Task Mismatch and Salary Penalties: Evidence from the Biomedical PhD Labor Market*

Holden A. Diethorn† Gerald R. Marschke‡

October 17, 2024

Click here for the latest version.

Abstract

We employ a task-based framework to study the impact of postdoctoral training on the earnings of US-trained biomedical PhDs. Using longitudinal person-level data on both job tasks and salary, we find that a positive postdoc salary premium emerges when the difference between the tasks performed during training and future employment is low and a negative premium emerges when task mismatch is high. Early career biomedical doctorates working in industry perform a greater variety of tasks than those employed as postdoctoral researchers, leading to differences in task-specific human capital that explain the persistent negative returns to postdoctoral training in industry. (*JEL* J24, I26, J31, J44)

*This paper is based upon work supported by the National Science Foundation under grant numbers DGE-1661278 and NSCE-1918445. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. This paper uses restricted-use microdata provided by the NSF and accessed via the NORC Data Enclave. The use of NSF data does not imply NSF endorsement of the research, research methods, or conclusions contained in this paper. All results have been reviewed to ensure that no confidential information is disclosed. For their helpful comments and suggestions, we thank David Bogle, Pinka Chatterji, Stephanie Cheng, Navid Ghaffarzadegan, Laurel Haak, Chun-Yu Ho, Xuan Jiang, Bruce Weinberg, and especially Donna Ginther, as well as participants at the 2021 NBER Conference on Investments in Early Career Scientists, the 2022 Society of Labor Economists meetings, the 2022 NBER Summer Institute Science of Science Funding meeting, the 2022 European Association of Labor Economists meetings, and the 2023 Academy of Management meetings. All remaining errors are our own.

†Kenyon College. Email: diethorn1@kenyon.edu

‡SUNY Albany and NBER. Email: marschke@albany.edu

"[A] set of issues ... concerns the nature of the training young scientists receive, and the mismatch between that training and their career prospects. The focus of young scientists on securing an academic research faculty position can lead them to overlook opportunities ... in other areas, such as in start-up and established industries, foundations, and government. Significantly, these opportunities may require training experience different from those associated with traditional academic careers."

- National Academies of Sciences, Engineering, and Medicine (2018)

1 Introduction

A growing literature in labor economics treats skill as multidimensional, with workers performing tasks to produce output while acquiring human capital in the skills used by these tasks via learning-by-doing. This model has important implications for life-cycle earnings and worker mobility: First, workers whose skills acquired through prior employment more closely align with the skill requirements of their current job will be more productive—and thus earn higher wages—than workers whose previous roles involved skills less relevant to the current job. Second, job losers will tend to seek out new jobs that require similar tasks as their prior jobs to avoid the lower salary that accompanies *task mismatch*—an empirical measure of the degree to which the skills acquired in past employment differ from those used in current employment (Gathmann and Schönberg, 2010; Guvenen et al., 2020; Lise and Postel-Vinay, 2020).[1] In the context of worker training programs, this model implies that the pay impacts of such programs will rely upon how well the tasks performed during training align with those performed in future jobs.

In this paper, we employ a multidimensional human capital framework to examine the impact of postdoctoral training on the life-cycle earnings of biomedical doctorates.[2] Biomedical science is a significant source of knowledge creation in the United States, garnering both a higher share of federally-funded R&D and more citations from US granted patents than any other field of study (National Science Board, 2024*b*,*a*). Graduate students and postdoctoral researchers conduct the "great majority" of biomedical research, serving as key inputs to the labs of tenure-track research faculty (Alberts et al., 2014). Since 1980, the number of newly-graduated US-trained biomedical PhDs per cohort has doubled (Figure A.1), with over two-thirds of each cohort starting their careers in postdoc positions for an average of five years (Figure A.2). Postdoc training is all but

[1]See Sanders and Taber (2012) for a review of the task-specific human capital literature and the related literature on industry-specific and occupation-specific human capital. See Deming (2023) for a review focused on the implications of multidimensional human capital for the wage structure. We view task mismatch as an empirical analog for multidimensional skill mismatch in the same way that task-specific human capital is an empirical analog for multidimensional human capital in the absence of direct measures of worker skill (Sanders and Taber, 2012).

[2]We use the term "doctorate" as shorthand for "doctoral degree holder" and use "postdoc" to refer to a doctorate employed as a postdoctoral researcher ("postdoctorate") or in reference to postdoctoral training itself depending on the context. Doctorates with prior postdoctoral training are referred to as "postdoc-trained" or "ex-postdocs" and those without prior postdoctoral training are referred to as "nonpostdoc-trained" or "never-postdocs."

necessary for those aspiring to a tenure-track ("TT") research job, with ex-postdocs comprising 90% of both new tenure-track and newly-tenured biomedical research faculty (Figure A.3). While most biomedical PhDs say they are in a postdoc position to obtain a tenure-track research job (Sauermann and Roach, 2016), the majority are unlikely to obtain one, with a higher share of such doctorates working in industry (28%) than in academic TT research positions (23%) ten years after graduation. This growth in the number of biomedical postdocs paired with the declining chance of obtaining a TT research position has attracted concern from economists, biomedical researchers, and federal funders alike, buttressed by evidence that ex-postdocs earn significantly less than their nonpostdoc-trained counterparts within the same sector of employment (Kahn and Ginther, 2017).[3]

The widespread availability of postdoc positions paired with the low chances of obtaining an academic TT research position leads a sizable share of the early career biomedical PhD workforce to experience large, abrupt dislocations between the task-specific skills they acquire during postdoctoral training and the skills most relevant to their future employment. We find that three-fourths of all postdocs are engaged in basic research as their primary task during training regardless of their future sector of employment (i.e., academia, industry, or government/nonprofits); in contrast, post-postdoc employment shows considerable heterogeneity in work activities across sectors, with basic research being the most important task only in academic jobs. Managing people or projects, applied research, development, and professional services are all reported as more important activities in industry jobs, with only 10% of ex-postdocs in industry primarily engaged in basic research. Compared to other sectors, postdoc-trained biomedical doctorates transitioning to careers in industry face the highest degree of task mismatch which, as we show, leads to a sizable postdoc salary penalty in industry.

Our study uses data from the NSF's Survey of Doctorate Recipients (SDR)—the largest nationally-representative and longitudinal survey of US-trained doctorates—to analyze earnings disparities between biomedical PhDs with and without prior postdoctoral training. As a baseline analysis, we estimate a Mincerian earnings function where the key variable of interest is an indicator for whether a biomedical PhD previously worked as a postdoctoral researcher. In specifications where experience is measured as the number of years since PhD graduation, we find that ex-postdocs earn 12% less annually than nonpostdoc-trained doctorates within the same biomedical field and cohort. Consistent with prior research by Kahn and Ginther (2017), we find that this postdoc salary penalty persists for over 15 years post-PhD and is greatest in industry, where ex-postdocs earn 16% less annually than those without postdoctoral training. In contrast, we find no general

[3]See Freeman et al. (2001*a*,*b*) for early concerns regarding the tournament-style of biomedical research, and see Stephan (2012) for a recent and comprehensive view of the scientific research environment, including the role played by postdoctoral researchers. Members of the biomedical research community have expressed concern that postdoc training is narrowly tailored towards academic pursuits (National Academies of Sciences, Engineering, and Medicine, 2018; Hayter and Parker, 2019) and that the small chance of a young biomedical scientist achieving a career as an independent researcher in academia, even after a prolonged period of postdoctoral training, hampers their ability to attract the best and brightest students to the field (e.g., National Institutes of Health, 2012; Alberts et al., 2014).

postdoc salary penalty in academia, instead finding a 16% postdoc salary *premium* for those in non-TT academic research (e.g., staff scientist) positions.

We exploit novel information in the SDR on the evolution of each doctorate's job tasks to directly assess the degree to which differences in task-specific human capital can explain the sizable postdoc salary penalty in industry. To do this, we construct task tenure variables for each of the 14 work activities included in the SDR which count the number of years a given PhD has performed each task as part of their prior employment or postdoc training. Controlling for these individual-level measures of accumulated task-specific human capital reduces the estimated industry postdoc salary penalty by 66%, rendering it statistically insignificant at conventional levels. In contrast, we find no evidence that differences in unobserved abilities or preferences, current job tasks, seniority, or employer size explain the postdoc salary penalty in industry. Instead, bias-adjusted treatment effect estimates reveal that the within-field-by-cohort 16% postdoc salary penalty in industry is a lower bound of the true penalty under plausible assumptions regarding selection into postdoc training (Oster, 2019), consistent with prior findings by Sauermann and Roach (2016) that higher ability biomedical PhDs tend to sort into postdoc training.

Next, we construct a single worker-level measure of task mismatch that compares the tasks performed during postdoctoral training to those in current employment. This measure allows us to explore whether differences in the degree of task mismatch faced by ex-postdocs across sectors might explain the sizable variation in postdoc salary premia, which ranges from positive in academic non-TT research jobs to negative in industry employment. Similar to the occupation-level measure devised by Gathmann and Schönberg (2010), our person-level measure of task distance/mismatch ranges continuously between zero and one where a value of zero indicates a perfect alignment between current and prior tasks and a value of one indicates a perfect misalignment. Augmenting our baseline Mincerian specification with sector fixed effects and an interaction term between postdoc-trained status and task distance, we find that a positive postdoc salary premium emerges when the difference between a PhD's current job tasks and those performed as part of postdoc training is low and a negative postdoc premium (or penalty) emerges when task mismatch is high. When limiting our sample to biomedical PhDs working in industry, we find evidence that ex-postdocs would earn the same as those without postdoc training in the absence of task mismatch, complementing the results of our task tenure approach. Altogether, our findings suggest that task mismatch explains both 1) the sizable postdoc salary penalty in industry and 2) the variation in postdoc salary premia across sectors.

Beyond Stinebrickner, Stinebrickner, and Sullivan (2019) who track the tasks and salary of two cohorts of Berea College students over time, ours is the only other paper to our knowledge to track both the outcomes and tasks performed by the same workers longitudinally, rather than relying on external occupation- or job-level survey data to infer the history of tasks performed by each worker. Job tasks vary substantially within occupations (Autor and Handel, 2013; Deming and Kahn, 2018), with workers tending to match to jobs within a given occupation that minimize the distance between the tasks of the job and the skill of the worker. Thus, assigning tasks to a worker based on occupation or job title will tend to overstate the distance between the worker's skill set and the tasks actually performed. Workers within a given occupation may also perform different tasks over their career, such as taking on more managerial tasks while retaining the same occupational title, and so longitudinal measures of worker-level tasks carry additional appeal. Other task-based studies analyze workers across a broad spectrum of occupations and categorize tasks into coarse categories such as abstract, routine, and manual. We show that skill-job mismatch is important even within a narrowly-defined education category and within jobs that make nearly exclusive use of abstract/analytical ability. Our findings suggest the importance of distinguishing between different types of analytical tasks when considering task-specific human capital models of wage determination, especially among highly-educated workers.

Our study also contributes to the literature measuring the effect of postdoctoral training in biomedical science on future career outcomes (e.g., Jacob and Lefgren, 2011; Su, 2013; Kahn and Ginther, 2017; Heggeness et al., 2018; Hayter and Parker, 2019; Cheng, 2023). First, we show that postdoc salary premia range from significantly positive in academic non-TT research positions to negative in industry jobs. Second, we resolve an empirical puzzle in the literature: despite evidence that biomedical PhDs who pursue postdoctoral training are of higher ability at graduation (Sauermann and Roach, 2016), they typically earn *less* than their nonpostdoc-trained counterparts, even within the same sector of employment (Kahn and Ginther, 2017). The negative pecuniary returns to postdoctoral training had brought into question whether such training was consistent with a model of human capital, and our study provides affirmative evidence, showing that a task-specific human capital framework explains both the sizable postdoc salary penalty in industry and heterogeneity in the returns to postdoctoral training across sectors. While the classic view of postdoctoral training is as an apprenticeship for one's future vocation (Bravo and Olsen, 2007), our findings suggest that entering postdoc employment might more usefully be viewed as purchasing a lottery ticket whose value is the enhanced probability of securing a rare tenure-track academic research position (the lottery prize)[4] and where the price of the ticket includes two instances of foregone earnings: 1) the foregone earnings from alternative employment not undertaken during the postdoc and 2) lower future earnings when the skills acquired during postdoc training do not match the requirements of the job obtained thereafter. Nevertheless, ex-postdocs who leave academia after losing the tenure-track researcher lottery are likely to receive a consolation prize in the form of higher earnings associated with industry jobs.

[4]In a supplemental person-level jobs analysis described in Appendix C, we find that biomedical doctorates with postdoc training are 21 percentage points more likely to obtain a tenure-track research position and 27 percentage points more likely to work in any academic research job. Industry-employed ex-postdocs are also 12 percentage points more likely than industry-employed never-postdocs to work in a research job.

2 Data

The NSF's Survey of Doctorate Recipients (SDR) is a biennial survey of a representative sample of US-trained Science, Engineering, and Health doctorates under the age of 76 and contains information on each doctorate's salary, employment sector, and whether their current employment is as a postdoc, in addition to many demographic and economic variables. SDR respondents are sampled from the NSF's Survey of Earned Doctorates (SED)—an annual census of newly-graduated PhD recipients from US institutions that contains information on each PhD's field of study, demographic and educational background, and whether they intend to take a postdoc position after graduation. Once a PhD is sampled in the SDR, they are typically surveyed in all future survey waves until aging out due to terminal illness or reaching 76 years of age. Each person surveyed in the SDR receives a unique identifier allowing for linkage to the SED and longitudinal linkage across different waves of the SDR. To construct a longitudinal dataset of biomedical doctorates, we begin by appending all SDR waves between 1993 and 2017 and linking to the SED. We then limit our sample to biomedical doctorates obtaining a PhD between 1981 and 2007, who were first surveyed in the SDR prior to 2010, for whom we can identify whether they were employed as a postdoc in each year post-PhD, and who are observed at least once after their first six years post-PhD and at least once in a job after any postdoc training.[5] As in Kahn and Ginther (2017), we group observations into one of three employment sectors: academia, industry, or government and nonprofits; we also consider subsectors within academia and industry: academic TT research, academic non-TT research, academic nonresearch, industry research, and industry nonresearch.[6]

Table 1 breaks down the person-year observations in the analytical sample by sector and subsector of employment and whether biomedical doctorates within each sector are postdoc-trained. The first three columns assign each doctorate to the employment sector they occupy at ten years post-PhD while the last three columns assign each person-year observation to the actual sector of employment in each given year. Ex-postdocs make up the majority of biomedical PhDs working in each sector and subsector, reflecting the high prevalence of postdoctoral training in biomedical science. Academia employs the highest share of biomedical doctorates by ten years post-PhD (53%), followed by industry (31%) and government/nonprofits (16%). Within academia and industry, jobs

[5] Appendix B details a similar strategy to that of Kahn and Ginther (2017) in determining whether a PhD has ever been employed as a postdoc and for how many years. In 2010, the SDR began sampling US-trained PhDs who reside outside of the United States, whereas previous waves only included US-trained PhDs residing in the US after graduation. Due to this sampling change, the NSF recommends caution when analyzing and interpreting pre- and post-2010 trends. The SDR 2010 wave also introduced new sample members that had graduated as far back as 2001; we are not able to reliably identify whether these individuals were ever employed as postdocs given that they are first sampled in the SDR many years after graduation and were not part of the SDR 2006 wave where doctorates were asked whether they had previously worked as a postdoc. We therefore restrict the sample to those first appearing in the SDR data prior to 2010. We also limit the sample to individuals who appear in the SDR in 1993 at the earliest due to survey format changes in 1993 and sampling changes in 1991. For all regressions, we use SDR survey weights designed by NCSES to minimize the potential for nonresponse bias in SDR estimates.

[6] A "research job" is defined as a job where the primary activity is reported as either basic research, applied research, development, or design, following the NSF's categorization of "research and development" activities. Tenure-track workers include those on the tenure-track and those who have received tenure.

that require research as the primary work activity have greater shares of postdoc-trained workers. Interestingly, the share of ex-postdocs employed in industry at ten years post-PhD (28%) exceeds the share employed in TT research positions (23%). Differences in the person counts between the third and last columns reflect nontrivial rates of PhD worker mobility across sectors over time: for example, 1468 biomedical PhDs in our sample are employed in industry at ten years post-PhD, which reflects only 82% of the 1786 sample members who work in industry for at least one year post-PhD; similarly, only 58% of sample members who ever work in academic non-TT research do so at ten years post-PhD, indicating strong mobility in and out of this subsector.

Table 2 shows considerable differences in the distribution of earnings across sectors, with PhDs in academia earning the least and those in industry earning the most at every percentile. Table 1 shows that ex-postdocs are more likely than never-postdocs to work in academia at ten years post-PhD (57% vs. 44%) and less likely to work in industry (28% vs. 37%) and government/nonprofits (15% vs. 19%). Given sector pay differences and the greater likelihood that ex-postdocs work in academia at ten years post-PhD, it is not surprising that they typically earn less than their nonpostdoc-trained counterparts. However, as we show in the next section, significant disparities in pay exist even between postdoc-trained and nonpostdoc-trained biomedical doctorates working within the same sector, with a particularly large penalty associated with those working in industry. Table A.1 shows that, compared to biomedical doctorates without postdoc training, ex-postdocs are more likely to be foreign born and temporary residents at time of PhD; to have been funded by research assistantships as graduate students; to have finished the PhD more quickly and at a younger age; and are less likely to have been married and to have children living at home at the time of graduation.

3 Baseline Estimation of Postdoc Salary Premia

3.1 Baseline Empirical Specification

We estimate postdoc salary premia by employment sector (i.e., academia, industry, or government/nonprofit) using the following person-level Mincer equation as our preferred specification:

$$\log(earn_{ifsct}) = \boldsymbol{X_i\beta} + \theta Postdoc_i + \boldsymbol{Exp_{it}\lambda} + \boldsymbol{\gamma_{fc}} + \boldsymbol{\gamma_s} + \boldsymbol{\gamma_t} + \varepsilon_{ifsct}, \tag{1}$$

where $earn_{ifsct}$ is the year t inflation-adjusted annualized salary of doctorate i who graduated with a PhD in field f from university s in year c, $\boldsymbol{X_i}$ is a vector of pre-determined individual-level controls, $Postdoc_i$ is an indicator variable for if doctorate i is postdoc-trained, $\boldsymbol{Exp_{it}}$ is a vector containing a quartic polynomial in experience, $\boldsymbol{\gamma_{fc}}$ are field-by-cohort fixed effects, $\boldsymbol{\gamma_s}$ are PhD institution (i.e., *alma mater*) fixed effects, $\boldsymbol{\gamma_t}$ are normalized year fixed effects, and ε_{ifsct} is an idiosyncratic error term.[7] We cluster standard errors at the person-level as each biomedical doctorate may appear more than once in the estimation sample and the regressor of interest, $Postdoc_i$, is fixed for each doctorate. For each person-year observation, we use the sample weight associated with the SDR wave in which the observation appears and include controls for race, sex, age at PhD, number of years spent in graduate school, source of PhD study financial support, whether completed a professional degree in conjunction with the PhD, marital status at time of graduation, whether had a child at home at time of graduation, foreign-born status, and whether the individual was a temporary resident. Field-by-cohort fixed effects ($\boldsymbol{\gamma_{fc}}$) control for field-cohort specific shocks that could influence both a doctorate's decision to pursue a postdoc and future career outcomes: these include the number of PhDs and postdocs in one's own field of study, the level of NIH funding allocated to one's field, and field-specific research agendas and breakthroughs (e.g., the Human Genome Project, the use of MRI and fMRI), as well as technological and methodological progress that open up both new avenues for research and new economic opportunities (e.g., advances in semiconductor technology leading to the proliferation of AI and machine learning methods in biomedical research). PhD institution fixed effects ($\boldsymbol{\gamma_s}$) capture the impact of PhD institution—and any unobserved characteristics of the doctorate that led to his or her acceptance into that institution and that may be correlated with the decision to do a postdoc—on future career outcomes.

