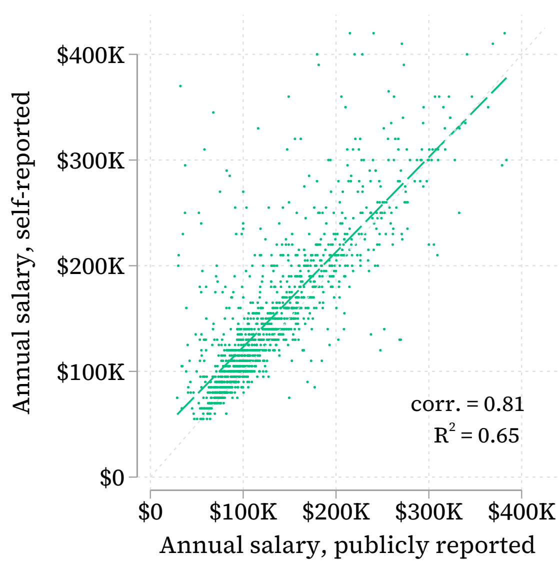
Note: Shows annual publication and grant outcomes for researchers e-mailed and survey respondents on logarithmic axes. Display excludes zero counts, normalized citations below one, annual grant funding below $1,000, and the top 1%. Replicated from Myers et al. (2023).

FIGURE C2. Sample Representativeness per Institutional Funding Measures



Note: Shows institutional R&D spending ranks for researchers e-mailed and survey respondents. Ranks above 150 are omitted. Replicated from Myers et al. (2023).

FIGURE C3. Correlation of Self- and Publicly-reported Annual Salaries



Note: Shows self-reported versus public annual salaries in the matched sample of 1,369 researchers, excluding values outside either measure's 1st–99th percentiles. The long-dashed line is the linear fit; the light dashed line denotes equality.

TABLE C2. 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
Field-normalized citations per year	23.07	49.10
Authors per publication	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: Reports summary statistics for the estimation sample. HERD variables describe institutions; Dimensions variables come from matched publication records.

TABLE C3. Summary Statistics of Other Variables (*continued*)

	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: Reports summary statistics for the estimation sample. N.R. denotes not reported.

TABLE C4. Pairwise Correlations of Subjective Output Measures

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

Note: Reports pairwise correlations of the intended output-type and audience measures summarized in Table 2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.2. Reducing Dimensions of Heterogeneity: k-means Clustering Results

Ideally, our model would allow for rich heterogeneity in researchers' preferences (i.e., the $m_i(\cdot)$, $b_i(\cdot)$, and $c_i(\cdot)$ functions). This would help ensure that (true) variation in preferences would not be mistaken for (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 $\mathbf{Z}_i$.[34] This motivates the following approach. We use k-means clustering to collapse the multi-dimensional heterogeneity 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 the standardized log Euclidean distance from one cluster centroid as T_i, which provides a smooth, continuous measure of heterogeneity. Thus, for the preference function parameter σ_i, we assume that $\sigma_i = \exp(\delta_{\sigma,0} + T_i \times \delta_\sigma)$. Figure C4 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.

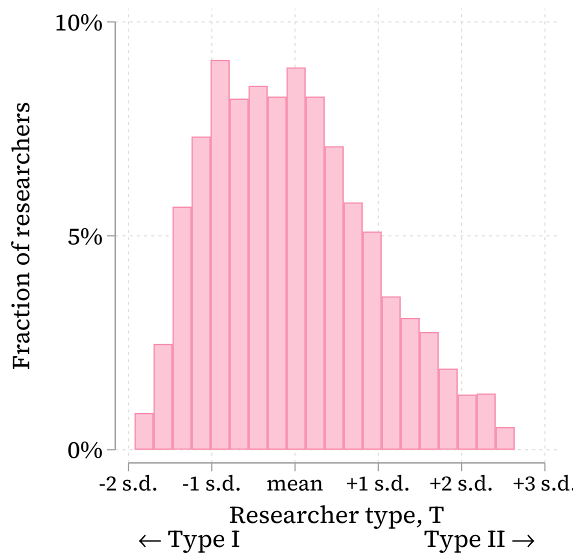
TABLE C5. Proportion of Response Variation Correlated with Model and Non-model Variables

	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: Reports R^2 from regressions of responses to each WTP experiment on model variables, with and without $\mathbf{Z}_i$.

