DRUID Working Paper No. 11-03

Not All Scientists pay to be Scientists: Heterogeneous Preferences for Publishing in Industrial Research

By

Henry Sauermann and Michael Roach

Danish Research Unit for Industrial Dynamics

www.druid.dk

Not All Scientists pay to be Scientists: Heterogeneous Preferences for Publishing in Industrial Research

Henry Sauermann
Georgia Institute of Technology
E-mail: herny.sauermann@mgt.gatech.edu

Michael Roach
University of Northe Carolina, Chapel Hill
e-mail: Michael.roach@unc.edu

Abstract:
A growing body of research on firms' "open science" strategies rests on the notion that scientists have a strong preference for publishing and that firms are able to extract a wage discount if they allow scientists to publish. Drawing on a survey of 1,400 life scientists about to enter the job market, we suggest an alternative view. First, we show significant heterogeneity in the price scientists assign to the opportunity to publish in firms, and those scientists who seek industry careers have particularly low preferences for publishing. Thus, many job applicants are not willing to accept lower wages for jobs that let them publish and firms pursuing open science strategies may instead have to pay publishing incentives that fulfill both sorting and incentive functions. Second, we show that scientists with higher ability have a higher price of publishing but also expect to be paid higher wages regardless of the publishing regime. Thus, they are not cheaper to hire than other scientists if allowed to publish, but they are more expensive if publishing is restricted. Finally, we show that scientists publish not simply for "peer recognition" but also for more specific reasons, including the opportunity to advance science or to move to higher-paying jobs. Different reasons predict what price a scientist assigns to the opportunity to publish and may also have very different implications for the sustainability of competitive advantages derived from open science strategies.

Keywords: Scientists; publishing; competitive advantage
Jel codes: O31 ; L82

ISBN 978- 87-7873-314-6

Acknowledgments
We thank participants in various seminars and conferences for comments. We thank especially Howard Aldrich, Rudi Bekkers, Iain Cockburn, Annamaria Conti, Jeff Edwards, Jeff Furman, Kostas Grigoriou, Bart Hamilton, Kenneth Huang, Nico Lacetera, Cornelia Meissner, Scott Rockart, Tim Simcoe, Deepak Somaya, Paula Stephan, Scott Stern, Peter Thompson, Jerry Thursby, Long Vo, and Dave Waguespack. We thank the Georgia Research Alliance and the Marion Ewing Kauffman Foundation for financial support.

1 Introduction

A growing body of research on the organization of industrial science rests on the notion that scientists have a strong desire to publish their research results and to participate in open science. For example, it has long been argued that the norms and values of academically trained scientists conflict with the closed and secretive logic of the private sector (Kornhauser, 1962; Miller, 1976; Shepard, 1956). More recently, it has been suggested that firms can attract and retain academically trained scientists by allowing them to publish research results. In return, firms may be able to extract a wage discount from scientists (Gans et al., 2010; Stern, 2004) while also garnering potential productivity benefits from interacting with the scientific community (Cockburn & Henderson, 1998; Gittelman & Kogut, 2003; Lim, 2004; Penin, 2007; Simeth & Raffo, 2011).

Even though the assumption that scientists value publishing is intuitive, empirical evidence is scarce and typically relies on indirect inferences from observed firm policies. For example, Stern (2004) showed that jobs that allow scientists to publish offer lower base wages, suggesting that scientists may be willing to "pay" to publish by taking a wage cut. Yet, many firms also provide explicit incentives for publishing (Cockburn & Henderson, 1998; Stuart & Liu, 2010), suggesting that scientists' own desire for publishing may be low. Given these conflicting empirical observations, it is not clear whether open science policies are associated with lower or higher levels of total pay.[1] More importantly, observed firm policies provide little insight into scientists' preferences because these policies are consistent with very different distributions of underlying preferences (cf. Lacetera & Zirulia, 2011; Lazear, 2000). Thus, our empirical understanding of scientists' preferences for publishing remains limited and important assumptions remain untested.

In this paper, we draw on survey data from 1,400 U.S. life sciences PhD candidates to provide insights into the price junior scientists assign to the opportunity to publish in industrial science. Our results challenge common assumptions regarding scientists' preferences, provide a deeper understanding of the distribution and nature of these preferences, and allow us to reconcile and re-interpret observed firm policies. First, we show that there is significant heterogeneity in scientists' preferences and that many scientists care little about publishing. Moreover, sorting effects are such that scientists seeking jobs in industry have particularly low preferences for publishing; in our sample 33% of them assign a price of zero to the opportunity to publish. This significant heterogeneity suggests that firms' publishing policies may not simply serve to accommodate scientists' taste for publishing. Instead, firms that seek to garner productivity benefits from engaging in open science may combine lower base wages (Stern, 2004) and

[1] The notion of "open science policies" may capture a wide range of aspects, including not only publishing but also conference attendance, industry-academia collaborations, etc. In this paper, we focus on publishing as a particularly important aspect.

publication-based incentives (Stuart & Liu, 2010) to serve both sorting and incentive functions (cf. Lacetera & Zirulia, 2011; Lazear, 2000).

Second, we show that scientists with higher ability have a higher price of publishing but also expect to be paid higher wages regardless of the publishing regime. Thus, they are not cheaper to hire than other scientists if allowed to publish, but they are more expensive if publishing is restricted because they require a larger wage premium to offset the lack of publishing opportunities. The relationships between ability, preferences for publishing, and wage expectations suggest important trade-offs for employers who need to consider disclosure policies in the context of scientific productivity.

Finally, we show that despite the attention placed on publishing for peer recognition (cf. Gans et al., 2010; Merton, 1973), scientists also publish for a variety of other and more specific reasons such as the ability to contribute to the advance of science or to obtain higher future pay. More importantly, these reasons have different relationships with the price scientists assign to publishing, possibly because they imply different degrees of substitutability between publishing and money. In addition, these reasons may have different implications for firms' ability to retain productive scientists and to gain a sustainable competitive advantage from open science strategies.

In the following section, we provide background on the role of publishing in the scientific system, discuss determinants of the price scientists assign to publishing opportunities in industry, and consider potential selection effects associated with the choice to pursue jobs in industry versus academia. In section 3, we describe the data and measures. In section 4, we present descriptive insights into the distribution of the price to publish and examine sources of heterogeneity. In section 5, we examine potential self-selection into industrial careers and outline implications for firms. Section 6 provides a summary and a discussion of opportunities for future research.

2 Conceptual background

2.1 Publishing and the institution of science

The sociology and economics of science highlight the importance of publishing for the progress of the scientific enterprise. Publishing facilitates the evaluation and diffusion of new knowledge in the scientific community, thus allowing researchers to build on existing knowledge in a cumulative fashion (Murray & O'Mahony, 2007; Nelson, 2004; Sorenson & Fleming, 2004). For the individual scientist, publishing is a mechanism to establish priority in discovery and to gain recognition from scientific peers (Dasgupta & David, 1994; Merton, 1973; Stephan, 2004). While these literatures have focused on science as exemplified in academia, publishing also plays an important role for firms in industry. For example, publishing enables firms to connect with the broader scientific community and to benefit from extramural knowledge (Cockburn & Henderson, 1998; Gittelman & Kogut, 2003; Lim, 2004), to pre-empt

competitors from patenting research results (Parchomovsky, 1999), and to signal scientific capabilities to outside stakeholders (Penin, 2007). Most importantly, pro-publishing policies may allow firms to accommodate scientists' taste for publishing and to extract wage discounts compared to a secretive regime (Gans et al., 2010; Stern, 2004). Despite these potential benefits, however, publishing also comes at a cost to the firm if it limits the ability to exclude others from using new knowledge and to appropriate financial returns (Arrow, 1962; Cohen et al., 2000; Dasgupta & David, 1994).

Given these trade-offs, firms' costs and benefits from publishing as well as the way in which disclosure strategies are implemented may depend importantly on the preferences of the scientists a firm employs or seeks to attract. In particular, if scientists assign a high price to the ability to publish research results, then firms pursuing an open science strategy may benefit from a large applicant pool, extract significant wage discounts, and more effectively absorb extramural knowledge. If scientists do not care about publishing, however, the permission to publish is unlikely to result in wage discounts. Moreover, firms that seek to engage in open science to garner knowledge-related benefits may have to provide explicit incentives for publishing.

Prior work typically assumes that all scientists share a strong preference for publishing because they were socialized during their academic training or because the reward system of science (Merton, 1973) spans both the industrial and the academic realm (Gans et al., 2010; Shepard, 1956; Stern, 2004). In contrast, the basic premise of this paper is that scientists may be heterogeneous with respect to their preferences for publishing. While some scientists place a high value on publishing when employed in a firm, others may not. In the words of one of our respondents, "I honestly find publications to be a distraction. That said, if you can find a job where I don't have to publish either papers, or reports, or reviews, or summaries, or some other periodic distraction from doing science then you are better at this game than I am." [2] In the following section, we discuss potential drivers of heterogeneity in the price scientists assign to publishing opportunities in industrial science.