3.2 Baseline Results

Table 3 reports postdoc salary premia estimates by employment sector (Panel I) and by academic and industry subsectors defined by whether the job is research-focused or a tenure-track/tenured position (Panel II). Following Kahn and Ginther (2017), columns (1) and (2) report estimates for specifications where a PhD's employment sector is that observed at 10 years post-PhD and experience is defined as years since PhD graduation.[8] Reassuringly, our results follow Kahn and Ginther (2017) in showing a negative postdoc salary premium in most sectors and finding no evidence of a postdoc salary premium in any sector or subsector.

Next, we explicitly estimate the effect of postdoc training on *future* salary by including only those observations associated with years after any postdoc training has ended and where we define a PhD's employment sector as that which they occupy in each given year, rather than at a single point in time.[9] Column (4) shows that postdocs who go on to work in academia after their training

[7] Salary is adjusted using the CPI-U. We follow Murphy and Welch (1990) and Lemieux (2006) by including a quartic polynomial in experience. To address the issue of collinearity between cohort fixed effects, year fixed effects, and experience, we normalize year fixed effects as in equation 2.95 of Deaton (1997) which, as discussed in Aguiar and Hurst (2013) and Lagakos et al. (2018), results in salary growth over time being attributed to experience and cohort effects and restricts year fixed effects to capturing only cyclical fluctuations in salary.

[8] Since the SDR is biennial, a doctorate may not be observed at exactly 10 years post-PhD. For such PhDs, we follow Kahn and Ginther (2017) by imputing their after-postdoc employment sector using 11 years, 12 years, and then 9 years post-PhD.

[9] Given that the average postdoc duration is between five and six years in biomedical science (Figure A.2), we also drop observations corresponding to a doctorate's first six years post-PhD regardless of postdoc status so that ex-postdoc and never-postdoc observations are comparable. There are very few after-postdoc observations for ex-

do not, in general, face a postdoc penalty on future salary; in contrast, ex-postdocs working in industry or government/nonprofits face a 15.8% and 10.6% penalty, respectively. Unlike earlier results, we now find a 15.9% postdoc salary *premium* in academic non-TT research jobs rather than a postdoc salary penalty. Only 58% of sample members who ever work in academic non-TT research do so at ten years post-PhD (Table 1), indicating strong mobility in and out of this subsector over time that likely biases results in columns (1) and (2) but not column (4). For this reason, column (4) represents our preferred specification. Predicted salary profiles generated from a dynamic version of specification (4) which includes interactions between postdoc-trained status and the quartic polynomial in experience show that the postdoc salary penalty in industry—particularly in nonresearch jobs—persists for over 15 years post-PhD, consistent with long-lasting scarring effects of task mismatch (Figure 1.A and Figure 2.A). In academia, postdoc-trained biomedical doctorates appear to earn less than their nonpostdoc-trained counterparts early in their careers; however, by 15 years post-PhD, the steeper salary profiles of ex-postdocs lead to a positive postdoc salary premium in academia.

Specifications (1) through (4) treat postdoc training like other forms of employment that add to labor market experience. Given that postdoc training and industry employment emphasize different skills, we would expect ex-postdocs and never-postdocs working in industry to differ in accrued task-specific human capital, thus leading to within-cohort differences in pay. However, we would not expect the same magnitude of disparity in the *entry-level* salaries of ex-postdocs and never-postdocs if differences in task-specific human capital accrual are the main driver of column (4) results. To gain insight into the disparities in pay between ex-postdocs and never-postdocs since year of entry, we estimate specifications (5) and (6) where experience is defined as years of *nonpostdoc* employment which, in a Mincerian framework, treats postdoc training like schooling.[10] Figure 1.B plots predicted salary profiles from dynamic versions of specification (6), showing that starting salaries of biomedical PhDs with and without postdoc training are largely similar across sectors except in academia where ex-postdocs persistently earn more. In particular, ex-postdocs earn a 23.2% entry-level premium relative to biomedical PhDs without postdoc training in non-TT research jobs (Table 3.II.B), suggesting that postdoc training increases the productivity of non-TT researchers due to the similar tasks emphasized in both types of jobs.

3.3 Paying to Do Science in Industry and Selection on Unobservables

Stern (2004) finds that industry-employed biomedical postdocs pay a negative compensating differential to participate in science and Sauermann and Roach (2014) find that PhD candidates inclined to pursue industrial R&D careers vary in the price they are willing to pay to be allowed to publish.

postdocs at the lowest levels of experience, and so failing to drop the first six years post-PhD for all doctorates would lead to salary-experience profiles that at the lowest levels of experience would be driven almost exclusively by never-postdoc observations.

[10]These specifications are otherwise identical to specifications (3) and (4) respectively, except that we now retain nonpostdoc observations during the first six years post-PhD.

Among biomedical doctorates who work in industry, we find that those with postdoc training are 12 percentage points more likely to be employed in a research-focused job (Appendix C). Thus, one might wonder whether the 15.8% postdoc salary penalty in industry reported in column (4) is driven by ex-postdocs in biomedicine "paying to do science" rather than due to differences in task-specific human capital. While we cannot explicitly rule out this explanation, the results in Stern (2004) and Sauermann and Roach (2014) cast doubt that paying to do science can explain cross-sectional differences in pay between doctorates working in industry. First, Stern notes that his finding a negative compensating differential to participate in science depends critically on the inclusion of individual fixed effects made possible by the structure of his survey data which include the observation of multiple job offers for each postdoc at a given point in time. Omitting individual fixed effects results in a positive and statistically insignificant coefficient estimate. Second, Sauermann and Roach note that scientists who are willing to pay the highest price to be able to publish in industry are scientists of perceived higher ability and from top-tier institutions, and so tend to be more expensive to hire even if publishing is allowed.

Since postdoc-trained status is clearly endogenous, our estimates of postdoc salary premia may not represent the true causal effects of postdoc training on earnings. One could imagine that biomedical doctorates with greater industry-relevant skills or ability at the time of graduation are more likely to forgo postdoc training, and so it is this greater endowment of industry-relevant skills—rather than differences in skill accumulation after PhD graduation—that generates the 15.8% postdoc salary penalty in industry. To examine the plausibility of this explanation, we estimate bias-adjusted treatment effects (Oster, 2019) which use selection on observables to model the plausible degree and direction of selection on unobservables and report detailed results in Appendix D. In all sectors—with the exception of academic non-TT research jobs—we find that bias-adjusted treatment effects are less than the corresponding estimates reported in Table 3 column (4). Thus, rather than being driven by unobserved abilities or preferences, we instead find that the 15.8% industry postdoc penalty is a likely lower bound of the true penalty, implying that biomedical doctorates who pursue postdoc training tend to be of higher ability than those who forgo it.

4 Evidence for a Task-Specific Human Capital Explanation

A novel feature of the SDR is that it provides individual-level, longitudinal measures of tasks that are directly linked to the salary of the job for which these tasks are performed.[11] For the task-based

[11] Primary and secondary tasks reflect the two tasks that each doctorate reports as occupying the most and second-most time during the typical work week. The list of activities/tasks that respondents may select are as follows: 1) Accounting, finance, contracts, 2) Applied research— study directed toward gaining scientific knowledge to meet a recognized need, 3) Basic research—study directed toward gaining scientific knowledge primarily for its own sake, 4) Computer programming—including systems or applications development, 5) Development—using knowledge gained from research for the production of materials, devices, 6) Design—of equipment, processes, structures, models, 7) Human resources—including recruiting, personnel development, training, 8) Managing or supervising people/projects, 9) Production, operations, maintenance—including chip production, operating lab equipment, 10) Quality or productiv-

analyses that follow, we limit our analytical sample to those doctorates whose tasks we observe at least two times during the first six years of post-PhD employment including any postdoc training.[12]

4.1 Task Differences Between Postdoc Training and Other Employment

Table 4 shows substantial differences between postdocs and nonpostdocs in the tasks reported as primary work activities at least once in the first six years post-PhD. Approximately three-fourths of all postdocs report basic research as their primary work activity within the first six years after graduation regardless of their subsequent sector of employment; in contrast, only 6%–15% of nonpostdocs are primarily engaged in basic research depending on employment sector. Applied research, professional services, development, and management are much more likely to be reported as the primary work activity of nonpostdocs as opposed to postdocs early in their career, especially in industry. Beyond primary and secondary work activities, SDR respondents also report any tasks that comprise at least 10% of work time. Figure 3 shows that biomedical postdocs are much more likely to be engaged in basic research and slightly more likely to be engaged in applied research during their postdoc employment compared to nonpostdoc-trained biomedical doctorates working in industry during their first six years post-PhD. Meanwhile, postdocs are considerably less likely to be engaged in development, design, management, and professional services during their postdoc training, giving nonpostdocs working in industry a better opportunity to develop skills in these tasks early in their career.

One empirical implication of a task-specific model of human capital is that, other things equal, a worker who moves to a new job that requires substantially different tasks than their previous job will typically be paid less than a worker whose previous job had more similar task requirements (Gathmann and Schönberg, 2010). Thus, in Figure 3 we show the percentage of postdoc-trained and nonpostdoc-trained biomedical doctorates working in industry who, in any year *after* the first six years post-PhD, report working in a job where they spend at least 10% of their time engaged in each given task. We then take the difference between the share performing each task during and after the first six years post-PhD and report this percentage-point difference as the "Task Change." Figure 3 shows that ex-postdocs in industry experience larger changes in each task relative to nonpostdoc-trained biomedical doctorates (except for computer applications). Figure 4 shows task changes by employment sector for ex-postdocs and never-postdocs in biomedical science. Comparing the magnitudes in the left and right panels, we see that ex-postdocs tend to face greater task changes than biomedical PhDs without postdoc training regardless of employment sector. Task changes facing postdocs are larger in industry than in the other sectors for 10 of 14 tasks, while

ity management, 11) Sales, purchasing, marketing—including customer service and public relations, 12) Professional services—including health care, counseling, financial services, legal services, 13) Teaching, and 14) Other.

[12] Since SDR 1993 is the first survey wave of our analytical sample, this restriction implicitly excludes doctorates graduating prior to 1989, as these doctorates would only be observed at most once in their first six years post-PhD in the SDR. See Table A.2 for observation and person counts of biomedical doctorates in each employment sector for this "task regression sample."

task changes for nonpostdoc-trained workers in industry are only the largest for 2 of 14 tasks, underscoring the greater degree of mismatch faced by postdocs transitioning to industry—both in comparison to other sectors and to nonpostdoc-trained doctorates working in industry.

4.2 Task-Specific Human Capital and the Industry Postdoc Salary Penalty

To fix ideas, we sketch a conceptual framework based on the task-based wage determination model of Autor and Handel (2013), made dynamic by relating task-based human capital to past task experiences. Details can be found in Appendix E. This framework implies that log wages of worker type i employed in sector k in year t (w_{ikt}) can be written as a function of time-invariant abilities in each task j (H_i^j) and task-specific human capital accrued as part of learning-by-doing in prior employment or training (H_{it}^j):

$$w_{ikt} = \alpha_k + \sum_j \lambda_k^j H_i^j + \sum_j \lambda_k^j H_{it}^j, \tag{2}$$

where α_k is a sector-specific intercept and λ_k^j represents the sector-specific productivity of task j. Suppose that the task-specific human capital accrued by a worker depends on their prior sector of employment such that θ_k^j units of human capital in task j are acquired for each year employed in sector k. Consider the case where there are two sectors of employment—academia (A) and industry (I)—and two workers—a postdoc-trained (p) and a nonpostdoc-trained (n) doctorate. To simplify, suppose that both workers graduated τ_t years ago (i.e., are of the same cohort), are presently employed in industry, and that worker p has spent all prior years post-PhD as a postdoc working in academia while worker n has worked in industry since graduation. Defining $\theta_\Delta^j \equiv \theta_A^j - \theta_I^j$ as the difference between annual task j specific human capital accrual in academia and industry and $H_\Delta^j \equiv H_p^j - H_n^j$ as the difference in endowed task j specific ability of worker p and n, log wages can be written as:

$$w_{iIt} = \alpha_I + \sum_j \lambda_I^j H_n^j + \sum_j \lambda_I^j \theta_I^j \tau_t + 1[i=p] * \left\{ \sum_j \lambda_I^j H_\Delta^j + \sum_j \lambda_I^j \theta_\Delta^j \tau_t \right\}, \tag{3}$$

where $1[i = p] = 1$ if worker i is type p and $1[i = p] = 0$ if worker is type n. Equation (3) implies that industry wage differences between postdoc-trained and nonpostdoc-trained workers of the same cohort are due to differences in endowed abilities (H_Δ^j) and between-sector differences in the rate of task j specific human capital accumulated as part of prior employment (θ_Δ^j).

Given this task-specific human capital framework, we would expect the estimated postdoc salary penalty in industry to shrink towards zero when controlling for task-specific work experience ("task tenure" or "task history")—the number of years spent performing each task as part of prior employment. To directly test this using biennial SDR data, we measure each doctorate's task tenure by calculating the percentage of previous jobs where the PhD reports performing the given task and multiplying this value by the number of years since PhD minus one. We calculate three sets of task history variables: one set each for the number of years where a given work activity was performed as the primary job task, the primary or secondary job task, and for at least 10% of work time. We estimate specifications using different combinations of these three sets of task tenure variables, where each set of task history variables is comprised of 14 variables (i.e., one for each task). While SDR data do not include the exact proportion of time spent on each task, including primary or secondary task histories alongside the history of tasks performed for at least 10% of work time allows us to account in some way for differences in the time allocated to different tasks.

Table 5 Panel A shows that the estimated industry postdoc salary penalty is substantially reduced when controlling for task-specific work experience: when controlling for both the history of primary tasks performed and those tasks performed for at least 10% of work time in column (6), we obtain a statistically insignificant estimate of the industry postdoc salary penalty that is roughly one-third the magnitude of the baseline estimate reported in column (1).[13] Results are consistent across specifications, with point estimates declining by between 43% and 66% across all specifications. Current tasks and task histories are likely correlated, and so we consider specifications that exclude task tenure and instead control for the tasks performed as part of current employment. Table 5 Panel B shows that adding controls for current job tasks—rather than the history of job tasks—does little to change the estimated industry postdoc salary penalty; in fact, including controls for current job tasks increases (albeit slightly) the estimated postdoc salary penalty across all specifications. Together, Panel A and B results suggest that differential task-specific human capital accumulation—and not differential preferences to perform current job tasks—explains the postdoc salary penalty in industry. Results in Table 5 Panel C where we control for both current tasks and task-specific work experience show that differences in accrued task-specific human capital are important for explaining the pay disparity between biomedical PhDs with and without postdoc training, even among those who work in jobs requiring the performance of the same set of tasks.

For insight into the importance of different types of accumulated task-specific human capital to industry salary determination, Table 6 reports coefficient estimates for the primary task history controls included in column (2) of Table 5, where each estimate measures the effect of spending an additional year engaged in the given primary task relative to if one spent an additional year primarily engaged in applied research. We find that substituting a year where one could primarily be engaged in applied research with a year where one is primarily engaged in basic research results in an approximate 4% decline in salary. Assuming constant returns to task tenure, this implies that a biomedical doctorate primarily engaged in basic research for five years (e.g., as a postdoc) stands to lose 20% of their industry earnings capacity compared to the case where they obtain an applied research-focused job in industry upon graduation. Substituting a year of primary focus in applied research for a focus on teaching, sales/marketing, or accounting are all associated with declines

[13] Specification (1) in Table 5 is identical to specification (4) in Table 3 Panel I.C, but is estimated on the set of doctorates in our analytical sample for whom we observe job tasks at least two times in the first six years post-PhD.

in salary, while substituting for a year focused on managing people or projects is associated with an increase in salary. Experience in development, design, production, and professional services all yield similar returns to a year spent in applied research.

4.3 Task Mismatch and Between-Sector Variation in Postdoc Salary Premia

In Table 3 Column (4), we detect sizable variation in postdoc salary premia across sectors that range from significantly positive in academic non-TT research jobs to significantly negative in industry employment. To test whether task mismatch explains this between-sector variation in postdoc salary premia, we construct two alternative person-level measures of the distance between the tasks performed as part of current employment and those performed during the first six years of post-PhD employment. These two measures differ only in the degree to which they weigh primary, secondary, and other tasks on which doctorates report spending at least 10 percent of their work time ("≥10% tasks"). Both measures of task distance are constructed to range continuously from zero to one, where a value of zero corresponds to a doctorate whose proportion of time spent on each task during their first six years post-PhD exactly matches the percentage of time spent on each task in current employment and a value of one corresponds to a PhD whose current tasks are completely different from the tasks performed in their first six years post-PhD.

Our first measure—which we term "extensive (margin) task mismatch"—treats all ≥10% tasks performed equally and is constructed as follows: 1) In each year, we calculate the percentage of time spent on each task under the assumption that the doctorate spends equal time on each task reported to take at least 10% of work time; 2) We average the amount of time allocated to each task across the first six years of post-PhD employment (including any postdoc training) to ensure that tasks performed across more years within the first six years post-PhD are allocated a greater amount of time; 3) We calculate the distance between tasks performed in the first six years post-PhD versus those performed as part of current employment using the same angular separation measure as in Gathmann and Schönberg (2010) subtracted from one.[14] Our second measure—which we term "intensive (margin) task mismatch"—differs only in that, in step 1) above, it weighs more heavily those ≥10% tasks that are reported to be the primary or secondary work activities rather than applying equal weight, thus taking into account the intensive margin of task performance in each year. Letting n be the number of ≥10% tasks reported by the PhD in a given year, we allocate a $\frac{3}{n+1}$ share of work time to be divided 60%-40% between the primary and secondary tasks, respectively. The remaining work time is then evenly allocated among the other ≥10% tasks. In

[14] Letting θ_{i1}^{j} and θ_{it}^{j} denote the share of time biomedical doctorate i spends performing task j as part of employment in their first six years post-PhD and as part of current employment, respectively, the degree of task mismatch (or task distance) between the two measures is calculated as

$$1 - \frac{\sum_{j=1}^{J}\left(\theta_{i1}^{j} * \theta_{it}^{j}\right)}{\left\{\left[\sum_{j=1}^{J}(\theta_{i1}^{j})^2\right] * \left[\sum_{j=1}^{J}(\theta_{it}^{j})^2\right]\right\}^{1/2}}.$$

all instances, the time allocated to the primary and secondary tasks exceeds the time allocated to any other $\geq$10% tasks.[15]

To test whether task mismatch explains the difference in postdoc salary premia across sectors, we estimate a version of specification (4) in Table 3 Panel I.A where we add 1) sector fixed effects to control for average salary differences between academia, industry, and government/nonprofits and 2) an interaction between the postdoc indicator and our measure of task mismatch/distance. The coefficient estimate corresponding to the postdoc-trained indicator then gives the postdoc salary premium associated with a postdoc whose tasks performed in the first six years post-PhD—including any postdoc training—perfectly matches the tasks performed as part of current employment (i.e., when there is no task mismatch). The coefficient on the interaction between the postdoc indicator and task mismatch indicates the degree to which task mismatch drives heterogeneity in the returns to postdoc training across sectors; if task mismatch drives this heterogeneity, we would expect the effect of postdoc training in the absence of task mismatch to be positive and the interaction between postdoc training and task mismatch to be negative.