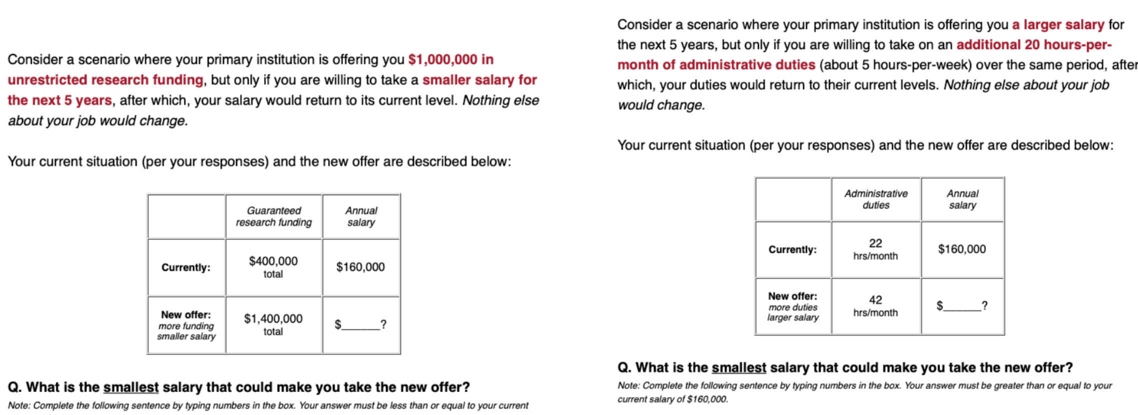
[34] We could allow the parameters that govern the preference functions to depend on deep parameters and these observables. For example, for the preference function parameter σ_i, we could assume that $\sigma_i = \exp(\delta_{\sigma,0} + \sum_z Z_{iz}\delta_{\sigma,z})$. However, this presents some practical estimation challenges (i.e., estimating approximately (35×3=) 105 additional parameters; the $\delta_{\sigma,z}$ parameters in the previous example).

FIGURE C4. Researcher Heterogeneity



Note: Shows the distribution of the researcher-type index T_i, the standardized log distance generated by the two-cluster k-means procedure. Values above the 99th percentile are omitted for display.

FIGURE C5. 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:

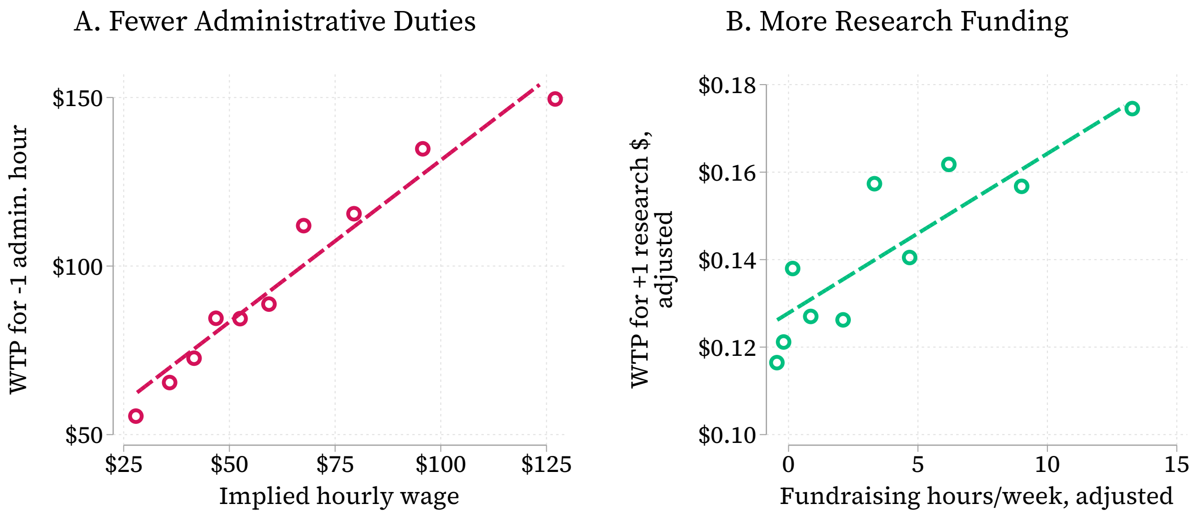
	Administrative duties	*Annual salary*
Currently:	22 hrs/month	$160,000
New offer: *more duties larger salary*	42 hrs/month	$______?