2.2 Sources of heterogeneity in the price of publishing

Our key construct of interest is the price a scientist assigns to the opportunity to publish when employed in a firm. We conceptualize the price of publishing ("PricePub") as the additional amount of pay at which the particular scientist is indifferent between a research position in industry that allows publishing and a position that restricts publishing but pays a wage premium, holding all other job attributes constant. Equivalently, PricePub can be interpreted as the wage discount the scientist is willing to accept if allowed to publish versus not allowed to publish. Empirically demonstrating heterogeneity in

[2] In addition to the quantitative measures used in the empirical analysis, we also asked respondents why – if at all – they valued publishing opportunities in firms. Detailed results from our analysis of open ended answers are reported in Appendix A.

PricePub is a key objective of this paper. In addition, we seek to understand potential sources of heterogeneity in the price of publishing, including differences in scientists' fundamental preferences for publishing and pay, reasons for publishing, and research ability.

2.2.1 Preferences for publishing and money

The price a scientist assigns to the opportunity to publish should reflect two different underlying preferences: the preference for publishing and the preference for money. A scientist with a stronger preference for publishing realizes greater utility from a job that offers publishing, holding other preference and job attributes constant (cf. Goddeeris, 1988; Hwang et al., 1992; Stern, 2004).[3] As such, a larger amount of money is needed to offset the lack of publishing for a scientist with a strong preference for publishing. A stronger preference for money, on the other hand, implies a greater utility from a given unit of money. Thus, a smaller wage premium will be sufficient to offset a given amount of utility lost due to publishing restrictions. Overall, we predict that PricePub increases with the preference for publishing, but decreases with the preference for money.[4] While straightforward, this discussion highlights that the price scientists assign to publishing is likely to reflect not only their preferences for publishing but also those for pay.

2.2.2 Reasons to publish

Drawing on the seminal work of Merton, the prevailing view in the literature is that scientists' primary motive for publishing is the desire to gain peer recognition or "kudos" (Gans et al., 2010; Merton, 1973; Stern, 2004). We suggest that it is useful to also consider additional and more specific reasons why scientists may value publishing because different reasons may have quite different implications for the price scientists assign to publishing. We will focus on two such reasons. First, scientists may value the open dissemination of knowledge because it facilitates the advancement of science (Dasgupta & David, 1994; Merton, 1973; Sorenson & Fleming, 2004). As such, scientist may feel a "repugnance" (Gans & Stern, 2010) towards attempts to limit access to knowledge for private benefit or profit. To quote one of our respondents, "Knowledge belongs to everyone, and the growth of scientific knowledge cannot occur if it is hoarded for money and collaboration and conversation is discouraged." A second reason for publishing we consider is that publications may increase a scientist's labor market value and future pay,

[3] We conceptualize preferences as parameters in the utility function such that a stronger preference for publishing increases the utility derived from the opportunity to publish. Building on the model used by Stern (2004), the utility scientist i derives from job j could be written as $U_{ij} = \lambda_{ij} + \alpha_i PUB_j + \beta_i PAY_j$, where λ_{ij} is the utility derived from a set of unobserved job attributes, α_i is the scientist's preference for publishing, PUB_j indicates whether or not job j allows publishing, β_i is the scientist's preference for pay, and PAY_j is the amount of pay offered by job j.

[4] Individuals with a strong preference for money may well expect higher levels of overall pay. However, we are not concerned with wage levels per se but with the amount of money that creates the same amount of utility as the opportunity to publish, i.e., with the trade-off scientists make between pay and publishing.

as documented in a long line of empirical research in both industry and academia (Konrad & Pfeffer, 1990; Stuart & Liu, 2010; Tuckman & Leahey, 1975). Thus, publications may be seen by scientists as a "currency" in the job market, and giving up current pay in order to build a stock of publications may be seen as an investment into future pay (cf. Becker, 1962; Levin & Stephan, 1991; Rosen, 1986). As described by one of our respondents, "Publishing builds a base where you can get better (i.e. higher-paying) jobs later. Less money now for the chance to make more later."[5]

How important each of these reasons is to the average scientist is an empirical question. Similarly, we cannot predict whether the preference for publishing is stronger when scientists value publishing primarily for its career benefits versus as a mechanism to contribute to the advancement of science. However, we conjecture that for a *given* level of an individual's preference for publishing, *why* a scientist values publishing affects the degree to which money can serve as a substitute for publishing, and thus how much money the scientist needs to be paid to give up publishing. Our argument draws on research in behavioral decision making (Baron & Spranca, 1997; Beattie & Barlas, 2001; Luce et al., 1999), where experimental studies have shown that subjects more easily make trade-offs between similar kinds of attributes, particularly if those attributes are "commodities" such as money or physical goods. On the other hand, individuals are less willing to make trade-offs between dissimilar attributes, especially when "protected" values such as honesty or human lives are involved. In our context, scientists who subscribe to the norm of openness may consider publishing to be more "sacred" and see restrictions on publishing as a violation of their moral code as scientists, leading to a low substitutability between publishing and money. In contrast, substitutability may be high if publishing is valued as a job market "currency" that leads to higher future pay. That is, a scientist who publishes primarily to gain higher future earnings should be more willing to give up publishing in exchange for higher current pay. Overall, we expect that, controlling for preferences for publishing and money per se, those scientists who value publishing primarily as a means to obtain higher future pay assign a lower price to publishing than scientists who value publishing primarily as a means to advance science.

2.2.3 Research ability

Notwithstanding the important role of serendipity in scientific research (Merton & Barber, 2004), scientists with strong research ability are likely to generate more and better knowledge than low-ability

[5] We conceptualize peer recognition as a general outcome that may be related to other more specific outcomes about which scientists ultimately care. For example, some scientists may value peer recognition because it results in better job offers and higher pay, while others value peer recognition because it signals that they have made important contributions to the advance of science. At the same time, some scientists may care about advancing science and not need peer recognition as a validation of their work. Some scientists may also derive direct utility from peer recognition rather than some related indirect outcome. Our empirical analysis includes measures for all three reasons (recognition, pay, and advancing science) and does not rely on assumptions about their relationships. While we focus on these three reasons in the main analysis, our exploratory analysis of open ended responses (Appendix A) suggests several additional interesting reasons that could be considered in future work.

scientists and should also expect more and better publications if they are allowed to publish.[6] The positive effect of ability on publishing output may be reinforced by the "winner-takes-all" nature of the scientific reward system (Freeman et al., 2001; Stephan, forthcoming). A larger quantity and quality of publications, in turn, will translate into higher levels of indirect benefits for the scientist, whether he values publications because they result in peer recognition, higher future income, or allow him to advance science. As a consequence, high-ability scientists should have a stronger preference for publishing, which, in turn, should increase the wage premium they require to give up the permission to publish. Thus, we expect a positive relationship between research ability and the price assigned to publishing, mediated by the preference for publishing (cf. Baron & Kenny, 1986).

A positive relationship between ability and the willingness to pay for publishing has also been suggested by Stern (2004) who argues that this relationship would lead to biases in cross-sectional estimates of compensating differentials between jobs that allow versus prohibit publishing. His novel empirical approach to compare multiple job offers received by a given scientist mitigates this bias by including individual fixed effects for each individual, thus controlling for unobserved heterogeneity in ability as well as other characteristics. However, Stern's analysis does not employ measures of ability and the empirical relationship between research ability and the price scientists assign to publishing remains an important open question.

2.3 Self-selection into industry versus academia

If there is heterogeneity in scientists' preferences for publishing and pay, then scientists may self-select into industrial versus academic research to best satisfy their preferences, potentially resulting in quite different populations of scientists in the two sectors (Agarwal & Ohyama, 2010; Roach & Sauermann, 2010; Rosen, 1986). Industrial research positions generally pay higher wages than faculty positions (National Science Board, 2010). One the other hand, publishing opportunities are likely to be more restricted in firms than in academia, despite the increasing role of "open science" strategies in industry (Dasgupta & David, 1994; Penin, 2007). Thus, we expect that scientists with a strong preference for pay are more likely to prefer a career in industry, while scientists with a strong preference for publishing are more likely to prefer a career in academia.[7] Such self-selection effects would raise further

[6] The relationship between research ability and publishing success may be attenuated if successful publishing depends not only on the quality of research results but also on certain distinct "publishing skills" such as writing an appealing framing for the paper, or conveying the key logic of a conceptual model. While such "publishing skills" play an important role in the social sciences, they are likely less important in the life sciences where research questions are typically well-defined and widely shared, papers tend to be short, and where the focus is on succinctly reporting empirical findings.