Column (1) of the "All Sectors" results in Table 7 shows that postdoc-trained doctorates tend to earn 8.2% less than their nonpostdoc-trained counterparts after controlling for average differences in salary across sectors. Allowing for the impact of postdoc training to vary by the degree of extensive task mismatch in column (2), we find that ex-postdocs who perform a set of tasks identical to those performed during postdoc training earn 7.6% more than their nonpostdoc-trained counterparts whereas those performing a completely different set of tasks earn 34.8% less. We obtain similar results in column ($2'$) when using intensive task mismatch to measure task distance. Altogether, these results indicate that a positive postdoc salary premium emerges when the difference between a doctorate's current job tasks and those performed as part of postdoc training is low and a negative postdoc premium (or penalty) emerges when task mismatch is high. Column (2) and ($2'$) results for academia are qualitatively similar—which we would expect given the positive postdoc salary premium in academic non-TT jobs and the null or negative results in other academic subsectors that we find in Table 3 column (4). For industry, we find that ex-postdocs are paid no differently than their nonpostdoc-trained counterparts in the absence of task mismatch, and that the mismatch between current tasks and those performed during the first six years post-PhD, *in particular*, drives the postdoc salary penalty in industry. Next, in columns (3) and ($3'$) we report estimates from specifications that include controls for primary and $\geq$10% tasks performed as part of current employment, allowing us to assess the degree to which task mismatch leads ex-postdocs working in the same types of jobs within the same sector to earn less than biomedical PhDs without postdoc training. In all panels, we find that task mismatch maintains a strong negative effect on salary.[16]

[15] For instance, if a person reports five $\geq$10% tasks, then the primary task is allocated 30% of work time, the secondary task is allocated 20% of work time, and the three other tasks are each allocated 16.67% of work time.

[16] Results are robust to including controls for each PhD's self-reported preferences for various job amenities including the "intellectual challenge", "degree of independence", and "contribution to society" of the job.

In Table 7, we show that the distance between tasks performed in the first six years of post-PhD employment and those performed as part of current employment explain the postdoc salary penalty in industry. For those with postdoc training, employment during the first six years may include both years employed as a postdoc and years employed in one's first post-postdoc job. To verify that the postdoc salary penalty in industry is largely due to a mismatch between the tasks associated with postdoc training not matching those performed as part of future industry employment, in Table A.5 we re-estimate the regression specifications included in Table 7 using task mismatch measures which, for ex-postdocs, only considers mismatch between tasks performed during postdoc training and current job tasks.[17] The results in Table A.5 largely mirror those in Table 7, suggesting that mismatch between current jobs tasks and postdoc tasks—and not other employment during the first six years post-PhD—largely explains the postdoc salary penalty in industry.

Lastly, we add task distance itself—rather than just its interaction with the postdoc indicator—to the specification, the results of which appear in Table 8. In the specifications included in Table 8, the coefficient on task distance shows the effect of task mismatch on nonpostdoc-trained biomedical doctorates while the coefficient on the interaction between task distance and the postdoc indicator tells us whether the effect of task distance varies by postdoc-trained status. The coefficient on the postdoc indicator then gives the residual difference in salary between postdoc-trained and nonpostdoc-trained biomedical doctorates, holding task distance constant. Whereas Table 7 shows the degree to which task mismatch explains postdoc salary premia, these regressions are useful for estimating the degree to which task mismatch leads to salary penalties in general. Across all specifications, we are unable to reject the hypothesis that postdoc-trained and nonpostdoc-trained biomedical doctorates are paid similarly once controlling for task mismatch. Results for academia suggest that task mismatch leads to salary penalties regardless of the postdoc-trained status of the biomedical doctorate. In contrast, results for industry and government/nonprofits indicate that ex-postdocs, in particular, experience salary penalties as a result of task mismatch. While this could suggest that postdoc training reduces one's match for industry or government/nonprofit jobs in ways that are correlated but not completely captured by our measure of task mismatch, it might also reflect a higher-chance that nonpostdoc-trained biomedical doctorates in these sectors experience a "positive" form of task mismatch that accompanies promotions into higher-paying jobs entailing different tasks than those performed during the first six years post-PhD.

5 Discussion of Results

We find that postdoc-trained biomedical doctorates in industry earn 16% less than their non-postdoc-trained counterparts annually, controlling for individual-level characteristics, a quartic

[17] There are some ex-postdocs whose training is completed prior to their first observation in the SDR—these individuals are dropped from the sample since we do not observe the tasks they perform during postdoc training. Table A.6 shows that the mean level of mismatch among postdoc-trained biomedical doctorates increases when explicitly measuring the mismatch between current job tasks and tasks performed during postdoc training.

polynomial in experience, PhD university (i.e., *alma mater*) fixed effects, and field-by-cohort fixed effects. We find no evidence that this industry postdoc salary penalty is explained by selection on unobservable ability or preferences at time of PhD, differential sorting into firms and occupations, seniority, or differences in current job tasks.[18] Instead, we find evidence consistent with a task-specific human capital model of wage determination where differences in salary between postdoc-trained and nonpostdoc-trained biomedical doctorates are the result of differences in the history of tasks performed as part of previous employment. First, we find substantial differences in the tasks emphasized as part postdoc training and industry employment: postdoc training is primarily focused on basic research, with little focus on development, design, management, professional services, and other tasks that are valued in industry employment. Second, inclusion of task tenure measures as mediating controls substantially reduces the magnitude of the estimated postdoc salary penalty in industry (by 66%), rendering the estimate statistically insignificant. We find that those who participate in postdoc training the longest suffer the largest postdoc salary penalties in industry, which is expected if differences in salary are largely due to postdocs deferring accrual of industry-relevant task-specific human capital while employed as a postdoc.[19] Evidence suggests that the estimated 16% postdoc salary penalty in industry is a lower-bound for the true impact of postdoc training on industry salary, which suggests that biomedical doctorates who first pursue postdoc training prior to industry employment are of greater ability at time of PhD graduation compared to their nonpostdoc-trained counterparts.

More generally, we find that a multidimensional human capital model does well to explain the differences in estimated effects of postdoc training across sectors, which range from a positive postdoc premium of 16% in academic non-TT research to a postdoc salary penalty (or negative premium) of 16% in industry. Evidence suggests that the positive postdoc premium for academic non-tenure-track research is not the result of selection on unobserved ability at time of PhD graduation; while biomedical doctorates who pursue postdoc training tend to be of greater ability than those that do not (including those that later sort into industry positions), this appears not to be the case for biomedical doctorates who sort into academic non-tenure-track research jobs. Bias-adjusted estimates of the postdoc salary premium exceed 16%, implying that selection bias attenuates rather than exaggerates the impact of postdoc training on salary in academic non-TT research positions. This result is consistent with postdoc training being an effective way to augment skills relevant to academic research, in addition to any other roles it may serve including as a holding tank for job market candidates (Andalib, Ghaffarzadegan, and Larson, 2018) or as a signal of pre-existing ability. Indeed, postdoc-trained biomedical doctorates are better able to secure research-focused positions, both in academia and industry, compared to their nonpostdoc-trained counterparts, with estimates robust to selection on unobserved ability at time of PhD graduation.

[18]See Appendix F for results pertaining to occupational sorting, sorting into firms, and seniority.

[19]See Appendix G for baseline salary regressions that allow for differences in postdoc salary premia based on the length of postdoc training.

A task-specific human capital explanation is consistent with the views of those within the biomedical community who argue both that postdoc training is specialized academic training and that initiatives to broaden the types of training and career preparation available to postdocs may be desirable given the growing number who go on to work outside academia. Towards this end, in 2013 the NIH initiated the Broadening Experiences in Scientific Training (BEST) grant program aimed at supporting institutions seeking to provide biomedical graduate students and postdocs with career development opportunities to facilitate an easier transition to nonacademic jobs (Meyers et al., 2016; Lenzi et al., 2020). Programs designed to expose biomedical postdocs to other career paths, such as research funding that requires postdocs to participate in a two to three month industrial internship, may better prepare biomedical postdocs for jobs in industry. While our results suggest that increasing the exposure of biomedical postdocs to skills valued in industry would be effective at lessening the postdoc salary penalty in industry, it is unclear whether a more "industry-oriented" postdoc would be of net benefit to society without a rigorous welfare analysis that lies outside the scope of the present paper. Such an analysis would need to quantify the impact of postdocs on academic output compared to industry output and quantify the benefit of each type of output to society while factoring in knowledge spillovers.

Whether pursuing postdoctoral training is worth it for a biomedical PhD depends on many factors, but having the requisite information to make such a decision is important. The findings in this paper suggest that postdoc training increases one's chances of obtaining an academic tenure-track research position by about 20% and an industry research position by 12%. Back-of-the-envelope calculations suggest that, on average, those headed to a career in industry after their postdoc will be paid $478,000 (undiscounted 2018 USD)—or $366,000 discounted annually at 3%—less in their first 20 years post-PhD than their nonpostdoc-trained counterparts who entered industry after graduation. However, a postdoc who lands a job in industry will still be paid $489,000 ($339,000 discounted) more than the average postdoc who subsequently works in academia.[20] Combining results, the average postdoc-trained biomedical doctorate who works in academia earns $967,000 (undiscounted) less than the average nonpostdoc-trained biomedical doctorate working in industry in their first 20 years post-PhD, for an average of $48,350 less per year.

6 Conclusion

This paper contributes to the growing empirical literature in labor economics that views tasks as fundamental to human capital acquisition and wage determination. We show that the structure of biomedical science leads a sizable share of its early career workforce to experience high degrees of skill mismatch when transitioning from postdoctoral training into industry employment. This leads

[20] An ex-postdoc in industry is paid $213,000 ($152,000 discounted) more in their first 20 years post-PhD than an ex-postdoc working as a tenure-track researcher. All calculations are based on combining the predicted salary profiles for the first six years post-PhD in Figure A.4 with the predicted salary profiles for subsequent years given in Figure 1.

ex-postdocs to earn 16% less annually than their nonpostdoc-trained counterparts, underscoring an important trade-off between acquiring skills via postdoc training versus learning-by-doing in industry. More generally, we find that the degree of mismatch between the tasks performed during postdoc training and those performed in future employment explains the variation in postdoc salary premia across sectors which ranges from positive when mismatch is low to negative when mismatch is high. Our findings resolve an empirical puzzle in the literature on the labor market outcomes of early career scientists, showing that task mismatch explains the tendency of postdoc-trained biomedical doctorates to earn persistently less than their nonpostdoc-trained counterparts employed in the same sector (Kahn and Ginther, 2017) despite evidence of higher ability at time of PhD graduation (Sauermann and Roach, 2016).

We show that distinguishing between different types of analytical tasks is valuable to explaining wage determination among highly-educated workers, with our results speaking specifically to the wage effects of "analytical task mismatch." Analytical task mismatch may be important for explaining between-sector heterogeneity in the returns to other forms of higher education of a given type (e.g., college degrees in a particular field), and so data collection on the type and intensity of analytical tasks performed by students during their academic training and subsequent career could be fruitful in examining within-field wage dispersion. Recent studies have examined the earnings consequences of worker-job mismatch by evaluating whether a worker's job is typically held by others with the same college major (Nordin, Persson, and Roof, 2010) or by relying on subjective measures of mismatch (Robst, 2007). To our knowledge, ours is the first work to measure education-job mismatch by comparing individual-level variation in the task content of academic training with the task requirements of future employment. An advantage of an individual-level task-based measure of mismatch is that it allows for differences in mismatch among workers with training in the same field of study and employed in the same sector or occupation without relying on subjective measures.

Lastly, our work further demonstrates the value of longitudinal data with information on both individual-level tasks and labor market outcomes. Most studies in the task literature rely on external occupation-level data to infer tasks performed by individual workers. Previous research shows that workers sharing the same occupational code are likely to be paid differently when each performs a different set of tasks (Autor and Handel, 2013). Among workers performing similar sets of tasks in their jobs, we find that the history of tasks performed as part of prior employment or training is an important factor in wage determination, showcasing the value of longitudinal measures of worker-level tasks. Previous work by Stinebrickner, Stinebrickner, and Sullivan (2019) utilized person-level and longitudinal measures of tasks to study wage variation among two cohorts of students graduating from a Kentucky liberal arts college. Our study of a nationally-representative sample of biomedical doctorates graduating over the course of two decades helps to demonstrate the broader relevance of task tenure to wage determination.

References

Aguiar, Mark and Erik Hurst. 2013. "Deconstructing Life Cycle Expenditure." *Journal of Political Economy*, 121(3): 437–492.

Alberts, Bruce, Marc W. Kirschner, Shirley Tilghman, and Harold Varmus. 2014. "Rescuing US Biomedical Research from Its Systemic Flaws." *Proceedings of the National Academy of Sciences of USA*, 111(16): 5773–5777.

Altonji, J. G., P. Arcidiacono, and A. Maurel. 2016. "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects." In *Handbook of the Economics of Education*. Vol. 5, 305–396. Elsevier.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113(1): 151–184.

Andalib, Maryam A., Navid Ghaffarzadegan, and Richard C. Larson. 2018. "The Postdoc Queue: A Labour Force in Waiting." *Systems Research and Behavioral Science*, 35: 675–686.

Autor, David H. and Michael J. Handel. 2013. "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages." *Journal of Labor Economics*, 31(2): S59–S96.

Barth, Erling. 1997. "Firm-Specific Seniority and Wages." *Journal of Labor Economics*, 15(3): 495–506.

Bravo, Norka Ruiz and Kathie L. Olsen. 2007. "Letter to National Postdoctoral Association." National Science Foundation.

Cheng, Stephanie D. 2023. "What's Another Year? The Lengthening Training and Career Paths of Scientists." *PLoS ONE*, 18(5).

Davis, James C., Holden A. Diethorn, Gerald R. Marschke, and Andrew J. Wang. 2021. "A Machine Learning Approach to Merging Datasets of Disparate Size: An Application to Identify Postdocs in LEHD Data." Manuscript.

Deaton, Angus. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* John Hopkins University Press (for World Bank).

Deming, David and Lisa B. Kahn. 2018. "Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals." *Journal of Labor Economics*, 36(S1): S337–S369.

Deming, David J. 2023. "Multidimensional Human Capital and the Wage Structure." In *Handbook of Economics of Education*. Vol. 7, 469–504. Elsevier.

Freeman, Richard B., Eric Weinstein, Elizabeth Marincola, Janet Rosenbaum, and Frank Solomon. 2001*a*. "Careers and Rewards in Bio Sciences: the Disconnect Between Scientific Progress and Career Progression." American Society for Cell Biology.

Freeman, Richard, Eric Weinstein, Elizabeth Marincola, Janet Rosenbaum, and Frank Solomon. 2001*b*. "Competition and Careers in Biosciences." *Science*, 294: 2293–2294.

Gathmann, Christina and Uta Schönberg. 2010. "How General is Human Capital? A Task-Based Approach." *Journal of Labor Economics*, 28(1): 1–49.

Guvenen, Fatih, Burhanettin Kuruscu, Satoshi Tanaka, and David Wiczer. 2020. "Multidimensional Skill Mismatch." *AEJ: Macroeconomics*, 12(1): 210–244.

Hayter, Christopher S. and Marla A. Parker. 2019. "Factors that Influence the Transition of University Postdocs to Non-Academic Scientific Careers: An Exploratory Study." *Research Policy*, 48: 556–570.

Heggeness, Misty L., Donna K. Ginther, I. Maria Larenas, and Frances D. Carter-Johnson. 2018. "The Impact of Postdoctoral Fellowships on a Future Independent Career in Federally Funded Biomedical Research." NBER Working Paper 24508.

Jacob, Brian A. and Lars Lefgren. 2011. "The Impact of NIH Postdoctoral Training Grants on Scientific Productivity." *Research Policy*, 40: 864–874.

Kahn, Shulamit and Donna K. Ginther. 2017. "The Impact of Postdoctoral Training on Early Careers in Biomedicine." *Nature Biotechnology*, 35(1): 90–94.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth across Countries." *Journal of Political Economy*, 126(2): 797–849.

Lemieux, Thomas. 2006. "The 'Mincer Equation' Thirty Years After 'Schooling, Experience, and Earnings'." In *Jacob Mincer: A Pioneer of Modern Labor Economics*. 127–145. Springer.

Lenzi, Rebecca N., Stephen J. Korn, Madeleine Wallace, Nancy L. Desmond, and Patricia A. Labosky. 2020. "The NIH "BEST" Programs: Institutional Programs, the Program Evaluation, and Early Data." *The FASEB Journal*, 34: 3570–3582.

Lise, Jeremy and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation." *American Economic Review*, 110(8): 2328–2376.

Meyers, Frederick J., Ambika Mathur, Cynthia N. Fuhrmann, Theresa C. O'Brien, Inge Wefes, Patricia A. Labosky, D'Anne S. Duncan, Avery August, Andrew Feig, Kathleen L. Gould, Michael J. Friedlander, Chris B. Schaffer, Audra Van Wart, and

Roger Chalkley. 2016. "The Origin and Implementation of the Broadening Experiences in Scientific Training Programs: An NIH Common Fund Initiative." *The FASEB Journal*, 30(2): 507–514.

Murphy, Kevin M. and Finis Welch. 1990. "Empirical Age-Earnings Profiles." *Journal of Labor Economics*, 8(2): 202–229.

National Academies of Sciences, Engineering, and Medicine. 2018. *The Next Generation of Biomedical and Behavioral Sciences Researchers: Breaking Through.* National Academy Press.

National Institutes of Health. 2012. "Biomedical Research Workforce Working Group Report." National Institutes of Health.

National Science Board. 2024*a*. "Invention, Knowledge Transfer, and Innovation. Science and Engineering Indicators 2024." National Science Foundation, NSB-2024-1.

National Science Board. 2024*b*. "Science and Engineering Indicators 2024: The State of U.S. Science and Engineering." National Science Foundation, NSB-2024-3.

Nordin, Martin, Inga Persson, and Dan-Olaf Roof. 2010. "Education–occupation mismatch: Is there an income penalty?" *Economics of Education Review*, 29: 1047–1059.

Oster, Emily. 2019. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business & Economic Statistics*, 37(2): 187–204.

Robst, John. 2007. "Education and job match: The relatedness of college major and work." *Economics of Education Review*, 26(4): 397–407.

Sanders, Carl and Christopher Taber. 2012. "Life-Cycle Wage Growth and Heterogeneous Human Capital." *Annual Review of Economics*, 4: 399–425.

Sauermann, Henry and Michael Roach. 2014. "Not All Scientists Pay to be Scientists: PhDs' Preferences for Publishing in Industrial Employment." *Research Policy*, 43: 32–47.

Sauermann, Henry and Michael Roach. 2016. "Why Pursue the Postdoc Path?" *Science*, 352(6286): 663–664.

Stephan, Paula. 2012. *How Economics Shapes Science.* Harvard University Press.