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 survey questions offering additional guaranteed funding or additional administrative duties in exchange for a salary adjustment.

FIGURE C6. Correlations of Implied Valuations



Note: Panel A shows WTP for one less administrative hour against the implied hourly wage. Panel B shows WTP per additional research dollar against weekly fundraising hours, residualizing both variables on the implied hourly wage and restoring their means. Points are binned means with linear fits; variables are trimmed at the 1st and 99th percentiles.

TABLE C6. 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: Reports regressions of standardized WTP-experiment responses on model variables, the survey-participation inverse Mills ratio (IMR), and WTP for home internet (Benchmark). Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

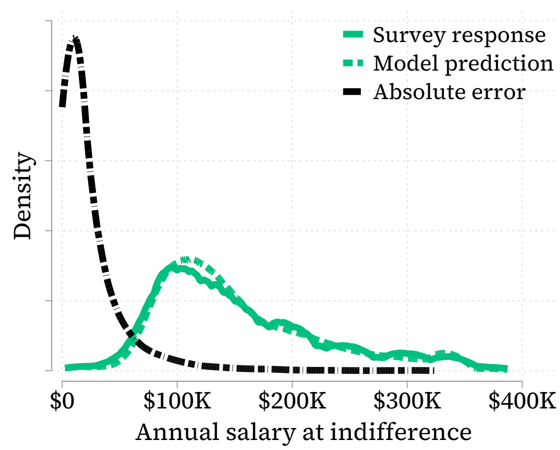
D. Additional Estimation and Counterfactual Results

Figure D1 compares survey responses with model predictions. Figure D2 shows the utility functions, and Figure D3 shows the upper tail of the productivity distribution.

Tables D1 and D2 report benchmark counterfactual results with $\chi_k = 1$ and under alternative specifications, respectively. Figure D4 compares observed and benchmark allocations, and Figure D5 shows marginal-product wedges. Table D3 reports market-level correlates of allocative efficiency, and Figure D6 shows allocative efficiency against the dispersion of guaranteed funding.