[7] Using a small sample of scientists, Roach & Sauermann (2010) provide earlier evidence of sorting based on publishing preferences by showing a positive correlation between the rated importance of publishing and the perceived attractiveness of an academic career. However, that study provides no insights into the price scientists assign to publishing opportunities and thus into wage discounts scientists would be willing to accept in return for publishing opportunities. Moreover, that study did not examine the relationships between preferences and other individual characteristics.

concerns about the common assumption that all scientists have a strong preference for publishing; rather, they would imply that firms face applicants who assign a relatively low value to the opportunity to publish and may not be willing to take a wage cut in return for publishing opportunities.[8]

3 Data and Measures

3.1 Sample and data collection

Our empirical analysis uses data from a national survey of junior scientists in PhD programs at tier one U.S. research universities. Using data from scientists prior to their initial career transition rather than employed scientists has two key advantages. First, the trade-offs between financial and non-financial job attributes such as pay and publishing are particularly salient in the context of initial job and career choices (cf. Agarwal & Ohyama, 2010; Aghion et al., 2008; Stern, 2004), making scientists facing such choices the more relevant sample to study these issues. Moreover, since virtually all scientists go through extensive academic training, our sampling strategy provides insights into a large part of the distribution of scientists' preferences in a given cohort, avoiding potential selection biases associated with samples of scientists who have already entered particular career paths or types of jobs.[9]

To obtain a sampling frame, we identified U.S. research universities with large doctoral programs in science and engineering fields by consulting the National Science Foundation's (2008) reports on earned doctorates. We selected a subset of 39 universities based primarily on program size while also ensuring variation in private/public status and geographic region. We pre-tested the survey in interviews with junior scientists similar to our target population and collected responses in the spring of 2010, approaching individual PhD students and postdoctoral researchers in two ways.

First, we collected roughly 30,000 names and email addresses from listings provided on department websites. We invited these individuals to participate in the survey using a four-contact strategy (one invitation, three reminders). When individual contact information was not available, we approached respondents through their department administrators. In those cases, we emailed administrators with the request to forward a survey link to their graduate students, and our research assistants additionally called administrators on the telephone to encourage their cooperation. Overall, 89% of our responses were obtained directly from respondents and 11% were obtained through administrators.

A concern with survey research is that the particular way in which respondents are approached may lead to sample selection or biased responses (Groves & Peytcheva, 2008). In our context, offering

[8] We follow the established literature in conceptualizing preferences for job attributes as drivers of job and career choices (Agarwal & Ohyama, 2010; Killingsworth, 1987; Rosen, 1986). However, it is possible that pre-existing career intentions also shape preferences for pay and publishing. We address this possibility in the empirical part of the paper.

[9] Since scientists' preferences may change over the life cycle, our results speak to the distribution of preferences in a given cohort of junior scientists, but they do not necessarily reflect the preferences of scientists at different points in the life cycle. Changes in preferences over time are an important issue for future research and will be discussed in the final section of this paper.

financial incentives for responding may increase the likelihood that respondents with above-average preferences for money respond, while a survey without financial incentives may primarily attract respondents with above-average non-pecuniary preferences (e.g., interest in research, helping others, etc.). To address this concern, we randomly assigned respondents into different conditions and varied key aspects of the survey invitation including incentives. This strategy should mitigate selection biases by design and also allows us to explicitly examine the presence and magnitude of such biases. We did not find significant differences across conditions with respect to the key variables considered in this study. Adjusting for 6.3% undeliverable emails, the direct survey approach achieved a response rate of 30%.[10]

For this study, we focus on a subset of respondents as follows. First, we limit our sample to respondents in the life sciences. This focus addresses potential field differences in the function of publications or labor market conditions and also allows us to compare our results to prior work conducted in the life sciences (Cockburn & Henderson, 1998; Stern, 2004; Stuart & Liu, 2010). Second, we focus on PhD students in the advanced stage of their studies—those who report that they have successfully completed their qualifying exams or equivalent milestones—since they are likely to have better-formed preferences and to have given more consideration to job and career choices than early-stage students. The final sample used for this study comprises (exactly) 1,400 life scientists at 37 institutions.[11]

3.2 Measures

Price assigned to publishing. Our featured dependent variable is the price a respondent assigns to the opportunity to publish in a firm, PRICEPUB. To obtain this measure, we asked respondents:

> *Assume that you are offered the following two jobs in an established firm. The positions differ only with respect to your opportunities to publish. What would be the <u>minimum</u> starting compensation <u>for you to accept each position</u>?*
>
> *Job 1: Allowed to publish research results*
> *Job 2: NOT allowed to publish research results*

Note that our framing explicitly states that the two jobs differ only with respect to publishing opportunities, and thus holds all other job attributes—such as geographic location, work activities, or resources for research—constant across the two options. Respondents indicated their reservation wage for each job using a sliding scale measure with anchors ranging from \$0 to \$200k.[12] Our measure of the price

[10] We conducted a non-response analysis by comparing responses of early and late respondents (Rogelberg & Stanton, 2007). We found no significant differences except that REASON_RECOG (see measurement section below) is somewhat lower for late respondents. We include controls for survey mode and response time in all regressions.

[11] The institutions with the largest number of cases are UC Davis (8.14%), Johns Hopkins (7.29%), U of Washington (6.79%), U of Michigan (4.71%), U of North Carolina, Chapel Hill (4.65%), U of Illinois (4.43%), U of Wisconsin (4.43), Cornell U (4.52%), U of Florida (4.07%), Emory U (4.00%), U of Chicago (3.71%), Yale U (3.57%) and Duke U (3.57%).

[12] A preceding question in our survey explicitly defined compensation as "including salary, bonuses and stock options". Also, note that our question elicits reservation wages on an annual basis; the price assigned to (not) publishing over longer periods is likely to be much larger than the PRICEPUB measured in our study.

assigned to publishing is computed as the difference between the wage for the job that does not allow publishing (WAGE_NOPUB) and the wage for the job that does allow publishing (WAGE_PUB). Thus, PRICEPUB = WAGE_NOPUB - WAGE_PUB, which is positive when respondents require additional pay to give up publishing or when they are willing to take a wage discount for having the opportunity to publish, and zero when they are indifferent between the two jobs.[13] An attractive feature of using the first-difference between two reservation wages for the same individual as the dependent variable is that it is equivalent to a fixed-effects estimation. As such, this approach controls for individual-level characteristics that affect both reservation wages such as otherwise unobserved differences in ability, outside job market opportunities, or general attitudes towards industry employment. In addition to the absolute measure of PRICEPUB, we also compute a measure that expresses PRICEPUB as a percentage of the reservation wage for the job with publishing: PRICEPUB% = (PRICEPUB/WAGE_PUB)*100.

Preferences for publishing and income. In a second question block we asked, "When thinking about an ideal job, how important is each of the following factors to you?" Respondents rated the items "Ability to publish research results" and "Financial income (e.g., salary, bonus)", which were presented in randomized order. Respondents used 5-point scales ranging from 1 (not at all important) to 5 (extremely important), resulting in the measures PREF _PUB and PREF_INCOME.

Reasons to publish. We asked respondents "To what extent are the following functions of publishing important or unimportant to you personally?". Respondents rated the following reasons on a 5-point scale ranging from 1 (extremely unimportant) to 5 (extremely important) in randomized order:

- Publications are a way to earn recognition from my peers and colleagues (REASON_RECOG)
- Having publications will lead to higher pay in the future (REASON _PAY)
- Publishing research results allows me to contribute to the advancement of knowledge (REASON _CONTR)

To capture the relative importance of our two focal reasons - pay versus contribution to the advancement of knowledge - we created dummy variables indicating which reason was judged as more important (REASON_PAY>REASON_CONTR, REASON_PAY=REASON_CONTR, and REASON_PAY<REASON_CONTR). Note that a simple ratio of the importance measures would be inappropriate because the measures are interval scales rather than ratio scales (Nunnally, 1978).

[13] We dropped from the sample 3.8% of respondents who indicated a negative price of publishing. Our analysis of open-ended responses suggests that some of them thought that the ability to publish comes with pressure to publish. Other respondents may have interpreted the response scales in the opposite direction, i.e., they thought of the dollar figures as indicators of the value they assign to these jobs rather than the wage they require to take the positions. While we are cautious to interpret negative PRICEPUB as valid, it could indicate that some scientists indeed place a negative value on the ability to publish, perhaps because they conceive of this "option" as a burden (Chubin & Hackett, 1990). Also, a small number of individuals indicated a WAGE_NOPUB that was multiple times larger than WAGE_PUB, likely signaling that they were essentially unwilling to take a job that does not allow publishing. To reduce the effect of such outliers on our analysis, we dropped cases where WAGE_NOPUB was more than three times as large as WAGE_PUB (1.5% of cases).

Ability and PhD program quality. We asked respondents to assess their own ability in response to the following question "How would you rate your research ability relative to your peers in your specific field of study?" (ABILITYSELF). The scale ranged from 0 (least skilled, lowest percentile) to 10 (most skilled, highest percentile). The average rating in our sample (6.48) is somewhat higher than the mean of the scale (5), which could reflect higher average objective ability in our sample of students at tier 1 research institutions. At the same time, self-ratings of ability may also reflect overconfidence (cf. Camerer & Lovallo, 1999). Even if ABILITYSELF partly reflects overconfidence, however, it should have a positive association with PRICEPUB since scientists' own perception of their ability (biased or not) should be a primary determinant of the value they assign to the opportunity to publish.