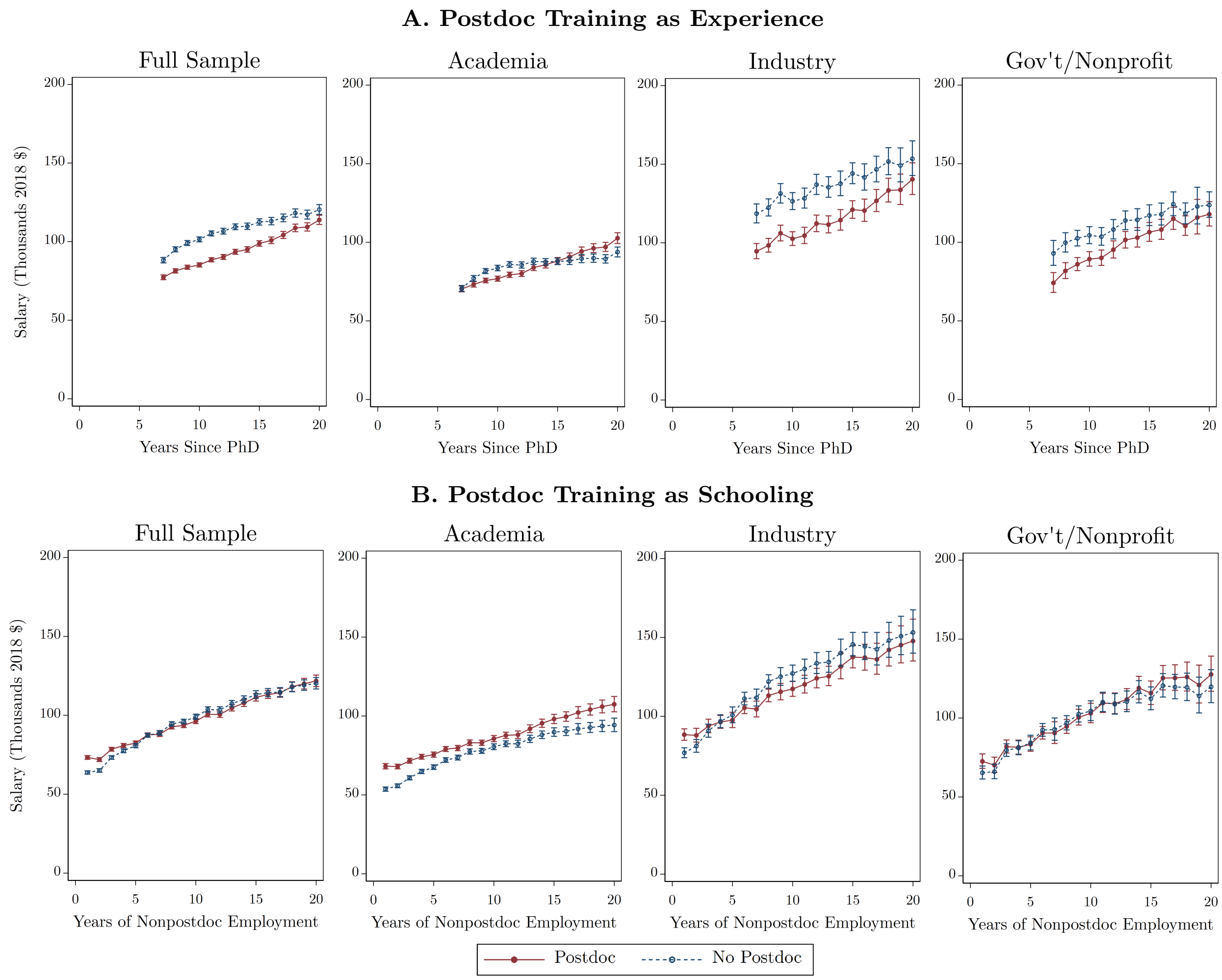
Stern, Scott. 2004. "Do Scientists Pay to be Scientists?" *Management Science*, 50(6): 835–853.

Stinebrickner, Ralph, Todd Stinebrickner, and Paul Sullivan. 2019. "Job Tasks, Time Allocation, and Wages." *Journal of Labor Economics*, 37(2): 399–433.

Su, Xuhong. 2013. "The Impacts of Postdoctoral Training on Scientists' Academic Employment." *The Journal of Higher Education*, 84(2): 239–265.

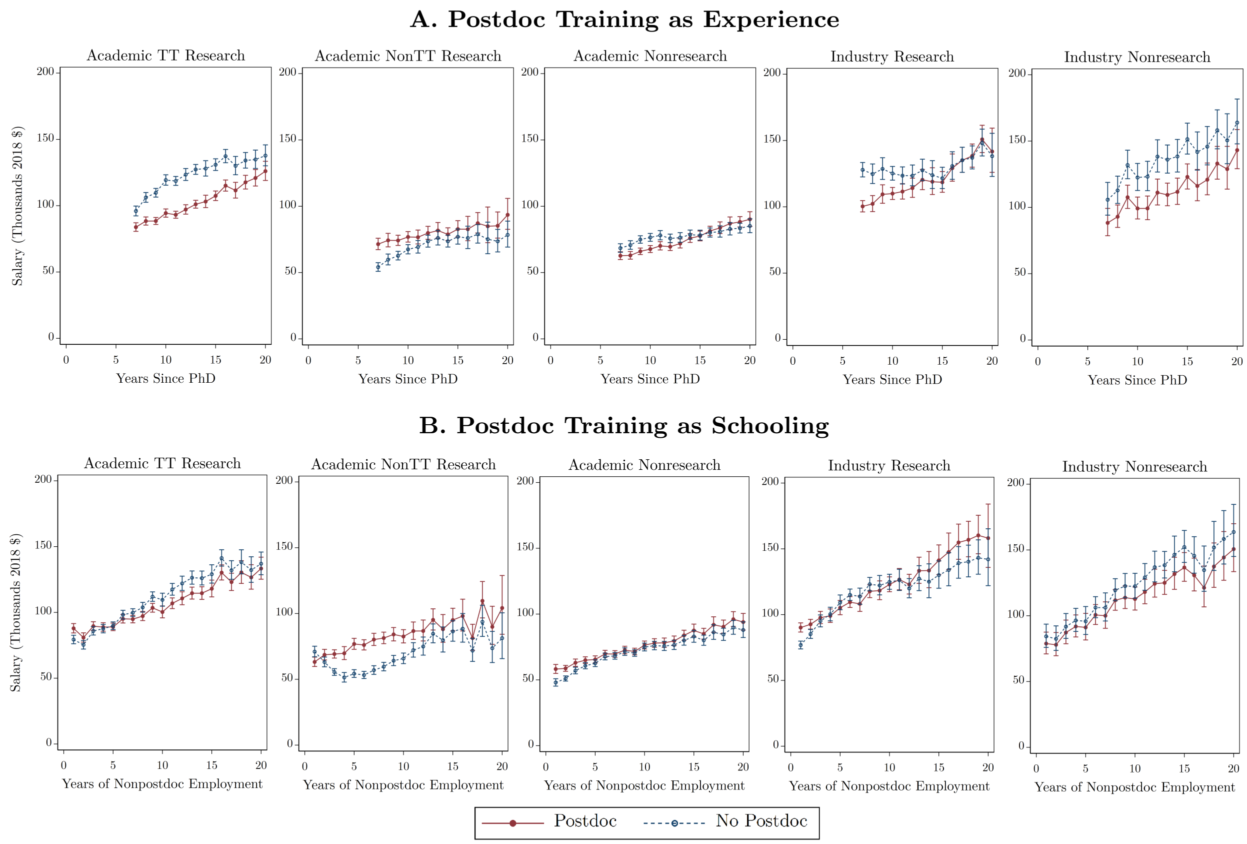
Figures and Tables

Figure 1: Predicted Salary Profiles by Postdoc-Trained Status and Employment Sector



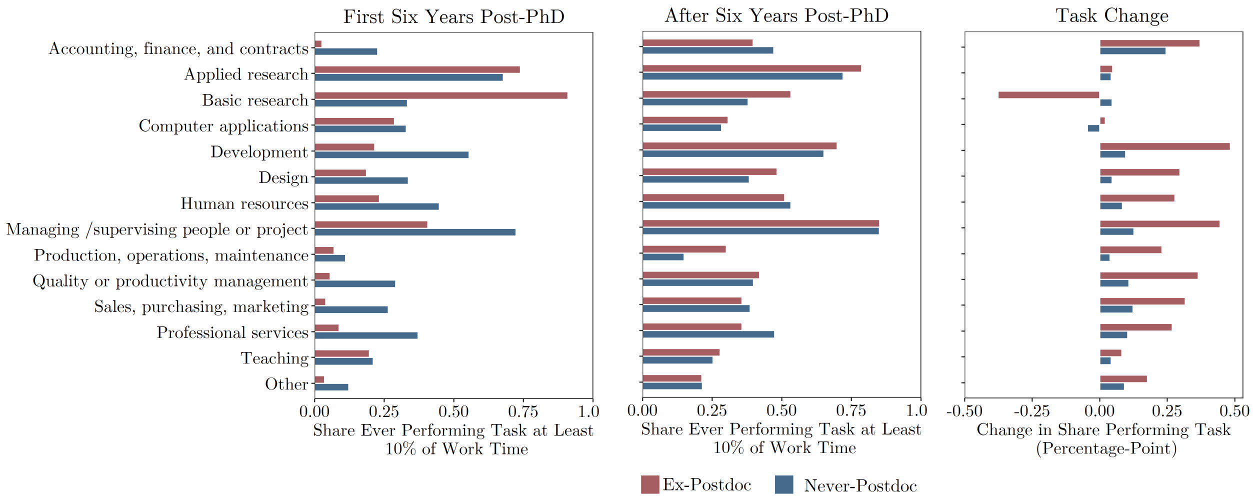
Notes: Figure 1 shows the average of predicted salary profiles for biomedical doctorates with and without postdoc training generated by augmented versions of specifications underlying results in columns (4) and (6) of Table 3 Panel I that include interactions between postdoc-trained status and the quartic polynomial in experience. The plots are generated by the following process: For each doctorate in the given employment sector sample, we generate two predictions of log(salary) in each year since PhD. The first prediction gives the log(salary) predicted if the person is assumed to have done a postdoc and the second prediction gives the log(salary) predicted if the person did not do a postdoc. Then, we average the predicted log(salary) across individuals in the given employment sector in each year since PhD and apply the exponential function to translate log(salary) into salary. We then plot these average predicted salary profiles with 95% confidence intervals in Figure 1. The employment sector subsamples are based on each doctorate's sector of employment in the given year, rather than the sector of employment at ten years post-PhD, in the underlying specifications used to generate the predictions. Salary adjusted for inflation using the CPI-U.

Figure 2: Predicted Salary Profiles by Postdoc-Trained Status and Employment Subsector



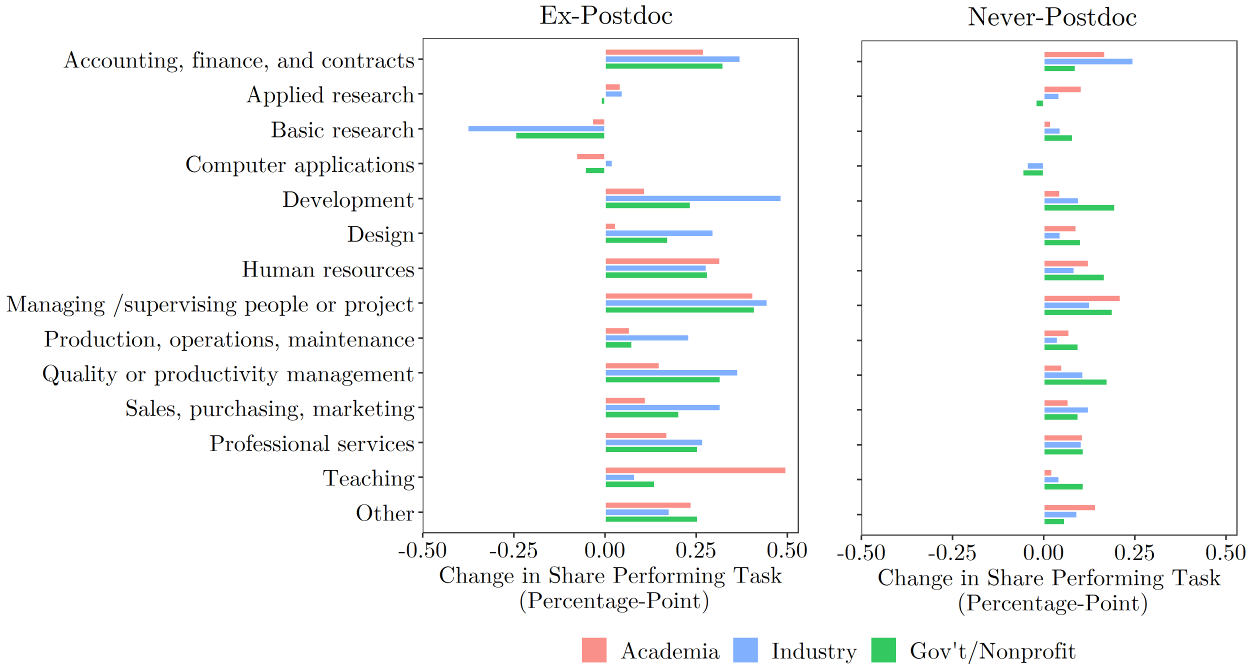
Notes: Figure 2 shows the average of predicted salary profiles for biomedical doctorates with and without postdoc training generated by augmented versions of specifications underlying results in columns (4) and (6) of Table 3 Panel II that include interactions between postdoc-trained status and the quartic polynomial in experience. See notes to Figure 1 for more detail.

Figure 3: Change in Tasks for Ex-Postdocs and Never-Postdocs Working in Industry



Notes: Figure 3 gives the change in the share of ex-postdocs and never-postdocs performing tasks for at least 10% of work time among biomedical doctorates working in industry. The tasks performed by ex-postdocs in their first six years represent the tasks performed as part of postdoc training. Greater magnitudes of task change represent greater degrees of mismatch in a given task. See Table A.3 for the underlying data used to construct this figure.

Figure 4: Change in Tasks for Ex-Postdocs and Never-Postdocs By Sector



Notes: Figure 4 gives the change in the share of ex-postdocs and never-postdocs performing tasks for at least 10% of work time among biomedical doctorates working in each sector. Greater magnitudes of task change represent greater degrees of mismatch in a given task. See Table A.3 and Table A.4 for the underlying data used to construct this figure.

Table 1: Analytical Sample Observation Counts (Person Counts) by Employment Sector

Employment Sector:	In Sector at 10 years post-PhD			In Sector in Year of Observation†		
Group:	Ex-Postdoc	Never-Postdoc	Total	Ex-Postdoc	Never-Postdoc	Total
All Sectors	21604 (3420)	7984 (1358)	29598 (4778)	16325 (3420)	6187 (1358)	22512 (4778)
Academia	12463 (1961)	3604 (593)	16067 (2554)	9221 (2192)	2720 (674)	11941 (2866)
TT Research	5092 (789)	529 (81)	5621 (870)	3630 (1111)	366 (132)	3996 (1243)
Non-TT Research	2422 (395)	494 (80)	2916 (475)	1625 (675)	363 (146)	1988 (821)
Nonresearch	4949 (777)	2581 (432)	7530 (1209)	3966 (1321)	1991 (577)	5957 (1898)
Industry	5964 (961)	2835 (507)	8799 (1468)	4519 (1193)	2189 (593)	6708 (1786)
Research	3179 (521)	1121 (188)	4300 (709)	2260 (805)	857 (292)	3117 (1097)
Nonresearch	2785 (440)	1714 (319)	4499 (759)	2259 (820)	1332 (474)	3591 (1294)
Gov't/Nonprofits	3187 (498)	1545 (258)	4732 (756)	2582 (809)	1278 (360)	3863 (1169)

Notes: Table 1 lists the number of person-year observations in our analytical sample by employment sector and the postdoc-trained status of the biomedical doctorate. Unique person counts in each cell appear in parentheses. Since a single worker may show up in different sectors at different times, the column sum of the person counts associated with the last three columns exceed the total number of persons included in the analytical sample. † = excludes salary observations corresponding to years when a biomedical doctorate is employed as a postdoc and any years within the first six years post-PhD.

Table 2: Earnings Distribution of Biomedical PhDs by Sector

Sector	Mean	P10	P25	P50	P75	P90
All Sectors	109008	55000	70000	95000	125000	165000
Academia	91700	55000	65000	80000	100000	135000
Industry	140705	65000	100000	125000	155000	200000
Gov't/Nonprofit	106275	60000	80000	100000	120000	165000

Notes: Table 2 reports the earnings distribution of biomedical PhDs at 10 years post-PhD by sector of employment. Percentile measures are rounded to the nearest $5000 for disclosure purposes. Earnings are reported in 2018 USD using the CPI-U.

Table 3: Postdoc Salary Premium by Employment Sector

Dependent Variable: log(salary)	(1)	(2)	(3)	(4)	(5)	(6)
Panel I. By Sector						
A. All Sectors	$N = 29598$		$N = 22512$		$N = 26312$	
Postdoc Training	-0.115***	-0.138***	-0.0842***	-0.117***	0.0253	0.000956
	(0.0202)	(0.0201)	(0.0236)	(0.0235)	(0.0209)	(0.0210)
R^2	0.181	0.272	0.130	0.246	0.143	0.244
B. Academia	$N = 16067$		$N = 11941$		$N = 13947$	
Postdoc Training	-0.0201	-0.0602**	0.0318	-0.00836	0.126***	0.0983***
	(0.0256)	(0.0277)	(0.0307)	(0.0337)	(0.0270)	(0.0294)
R^2	0.232	0.363	0.159	0.314	0.159	0.301
C. Industry	$N = 8799$		$N = 6708$		$N = 7898$	
Postdoc Training	-0.138***	-0.213***	-0.103**	-0.158***	-0.0102	-0.0450
	(0.0377)	(0.0376)	(0.0423)	(0.0410)	(0.0380)	(0.0385)
R^2	0.176	0.381	0.132	0.400	0.141	0.376
D. Gov't/Nonprofit	$N = 4732$		$N = 3863$		$N = 4467$	
Postdoc Training	-0.135***	0.00392	-0.0867**	-0.106**	0.0318	0.0177
	(0.0404)	(0.0542)	(0.0349)	(0.0450)	(0.0322)	(0.0396)
R^2	0.201	0.409	0.201	0.540	0.224	0.528
Panel II. By Academic/Industry Subsector						
A. Academic TT Research	$N = 5621$		$N = 3996$		$N = 4394$	
Postdoc Training	-0.176***	-0.381***	-0.0941*	-0.174***	0.00601	-0.0500
	(0.0515)	(0.0409)	(0.0495)	(0.0557)	(0.0455)	(0.0533)
R^2	0.358	0.516	0.168	0.349	0.169	0.349
B. Academic NonTT Research	$N = 2916$		$N = 1988$		$N = 2408$	
Postdoc Training	-0.102*	-0.130	-0.0244	0.159**	0.115**	0.232***
	(0.0591)	(0.130)	(0.0584)	(0.0788)	(0.0541)	(0.0678)
R^2	0.242	0.491	0.189	0.531	0.165	0.498
C. Academic Nonresearch	$N = 7530$		$N = 5957$		$N = 7145$	
Postdoc Training	-0.0253	0.0114	0.00369	-0.0416	0.0812**	0.0481
	(0.0333)	(0.0399)	(0.0396)	(0.0476)	(0.0346)	(0.0397)
R^2	0.208	0.445	0.189	0.453	0.174	0.419
D. Industry Research	$N = 4300$		$N = 3117$		$N = 3801$	
Postdoc Training	-0.101*	-0.176***	-0.00865	-0.0832*	0.0714*	0.0162
	(0.0540)	(0.0598)	(0.0490)	(0.0446)	(0.0430)	(0.0440)
R^2	0.183	0.390	0.138	0.482	0.149	0.453
E. Industry Nonresearch	$N = 4499$		$N = 3591$		$N = 4097$	
Postdoc Training	-0.153***	-0.207***	-0.160***	-0.155***	-0.0701	-0.0707
	(0.0499)	(0.0644)	(0.0570)	(0.0762)	(0.0520)	(0.0722)
R^2	0.221	0.522	0.177	0.499	0.180	0.473
Observations during postdoc included?:						
Yes	✓	✓				
No			✓	✓	✓	✓
Postdoc Training Treated As:						
Experience	✓	✓	✓	✓		
Schooling					✓	✓
Fixed Effects						
Field + Cohort + Year	✓		✓		✓	
Field-Cohort + PhD University + Year		✓		✓		✓

Notes: This table reports results from salary regressions based on equation (1). Columns (1) and (2) include all person-year observations and define a PhD's employment sector as that observed at 10 years post-PhD. Columns (3) through (6) keep only person-year observations corresponding to years after any postdoc training and define a PhD's employment sector as that held by the PhD in the given year. Columns (1) through (4) measure experience as years since PhD graduation. Columns (5) and (6) measure experience as years of *nonpostdoc* employment. Columns (3) and (4) drop all observations within the first six years after PhD. Robust standard errors clustered at individual-level are in parentheses. Estimates produced using survey weights. All specifications include the controls listed in the notes to Table A.1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Primary Tasks Performed by Doctorates Before and After the First Six Years Post-PhD by Postdoc-Trained Status

Employment Sector:	Academia				Industry				Gov't/Nonprofits			
Period (Years Post-PhD):	First Six Years		After Six Years		First Six Years		After Six Years		First Six Years		After Six Years	
Group:	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc
Acct., Finance, and Contracts	–	–	–	1.50%	–	–	4.21%	4.21%	–	–	–	3.91%
Applied Research	17.65%	21.58%	20.45%	22.28%	45.98%	28.86%	40.23%	52.91%	40.58%	28.52%	47.83%	41.02%
Basic Research	14.85%	79.32%	18.21%	64.34%	5.75%	75.15%	4.21%	10.22%	9.42%	75.00%	15.22%	40.63%
Computer Applications	–	–	–	1.20%	6.90%	2.00%	7.28%	6.21%	–	–	–	–
Development	–	–	–	1.70%	18.39%	–	27.59%	25.85%	–	–	10.14%	8.59%
Design	–	–	–	–	–	–	–	4.61%	–	–	–	–
Human Resources	–	–	3.36%	1.00%	–	–	–	–	–	–	–	–
Managing People or Projects	8.40%	1.20%	21.29%	23.48%	19.16%	–	38.70%	37.27%	23.19%	–	49.28%	40.23%
Production, Operations, Maint.	–	–	–	–	–	–	–	3.81%	–	–	–	–
Quality or Productivity Mgmt.	–	–	–	–	–	–	4.60%	5.81%	–	–	–	–
Sales, Purchasing, Marketing	–	–	–	–	4.98%	–	8.05%	8.82%	–	–	–	–
Professional Services	13.73%	5.00%	13.45%	6.49%	24.90%	4.21%	31.80%	20.64%	25.36%	5.86%	23.19%	15.23%
Teaching	62.18%	4.30%	64.43%	33.67%	–	–	–	–	–	–	–	–
Other	3.08%	1.30%	7.56%	3.70%	7.28%	–	10.34%	9.02%	15.22%	–	15.22%	17.58%
N	357	1001	357	1001	261	499	261	499	138	256	138	256

Notes: In this table, we calculate the proportion of postdoc-trained and nonpostdoc-trained biomedical doctorates that report each given task as their primary work activity at least once 1) in their first six years post-PhD and 2) after their first six years post-PhD. We restrict the sample to biomedical doctorates that are employed in each employment sector at 10 years post-PhD and for whom we observe at least two times in their first six years post-PhD. For ex-postdocs, we only consider observations in the first six years post-PhD that correspond to years employed as a postdoc; after six years post-PhD, we only consider observations corresponding to years after any and all years employed as a postdoc, and where the doctorate is employed in the given employment sector. For never-postdocs, we only consider observations corresponding to years where the person is employed in the given employment sector. "–" reported in cells of insufficient size to be disclosed. N reports person counts.