FIGURE D1. Model Fit

A. Answer, Prediction, and Error Distributions

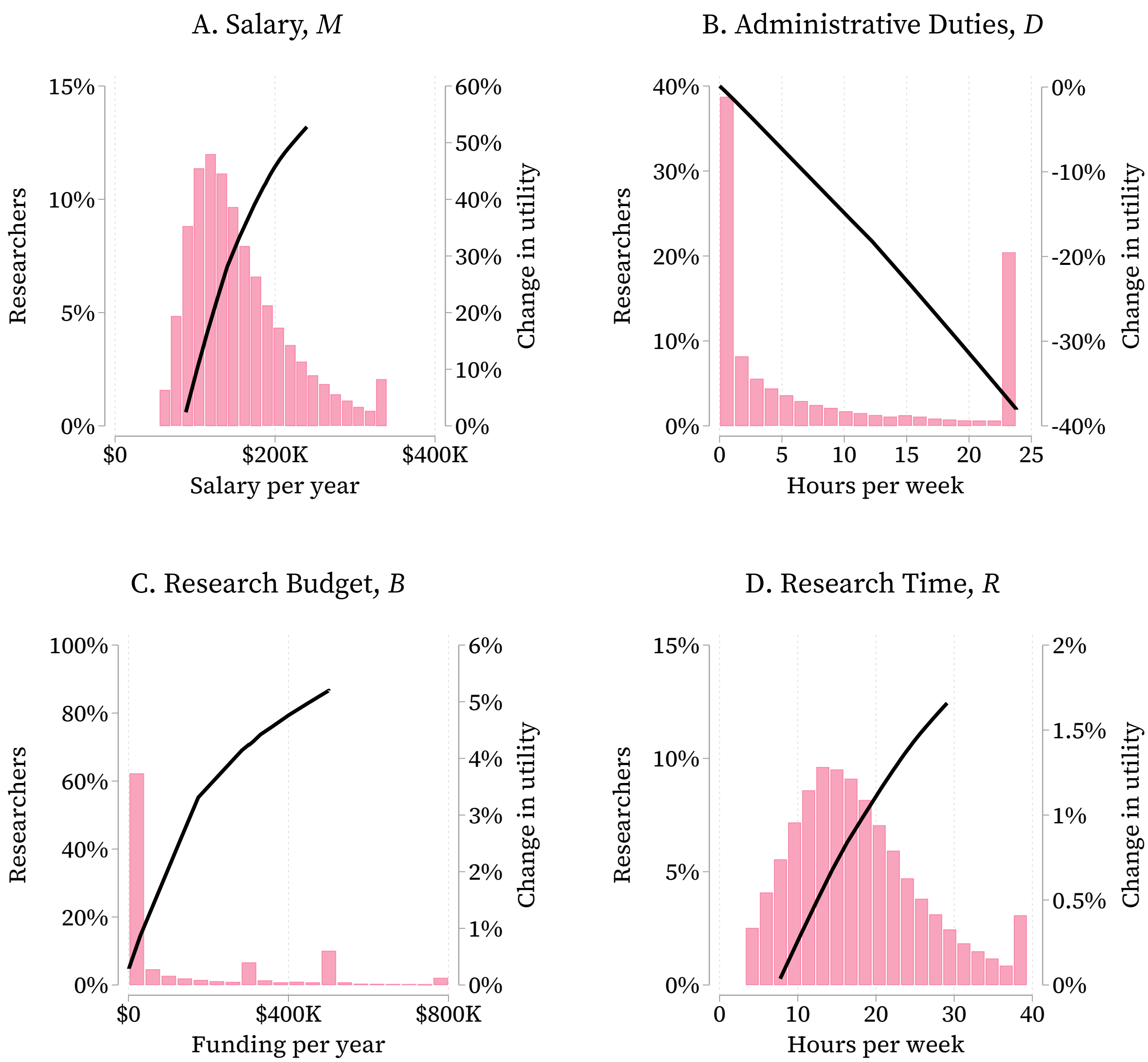


B. Distributions of Question-specific Errors in Percentages

	*p*10	*p*25	*p*50	*p*75	*p*90
Guaranteed funding +$250,000	0.11	0.88	6.18	15.54	29.58
Guaranteed funding +$1,000,000	0.63	3.44	12.25	27.04	53.70
Research time +20 hours/month	2.19	5.56	11.40	21.01	35.64
Research time reduced to zero	0.26	2.42	6.17	12.74	22.00