To obtain a measure of PhD program quality, we follow recent work (e.g., Agarwal & Ohyama, 2010; Sauermann & Cohen, 2010) and use the rankings published by the National Research Council (2010).[14] We employ the ranking of a program's "research activity", which reflects factors such as the average number of publications per faculty, citations, as well as grants and awards. We reverse coded the measure such that high scores reflect high program quality (NRC_RANKING).[15]

ABILITYSELF and NRC_RANKING are likely to provide complementary insights. For example, while the NRC ranking measures quality differences across programs, ABILITYSELF may reflect primarily ability differences within programs to the extent that respondents focus on a "local" reference group of peers when assessing their own ability (cf. Greenberg et al., 2007). In addition to differences in research activity, the NRC_RANKING may also capture differences in the quality of research training as well as other department characteristics such as resources or norms with respect to openness. Despite the multifaceted nature of the NRC_RANKING, this measure is of particular relevance because program quality is likely to serve as an important signal in the job market and firms have a strong interest in hiring PhD scientists graduating from top programs.

Preferences for careers in industry versus academia. To examine potential sorting effects, we asked respondents "Putting job availability aside, how attractive do you personally find each of the following careers?" Respondents separately rated "University faculty with an emphasis on research or development" and "Job in an established firm with an emphasis on research or development" on 5-point scales ranging from 1 (extremely unattractive) to 5 (extremely attractive). We used these measures to create three dummy variables reflecting the relative attractiveness of these two careers:

[14] The NRC data as well as detailed descriptions of the data collection and ranking procedure are available at http://www.nap.edu/rdp/. Note that NRC does not publish one single research ranking for each program but a probabilistic range including a 5th percentile and a 95th percentile ranking. We averaged the two rankings to obtain a single measure.

[15] Since graduate students in the life sciences tend to work in their advisor's lab, it would be desirable to use a measure of the advisor's research capabilities rather than the research quality of the broader department. Unfortunately, we do not have an objective measure of advisor quality. However, we asked respondents to assess the research accomplishments of their advisor on a scale equivalent to that used for ABILITYSELF. The results using this measure are very similar to those using NRC_RANKING. We feature the latter measure because it is more objective and stems from a different data source.

INDUSTRY>ACADEMIA, INDUSTRY=ACADEMIA, and INDUSTRY<ACADEMIA, respectively. While these measures do not tell us which scientists will eventually work in industry, they should capture which scientists are more likely to be in the pool of applicants from which firms can draw, which is of primary interest in the context of this study.

Control variables. We routinely control for age (AGE), gender (MALE), U.S. citizenship status (USCITIZEN), and field of study (10 life sciences subfields including cell/molecular biology, biochemistry/biophysics, neuroscience, ecology, genetics, microbiology, immunology, developmental biology, pharmacology, and other biological/life sciences). Since perceived labor market conditions may shape reservation wages and career preferences, we asked respondents "What do you think is the probability that a PhD in your field can find the following positions after graduation (and any potential Post-docs)?" Respondents indicated subjective probabilities for "University faculty with an emphasis on research or development" (JOBAVAIL_ACAD) and "Job in an established firm with an emphasis on research or development" (JOBAVAIL_IND), respectively. Finally, we also include a set of dummy variables indicating whether a survey response was elicited directly or via a departmental administrator and after which of the four direct survey contacts a response was received (RESPONSEMODETIME). We introduce additional measures used for supplementary analyses below.

3.3 Advantages and limitations of the empirical strategy

Our approach to elicit reservation wages for hypothetical job offers has several advantages over approaches that infer preferences from real job offers or from equilibrium compensating differentials emerging in the job market. First, if scientists are heterogeneous with respect to their preferences and self-select into organizations offering different job attributes, then compensating differentials observed in the labor market reflect the preferences of the marginal individual and provide limited insights into the larger distribution of preferences (Aghion et al., 2008; Charles & Guryan, 2008; Killingsworth, 1987; Rosen, 1986). In contrast, our approach provides information on the preferences of each individual, thereby allowing us to examine the full distribution of preferences in a cohort of scientists.

Second, our approach provides direct insights into the supply side of the labor market (scientists) while holding constant the influence of the demand side (employers). This aspect is particularly important given that equilibrium compensating differentials observed in the labor market confound the supply-side "preference effect" (i.e., scientists' preferences for publishing lead them to accept lower pay) with an offsetting demand-side "productivity effect" (i.e., science-oriented firms that allow publishing may be more productive and thus able to pay higher wages) (Stern, 2004).

Third, our approach identifies publishing policies as the only difference between the two jobs, allowing us to attribute wage differences to differences in publishing opportunities. In contrast, empirical

work analyzing labor market compensating differentials or real job offers faces the challenge of disentangling publishing policies from correlated job attributes such as freedom, the nature of the research activity, or access to resources, some of which may be unobserved (Brown, 1980).

Finally, a key challenge in empirical work is to account for unobserved ability, which may lead to an upward bias in the correlation between pay and publishing and to a downward bias in estimates of compensating differentials (Brown, 1980; Gibbons & Katz, 1992; Hwang et al., 1992). While individual fixed effects can be used to address this problem (Stern, 2004), that approach confounds ability with other unobservable factors. Our measures of ability and program quality allow us to explicitly examine the substantively important relationships between these measures and scientists' preferences.

Despite these advantages, our empirical approach also has its limitations. One concern is that stated reservation wages for hypothetical jobs may not be reliable indicators of the amount of money scientists would require in a realistic choice situation. Somewhat mitigating this concern, we show below that the stated reservation wages are close to actual starting salaries. Moreover, since our primary interest is in the distribution of PRICEPUB and in the drivers of individual differences, factors that influence all responses in similar ways (e.g., generally overstated reservation wages) should not bias our key results.

A second concern is that our question frames opportunities for publishing as dichotomous, while disclosure in firms may be a continuum that involves intermediate levels of openness such as the disclosure of some results but not others, disclosure with a delay, or combinations of patenting and publishing (Gans et al., 2010; Gittelman & Kogut, 2003; Hackett, 1990; Hicks, 1995). The open-ended responses (see Appendix A) do not indicate that this framing was a problem for respondents and we do not expect that it produces any biases in the relationships among the variables of interest. However, the price scientists assign to "partial" openness is likely to be lower than the PRICEPUB observed in our study. Thus, our finding that many scientists assign a relatively low price to publishing in industry should be reinforced if even jobs "with publishing" impose certain restrictions.

A general concern with self-reported preferences is that individuals may overstate preferences that they think are socially desirable (e.g., publishing) and give artificially low scores to preferences that may seem less socially desirable (e.g., money) (Moorman & Podsakoff, 1992). To mitigate this problem, we stated clearly in the survey invitation that responses would be kept strictly confidential and we emphasized that results would not be shared with the respondents' department or university. We also find that the average preference rating for income is higher than that for publishing, suggesting that social desirability bias is likely to be limited.

Finally, a concern with measures drawn from the same survey instrument is that relationships between variables may be spurious due to common methods bias (Podsakoff et al., 2003). To mitigate common methods bias, we used different question formats including rating scales, slider scales, as well as

open ended questions. We also placed questions regarding dependent and independent variables on different pages of the survey questionnaire and separated them by unrelated questions (Podsakoff et al., 2003). Common methods bias is not a concern regarding the NRC ranking since this measure was obtained from an independent data source.

4 Heterogeneity in the price of publishing

Table 1 provides summary statistics and table B1 (Appendix) shows correlations. The average reservation wage for a job that does not allow publishing (WAGE_NOPUB) is $93.65k, compared to $71.09k for the job that allows publishing (WAGE_PUB). The average price of publishing, the difference between the two wages, is $22.56k. In relative terms, the price of publishing amounts to a 33% premium over the wage with publishing for the average scientist.[16] Figure 1 plots the cumulative distribution of PRICEPUB%, showing the share of scientists who value publishing at or below a certain price.