Table 5: Controlling for Task History and Current Tasks in Industry Salary Regressions

Dependent Variable: log(salary)	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Task History Controls			$N = 3104$			
Postdoc Training	-0.228***	-0.126*	-0.112*	-0.130**	-0.0917	-0.0781
	(0.0634)	(0.0666)	(0.0633)	(0.0629)	(0.0641)	(0.0664)
R^2	0.498	0.518	0.527	0.524	0.536	0.537
Panel B. Current Job Task Controls			$N = 3104$			
Postdoc Training	-0.228***	-0.237***	-0.233***	-0.249***	-0.236***	-0.242***
	(0.0634)	(0.0644)	(0.0639)	(0.0645)	(0.0641)	(0.0643)
R^2	0.498	0.511	0.512	0.507	0.516	0.517
Panel C. Current & Task History Controls			$N = 3104$			
Postdoc Training	-0.228***	-0.138**	-0.114*	-0.136**	-0.100	-0.0932
	(0.0634)	(0.0669)	(0.0644)	(0.0630)	(0.0642)	(0.0639)
R^2	0.498	0.531	0.538	0.531	0.549	0.552
Included Task Control Sets						
Primary Activity		✓				✓
Primary or Secondary Activity			✓		✓	
Activity $\geq$ 10% of Work Time				✓	✓	✓

Notes: This table reports regressions results based on the specification given in equation (1) where our sample includes all biomedical doctorates in the SDR graduating between 1993 and 2006 who are observed in at least two of their first six years post-PhD. Experience is defined as years since PhD graduation for all biomedical doctorates. For each doctorate, we keep only those person-year observations corresponding to years after any and all years employed as a postdoc, and we drop all observations within the first six years after PhD. Subsamples are based on the employment sector associated with each person-year observation. In Panel A, we add controls for the history of tasks performed as part of previous employment. In Panel B, we add controls for the tasks associated with the current job. Robust standard errors clustered at individual-level are in parentheses. Estimates produced using survey weights. Specifications (1) - (6) include all controls listed in the notes to Table A.1 as well as PhD university fixed effects and field-by-cohort fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Coefficient Estimates on Task Tenure Controls

Dependent Variable: log(salary)		
Postdoc Training	-0.126*	(0.0666)
Accounting Experience	-0.0898**	(0.0394)
Basic Research Experience	-0.0394***	(0.0108)
Computer App Experience	-0.00224	(0.0116)
Development Experience	0.00149	(0.0125)
Design Experience	0.0202	(0.0293)
HR Experience	0.0295	(0.0118)
Management Experience	0.0204**	(0.0103)
Production Experience	-0.0286	(0.0273)
Quality/Productivity MGMT Experience	0.00277	(0.0309)
Sales/Marketing Experience	-0.0351**	(0.0152)
Professional Services Experience	0.00652	(0.00873)
Teaching Experience	-0.0641**	(0.0264)
Other Experience	-0.0268	(0.0237)
N	3104	
R^2	0.518	

Notes: Table 6 reports coefficient estimates on the (primary) task history controls included in the regression whose main results are report in Panel A column (2) of Table 5. Applied research is the base case and so estimates yield the value of spending an additional year in a job with the given primary task relative to a job where applied research is the primary task. Robust standard errors clustered at individual-level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Task Mismatch and Postdoc Salary Premia

Dependent Variable: log(salary)	(1)	(2)	(2′)	(3)	(3′)
Panel A. All Sectors ($N = 10215$)					
Postdoc	-0.0815***	0.0761**	0.0609*	0.0264	0.0231
	(0.0313)	(0.0349)	(0.0345)	(0.0366)	(0.0363)
Postdoc * Task Distance		-0.424***	-0.407***	-0.322***	-0.325***
		(0.0534)	(0.0514)	(0.0579)	(0.0573)
R^2	0.323	0.334	0.334	0.373	0.374
Panel B. Academia ($N = 5442$)					
Postdoc	-0.0185	0.104**	0.0938**	0.0147	0.0191
	(0.0415)	(0.0452)	(0.0442)	(0.0424)	(0.0427)
Postdoc * Task Distance		-0.383***	-0.392***	-0.238***	-0.261***
		(0.0711)	(0.0701)	(0.0703)	(0.0731)
R^2	0.422	0.431	0.432	0.482	0.482
Panel C. Industry ($N = 3104$)					
Postdoc	-0.228***	-0.0354	-0.0550	-0.0457	-0.0651
	(0.0634)	(0.0720)	(0.0731)	(0.0790)	(0.0795)
Postdoc * Task Distance		-0.427***	-0.384***	-0.392***	-0.351***
		(0.104)	(0.101)	(0.113)	(0.109)
R^2	0.498	0.506	0.504	0.521	0.520
Panel D. Gov't/Nonprofit ($N = 1669$)					
Postdoc	-0.103	0.0122	-0.00664	0.165*	0.153
	(0.0789)	(0.0881)	(0.0847)	(0.0976)	(0.0956)
Postdoc * Task Distance		-0.356**	-0.323**	-0.474***	-0.467***
		(0.151)	(0.141)	(0.156)	(0.147)
R^2	0.703	0.708	0.707	0.724	0.724
Task Mismatch Measure:					
Extensive		✓		✓	
Intensive			✓		✓
Additional Controls					
Current Tasks				✓	✓

Notes: For "All Sectors" regressions, we include sector fixed effects to control for average salary differences between academia, industry, and gov't/nonprofits. This table reports regressions results based on the specification given in equation (1) where our sample includes all biomedical doctorates in the SDR graduating between 1993 and 2006 who are observed in at least two of their first six years post-PhD. Experience is defined as years since PhD graduation for all biomedical doctorates. For each doctorate, we keep only those person-year observations corresponding to years after any and all years employed as a postdoc, and we drop all observations within the first six years after PhD. Subsamples are based on the employment sector associated with each person-year observation. Robust standard errors clustered at individual-level are in parentheses. Estimates produced using survey weights. Specifications include all controls listed in the notes to Table A.1 as well as PhD university fixed effects and field-by-cohort fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Task Mismatch and Salary Penalties

Dependent Variable: log(salary)	(1)	(2)	(2′)	(3)	(3′)
Panel A. All Sectors ($N = 10215$)					
Postdoc	-0.0815***	0.0147	-0.0137	-0.0259	-0.0554
	(0.0313)	(0.0461)	(0.0435)	(0.0450)	(0.0425)
Postdoc * Task Distance		-0.236**	-0.148	-0.159	-0.0492
		(0.115)	(0.115)	(0.109)	(0.110)
Task Distance		-0.190*	-0.262**	-0.172*	-0.291***
		(0.102)	(0.104)	(0.102)	(0.104)
R^2	0.323	0.335	0.336	0.374	0.375
Panel B. Academia ($N = 5442$)					
Postdoc	-0.0185	0.0445	0.0405	-0.0353	-0.0539
	(0.0415)	(0.0536)	(0.0515)	(0.0512)	(0.0502)
Postdoc * Task Distance		-0.187	-0.185	-0.0720	-0.0242
		(0.130)	(0.140)	(0.119)	(0.129)
Task Distance		-0.198*	-0.209*	-0.175*	-0.299***
		(0.114)	(0.124)	(0.104)	(0.114)
R^2	0.422	0.431	0.433	0.482	0.484
Panel C. Industry ($N = 3104$)					
Postdoc	-0.228***	-0.0245	-0.0826	-0.0158	-0.0727
	(0.0634)	(0.0971)	(0.0963)	(0.0998)	(0.0970)
Postdoc * Task Distance		-0.457**	-0.302	-0.478***	-0.327
		(0.204)	(0.209)	(0.211)	(0.209)
Task Distance		0.0306	-0.0833	0.0942	-0.0264
		(0.182)	(0.182)	(0.192)	(0.190)
R^2	0.498	0.506	0.504	0.521	0.520
Panel D. Gov't/Nonprofit ($N = 1669$)					
Postdoc	-0.103	0.00507	-0.0496	0.148	0.0980
	(0.0789)	(0.113)	(0.101)	(0.129)	(0.119)
Postdoc * Task Distance		-0.338	-0.201	-0.431*	-0.316
		(0.218)	(0.193)	(0.227)	(0.203)
Task Distance		-0.0185	-0.123	-0.0422	-0.155
		(0.168)	(0.137)	(0.195)	(0.170)
R^2	0.703	0.708	0.708	0.725	0.725
Task Mismatch Measure:					
Extensive		✓		✓	
Intensive			✓		✓
Additional Controls					
Current Tasks				✓	✓

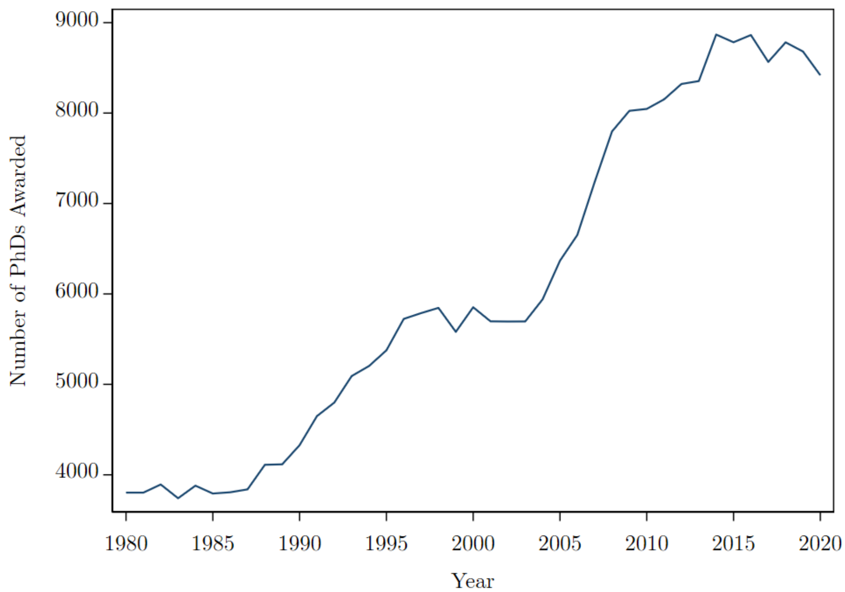
Notes: See notes to Table 7. In Table 8, we control for task distance in addition to its interaction with postdoc-trained status. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix

Table of Contents

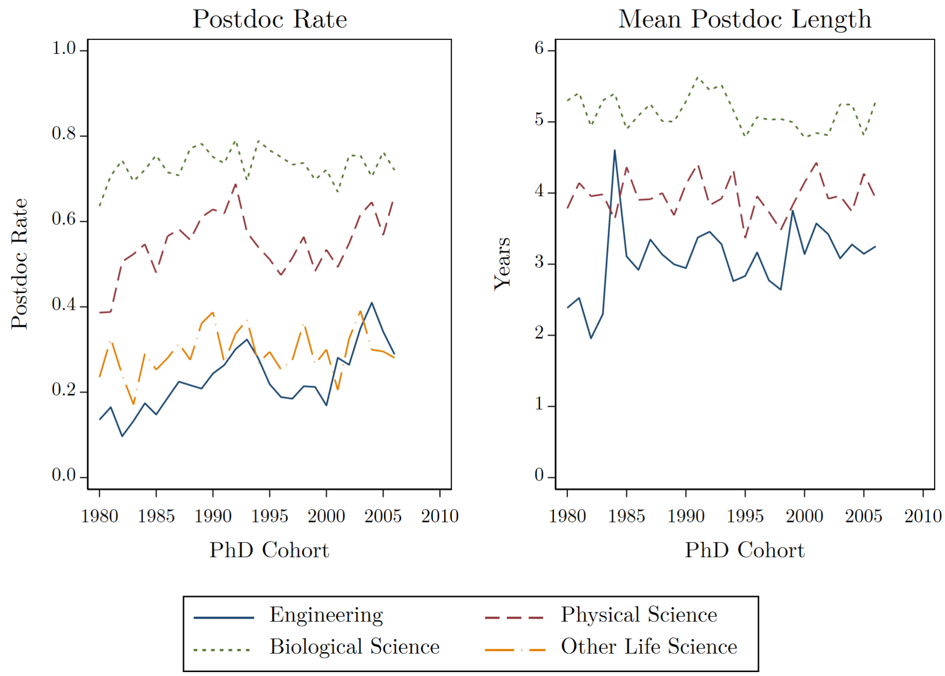
A Supplementary Figures and Tables

Figure A.1: Number of PhDs Awarded in Biomedical Fields by Year



Notes: Figure A.1 shows the number of PhDs awarded in Biological and Biomedical Sciences in each year. Data is from the NSF's Survey of Earned Doctorates (SED).

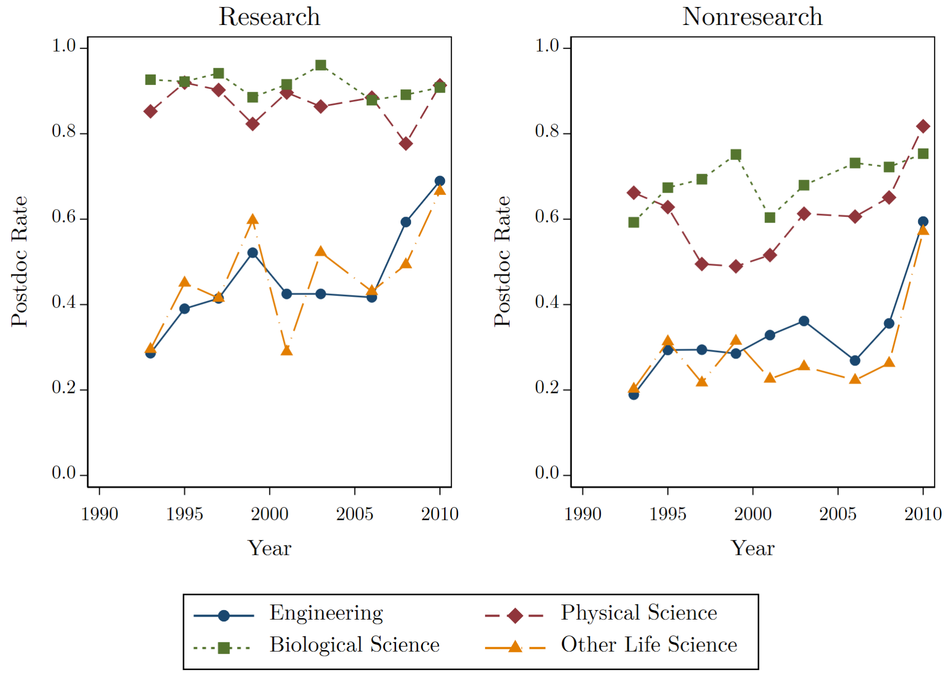
Figure A.2: Postdoc Rate and Length by S&E Field



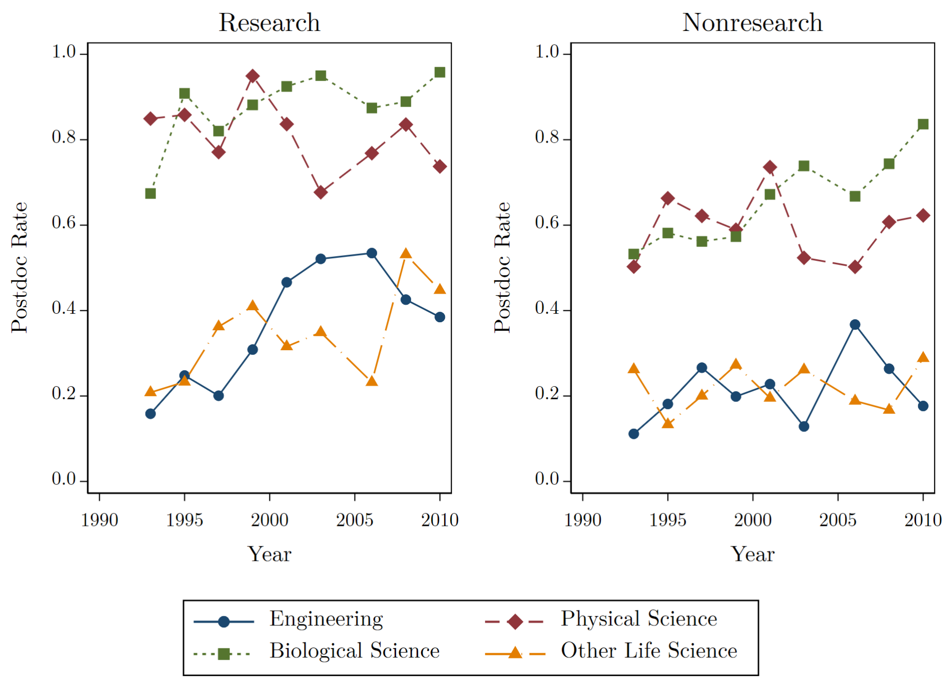
Notes: The left panel of Figure A.2 shows the proportion of doctorates in each PhD cohort that ever take a postdoc by broad field of study. The right panel show the mean length of postdoc training for all postdoc-trained PhD cohort members by broad field of study. Sample restricted to doctorates appearing in the NSF's Survey of Doctorate Recipients in any wave(s) between 1993 and 2015 and graduating as early as 1980. We restrict sample to doctorates who first appear in the SDR prior to 2010 due to SDR sampling changes starting in that year.