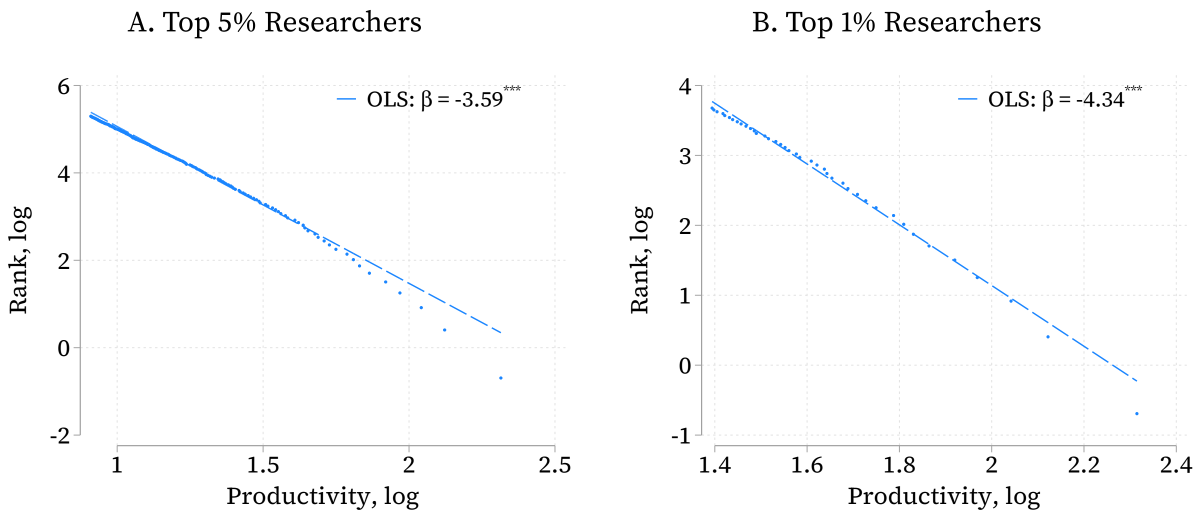
Note: Panel A shows stated and predicted annual salaries at indifference and absolute errors, pooled across the four experiments; values outside the 1st–99th percentiles are omitted for display. Panel B reports absolute-error percentiles as percentages of stated salaries.

FIGURE D2. Utility Function Visualization



Note: Shows utility changes (curves) and sample distributions (bars), holding other variables and all parameters at sample means. Changes are relative to the 10th-percentile value; curves end at the 90th percentile.

FIGURE D3. Power Laws in the Productivity Tail



Note: Shows log rank against log productivity, normalized by its science-wide mean, in the indicated upper tails. Lines fit $\log(\text{rank} - 1/2)$ on $\log(A_i)$ by OLS without trimming or winsorization. Stars summarize draw-specific regressions with heteroskedasticity-robust standard errors: $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, and $^{***}\ p < 0.01$.

TABLE D1. Observed versus Optimal Benchmark Comparison: $\chi_k = 1$

		Input(s) optimized		
		B, R (1)	R only (2)	B only (3)
Change in inputs & outcomes				
research budget: B_i	mean	~0%	~0%	~0%
	s.d.	–24.3%	~0%	–24.4%
research time: R_i	mean	~0%	~0%	~0%
	s.d.	+9.6%	+8.6%	~0%
fundraising time: F_i	mean	–100%	~0%	~0%
	s.d.	–100%	~0%	~0%
utility: $\mathcal{U}_i$	mean	+4.5%	~0%	+4.3%
	s.d.	+8.5%	~0%	+8.5%
output, private value: $b_i(Q_i)$	mean	+93.4%	–0.1%	+91.4%
	s.d.	+52.1%	~0%	+51.9%
output, social value: $\mathcal{V}_i(Q_i)$	mean	+93.2%	+0.6%	+90.3%
	s.d.	+51.7%	+0.1%	+51.4%
Private utility, compensating variation: Income				
% of own income		+5.4%	~0%	+5.1%
\$/year		+\$9,318	–\$13	+\$8,892
Social value, equivalent variation: Funding				
researcher-level, \$/year		+\$213,179	+\$1,214	+\$205,415
population-level, \$/year		+\$25.9B	+\$0.15B	+\$24.9B
Allocative efficiency				
initial allocation = 51.8%				
benchmark allocation		100%	52.1%	98.5%

Note: Reports the counterfactuals in Table 4 with $\chi_k = 1$ in every market; other definitions are unchanged. Compensating variation averages finite individual values. Population funding equivalents divide sample totals by the 3.3% response rate.