Figure 1: Cumulative distribution of PRICEPUB%

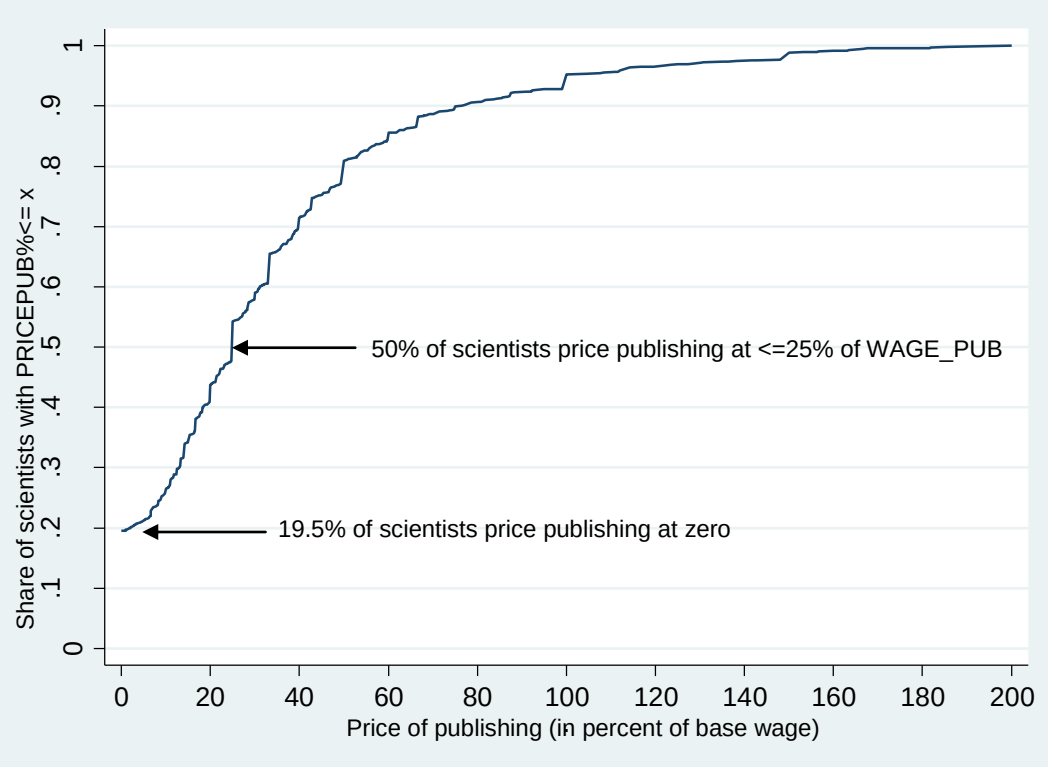


Figure 1 shows that 19.5% of respondents assign a price of zero, i.e., they would not accept a wage discount for the permission to publish, or, equivalently, they would not require a wage premium if

[16] The median WAGE_PUB is $70k, and the median WAGE_NOPUB is $85k. This compares to 2006 median starting salaries of $74k for life scientists in industry (National Science Board, 2010, table 3-20). Thus, the reservation wages reported by our respondents are consistent with actual salary figures.

publishing were restricted. At the top end of the distribution, a small share of scientists assign extremely high values to publishing, essentially being unwilling to take jobs that prohibit publishing (see also quotes in Appendix A). Thus, contrary to common assumptions, figure 1 illustrates that there is considerable heterogeneity in the price scientists assign to publishing in industrial research jobs. In the next section, we examine potential sources of this heterogeneity, focusing on preferences for publishing and income, reasons to publish, and research ability.

4.1 Predictors of the price of publishing

We use regression analysis to examine predictors of the price of publishing. Since PRICEPUB takes on only non-negative values, we employ Poisson quasi-maximum likelihood estimation. Poisson QML is consistent for integer and continuous non-negative data, so long as the conditional mean is correctly specified (Ding et al., 2010; Santos Silva & Tenreyro, 2006; Wooldridge, 1997). Robustness checks using alternative estimation techniques are reported below. We allow for heteroskedasticity and intraclass correlations by clustering standard errors at the level of the university.

The results are reported in table 2. In model 1 we examine the baseline relationship between PRICEPUB and scientists' preferences for publishing and pay, respectively. As expected, we find that the price a scientist assigns to the opportunity to publish in industry increases with his preference for publishing and decreases with his preference for income. A one-standard deviation increase in PREF_PUB is associated with a 56.8% higher PRICEPUB, while the same change in PREF_INCOME is associated with a 12.7% lower PRICEPUB. While the price of publishing is most strongly associated with a scientist's preference for publishing as we would expect, these results suggest that simultaneously considering the preference for money provides a more complete understanding.

We next include measures of the different reasons for publishing (model 2) while controlling for the preference for publishing and income per se. The extent to which publishing is valued as a mechanism to contribute to the advancement of knowledge is associated with a higher PRICEPUB, consistent with our expectation that money is less of a substitute for that function of publishing. In contrast, the extent to which publishing is valued because it may result in higher future pay has a significant negative coefficient, consistent with the idea that money and publications are better substitutes if publications are seen primarily as a way to get higher-paying jobs. The extent to which publishing is valued because it may result in peer recognition does not have a significant coefficient. The latter result may reflect that recognition is a more general outcome that relates to a variety of qualitatively different ultimate objectives, with no clear implications for the substitutability between money and publishing.[17] Model 3 includes dummy variables indicating which of our two focal reasons is relatively more important. Holding

[17] The qualitative results reported in this section hold when we enter the measures of reasons individually rather than jointly.

preferences for publishing and pay constant, scientists who value publishing primarily as a means to obtain higher-paying jobs assign a 27% lower price to publishing than those who value publishing primarily as a mechanism to contribute to public knowledge.

Model 4 shows that the price of publishing increases with scientists' subjective ability and with the quality of the PhD program. More concretely, a one-standard deviation higher score of ABILITYSELF is associated with a 7% higher PRICEPUB, while a one-standard deviation higher NRC_RANKING is associated with a 9.8% higher PRICEPUB. Model 5 additionally includes the measures of preferences and suggests that the effect of ABILITYSELF is largely mediated by preferences for publishing and pay. Thus, scientists who believe they are at the top of the ability distribution have a stronger preference for publishing, which in turn implies a higher price of the opportunity to publish.

4.2 Predictors of reservation wages

Our featured regressions inform us how independent variables are related to the price of publishing, which is the difference between the reservation wage for the job without publishing (WAGE_NOPUB) and the reservation wage for the job with publishing (WAGE_PUB). However, these regressions do not reveal the extent to which independent variables affect PRICEPUB via WAGE_NOPUB versus WAGE_PUB. To provide additional insights, we analyze the two reservation wages separately. Since we examine two dependent variables for the same individuals, we estimate these models using multivariate regression, i.e., two OLS models are estimated simultaneously and the error terms are allowed to have nonzero correlations (Edwards, 1995).

Model 1 shows that the preference for publishing has no significant effect on WAGE_PUB but it has a large positive effect on WAGE_NOPUB. This observation is consistent with our argument that individuals with a strong preference for publishing "lose" more utility when publishing is restricted, thus requiring a larger amount of pay to offset that utility loss. The importance of income has a positive association with WAGE_PUB, likely reflecting that individuals who care strongly about money generally ask for higher base levels of salary. The preference for income does not significantly affect WAGE_NOPUB, which likely reflects that individuals with a strong preference for money desire higher salaries generally but also require a smaller wage premium to offset the lack of publishing. The reasons why scientists value publishing affect the price of publishing primarily via WAGE_NOPUB (model 2).

Model 3 shows that the ability measures increase both reservation wages but have a larger effect on the reservation wage for the job that restricts publishing, resulting in a positive net effect on the price of publishing. Thus, even though high-ability individuals assign a higher price to publishing and may give up more pay if publishing is allowed versus restricted, they still require higher levels of pay than low-ability individuals in the same publishing regime.

4.3 Supplementary analyses and robustness checks

Our observation that ability is associated with higher reservation wages regardless of the firm's publishing policy raises the question whether high ability scientists assign a higher price to publishing only in absolute terms (PRICEPUB) or also relative to the base wage. To examine this question, we run our key models using PRICEPUB% (the price of publishing as a percentage of WAGE_PUB). The results are reported in Table 4, models 1-3. The NRC_RANKING measure has a significant positive coefficient; a one-SD higher NRC_RANKING is associated with a 9% higher PRICEPUB%. The coefficient on ABILITYSELF is positive, but it is not significant. Thus, while the relationship between measures of ability and PhD program quality and the price of publishing in relative terms is positive, it is weaker than that with the price of publishing in absolute terms.

Second, we further probe whether the preference for publishing mediates the effect of ABILITYSELF on PRICEPUB by regressing PREF_PUB on ABILITYSELF and on NRC_RANKING (cf. Baron & Kenny, 1986). ABILITYSELF strongly predicts the preference for publishing (model 4), consistent with the mediation mechanism suggested in our discussion. In contrast, we find no significant relationship between NRC_RANKING and the preference for publishing. While we are unable to determine the underlying source of the relationship between NRC_RANKING and PRICEPUB, it remains important from a managerial perspective. To address the concern that correlations between ABILITYSELF and PREF_PUB may reflect common methods bias, we also estimate a regression of PREF_INCOME but find no significant relationship with ABILITYSELF (model 5).

Table 4 reports a series of additional robustness checks. First, we re-estimated key models using tobit regression (models 6 and 7). The results are consistent with our earlier findings.

A potential concern is that our respondents have not thought much about career and job choices. To address this concern, we restrict the sample in two ways. First, we limit the sample to those individuals who indicated that they were looking for a job either at the time of the survey or anticipated entering the job market in the following year. Additionally, we asked respondents "Generally speaking, to what extent have you thought about your future career plans?", and keep only individuals who indicated that they had thought about their career plans at least to "some extent" (3 or above on a 5-point scale). We find that the coefficients are robust and in many cases larger than in the full sample (models 8 and 9).

A related concern is that some individuals may not have authored any publications and may not have thought about the importance and meaning of publishing to them. Note that publishing is quite common among life science PhD students; in our sample, 77% of individuals have at least one publication and the median respondent has two publications. The results using only cases with at least one publication (models 10 and 11) are qualitatively identical to our featured results.