Figure A.3: Postdoc Rate of New Tenure-Track and Newly-Tenured Faculty by S&E Field

A. New Tenure-Track

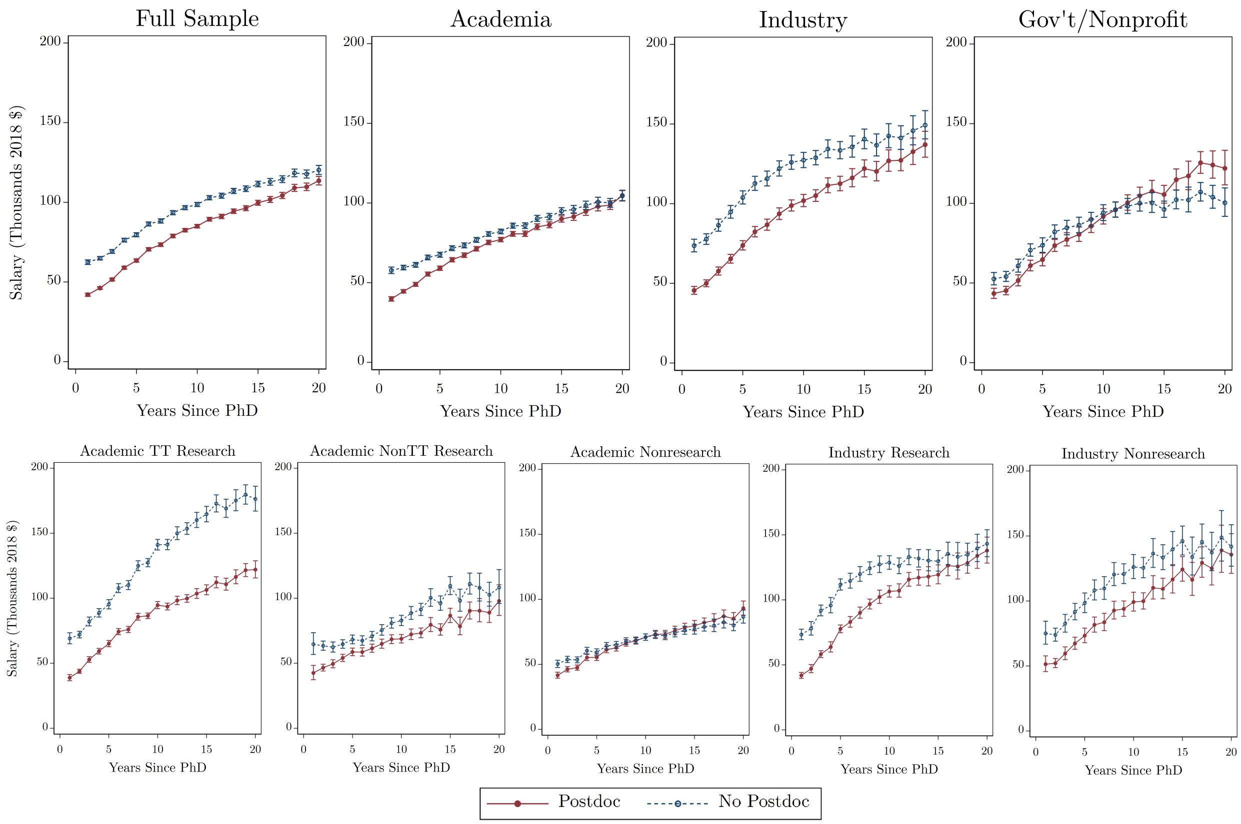


B. Newly-Tenured



Notes: Figure A.3 shows the postdoc-share of individuals who first report being employed in a tenure-track position (Panel A) or tenured position (Panel B) in a given SDR wave by broad field of study. Sample restricted to doctorates appearing in the NSF's Survey of Doctorate Recipients in any wave(s) between 1993 and 2015 and graduating as early as 1980. We restrict sample to doctorates who first appear in the SDR prior to 2010 due to SDR sampling changes starting in that year.

Figure A.4: Predicted Salary Profiles by Postdoc-Trained Status and Employment Sector



Notes: Figure A.4 shows the average of predicted salary profiles for biomedical doctorates with and without postdoc training generated by augmented versions of the specification underlying results in column (2) of Table 3 that include interactions between postdoc-trained status and the quartic polynomial in experience. See notes to Figure 1 for more detail.

Table A.1: Summary Statistics by Postdoc-Trained Status

Employment Sector:	Full Sample		Academia		Industry		Gov't/Nonprofit	
Group:	Ex-Postdoc	Never-Postdoc	Ex-Postdoc	Never-Postdoc	Ex-Postdoc	Never-Postdoc	Ex-Postdoc	Never-Postdoc
Foreign-born	0.25	0.20	0.25	0.17	0.27	0.22	0.23	0.17
Temp. Resident	0.13	0.07	0.13	0.06	0.14	0.09	0.10	0.06
Age at PhD	30.47	32.69	30.53	33.19	30.26	31.55	30.75	33.26
Female	0.39	0.38	0.39	0.40	0.39	0.36	0.40	0.36
Asian	0.18	0.13	0.17	0.10	0.21	0.17	0.17	0.10
Minority	0.08	0.10	0.08	0.09	0.06	0.11	0.17	0.10
PhD Length	6.69	7.75	6.77	7.97	6.57	7.45	6.81	7.96
Married at PhD	0.53	0.63	0.55	0.66	0.53	0.60	0.51	0.59
Child at PhD	0.30	0.45	0.32	0.47	0.28	0.41	0.30	0.40
Fellowship during PhD	0.17	0.17	0.17	0.17	0.15	0.15	0.19	0.15
RA during PhD	0.31	0.23	0.30	0.21	0.33	0.27	0.30	0.22
TA during PhD	0.12	0.14	0.12	0.16	0.11	0.10	0.11	0.15
Mother's Highest Education: BA	0.22	0.20	0.22	0.19	0.22	0.20	0.21	0.19
Mother's Highest Education: $>$ BA	0.19	0.16	0.20	0.16	0.18	0.18	0.21	0.19
Father's Highest Education: BA	0.23	0.21	0.24	0.20	0.22	0.21	0.20	0.24
Father's Highest Education: $>$ BA	0.34	0.30	0.34	0.30	0.32	0.30	0.35	0.27
N	3420	1358	2192	674	1193	593	809	360

Notes: Table A.1 reports weighted means for postdoc-trained and nonpostdoc-trained biomedical doctorates in the analytical sample by employment sector, where the weights used for each doctorate are those from the most recent SDR wave wherein each doctorate is observed. Sample counts are unweighted. For each cell, approximately 10% of PhD length calculations were imputed at the mean value (seven years) for the analytical sample. Since a single worker may show up in different sectors at different times, the sum of the person counts across subsectors exceeds the full sample person count. In addition to the variables above, all regressions include the following controls: a quartic polynomial in experience; an interaction between marital status and having a child in the home at time of PhD; female interactions with age, race, foreign-born status, temporary resident status, and marital status and child at home status (and its interaction); whether the PhD had already earned a professional degree such as an MD at time of graduation; whether PhD length was imputed; indicators for each parent's highest level of education (BA, MA, Professional degree, or PhD); normalized year fixed effects; field fixed effects; and cohort fixed effects. Our preferred specifications also include field-by-cohort fixed effects and PhD institution fixed effects. Field fixed effects represent the following fields of study: Anatomy; Bacteriology; Biochemistry; Biology/Biomedical Sciences, General; Biology/Biomedical Sciences, Other; Biomedical Sciences; Biometrics & Biostatistics; Biophysics; Biotechnology & Bioinformatics; Botany/Plant Biology; Cell/Cellular Biology & Histology; Developmental Biology/Embryology; Ecology; Endocrinology; Entomology; Evolutionary Biology; Genetics/Genomics, Human & Animal; Immunology; Microbiology; Molecular Biology; Neurosciences & Neurobiology; Nutrition Sciences; Parasitology; Pathology, Human & Animal; Pharmacology, Human & Animal; Physiology, Human & Animal; Plant Genetics; Plant Pathology/Phytopathology; Plant Physiology; Toxicology; Zoology.

Table A.2: Task Regression Sample Observation Counts (Person Counts) by Employment Sector

Employment Sector	In Sector in Year of Observation†		
	Ex-Postdoc	Never-Postdoc	Total
All Sectors	7541 (1804)	2674 (675)	10215 (2479)
Academia	4186 (1134)	1256 (333)	5442 (1467)
TT Research	1466 (509)	133 (58)	1599 (567)
Non-TT Research	776 (358)	185 (74)	961 (432)
Nonresearch	1944 (692)	938 (284)	2882 (976)
Industry	2211 (638)	893 (271)	3104 (909)
Research	1077 (412)	363 (137)	1440 (549)
Nonresearch	1134 (437)	530 (212)	1664 (649)
Gov't/Nonprofits	1144 (416)	525 (165)	1669 (581)

Notes: Table A.2 lists the number of person-year observations in our task regression sample by employment sector and the postdoc-trained status of the biomedical doctorate. Unique person counts in each cell appear in parentheses. Since a single worker may show up in different sectors at different times, the column sum of the person counts exceed the total number of persons included in the task regression sample. † = excludes salary observations corresponding to years when a biomedical doctorate is employed as a postdoc and any years within the first six years post-PhD.

Table A.3: Tasks Performed by Doctorates Working in Industry Before and After the First Six Years Post-PhD by Postdoc-Trained Status

Employment Sector:	Industry					
Period (Years Post-PhD):	First Six Years		After Six Years		Task Change	
Group:	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc
Accounting, Finance, and Contracts	22.61%	2.61%	47.13%	39.68%	24.52	37.07
Applied Research	67.82%	73.95%	72.03%	78.76%	4.21	4.81
Basic Research	33.33%	90.98%	37.93%	53.31%	4.60	-37.68
Computer Applications	32.95%	28.66%	28.35%	30.66%	-4.60	2.00
Development	55.56%	21.64%	65.13%	69.94%	9.58	48.30
Design	33.72%	18.64%	38.31%	48.30%	4.60	29.66
Human Resources	44.83%	23.25%	53.26%	51.10%	8.43	27.86
Managing People or Projects	72.41%	40.68%	85.06%	85.17%	12.64	44.49
Production, Operations, Maintenance	11.11%	7.01%	14.94%	30.06%	3.83	23.05
Quality or Productivity Management	29.12%	5.61%	39.85%	42.08%	10.73	36.47
Sales, Purchasing, Marketing	26.44%	4.01%	38.70%	35.67%	12.26	31.66
Professional Services	37.16%	8.82%	47.51%	35.67%	10.34	26.85
Teaching	21.07%	19.64%	25.29%	27.86%	4.21	8.22
Other	12.26%	3.61%	21.46%	21.24%	9.20	17.64
N	261	499	261	499	261	499

Notes: Table A.3 shows the proportion of postdoc-trained and nonpostdoc-trained biomedical doctorates that report spending at least 10% of their work time engaged in the given activity at least once 1) in their first six years post-PhD and 2) after their first six years post-PhD. We restrict the sample to biomedical doctorates that are employed in industry at 10 years post-PhD and for whom we observe at least two times in their first six years post-PhD. For both postdoc-trained and nonpostdoc-trained biomedical doctorates, we then report the percentage-point difference between the fraction of each performing each task within and after their first six years post-PhD, and refer to this measure as the "task change" of each group. For ex-postdocs, we only consider observations in the first six years post-PhD that correspond to years employed as a postdoc; after six years post-PhD, we only consider observations corresponding to years after any and all years employed as a postdoc, and where the doctorate is employed in industry. For never-postdocs, we only consider observations corresponding to years where the person is employed in industry.

Table A.4: Tasks Performed by Doctorates Before and After the First Six Years Post-PhD by Postdoc-Trained Status

Employment Sector:	Academia						Gov't/Nonprofit					
Period (Years Post-PhD):	First Six Years		After Six Years		Task Change		First Six Years		After Six Years		Task Change	
Group:	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc	No Pdoc	Pdoc
Accounting, Finance, and Contracts	10.92%	5.99%	27.73%	33.07%	16.81	27.07	33.33%	6.64%	42.03%	39.06%	8.70	32.42
Applied Research	56.02%	67.73%	66.39%	71.93%	10.36	4.20	77.54%	75.00%	75.36%	73.83%	-2.17	-1.17
Basic Research	64.15%	94.41%	66.11%	90.91%	1.96	-3.50	40.58%	92.97%	48.55%	68.36%	7.97	-24.61
Computer Applications	23.53%	29.77%	23.25%	21.88%	-0.28	-7.89	36.96%	34.38%	31.16%	28.91%	-5.80	-5.47
Development	18.21%	15.38%	22.69%	26.27%	4.48	10.89	28.26%	19.92%	47.83%	43.36%	19.57	23.44
Design	10.08%	21.18%	19.05%	24.08%	8.96	2.90	23.19%	19.53%	33.33%	36.72%	10.14	17.19
Human Resources	32.21%	26.17%	44.54%	57.74%	12.32	31.57	38.41%	22.27%	55.07%	50.39%	16.67	28.13
Managing People or Projects	63.03%	49.45%	84.03%	90.01%	21.01	40.56	71.74%	48.05%	90.58%	89.06%	18.84	41.02
Production, Operations, Maintenance	8.40%	10.89%	15.41%	17.58%	7.00	6.69	9.42%	8.59%	18.84%	16.02%	9.42	7.42
Quality or Productivity Management	9.52%	5.00%	14.57%	19.98%	5.04	14.99	23.19%	4.30%	40.58%	35.94%	17.39	31.64
Sales, Purchasing, Marketing	8.12%	3.10%	14.85%	14.19%	6.72	11.09	21.74%	6.64%	31.16%	26.95%	9.42	20.31
Professional Services	33.05%	9.99%	43.70%	26.97%	10.64	16.98	37.68%	10.55%	48.55%	35.94%	10.87	25.39
Teaching	90.48%	36.46%	92.72%	86.11%	2.24	49.65	30.43%	26.56%	41.30%	40.23%	10.87	13.67
Other	19.33%	4.40%	33.61%	28.07%	14.29	23.68	26.09%	5.86%	31.88%	31.25%	5.80	25.39
N	357	1001	357	1001	357	1001	138	256	138	256	138	256

Notes: This table shows the proportion of postdoc-trained and nonpostdoc-trained biomedical doctorates that report spending at least 10% of their work time engaged in the given activity at least once 1) in their first six years post-PhD and 2) after their first six years post-PhD. We restrict the sample to biomedical doctorates that are employed in the given employment sector at 10 years post-PhD and for whom we observe at least two times in their first six years post-PhD. For both postdoc-trained and nonpostdoc-trained biomedical doctorates, we then report the percentage-point difference between the fraction of each performing each task within and after their first six years post-PhD, and refer to this measure as the "task change" of each group. For ex-postdocs, we only consider observations in the first six years post-PhD that correspond to years employed as a postdoc; after six years post-PhD, we only consider observations corresponding to years after any and all years employed as a postdoc, and where the doctorate is employed in industry. For never-postdocs, we only consider observations corresponding to years where the person is employed in industry.

Table A.5: Task Mismatch and Postdoc Salary Premia (Postdoc Task Mismatch Only)

Dependent Variable: log(salary)	(1)	(2)	(2′)	(3)	(3′)
Panel A. All Sectors (N = 9031)					
Postdoc	-0.111***	0.0623*	0.0552	-0.0177	-0.00944
	(0.0321)	(0.0375)	(0.0371)	(0.0416)	(0.0413)
Postdoc * Task Distance		-0.370***	-0.368***	-0.235***	-0.250***
		(0.0493)	(0.0481)	(0.0540)	(0.0537)
R^2	0.345	0.355	0.355	0.393	0.393
Panel B. Academia (N = 4845)					
Postdoc	-0.0572	0.0904*	0.0941*	-0.0481	-0.0322
	(0.0447)	(0.0506)	(0.0486)	(0.0534)	(0.0536)
Postdoc * Task Distance		-0.359***	-0.402***	-0.162**	-0.196***
		(0.0718)	(0.0694)	(0.0733)	(0.0751)
R^2	0.449	0.457	0.460	0.504	0.504
Panel C. Industry (N = 2675)					
Postdoc	-0.231***	-0.00763	-0.0438	-0.000309	-0.0294
	(0.0690)	(0.0868)	(0.0946)	(0.111)	(0.119)
Postdoc * Task Distance		-0.375***	-0.307***	-0.350***	-0.299**
		(0.102)	(0.106)	(0.128)	(0.133)
R^2	0.506	0.511	0.510	0.530	0.529
Panel D. Gov't/Nonprofit (N = 1511)					
Postdoc	-0.0480	0.0932	0.0827	0.244*	0.267**
	(0.100)	(0.114)	(0.111)	(0.130)	(0.132)
Postdoc * Task Distance		-0.337**	-0.317**	-0.486***	-0.515***
		(0.147)	(0.135)	(0.168)	(0.162)
R^2	0.729	0.733	0.733	0.748	0.748
Task Mismatch Measure:					
Extensive		✓		✓	
Intensive			✓		✓
Additional Controls					
Current Tasks				✓	✓

Notes: See notes to Table 7. In Table A.5, we restrict our measures of task mismatch to capture the distance between tasks performed as part of current employment and postdoc training, in particular, rather than other jobs held within six years post-PhD. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: Task Distance Summary Statistics by Sector and Group

		Task Mismatch Measure			
		Extensive		Intensive	
Sector	Group	(1)	(2)	(1)	(2)
All Sectors	Never-Postdoc	0.322 (0.217)		0.289 (0.226)	
	Ex-Postdoc	0.380 (0.241)	0.474 (0.262)	0.355 (0.257)	0.452 (0.280)
Academia	Never-Postdoc	0.287 (0.198)		0.244 (0.201)	
	Ex-Postdoc	0.338 (0.213)	0.420 (0.228)	0.298 (0.222)	0.379 (0.241)
Industry	Never-Postdoc	0.359 (0.233)		0.334 (0.242)	
	Ex-Postdoc	0.449 (0.270)	0.587 (0.285)	0.450 (0.281)	0.599 (0.291)
Gov't/Nonprofit	Never-Postdoc	0.341 (0.219)		0.319 (0.235)	
	Ex-Postdoc	0.398 (0.248)	0.468 (0.269)	0.378 (0.267)	0.451 (0.290)

Notes: Table A.6 reports the mean and standard deviation (in parentheses) of the extensive and intensive mismatch measures by postdoc-trained status of the biomedical doctorate and sector. In column (1), task mismatch uses all tasks performed within the first six years for all doctorates regardless of postdoc-trained status. In column (2), only observations in the first six years that correspond to years employed as a postdoctoral researcher are included in the task mismatch calculation for postdoc-trained biomedical doctorates.

B Identifying Postdocs and Postdoc Length in SDR-SED Data

Our dataset contains three sources of information regarding a doctorate's postdoc status. The first source is the SED, wherein respondents are asked "What best describes your (within the next year) postgraduate plans?" and "What is the status of your postgraduate plans (in the next year)?" Starting in 2004, SED respondents are also asked "Do you intend to take a 'postdoc' position?". Using these questions, we assign a person as doing a postdoc if the respondent says that, post-graduation, he/she plans to do either a: 1) postdoc fellowship, 2) postdoc research associateship, 3) traineeship, or 4) internship/ clinical residency, and also states that he/she 1) will be either returning to present employment, 2) has accepted a position, or 3) is in negotiation with one or more specific organizations.

The second source containing information on postdoc status is the SDR. In each SDR wave, doctorates are asked whether they are currently working and whether their current job is a "post-doc." If a doctorate reports being in a postdoc job in any SDR wave, then we consider them to have done a postdoc. The third source comes from the Special Topic Module included on the SDR 1995 and 2006 waves wherein respondents are asked how many postdoc positions they have ever held and the starting and ending dates for their last three postdoc positions. We follow Kahn and Ginther (2017) in referring to these as the SDR Retrospective Surveys. If a doctorate reports having done at least one postdoc on either SDR Retrospective Survey, then we count that person as having done a postdoc. If a doctorate reports never having done a postdoc on the Retrospective Surveys, then we label the person as having never done a postdoc. In rare cases, sources disagree about whether a person has ever done a postdoc. If SED states that a person plans to do a postdoc, but then they never report doing a postdoc in any SDR wave and they claim to have never taken a postdoc position in the SDR Retrospective Surveys, then we label that person as never having done a postdoc. If a doctorate ever claims to have done a postdoc in any SDR wave (including the SDR Retrospective Surveys), then we label them as having done a postdoc.