TABLE D2. Observed versus Optimal Benchmark Comparison: Alternative Specifications

		Alternative specifications								
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Change in inputs & outcomes										
research budget: B_i	mean	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%
	s.d.	–24.3%	–18.4%	–24.5%	–23.0%	–22.3%	–24.5%	–26.7%	–23.7%	–28.3%
research time: R_i	mean	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%
	s.d.	+9.6%	+4.6%	+7.0%	+14.7%	+27.5%	+9.0%	+7.1%	+9.2%	+11.7%
fundraising time: F_i	mean	–100%	–100%	–100%	–100%	–100%	–100%	–100%	–100%	–100%
	s.d.	–100%	–100%	–100%	–100%	–100%	–100%	–100%	–100%	–100%
utility: $\mathcal{U}_i$	mean	+4.5%	+1.0%	+3.8%	+4.2%	+2.8%	+4.3%	+3.7%	+4.0%	+2.0%
	s.d.	+8.5%	+0.6%	+5.5%	+9.9%	+8.4%	+7.1%	+6.1%	+6.7%	–0.4%
output, private value: $b_i(Q_i)$	mean	+93.4%	+93.1%	+90.4%	+103%	+124%	+93.8%	+120%	+97.8%	+211%
	s.d.	+52.1%	+91.6%	+53.5%	+53.3%	+60.3%	+52.3%	+56.1%	+53.1%	+74.7%
output, social value: $\mathcal{V}_i(Q_i)$	mean	+108%	+85.0%	+98.4%	+128%	+167%	+111%	+148%	+110%	+114%
	s.d.	+64.7%	+73.4%	+62.6%	+71.3%	+107%	+65.8%	+81.9%	+65.3%	+63.6%
Private utility, compensating variation: Income										
% of own income		+5.4%	+1.6%	+4.8%	+4.9%	+3.3%	+5.4%	+4.8%	+5.1%	+2.6%
$/year		+$9,318	+$2,880	+$8,355	+$8,328	+$5,508	+$9,338	+$8,134	+$8,868	+$3,970
Social value, equivalent variation: Funding										
researcher-level, $/year		+$266,253	+$477,450	+$291,100	+$245,864	+$223,952	+$276,743	+$433,253	+$286,950	+$770,262
population-level, $/year		+$32.3B	+$57.9B	+$35.3B	+$29.8B	+$27.2B	+$33.6B	+$52.6B	+$34.8B	+$93.4B
Allocative efficiency										
initial allocation		48.1%	54.0%	50.4%	43.9%	37.4%	47.5%	40.4%	47.5%	46.7%
benchmark allocation		100%	100%	100%	100%	100%	100%	100%	100%	100%
Specification choices										
Measurement-error share: ρ		0.4	0.0	0.2	0.6	0.8	0.4	0.4	0.4	0.4
Posterior shock correlation		0	—	0	0	0	0	Empirical	0	1
Measurement error in T_i		Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Winsorization percentiles		1–99	1–99	1–99	1–99	1–99	1–99	1–99	5–95	1–99