Next, we examine whether the relationships among our key constructs reflect differences in scientists' research activities. For example, respondents engaged in basic research might place a higher value on publishing and also report stronger motives to contribute to the stock of public knowledge than those working on applied questions. We asked respondents to rate on three separate 5-point scales the extent to which their current work contributes fundamental insights (basic research), creates knowledge to solve practical problems (applied research), and uses knowledge to develop materials, devices, or software (development). As expected, we find that PRICEPUB increases with the extent to which respondents are engaged in basic or applied research (models 12 and 13). More importantly, however, including these additional controls does not affect our results regarding reasons to publish. The coefficient of ABILITYSELF decreases marginally, reflecting that research ability increases with the extent to which PhD candidates are actually engaged in research.

Although not presented in table 4, we also considered the possibility that respondents may associate differences in job attributes other than the opportunity to publish with the two jobs, even though the question stated that the positions differ only with respect to opportunities to publish. Our analysis of open-ended responses (Appendix A) shows that about 20 individuals associated publishing with other job attributes such as freedom, trust, more interesting research, or a generally more "open" atmosphere. Excluding these individuals from the analysis does not change our results (available upon request).

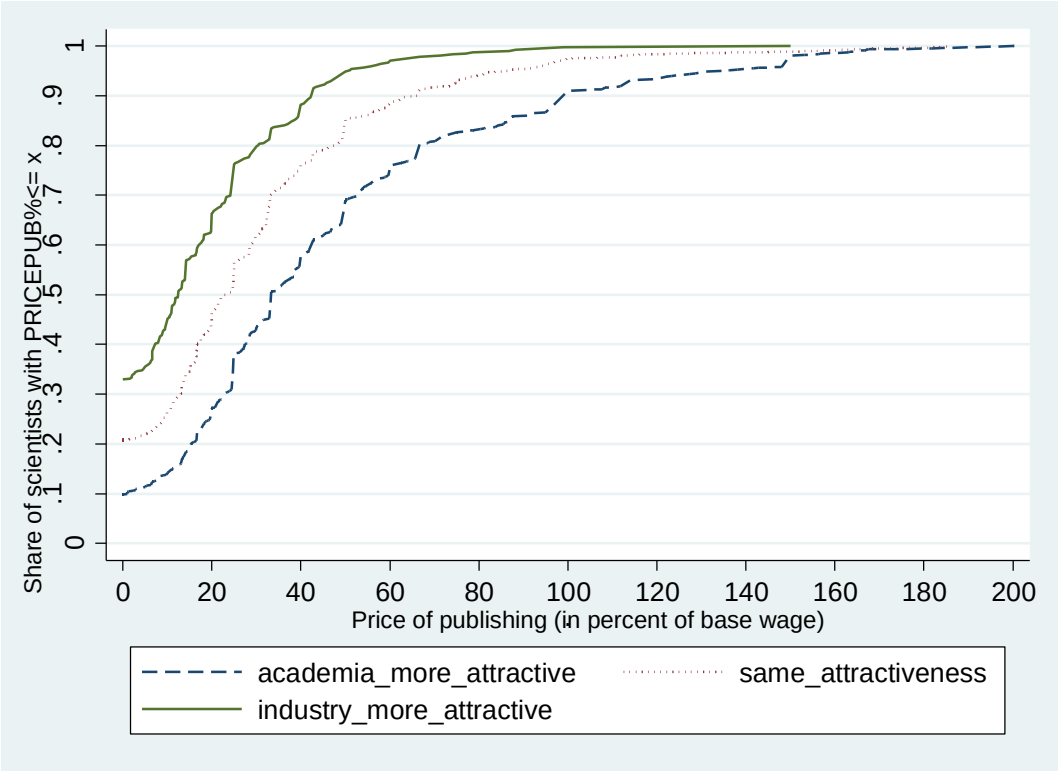
5 PricePub, preferences for industrial versus academic R&D, and implications for firms

5.1 PricePub and career preferences

We now examine whether the price of publishing is systematically related to preferences for careers in industrial versus academic science. Figure 2 shows the cumulative distribution of PRICEPUB% separately for three groups of scientists: those who rated a career in industry more attractive than a career in academia (30% of sample), those who rated both careers similarly attractive (26%), and those who rated a career in academia as more attractive (44%). As expected, we see that the price scientists assign to the opportunity to publish when employed in a firm is lowest among those scientists who prefer an industrial career. More specifically, 33% of these scientists state a PRICEPUB of zero, i.e., they would not require a wage premium to forego publishing opportunities and they would not accept a wage discount in return for having the opportunity to publish. The average PRICEPUB is \$11.3k or 17% of the base wage. In contrast, among those who find an academic career more attractive, only 10% have a PRICEPUB of zero, with an average price of \$32k, or 46% of the base wage.[18]

[18] Those scientists who prefer academic employment have a somewhat higher reservation wage for taking a job in industry than those who prefer industry employment (WAGE_PUB=\$73.43k vs. \$68.33k).

Figure 2: PricePub% by career preference



To examine more systematically which scientists are likely to self-select into industrial science, we estimate multinomial logit models predicting whether a respondent finds a career in industry more attractive than a career in academia, finds both careers similarly attractive, or prefers academia. Consistent with our expectation, table 5 shows that individuals with strong preferences for income are more likely to prefer industry over academia, while those with strong preferences for publishing are less likely to prefer industry. Given our interest in the price of publishing, we next include PRICEPUB rather than the preferences for publishing and income. Consistent with figure 2, we find that scientists who assign a higher price to publishing are less likely to seek a career in industry. Finally, ABILITYSELF is also negatively associated with a preference for industry, even controlling for the price of publishing.

Given the cross-sectional nature of our data, these regressions provide only limited insights regarding the causality of the observed relationships. One particular concern is that our respondents have pre-existing preferences for employment in industry or academia and that these career preferences shape their preferences for publishing and money, and thus their PRICEPUB. To address this concern, we asked respondents at the beginning of the survey to recall their career intentions at the time they matriculated into the PhD program. In model 4 we include variables indicating whether respondents thought they were more likely to pursue a career in industry (STARTPHD_IND>ACAD), more likely to pursue a career in academia (STARTPHD_IND<ACAD) or whether both careers were equally likely

(STARTPHD_IND=ACAD). As expected, we find that career intentions at the start of the PhD program are very strong predictors of current career preferences. More importantly, including these additional controls does not significantly reduce the coefficient of PRICEPUB.[19] Additional robustness checks addressing endogeneity concerns with respect to PRICEPUB are discussed in Appendix C.

A final consideration is that not all respondents will eventually work in industrial or academic research, and individuals who intend to enter other careers - such as in consulting, teaching, or government research - may state a particularly low price of publishing in firms. We restrict our sample to those individuals who rated research in industry or academia among their most attractive careers, dropping 28% of cases who found some other career more attractive. Dropping these cases leads to only small changes in the overall distribution of PRICEPUB, which now has a mean of $24.41k, with 17.5% of individuals stating a PRICEPUB of zero (compared to $22.56k and 19.5% for the full sample). Models 5 and 6 use only this smaller sample; the results are very similar to our featured regressions.

5.2 Implications for firms

The key insights of our empirical analysis are (1) not all scientists assign a high price to the opportunity to publish in industry, with many assigning a price of zero; (2) scientists who seek careers in industry have a considerably lower PRICEPUB than those seeking academic careers; (3) PRICEPUB depends on the particular reasons for publishing; and (4) PRICEPUB increases with ability. These results have important implications for science-based firms.

First, our results suggest that firms seeking to restrict publishing can avoid paying a wage premium if they hire those scientists who assign little or no value to the opportunity to publish.[20] However, these individuals tend have lower ability and may be less productive researchers, potentially offsetting higher value appropriation (via limited disclosure) with lower value creation. When research productivity is very important, firms seeking to limit disclosure may be better off hiring scientists with higher ability, even if those scientists also tend to assign a higher price to publishing and require a larger wage premium to give up publishing.

Firms seeking to implement an open science strategy may seek to attract scientists who place a high value on publishing in order to benefit from both wage discounts and higher research ability. However, these scientists tend to prefer employment in academia rather than industry.[21] Moreover, it is

[19] Responses to the retrospective career goals question may partly reflect career preferences at the time of the survey. However, any such bias would imply that this measure controls for "too much", making this robustness check conservative.
[20] Recent data suggest that approximately 25% of life sciences PhDs are employed in industry 5-6 years after graduation (Stephan, forthcoming) and that only about 30% of research-active life scientists in industry do not publish (Sauermann & Stephan, 2010). If those scientists at the low end of the PRICEPUB distribution (figure 2) are the first ones to take industry positions that restrict publishing, most of these positions may indeed be filled with scientists who do not require a wage premium.
[21] Firms seeking to hire scientists with a strong open science orientation may benefit from the scarcity of tenure track academic positions, potentially forcing some scientists who would prefer academia to go into industry (Cyranoski et al., 2011).

important to emphasize that the notion of a "wage discount" refers to the fact that some scientists are willing to accept a lower wage for a job that allows publishing versus a job that does not; it does not imply that scientists with a high preference for publishing are overall cheaper to hire than scientists with a low preference for publishing. In fact, our analysis of reservation wages (table 3) showed that the preference for publishing had a small positive (but not significant) relationship with WAGE_PUB and a large positive relationship with WAGE_NOPUB. Thus, scientists with a strong preference for publishing do not come cheaper if allowed to publish, but they are more expensive if publishing is restricted. Even when firms are primarily interested in research ability rather than publishing, our results suggest that they may have to offer higher wages (positive effect of ability on both reservation wages) *and* the opportunity to publish (to avoid the large wage premia required to keep high ability scientists from publishing).