Next, we seek to determine which years a person was employed as a postdoc. We create a variable ("pdoc_year") that equals one if the doctorate was in a postdoc in the given year and equals zero if the doctorate was not in a postdoc in the given year. Once we form this variable, we will take its sum across years for each doctorate to measure each doctorate's duration (or "length") of postdoc training. If a person was found to have never done a postdoc (pdoc==0), then pdoc_year==0 for all years. If the person could be identified as a postdoc based solely on information from the SED, then we labeled the year of PhD receipt as being a year that the doctorate was employed as a postdoc. For those who report currently being in a postdoc position in an SDR wave, we have the year that they began that current employment and so label all years from the start of employment to that SDR wave as years in a postdoc. For doctorates in the SDR 1995 and/or 2006 wave ("SDR Retrospective Surveys"), we have information on the start and end dates of a person's last three postdoc positions, and so label any years within any of the reported postdocs as postdoc years. Additionally, we consider all years after the end of the last reported postdoc on the SDR Retrospective Surveys as being years where a doctorate was not in a postdoc, assuming we have no other evidence to suggest the person took up an additional postdoc after that time. Similarly, for doctorates who report having done at most three postdocs throughout their career in the SDR Retrospective Surveys, we label years preceding the start of their first reported postdoc as years that the person was not in a postdoc, assuming no additional evidence to suggest otherwise. Additionally, we label any years 1) between the end of the 2nd most recent postdoc and the start of the most recent postdoc or 2) between the end of the 3rd most recent postdoc and the start of the 2nd most recent postdoc as "non-postdoc" years. Lastly, we label as non-postdoc years any SDR year where a doctorate reports not being currently employed in a postdoc position.

In addition, we impute whether a year is or is not a postdoc year in special cases to avoid sample attrition. The need for imputation is due to two features of the SDR. First, the SDR is typically biennial, and so there is usually one year in between SDR waves, although there are two cases where there are two-year gaps: between SDR 2003 and 2006 and between SDR 2010 and 2013. Second, new sample members to the SDR have typically been added between one and three years after PhD receipt. This means that some doctorates may have one or two years between their PhD graduation year and entry into the SDR where postdoc status is missing.[21]

Our imputation strategy is as follows: if a doctorate reports not being in a postdoc in both the SDR wave before and after the gap year(s), then those gaps years are considered as non-postdoc years. Similarly, if a person reports being in a postdoc in both the SDR wave before and after the gap year(s), then those gap years are considered postdoc years. If a doctorate reports doing a postdoc in the SDR wave before a gap year, but reports not doing a postdoc in the SDR wave after the gap year, then we split the difference for gap years by assigning a value of 0.5 to our postdoc year variable. If a doctorate is surveyed in the SDR within three years, but has gap years preceding appearance in the SDR, then we assign a value of 0.5 if the person reports a postdoc position in his/her first SDR wave and assign a value of 0 if the person reports no postdoc position in his/her first SDR wave.[22] For biomedical doctorates first sampled in the SDR prior to SDR 2010, we are able to identify if a doctorate was ever a postdoc in 99% of cases. In 86% of cases, we are able to identify or impute whether or not a biomedical doctorate is employed as a postdoc in each year since PhD graduation.[23]

[21]Starting with SDR 2010, doctorates obtaining PhDs more than three years prior to the survey date were newly sampled; for these cases, there are many years where we cannot determine postdoc status, and so we exclude these doctorates from our analytical sample.

[22]After our imputation strategy, the majority of doctorates who ever have a year where we fail to determine postdoc status are those who first appear in the SDR in the 2015 wave. The SDR 2015 wave was unique in that 80% of the SDR 2015 sample members were new to the survey, whereas in past cycles around 10% of the sample members were new. This was due to the SDR being expanded from 47,000 to 120,000 members, with members being added even when having graduated much earlier than 2015.

[23]In the analytical sample used in this study, we find that 77% of postdoc person-years occur in academia, 17% occur in government/nonprofits, and only 6% occur in industry.

C Research Job Regressions

We also analyze the relationship between postdoc training and the likelihood that biomedical doctorates obtain research jobs in academia and industry. Our empirical model is given by the following person-level linear probability model (LPM) specification:

$$job_{ifsc} = \boldsymbol{X_i \beta} + \theta Postdoc_i + \boldsymbol{\gamma_{fc}} + \boldsymbol{\gamma_s} + \varepsilon_{ifsc}, \tag{4}$$

where job_{ifsc} is an indicator variable for if doctorate i who graduated with a PhD in field f from university s in year c ever obtains a given research job and all other variables are defined as before. We consider five different indicator variables: The first is for whether a doctorate ever finds any type of nonpostdoc research position ("any"), the second is for whether a doctorate ever finds a nonpostdoc research position in academia ("academic"), the third is for whether a doctorate ever lands a tenure-track research job in academia ("tenure-track"), the fourth is for whether an individual obtains tenure in an academic research position ("tenured") conditional on having obtained a tenure-track research position, and the fifth is an indicator variable for if a doctorate ever obtains a research position in industry conditional on ever working in industry ("industry"). The analytical sample members for these regressions are the same as those in the salary regressions and robust standard errors are computed allowing for clustering at the field-cohort level.

Table C.1 reports the results using the LPM specification given by equation (4). We find that doing a postdoc increases the likelihood of working in any research job by 24.2 percentage points, an academic research position by 26.5 percentage points, and a tenure-track research position by 21.3 percentage points.[24] Among those that ever take a tenure-track job and whom we observe after they are up for their tenure decision, we find that postdoc training does not significantly impact the ability of tenure-track researchers to obtain tenure.[25] Lastly, among doctorates who ever work in industry, we find that postdoc training raises the probability of obtaining a research position in industry by 12.3 percentage points.

[24] Table C.2 shows that, more generally, postdoc-trained biomedical doctorates are more likely to land academic jobs, including tenure-track jobs, but that the estimated effects of landing *research-focused* academic and tenure-track jobs (as shown in Table C.1) are greater by comparison.

[25] This sample includes individuals who report being on the tenure track at some point and then later report either 1) being in a tenured position or 2) not in a tenured position and no longer on the tenure track.

Table C.1: Postdoc Training and the Likelihood of a Research Job

	Any	Academic	Tenure-Track	Tenured	Industry
Postdoc Training	0.242***	0.265***	0.213***	-0.0634	0.123***
	(0.0198)	(0.0193)	(0.0147)	(0.168)	(0.0435)
R^2	0.296	0.269	0.263	0.680	0.492
N	4778	4778	4778	798	1786
Fixed Effects					
Field-Cohort	✓	✓	✓	✓	✓
PhD University	✓	✓	✓	✓	✓

Notes: This table reports regressions results where the dependent variable for each column is an indicator variable for the type of research job given by the column name. Observations are person-level. The samples used for the "Academic" and "Tenure-Track" columns include biomedical doctorates in the SDR graduating between 1980 and 2007 for whom we have observed for at least 10 years post-PhD. The sample used for the "Tenured" column includes biomedical doctorates in the SDR graduating between 1980 and 2006 who report being on the tenure track at some point and then later report either 1) being in a tenured position or 2) not in a tenured position and no longer on the tenure track. The sample used for the "Industry" column includes biomedical doctorates in the SDR graduating in or after 1980 who ever report working in industry. Robust standard errors clustered at the field-cohort level are in parentheses. Specifications include all controls listed in the notes to Table A.1.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.2: Postdoc Training and the Likelihood of an Academic Job

	Academic (Any)	Tenure-Track	Tenured
Postdoc Training	0.169***	0.167***	-0.0149
	(0.0206)	(0.0198)	(0.0526)
R^2	0.249	0.267	0.459
N	4778	4778	1583
Fixed Effects			
Field-Cohort	✓	✓	✓
PhD University	✓	✓	✓

Notes: See notes to Table C.1. This table reports regressions results where the dependent variable for each column is an indicator variable for the type of job given by the column name which, unlike Table C.1, are not restricted to research-focused jobs.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

D Bias-Adjusted Estimates of the Effect of Postdoc Training

D.1 Method for Estimating Bias-Adjusted Treatment Effects

Oster's (2019) bias-adjusted treatment effect estimator is motivated by the following data generating process:

$$Y = \beta X + \Psi\omega^0 + W_2 + \varepsilon,$$

where Y is the outcome of interest, X is a scalar treatment variable, ω^0 is a vector of observed controls, and W_2 and ε are unobserved.[26] Letting $W_1 \equiv \Psi\omega^0$, a proportional selection relationship can be defined as $\delta\frac{\sigma_{1X}}{\sigma_1^2} = \frac{\sigma_{2X}}{\sigma_2^2}$, where $\sigma_{iX} \equiv \text{cov}(W_i, X)$ and $\sigma_i^2 \equiv \text{var}(W_i)$ for $i \in \{1, 2\}$, and where δ measures the level of selection on unobservables relative to observables. Let the coefficient and the R^2 obtained from a regression of Y on X ("uncontrolled regression") be denoted $\mathring{\beta}$ and $\mathring{R}$, respectively. Let the coefficient and the R^2 obtained from a regression of Y on X and ω^0 ("controlled regression") be denoted $\tilde{\beta}$ and $\tilde{R}$, respectively. Lastly, let the R^2 obtained from a hypothetical regression of Y on X, ω^0, and W_2 ("fully-specified regression") be denoted as R_{max}. Then, under some additional assumptions, Oster (2019) shows that a consistent bias-adjusted treatment effect (β^*) can be approximated by the following:

$$\beta^* \approx \tilde{\beta} - \delta\left[\mathring{\beta} - \tilde{\beta}\right]\frac{R_{max} - \tilde{R}}{\tilde{R} - \mathring{R}}.$$

Oster (2019) subsequently develops a consistent bias-adjusted treatment effect estimator that relaxes the additional restrictions used to derive the above approximation, and we use this more robust estimator to measure the sensitivity of our results to selection on unobservables.[27]

D.2 Bias-Adjusted Salary Regression Results

If postdoc-trained biomedical doctorates have lower ability at the time of PhD completion than those who forgo postdoc employment, then the postdoc salary penalty in industry reported in column (4) of Table 3 could potentially be explained by selection on unobserved ability at time of graduation. This explanation is unlikely for two reasons: First, Sauermann and Roach (2016) find that higher-ability biomedical doctorates plan on pursuing postdoc training, which would point to our estimates of a postdoc penalty being too conservative rather than too extreme.[28] Second, we include controls that are likely correlated with ability at time of graduation; these include field-by-cohort fixed effects, PhD university fixed effects, the education level of each biomedical doctorates'

[26] A key assumption in what follows is that W_2 is orthogonal to W_1; therefore, W_2 should be viewed as the *residualized portion* of the unobservables after a hypothetical regression of the unobservables on ω^0. See Appendix A.1 of Oster (2019) for a discussion of this assumption.

[27] This method is implemented using the user-created Stata command `psacalc` accessible via Emily Oster's website.

[28] Ability is proxied by four measures in Sauermann and Roach (2016): 1) number of peer-reviewed publications, 2) fellowships from a federal agency, 3) their PhD program's National Research Council (NRC) ranking, and 4) respondent's assessment of their own research ability relative to peers.

mother and father, length of time in a graduate program, graduate program funding source, and various background characteristics that are likely related to ability.[29]

Nevertheless, we test whether residual variation in unobserved ability at time of graduation might explain the postdoc salary penalty in industry by estimating bias-adjusted treatment effects as formulated in Oster (2019) and report the results of this, and the results for other sectors (and subsectors), as a robustness check in Panel A of Table D.1 (and Table D.2). We find that the inclusion of controls, which are plausibly correlated with ability, pushes the estimated impact of postdoc training on future salary in a negative direction for all sectors in Table D.1, which is consistent with postdoc-trained biomedical doctorates having higher ability than their nonpostdoc-trained counterparts. While we are not able to pinpoint the causal impact of postdoc training in the absence of a valid instrument for postdoc attainment, under the plausible assumption that selection on unobservables acts in the same direction as selection on observables, we can bound the value for the causal impact by using the Oster (2019) method for estimating bias-adjusted treatment effects. To do so, we must select an upper-bound for the level of selection on unobservables relative to selection on observables (δ) and the R^2 that we would expect from a fully-specified model that we would be able to estimate if the unobservables were instead observable (R_{max}). We follow Altonji, Elder, and Taber (2005) and Oster (2019) in treating $\delta = 1$ as an upper-bound for the level of selection on unobservables relative to observables.[30] Oster (2019) suggests that researchers arguing for the stability of their results consistent with that of randomized treatment should consider an upper bound value of $1.3\tilde{R}^2$ for R_{max}, where $\tilde{R}^2$ is the R^2 obtained from the controlled regression. Thus, we use this R_{max} and $\delta = 1$ to calculate an upper-bound value for the impact of postdoc training on after-postdoc salary in each employment sector and subsector, which we report as θ^* in Table D.1 and Table D.2.

We find that each point estimate in Panel A of Table D.1 is negative and of greater magnitude compared to the estimate in the corresponding controlled regression, suggesting that, under the plausible assumption that selection on unobservables runs in the same direction as selection on observables, the magnitude of each estimate in column (4) of Table 3 is a lower-bound for the causal impact of postdoc training on after-postdoc salary, while each estimate reported as θ^* represents an upper-bound.[31] Altogether, these results suggest that ability bias is unlikely to explain the

[29] Field-by-cohort fixed effects will be correlated with ability if individuals sort into different biomedical fields based on ability. PhD university fixed effects will be correlated with ability insofar as universities admit students to biomedical PhD programs based on individual ability (e.g., as measured by application materials including GRE scores and GPA) and insofar as different universities have different impacts on the human capital accumulation of PhD students. Parent's education level may proxy for socioeconomic background and possibly inherited traits impacting educational performance.

[30] As argued in Oster (2019), δ represents the relative degree of selection on the *residualized portion* of the unobservables (i.e., the variation in the unobservables unrelated to variation in the observables).

[31] The calculated upper-bounds all lie outside the 95% confidence interval of the corresponding estimate in column (4) of Table 3, indicating that correcting for selection on unobservables is potentially important. Altonji, Arcidiacono, and Maurel (2016) note that in the context of evaluating the impact of college field choice on future earnings, "much of the variance in earnings at a point in time is due to measurement error or permanent and transitory shocks that

existence of a postdoc penalty in industry, and that the true salary penalty in industry caused by postdoc training is somewhere between 15.8% and 26.2%, depending on the level of selection on unobservables and the degree to which inclusion of the unobservables as controls would increase the R^2 of the model.

When treating postdoc training as schooling in Panel B of Table D.1 and Table D.2, we find that the direction of selection bias is in the same direction as the results in Panel A when postdoc training is treated as experience. Of all the results in Table D.1 and Table D.2, only academic non-tenure-track research yields bias-adjusted estimates of the effect of postdoc training which push the estimate in a positive direction. This suggests that biomedical doctorates choosing a job in non-tenure-track research directly after graduation may be of higher ability compared to those who take a postdoc position, but our results suggest that postdoc training ultimately leads to higher earnings for those in this sector, which is consistent with postdoc training being an effective way to augment skills relevant to academic research.

D.3 Bias-Adjusted Research Job Regression Results

As with the impact of postdoc training on salary, unobservable ability at the time of graduation could potentially explain the impact of postdoc training on the ability of biomedical doctorates to obtain different types of research-focused jobs. Therefore, we test the robustness of our research job regression results reported in Table C.1 to selection on unobservables using the Oster's (2019) method as before and report the results in Table D.3. We find that the results in Table C.1 represent upper-bound estimates of the true impact of postdoc training on the likelihood of obtaining tenure-track and industry research jobs, whereas the bias-adjusted treatment effects represent lower-bounds. This finding, in conjunction with the direction of bias detected in the salary regressions in Panel A of Table D.1, is consistent with postdoc-trained biomedical doctorates having greater ability at the time of graduation compared to their nonpostdoc-trained counterparts, assuming that high-ability doctorates are more likely to obtain tenure-track and industry research positions. On the other hand, we find that correcting for selection on unobervables increases the positive effect of postdoc training on the chances that a biomedical doctorate works in any academic research job after-postdoc. This may indicate that doctorates of lower ability at time of graduation sort into postdoc training to augment their academic research skills in hopes of increasing their chance at nontenure-track research positions in academia, such as staff scientist positions. However, we find that, in all cases, the bias-adjusted treatment effect lies within one standard error of the esti-

occur after college decisions have been made" and thus are not a source of selection bias. The same argument can be made for the postdoc decision. It is important to note that the analysis in this section evaluates the sensitivity of our results to selection on unobserved ability *at the time of PhD graduation*, with the results based on movements in coefficients when controls determined by the time of PhD graduation are added to the regression specifications. It is not meant to test sensitivity to variables not determined by the time of PhD, such as tasks to be performed as part of future employment or as part of postdoc training that led to the accumulation of task-specific human capital (which is the focus of Section 4).

mates reported in Table C.1, indicating that the results are not especially sensitive to selection on unobservables.[32]

Table D.1: Sensitivity of Salary Regression Results to Selection on Unobservables by Sector

Sector:	All		Academia		Industry		Gov't/Nonprofit	
	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2
Panel A. Postdoc Training as Experience								
Uncontrolled	-0.0164	0.000	0.0815	0.003	-0.0675	0.001	0.0252	0.000
Controlled	-0.117	0.246	-0.00836	0.314	-0.158	0.400	-0.106	0.540
R_{max}	0.320		0.408		0.521		0.702	
θ^*	-0.174		-0.0775		-0.262		-0.510	
N	22512		11941		6708		3863	
Panel B. Postdoc Training as Schooling								
Uncontrolled	0.0212	0.000	0.118	0.006	-0.0384	0.000	0.0524	0.002
Controlled	0.001	0.245	0.0983	0.301	-0.0450	0.376	0.0177	0.528
R_{max}	0.317		0.391		0.488		0.686	
θ^*	-0.004		0.0835		-0.0518		-0.0835	
N	26312		13947		7898		4467	

Notes: We test if the results in columns (4) and (6) of Table 3 are robust to allowing for selection on unobservables using the methods developed in Oster (2019) in Panel A and Panel B, respectively; see notes to Table 3. We report both the estimated impact of postdoc training on log(salary) and the R^2 for regressions without any controls ("uncontrolled") and with all of the controls ("controlled") in our most general regression specification. We then calculate the estimated effect of postdoc training on after-postdoc salary (θ^*) given an equal degree of selection on unobservables as selection on observables ($\delta = 1$) and where we select R_{max} as equal to $1.3 * \tilde{R}^2$ where $\tilde{R}^2$ is the R^2 obtained from the controlled regression.

[32] We use the standard errors reported in Table C.1. The results for tenured positions are quite sensitive to selection on unobservables — this makes sense given the sensitivity of the results to selection on observables, paired with the fact that inclusion of the observable controls increases the R^2 drastically relative to the uncontrolled regression.

Table D.2: Sensitivity of Salary Regression Results to Selection on Unobservables by Subsector

Sector:	Academia						Industry			
Subsector:	TT Res.		Non-TT Res.		Nonres.		Res.		Nonres.	
	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2
Panel A. Postdoc Training as Experience										
Unctrld.	-0.0962	0.002	-0.00343	0.000	0.0318	0.001	-0.0254	0.000	-0.101	0.002
Ctrld.	-0.174	0.349	0.159	0.531	-0.0416	0.453	-0.0832	0.482	-0.155	0.499
R_{max}	0.454		$0.611^{\dagger}$		0.589		0.626		0.649	
θ^*	-0.339		$0.546^{\dagger}$		-0.135		-0.232		-0.273	
N	3996		1988		5957		3117		3591	
Panel B. Postdoc Training as Schooling										
Unctrld.	-0.00721	0.001	0.0364	0.000	0.0632	0.002	-0.00163	0.000	-0.0680	0.001
Ctrld.	-0.0500	0.0349	0.232	0.498	0.0481	0.419	0.0162	0.453	-0.0707	0.473
R_{max}	0.454		$0.572^{\dagger}$		0.544		0.589		0.615	
θ^*	0.00102		$0.573^{\dagger}$		0.0316		0.0519		-0.0756	
N	4394		2408		7145		3801		4097	

Notes: We test if the results in columns (4) and (6) of Table 3 are robust to allowing for selection on unobservables using the methods developed in Oster (2019) in Panel A and Panel B, respectively; see notes to Table 3. We report both the estimated impact of postdoc training on log(salary) and the R^2 for regressions without any controls ("uncontrolled") and with all of the controls ("controlled") in our most general regression specification. We then calculate the estimated effect of postdoc training on after-postdoc salary (θ^*) given an equal degree of selection on unobservables as selection on observables ($\delta = 1$) and where we select R_{max} as equal to $1.3 * \tilde{R}^2$ where $\tilde{R}^2$ is the R^2 obtained from the controlled regression. † = we set $R_{max} = 1.15 * \tilde{R}^2$ since $1.3 * \tilde{R}^2$ exceeds the R^2 obtained from a controlled regression with person fixed effects.