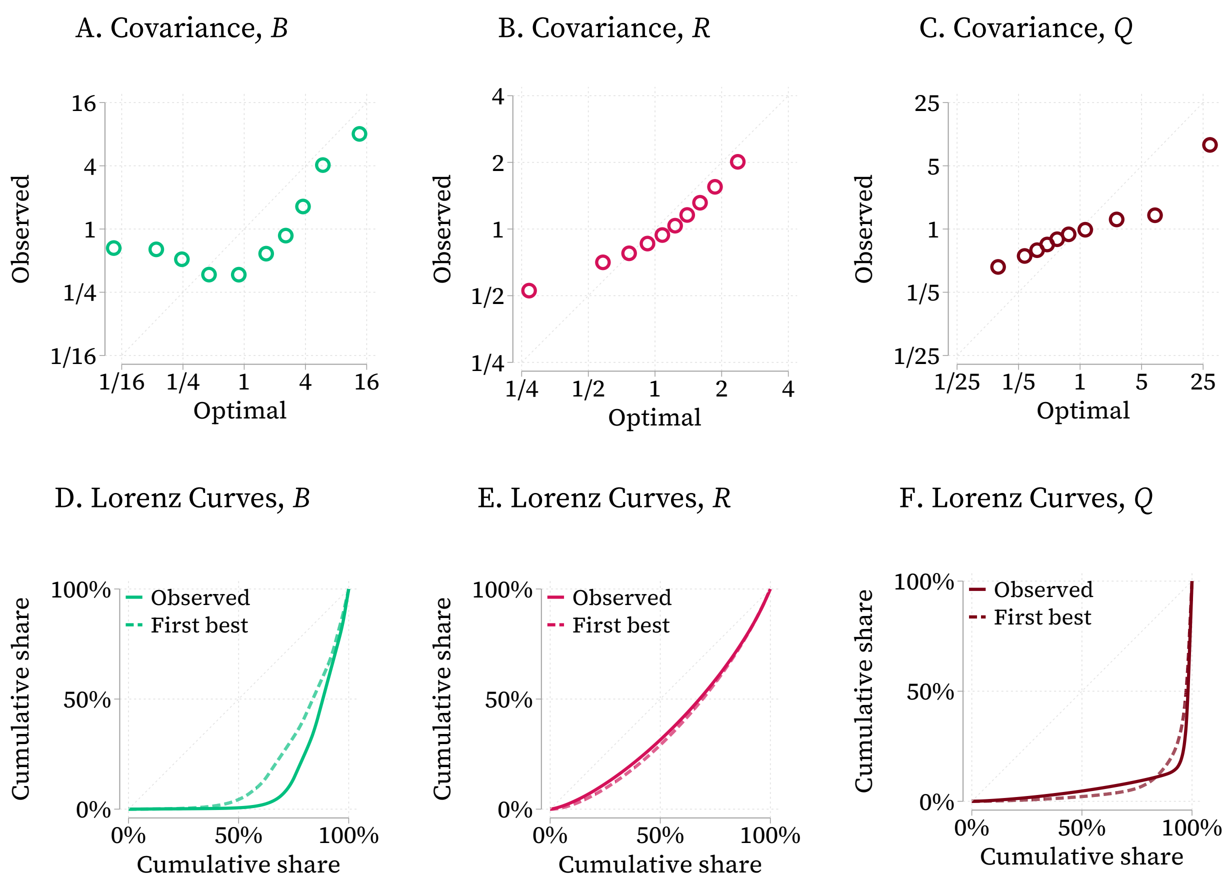
Note: Reports joint-input benchmark results under the specifications listed below; other definitions match Table 4. Correlations refer to posterior shocks ε_{ijb} across variables within a researcher; the empirical specification uses observed within-market correlations of transformed variables. Column 6 holds T_i and researcher-specific preference parameters at their original estimates. Winsorization applies to α_i, γ_i, and ϕ_i. The $\rho = 0$ specification is deterministic.

TABLE D3. Correlates of Market-Level Allocative Efficiency

candidate predictor	median std. coef.	MC-draw selection
Set 1: Variances: M, G, F, R, A, O, H		
log variance of guaranteed funding	0.567	72.0%
log variance of research time	0.240	16.0%
log variance of total work hours	−0.201	14.0%
log variance of administrative duties	−0.205	10.0%
log variance of fundraising time	−0.356	10.0%
log variance of salary	0.140	10.0%
log variance of teaching/other duties	0.267	8.0%
Set 2: Variances: $M, G, F, R, A, O, H, \phi, \gamma, \alpha$		
log variance of guaranteed funding	0.547	60.0%
log variance of funding intensity	0.386	40.0%
log variance of fundraising ability	−0.360	28.0%
log variance of total work hours	−0.178	18.0%
log variance of administrative duties	−0.193	16.0%
log variance of fundraising time	−0.263	12.0%
log variance of research productivity	0.201	12.0%
log variance of research time	0.300	12.0%
log variance of salary	−0.087	8.0%
log variance of teaching/other duties	0.137	8.0%