Even though a formal treatment of firm policies is beyond the scope of this paper, our results suggest that the employment contracts offered by firms seeking to implement an "open science" strategy should fulfill two functions. First, such contracts should induce sorting such that they attract those scientists with a strong preference for publishing and with high research ability. Second, these contracts should motivate employees to exert high research effort and to actually publish their results. Both goals can be achieved by providing incentives that are tied to publishing performance, coupled with lower base wages (cf. Boudreau & Lakhani, 2011; Elfenbein et al., 2010; Lazear, 2000). Such a contract will attract scientists who like to publish and who think that they are able to achieve high publishing performance, while deterring scientists who dislike publishing or feel less productive. The performance-contingent component will also increase employees' publishing effort regardless of their publishing preferences; moreover, recent work by Lacetera & Zirulia (2011) suggests that incentives for publishing may be particularly effective if scientists already have a strong "intrinsic" preference for publishing because preferences and incentives act as complements.

Finally, we suggest that firms also need to consider that scientists value publishing for different reasons. If a scientist accepts a wage discount because he truly values open knowledge disclosure and contributing to the advancement of science, his employer may indeed be able to enjoy the benefits of lower labor cost in the long term. In contrast, if the scientist accepts a wage discount in return for publishing because he expects that publications lead to higher pay in the future, he will seek to recoup that investment at later points. In the words of one of our respondents, "Publishing allows you to publicly document your research success. This is GREAT for resume building and can be used as leverage for salary negotiations". Thus, while attracting scientists who value publishing for non-financial reasons may allow the firm to appropriate rents vis-à-vis the scientist even in the long-term, pay-related advantages from attracting scientists who value publications for career reasons may not be sustainable.

6 Summary and discussion

Drawing on survey data from a sample of 1,400 PhD candidates prior to entering the job market, we examine the price scientists assign to publishing opportunities in industrial employment. Contrary to the common belief that all scientists have a strong preference for publishing, we find significant heterogeneity in the price of publishing. Almost 20% of scientists value publishing at a price of zero, and that share rises to 33% among those who find industrial employment particularly attractive. Second, we show that scientists with higher ability have a higher price of publishing but also expect to be paid higher wages regardless of the publishing regime. Finally, we show that scientists value publishing not just for peer recognition but also for other and more specific reasons, such as the opportunity to advance science or to move to higher-paying jobs. More importantly, different reasons are related in distinct ways to the price assigned to publishing and may also have very different implications for the sustainability of competitive advantages derived from open science strategies.

In addition to implications for firms, which we discussed in section 5.2, our results also have more general implications and point towards promising avenues for future research. First, our results suggest an explanation for seemingly conflicting findings reported in prior research, namely that scientists pay for the opportunity to publish (Stern, 2004) while also getting paid for publishing (Cockburn & Henderson, 1998; Stuart & Liu, 2010). We suggest that the interpretation of these results requires distinguishing between base wages (the dependent variable in Stern's (2004) study) and variable pay. Empirical work is needed to resolve the question whether – and how much – scientists "pay" in terms of total compensation. More interestingly, our results suggest that the observed firm policies may serve as sorting and incentive mechanisms that address considerable heterogeneity in applicants' preferences and ability. More research is needed to gain a better understanding of heterogeneity in the characteristics of research positions offered by firms, and how scientists and firms eventually match up (cf. Charles & Guryan, 2008; Hwang et al., 1998; Killingsworth, 1987).

Future theoretical work is needed to study optimal incentive contracts when preferences are heterogeneous and correlated with ability. Particularly interesting are the dynamics as scientists build a publication record, increasing outside job opportunities over time. Theorists as well as empiricists could also examine dysfunctional effects of publishing incentives. While scholars of (academic) science have suggested that strong publishing incentives may lead to incrementalism, the withholding of research tools, and outright cheating (Lacetera & Zirulia, forthcoming; Merton, 1973), it is not clear whether publishing incentives have similar effects in the context of industrial science. Moreover, we suggest that such behaviors may be conditioned by the reasons why scientists value publishing, e.g., whether they seek to publish primarily out of career concerns or to advance science.

Although our study focuses on industry employment where choices between publishing and pay are most salient, such choices may be increasingly important in academia as well. While it is unlikely that academics will give up publishing completely, academic scientists' increasing attention toward research commercialization has raised concerns over publication delays and the incomplete disclosure of research results (cf. Blumenthal et al., 1986; Hackett, 1990; Slaughter & Rhoades, 2004). Our discussion may provide a useful starting point for future research on the trade-offs scientists are willing to make between disclosure and pay in the particular context of academia. For example, to the extent that patents also result in recognition from scientific peers (Azoulay et al., 2007), scientists who care about publishing primarily as a means to achieve peer recognition could publish less and patent more, potentially garnering additional financial returns. In contrast, scientists who value publishing primarily as a mechanism to advance science and who value the open disclosure of results may not see patents as a viable substitute for publications.[22]

While we focus on the price scientists assign to publishing opportunities, future work may apply our approach to study preferences for other factors such as the opportunity to do social good (Goddeeris, 1988) or autonomy on the job (Aghion et al., 2008; Lacetera, 2009). With respect to the latter, Aghion et al. (2008) develop a theoretical model in which industry has to offer a wage premium relative to academia because industry jobs provide less freedom. Generalizing our discussion in section 5, we expect that the size of that wage premium depends critically on the distribution of scientists' preferences for freedom and on selection effects associated with career choice.

Finally, while our results speak to the nature of scientists' preferences at a given point in time, future work is needed to examine how preferences might change over time. To some extent, preferences may be trait-like or shaped primarily by educational experiences prior to scientists entering the labor market (Austin, 2002; Ding, 2011; Lykken et al., 1993). On the other hand, preferences may also be shaped by socialization at work (Allen & Katz, 1992) or change over time due to life cycle effects (Levin & Stephan, 1991). Insights into the extent and direction of such changes may have important implications for future research as well as for managers and policy makers who often interact with scientists over extended periods of time.

[22] If scientists think of patents and publications as substitutes, a low price of publishing observed in our study may reflect that scientists expect to patent rather than publish when employed in a firm. We do not believe that patenting is seen as a viable substitute for publishing by a significant number of our respondents, and we find no evidence of this thinking in our open ended questions. More importantly, even if such a mechanism were at play, it would not change the fact that scientists assign a low price to publishing, with the associated implications for firms considering their participation in "open science".

Table 1: Summary Statistics

Variable	Type	Mean	SD	Min	Max
WAGE_PUB	continuous	71.09	22.12	30	200
WAGE_NOPUB	continuous	93.65	34.84	30	200
PRICEPUB	continuous	22.56	23.82	0	129
PRICEPUB%	continuous	33.00	34.52	0	200
PREF_INCOME	5-point	3.94	0.71	1	5
PREF_PUB	5-point	3.78	0.98	1	5
REASON_PAY	5-point	3.59	0.96	1	5
REASON_CONTR	5-point	4.38	0.70	1	5
REASON_RECOG	5-point	4.10	0.85	1	5
ABILITYSELF	continuous	6.48	1.59	1.9	10
NRC_RANKING	continuous	-40.29	25.46	-136	-1
IND<ACAD	dummy	0.44		0	1
IND=ACAD	dummy	0.26		0	1
IND>ACAD	dummy	0.30		0	1
JOBAVAIL_ACAD	continuous	39.56	24.77	0	100
JOBAVAIL_IND	continuous	55.92	22.10	1	100
MALE	dummy	0.41		0	1
AGE	continuous	28.15	3.31	21	45
USCITIZEN	dummy	0.81		0	1
STARTPHD_IND<ACAD	dummy	0.52		0	1
STARTPHD_IND=ACAD	dummy	0.25		0	1
STARTPHD_IND>ACAD	dummy	0.23		0	1
BASIC_RESEARCH	5-point	4.41	0.77	1	5
APPLIED_RESEARCH	5-point	3.23	1.13	1	5
DEVELOPMENT	5-point	2.16	1.15	1	5