Table D.3: Sensitivity of Research Job Regression Results to Selection on Unobservables

Research Job Type:	Any		Academic		Tenure-Track		Tenured		Industry	
	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2	$\hat{\theta}$	R^2
Uncontrolled	0.258	0.062	0.258	0.056	0.228	0.060	0.120	0.005	0.153	0.023
Controlled	0.242	0.296	0.265	0.269	0.213	0.263	-0.0634	0.680	0.122	0.492
R_{max}	0.384		0.349		0.342		0.884		0.640	
θ^*	0.231		0.271		0.202		-1.47		0.090	
N	4778		4778		4778		798		1786	

Notes: We test if the results in Panel B of Table C.1 are robust to allowing for selection on unobservables using the methods developed in Oster (2019); see notes to Table C.1. We report both the estimated impact of postdoc training on obtaining research jobs and the R^2 for regressions without any controls ("uncontrolled") and with all of the controls ("controlled"). We then calculate the estimated effect of postdoc training (θ^*) given an equal degree of selection on unobservables as selection on observables ($\delta = 1$) and where we select R_{max} as equal to $1.3 * \tilde{R}^2$ where $\tilde{R}^2$ is the R^2 obtained from the controlled regression.

E A Task-Based Framework of Wage Determination

Our conceptual framework represents a dynamic extension of the model in Autor and Handel (2013) where workers augment their skills over time through the performance of tasks. We write worker i's stock of skills at time t as $\Phi_{it} = \{\phi_{it}^1 \dots \phi_{it}^J\}$ where each $\phi_{it}^j > 0$ gives worker i's stock of task j specific human capital at time t which is measured in the units of task j that worker i can perform in a unit of time ("task efficiency"). Assume worker i produces output in sector $k \in \{1, \dots, K\}$ by utilizing task-specific skills ϕ_{it}^j for $j \in \{1, \dots, J\}$ as follows:

$$Y_{ikt} = e^{\alpha_k + \sum_j \lambda_k^j \phi_{it}^j}, \tag{5}$$

where $\lambda_k^j \geq 0 \; \forall \; j, k$ measures the productivity of task j in producing output in sector k and where all tasks are performed simultaneously as part of production in each unit of time. As in Autor and Handel (2013), we normalize the output price for each sector to unity, and also note that α_k is not constrained to be positive, thus allowing for a worker's marginal productivity in sector k to be negative in the case of insufficient skills (e.g., an untrained air pilot).

If workers are paid their marginal product, then the log wage of worker i in sector k is:

$$w_{ikt} = \alpha_k + \sum_j \lambda_k^j \phi_{it}^j. \tag{6}$$

We write task j specific human capital as the sum of endowed task j specific ability and task j specific human capital accrued over time (through training or labor market experience):

$$\phi_{it}^j = H_i^j + H_{it}^j. \tag{7}$$

Then plugging (7) into (6) we get:

$$w_{ikt} = \alpha_k + \sum_j \lambda_k^j H_{it}^j + \sum_j \lambda_k^j H_i^j, \tag{8}$$

which shows that wage differences between workers in sector k are the result of differing levels of endowed and/or accrued task-specific human capital.[33]

We assume that task j specific human capital accrual is the result of learning in previous employment (including postdoctoral training) such that:

$$H_{it}^j = \theta_{it}^j \tau_t, \tag{9}$$

[33] We note that it is possible that differences in task-specific human capital do not lead to differences in wages, depending on the relative productivity of each task j in production of output in sector k; that is, differences in task-specific human capital could be perfectly offset by differences in the productivity of each task.

where τ_t gives the number of years spent in previous employment as of year t and θ_{it}^{j} denotes the amount of task j specific human capital accrued per each unit of time performing task j multiplied by the share of years of previous employment spent performing task j.[34] Substituting (9) into (8), we get:

$$w_{ikt} = \alpha_k + \sum_j \lambda_k^j \theta_{it}^j \tau_t + m_{ik}, \tag{10}$$

where $m_{ik} = \sum_j \lambda_k^j H_i^j$ represents worker-sector match quality which is a function of worker skill endowments and sector-specific returns to skills. Equation (10) implies that workers with greater levels of accumulated task-specific human capital in those tasks that are most productive to their current employer will tend to be paid more.

Suppose now that there are two tasks: research (R) and nonresearch (N). Also suppose there are two sectors—academia (A) and industry (I)— and for simplicity assume that all workers in the same sector k accrue task j specific human capital at the same rate so that $\theta_{it}^j \equiv \theta_k^j$. We index sectors of *previous* employment by k' and index the current sector of employment by k as before. Letting $\tau_{ik't}$ give the number of years worker i spent in sector k' as part of previous employment as of year t, equation (10) can be written as:

$$w_{ikt} = \alpha_k + \lambda_k^R \left[\theta_{A'}^R \tau_{iA't} + \theta_{I'}^R \tau_{iI't}\right] + \lambda_k^N \left[\theta_{A'}^N \tau_{iA't} + \theta_{I'}^N \tau_{iI't}\right] + m_{ik}. \tag{11}$$

Also suppose that there are two types of workers p and n of the same level of overall experience (i.e., $\sum_{k'} \tau_{pk't} \equiv \tau_{pt} = \tau_{nt} \equiv \tau_t$) and who both work in industry. Suppose worker p spent all previous years in the academic sector as a postdoc while worker n has worked in industry ever since PhD graduation. Then we have the following:

$$w_{pIt} = \alpha_I + \lambda_I^R \theta_{A'}^R \tau_t + \lambda_I^N \theta_{A'}^N \tau_t + m_{pI},$$
$$w_{nIt} = \alpha_I + \lambda_I^R \theta_{I'}^R \tau_t + \lambda_I^N \theta_{I'}^N \tau_t + m_{nI},$$

where $m_{ik} = \lambda_k^R H_i^R + \lambda_k^N H_i^N$. Then wage differences between workers are due to differences in endowed task-specific human capital and differences in accrued task-specific human capital caused by $\theta_{A'}^R \neq \theta_{I'}^R$ or $\theta_{A'}^N \neq \theta_{I'}^N$.[35]

Let $\theta_\Delta^j \equiv \theta_{A'}^j - \theta_{I'}^j$ and $m_{\Delta I} \equiv m_{pI} - m_{nI}$. Then wages for both types of workers can be written as the following:

$$w_{iIt} = \alpha_I + \lambda_I^R \theta_{I'}^R \tau_t + \lambda_I^N \theta_{I'}^N \tau_t + m_{nI} + 1[i = p] * \left\{\lambda_I^R \theta_\Delta^R \tau_t + \lambda_I^N \theta_\Delta^N \tau_t + m_{\Delta I}\right\}, \tag{12}$$

where $1[i = p] = 1$ if worker i is type p and $1[i = p] = 0$ if worker is type n. Equation (12) implies

[34] A simple proxy for θ_{it}^j is the share of years of previous employment spent performing task j.

[35] A reasonable assumption might be that $\theta_{A'}^R > \theta_{I'}^R$ and $\theta_{A'}^N < \theta_{I'}^N$.

that industry wage differences between postdoc-trained (type p) and nonpostdoc-trained (type n) workers of the same cohort are due to differences in worker-sector match quality $m_{\Delta I}$—which is governed by differences in endowed ability in each task (i.e., differences in H_i^j)—and between-sector differences in the rate of task j specific human capital accumulated as part of production (θ_Δ^j). In this simplified example, we considered the case where a postdoc-trained doctorate is entering the first year of employment in industry. Under the assumption that θ_k^j and λ_k^j remain fixed over time for each sector and do not differ by worker type, differences in task-specific human capital, and thus wage differences, will persist between postdoc-trained and nonpostdoc-trained workers in industry.[36]

[36] Note that the magnitude and direction of the difference is an empirical question: if pure research abilities are more valuable than other types of abilities in industry, then postdoc training could potentially lead to postdoc-trained biomedical doctorates earning more than their nonpostdoc-trained counterparts, assuming that postdoc training is primarily focused on pure research. However, it could be the case that nonresearch skills are sufficiently valued in industry that nonpostdoc-trained workers in industry tend to earn more; allowing for more than two tasks, it could be that the type of research conducted in academia is qualitatively different from that in industry. Lastly, differences in task-specific human capital accrual between postdoc-trained and nonpostdoc-trained biomedical doctorates working in industry could be perfectly offset by differences in the productivity of each task, resulting in equal wages.

F Exploring Alternative Mechanisms for the Industry Postdoc Salary Penalty

Sorting by Occupation or Employer A possible explanation for the postdoc salary penalty in industry is that industry-employed biomedical doctorates with postdoc training tend to sort into different firms or occupations than biomedical doctorates without postdoc training. The SDR contains information on occupation, as well as a limited set of employer characteristics including size, location (state/country code), and type. We therefore estimate regressions where worker occupation, employer size, employer location, and employer type are included as controls.[37] Column ($4''$) of Table F.1 shows that including these controls does not eliminate the industry postdoc salary penalty. While we find no evidence that employer characteristics are a driver of the industry postdoc salary penalty, we cannot rule out this mechanism entirely as employer information in the SDR is limited, and so a linked employer-employee dataset of the doctoral workforce is necessary for a stronger test of this mechanism.[38]

Seniority Pay Biomedical doctorates who forgo postdoc training to enter industry directly after graduation can build up seniority at the firm where they work earlier in their career than their postdoc counterparts. The existence of a return to employer-specific seniority would mean that when postdoc-trained biomedical doctorates enter a firm, they will tend to be paid less than nonpostdoc-trained colleagues, even if they are otherwise identical in terms of skill.[39] In each SDR wave, respondents are asked if they have the same employer as in the last SDR wave. Using responses to these questions, we construct a variable that measures seniority (i.e., how many years an individual has been at their current employer as of the given year) and augment our specification by including a quartic polynomial in seniority. Column ($4'''$) of Table F.1 gives the results: we find that including seniority as a control in the regressions does not diminish the estimated postdoc penalty in industry.

[37] Employer types in the industry employment sector include the following: 1) Private-for-profit, 2) Self-employed, not incorporated, 3) Self-employed, incorporated, and 4) Other. See SDR survey questionnaire for list of occupation codes. We use occupation-by-year fixed effects to control for occupation as this both allows the impact of a given occupation to change over time and also is robust to changes in occupational codes in the SDR that have occurred over time.

[38] Davis et al. (2021) uses American Community Survey (ACS) and Longitudinal Employer-Household Dynamics (LEHD) data to create a new employer-employee linked dataset of the doctoral workforce. Davis et al. (2021) contains a preliminary analysis of the returns to postdoc training for biomedical doctorates and finds that the postdoc salary penalty for nonacademic jobs remains after including both firm fixed effects and occupation fixed effects, although the magnitude of the penalty is reduced relative to specifications not including these controls. Given the differences in the data sources, and thus samples, used in this paper and in Davis et al. (2021), the results are not directly comparable—see Davis et al. (2021) for a fuller discussion.

[39] Barth (1997) finds evidence of within-firm seniority pay not explained by firm-specific human capital accumulation using Norwegian microdata.

Table F.1: Industry Postdoc Salary Premium with Alternative Mechanisms as Controls

Dependent Variable: log(salary)	(4)	(4′)	(4″)	(4‴)
	$N = 6708$		$N = 6392$	
Postdoc Training	-0.158***	-0.180***	-0.193***	-0.190***
	(0.0410)	(0.0426)	(0.0400)	(0.0402)
R^2	0.400	0.403	0.522	0.522
Postdoc Training Treated As:				
Experience	✓	✓	✓	✓
Schooling				
Controls				
Baseline	✓	✓	✓	✓
Research and Management Job indicators		✓	✓	✓
Firm Characteristics & Occupation FE			✓	✓
Seniority				✓

Notes: See notes for column (4) in Table 3. Here we add controls for potential mechanisms that could drive the relationship between postdoc training and after-postdoc salary. All specifications include field-cohort fixed effects, year fixed effects, and PhD university fixed effects. Postdoc training is treated as experience such that experience is defined as years since PhD graduation for all biomedical doctorates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

G Does Postdoc Spell Duration Matter?

The results reported in column (4) of Table 3 estimate the impact of postdoc training on future salary, regardless of the length of postdoc training. If differences in salary between ex-postdocs and nonpostdocs in industry are driven by differences in task-specific human capital, we would expect ex-postdocs who spent the longest time in postdoc training—and therefore deferred on-the-job training in industry the longest—to suffer the largest after-postdoc salary penalties. To test this, we repeat the analysis in Table 3 after replacing the single indicator variable for if a biomedical doctorate is postdoc-trained with three indicator variables based on whether a doctorate participated in postdoc training for 1) no longer than three years, 2) greater than three years but less than six years, and 3) exceeding six years. Table G.1 reports the results. We first focus attention to specification (4) where postdoc training is treated as employment experience. The results suggest that postdocs finding a job in academia do not suffer a salary penalty regardless of how long they are employed as a postdoc. However, biomedical doctorates who spend any number of years employed as a postdoc experience a salary penalty in excess of 10% in industry, with those who spend the most time working as a postdoc suffering the largest penalty. In specification (6) we treat postdoc training as a form of schooling and find that the postdoc penalty in industry is no longer statistically significant for postdocs of any length. We also detect increases in after-postdoc salary for biomedical doctorates that spend greater than three years in postdoc positions and who find employment in academia; those with the longest postdocs tend to earn more, possibly due to postdoc employment serving as a holding position as one waits for an academic position at a research-intensive university, which are typically higher-paying than other entry-level positions in academia.[40]

To test whether the chances of obtaining a research job in academia, including a tenure-track research position, are increasing in the length of postdoc training, we repeat the analysis in Table C.1 after replacing the single indicator variable for if a biomedical doctorate is postdoc-trained with the three indicator variables based on postdoc length. Panel B of Table G.2 shows that biomedical doctorates employed in postdoc positions of any length have greater chances than nonpostdocs in obtaining academic research and tenure-track research positions, with those with postdoc lengths exceeding three years having the greatest chances on landing these positions. Additionally, biomedical doctorates with postdoc lengths greater than three years are also more likely to obtain a research position in industry than those without any postdoc experience. The likelihood that a tenure-track researcher obtains tenure does not appear to be impacted by postdoc length. In general, doing a postdoc longer than three years leads to significantly greater chances of landing an academic research position, a tenure-track research position, and an industry research position.

[40] Andalib, Ghaffarzadegan, and Larson (2018) model postdoc positions using a queuing model. Cheng (2023) finds that remaining in postdoc training for longer periods increases the chances of securing a non-tenure-track academic position at research-intensive institutions.

Table G.1: Postdoc Salary Premia by Postdoc Length

Dependent Variable: log(salary)	(3)	(4)	(5)	(6)
Panel A. All Sectors	$N = 22512$		$N = 26312$	
0 years < Postdoc Length ≤ 3 years	-0.0245	-0.0535*	0.0142	-0.00439
	(0.0293)	(0.0291)	(0.0259)	(0.0262)
3 years < Postdoc Length ≤ 6 years	-0.0728***	-0.115***	0.0430*	0.00877
	(0.0252)	(0.0254)	(0.0225	(0.0228)
Postdoc Length > 6 years	-0.222***	-0.231***	-0.0135	-0.0134
	(0.0290)	(0.0293)	(0.0270)	(0.0277)
Panel B. Academia	$N = 11941$		$N = 13947$	
0 years < Postdoc Length ≤ 3 years	0.0609*	0.00406	0.0904***	0.0466
	(0.0361)	(0.0404)	(0.0323)	(0.0358)
3 years < Postdoc Length ≤ 6 years	0.0453	0.00226	0.155***	0.122***
	(0.0327)	(0.0360)	(0.0290)	(0.0314)
Postdoc Length > 6 years	-0.0537	-0.0517	0.133***	0.156***
	(0.0361)	(0.0408)	(0.0333)	(0.0395)
Panel C. Industry	$N = 6708$		$N = 7898$	
0 years < Postdoc Length ≤ 3 years	-0.0435	-0.122**	-0.0129	-0.0628
	(0.0523)	(0.0482)	(0.0459)	(0.0451)
3 years < Postdoc Length ≤ 6 years	-0.0942**	-0.139***	0.00540	-0.0238
	(0.0468)	(0.0458)	(0.0428)	(0.0433)
Postdoc Length > 6 years	-0.264***	-0.283***	-0.0791	-0.0736
	(0.0587)	(0.0620)	(0.0565)	(0.0595)
Panel D. Gov't/Nonprofit	$N = 3863$		$N = 4467$	
0 years < Postdoc Length ≤ 3 years	-0.0713	-0.112*	-0.0216	-0.0412
	(0.0483)	(0.0678)	(0.0440)	(0.0586)
3 years < Postdoc Length ≤ 6 years	-0.0329	-0.0762	0.0945***	0.0450
	(0.0370)	(0.0480)	(0.0347)	(0.0432)
Postdoc Length > 6 years	-0.267***	-0.171**	-0.00833	0.0918
	(0.0548)	(0.0681)	(0.0544)	(0.0616)
Postdoc Training Treated As:				
Experience	✓	✓		
Schooling			✓	✓
Fixed Effects				
Field + Cohort + Year	✓		✓	
Field-Cohort + PhD University + Year		✓		✓

Notes: See notes for columns (3) through (6) in Table 3. The only change relative to Table 3 is that we replace a single indicator variable for postdoc training with a set of three indicator variables based on a doctorate's length of postdoc training. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G.2: Impact of Postdoc Length on Securing Job Type

	Any	Academic	Tenure-Track	Tenured	Industry
Panel A. Any Job					
0 years < Postdoc Length ≤ 3 years	...	0.0867***	0.111***	-0.0231	...
	...	(0.0258)	(0.0251)	(0.0634)	...
3 years < Postdoc Length ≤ 6 years	...	0.199***	0.201***	-0.00460	...
	...	(0.0232)	(0.0223)	(0.0533)	...
Postdoc Length > 6 years	...	0.234***	0.177***	-0.0299	...
	...	(0.0275)	(0.0298)	(0.0728)	...
R^2	...	0.256	0.270	0.460	...
N	...	4778	4778	1583	...
Panel B. Research Job					
0 years < Postdoc Length ≤ 3 years	0.138***	0.139***	0.105***	0.106	0.0578
	(0.0253)	(0.0245)	(0.0180)	(0.182)	(0.0518)
3 years < Postdoc Length ≤ 6 years	0.285***	0.321***	0.260***	-0.0286	0.165***
	(0.0215)	(0.0228)	(0.0179)	(0.177)	(0.0487)
Postdoc Length > 6 years	0.312***	0.340***	0.281***	-0.0949	0.131**
	(0.0260)	(0.0282)	(0.0248)	(0.197)	(0.0609)
R^2	0.308	0.285	0.280	0.682	0.496
N	4778	4778	4778	798	1786
Fixed Effects					
Field-Cohort	✓	✓	✓	✓	✓
PhD University	✓	✓	✓	✓	✓

Notes: See notes to Table C.1. The only change relative to Table C.1 is that we replace a single indicator variable for postdoc training with a set of three indicator variables based on a doctorate's length of postdoc training. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$