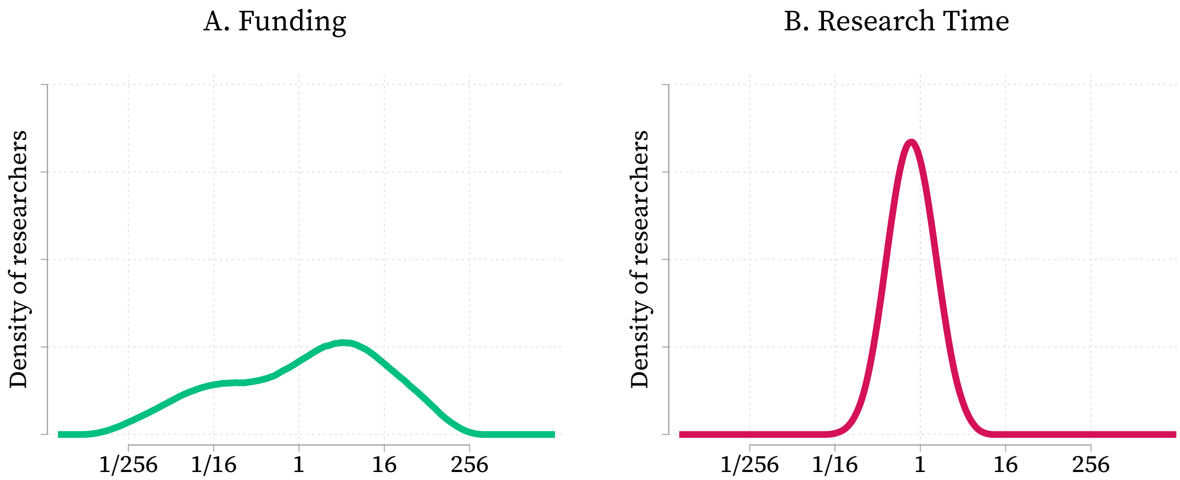
Note: Reports adaptive-lasso results with markets equally weighted. Coefficients summarize draws in which a predictor is selected; selection rates give the share of draws selecting it. These are descriptive selection summaries, not selection-adjusted inference.

FIGURE D4. Observed and optimal input and output levels



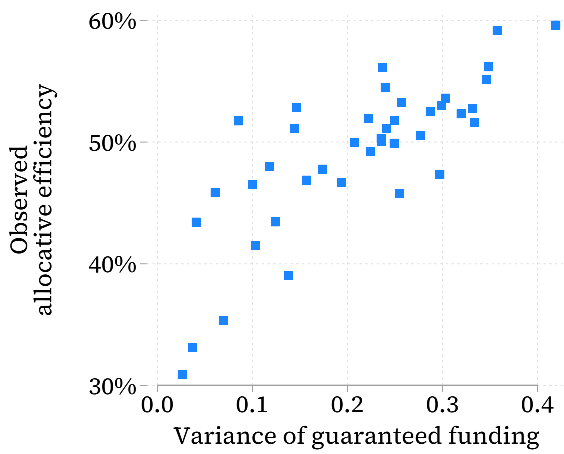
Note: Panels A–C show observed versus joint-benchmark funding, research time, and private output value in logs after removing market fixed effects. Points are ten binned means; gray lines denote equality. Panels D–F show the corresponding equally weighted Lorenz curves. Output in Panels C and F is $b_i(Q_i)$.

FIGURE D5. Marginal Product Wedges



Note: Panels A and B show log wedges for funding and research time, respectively, relative to the joint-benchmark market shadow prices. Zero denotes equality. Densities exclude the bottom and top 1% and use a common bandwidth.

FIGURE D6. Allocative Efficiency and the Dispersion of Guaranteed Funding



Note: Shows observed allocative efficiency against within-market guaranteed-funding variance, measured in squared millions of dollars. Points are markets; the correlation gives each market equal weight.