Table 2: Price assigned to publishing in industry

	1	2	3	4	5	6
	poisson	poisson	poisson	poisson	poisson	poisson
	PRICEPUB	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**
PREF_PUB	0.457**	0.429**	0.447**		0.458**	0.448**
	[0.024]	[0.030]	[0.026]		[0.024]	[0.026]
PREF_INCOME	-0.191**	-0.136**	-0.157**		-0.188**	-0.155**
	[0.027]	[0.032]	[0.029]		[0.028]	[0.029]
REASON_PAY		-0.085**				
		[0.029]				
REASON_CONTR		0.190**				
		[0.046]				
REASON_RECOG		-0.008				
		[0.036]				
REASON_PAY=REASON_CONTR			-0.167**			-0.161**
			[0.058]			[0.058]
REASON_PAY>REASON_CONTR			-0.318*			-0.313*
			[0.134]			[0.133]
ABILITYSELF				0.043**	0.006	0.007
				[0.014]	[0.015]	[0.015]
NRC_RANKING				0.004**	0.004**	0.004**
				[0.001]	[0.001]	[0.001]
SUBFIELD	incl.	incl.	incl.	incl.	incl.	incl.
JOBAVAIL_ACAD	0.001	0.001	0.001	0.003*	0.002	0.002
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
JOBAVAIL_IND	0.001	0.000	0.000	0.001	0.000	0.000
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
MALE	0.164**	0.164**	0.165**	0.120*	0.161**	0.162**
	[0.050]	[0.051]	[0.050]	[0.056]	[0.052]	[0.052]
AGE	-0.003	-0.004	-0.003	0.009	-0.001	-0.001
	[0.009]	[0.009]	[0.010]	[0.011]	[0.009]	[0.009]
USCITIZEN	-0.041	-0.035	-0.042	-0.116	-0.058	-0.058
	[0.083]	[0.071]	[0.079]	[0.077]	[0.083]	[0.079]
RESPONSEMODETIME	incl.	incl.	incl.	incl.	incl.	incl.
Constant	2.025**	1.566**	2.092**	2.531**	1.976**	2.039**
	[0.593]	[0.565]	[0.602]	[0.524]	[0.600]	[0.607]
Observations	1,400	1,400	1,400	1,400	1,400	1,400
Chi-square	2527.081	3225.933	3747.495	272.198	3248.227	4272.832
df	20	23	22	20	22	24

Note: PRICEPUB is the price of publishing, computed as PRICEPUB=WAGE_NOPUB-WAGE_PUB. Models estimated using Poisson QML regression with standard errors clustered at the level of the university. Standard errors in brackets; *=significant at 5%, **=significant at 1%. REASON_PAY<REASON_CONTR is the omitted category in models 3 and 6.

Table 3: Reservation Wages

	1		2		3		4	
	a	b	a	b	a	b	a	b
	mvreg		mvreg		mvreg		mvreg	
	WAGE_PUB	**WAGE_NOPUB**	**WAGE_PUB**	**WAGE_NOPUB**	**WAGE_PUB**	**WAGE_NOPUB**	**WAGE_PUB**	**WAGE_NOPUB**
PREF_PUB	0.756	9.175**	0.603	8.868**			0.037	8.262**
	[0.611]	[0.934]	[0.620]	[0.947]			[0.619]	[0.948]
PREF_INCOME	2.593**	-1.834	2.695**	-0.971			2.485**	-1.133
	[0.838]	[1.280]	[0.873]	[1.333]			[0.862]	[1.321]
REASON_PAY=REASON_CONTR			0.725	-3.067			0.703	-3.010
			[1.320]	[2.014]			[1.301]	[1.993]
REASON_PAY>REASON_CONTR			-2.862	-7.255*			-2.539	-6.928*
			[2.214]	[3.378]			[2.182]	[3.343]
ABILITYSELF					2.326**	3.283**	2.249**	2.452**
					[0.374]	[0.590]	[0.378]	[0.580]
NRC_RANKING					0.074**	0.154**	0.078**	0.152**
					[0.026]	[0.041]	[0.026]	[0.040]
SUBFIELD	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
JOBAVAIL_ACAD	-0.039	-0.010	-0.041	-0.013	-0.030	0.032	-0.034	0.003
	[0.025]	[0.038]	[0.025]	[0.038]	[0.025]	[0.040]	[0.025]	[0.038]
JOBAVAIL_IND	0.148**	0.166**	0.148**	0.161**	0.126**	0.149**	0.129**	0.133**
	[0.029]	[0.044]	[0.029]	[0.044]	[0.028]	[0.045]	[0.028]	[0.044]
MALE	4.133**	7.914**	4.176**	7.968**	3.283**	5.930**	3.212**	6.951**
	[1.208]	[1.845]	[1.208]	[1.843]	[1.202]	[1.898]	[1.202]	[1.842]
AGE	-0.052	-0.125	-0.055	-0.137	0.011	0.225	0.025	-0.008
	[0.184]	[0.280]	[0.184]	[0.280]	[0.181]	[0.286]	[0.182]	[0.278]
USCITIZEN	-1.068	-2.211	-0.983	-2.119	-0.636	-3.334	-0.263	-1.723
	[1.534]	[2.343]	[1.535]	[2.342]	[1.526]	[2.409]	[1.532]	[2.348]
RESPONSEMODETIME	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	56.331**	65.564**	55.941**	66.285**	56.454**	66.505**	45.271**	54.452**
	[12.180]	[18.608]	[12.186]	[18.596]	[11.442]	[18.067]	[12.132]	[18.593]
Observations	1,400		1,400		1,400		1,400	
Parameters	21		23		21		25	

Note: Columns "a" regress the reservation wage for a job with publishing and columns "b" regress the reservation wage for a job without publishing. All models are estimated using multivariate regression. Standard errors in brackets; *=significant at 5%, **=significant at 1%.

Table 4: Supplementary Analyses and Robustness Checks

	Full Sample							Final Years & Thought		Publications>0		Full Sample	
	1	2	3	4	5	6	7	8	9	10	11	12	13
	poisson	poisson	poisson	oprobit	oprobit	tobit	tobit	poisson	poisson	poisson	poisson	poisson	poisson
	PRICEPUB%	**PRICEPUB%**	**PRICEPUB%**	**PREF_PUB**	**PREF_INCOME**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**
PREF_PUB	0.452**		0.461**				11.250**		0.397**		0.440**		0.443**
	[0.027]		[0.026]				[0.702]		[0.035]		[0.029]		[0.028]
PREF_INCOME	-0.196**		-0.191**				-4.253**		-0.128**		-0.155**		-0.163**
	[0.030]		[0.030]				[0.910]		[0.042]		[0.030]		[0.030]
REASON_PAY=REASON_CONTR	-0.181**		-0.172**				-4.234**		-0.257**		-0.164**		-0.166**
	[0.058]		[0.058]				[1.456]		[0.084]		[0.061]		[0.059]
REASON_PAY>REASON_CONTR	-0.228		-0.228				-7.702**		-0.456*		-0.265*		-0.303*
	[0.120]		[0.119]				[2.625]		[0.189]		[0.111]		[0.131]
ABILITYSELF		0.010	-0.024	0.114**	0.028	1.058**	0.063	0.073**	0.031	0.051**	0.011	0.032*	0.003
		[0.014]	[0.014]	[0.020]	[0.022]	[0.355]	[0.386]	[0.022]	[0.027]	[0.017]	[0.019]	[0.013]	[0.015]
NRC_RANKING		0.003*	0.003*	0.000	-0.002	0.085*	0.081*	0.005*	0.004*	0.005**	0.004**	0.004**	0.004**
		[0.001]	[0.001]	[0.001]	[0.001]	[0.038]	[0.035]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]	[0.001]
BASIC_RESEARCH												0.072**	-0.004
												[0.027]	[0.027]
APPLIED_RESEARCH												0.076**	0.056
												[0.029]	[0.029]
DEVELOPMENT												0.026	0.016
												[0.027]	[0.024]
SUBFIELD	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
JOBAVAIL_ACAD	0.002	0.003*	0.002*	0.004*	0.001	0.082*	0.049	0.002	0.001	0.003	0.002	0.003*	0.002
	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	[0.032]	[0.028]	[0.002]	[0.001]	[0.002]	[0.001]	[0.001]	[0.001]
JOBAVAIL_IND	-0.002	-0.001	-0.002	0.002	-0.001	0.039	0.014	0.002	0.001	0.001	0.000	0.001	0.000
	[0.001]	[0.001]	[0.001]	[0.002]	[0.002]	[0.036]	[0.033]	[0.002]	[0.002]	[0.002]	[0.002]	[0.001]	[0.001]
MALE	0.082	0.048	0.094*	-0.108	0.095	2.664	4.121**	0.077	0.121	0.052	0.077	0.110	0.158**
	[0.045]	[0.049]	[0.045]	[0.067]	[0.065]	[1.611]	[1.409]	[0.082]	[0.083]	[0.059]	[0.061]	[0.058]	[0.052]
AGE	0.000	0.012	0.002	0.026**	-0.014	0.234	-0.096	0.023*	0.007	0.014	0.004	0.009	-0.002
	[0.009]	[0.011]	[0.009]	[0.008]	[0.012]	[0.279]	[0.265]	[0.010]	[0.010]	[0.012]	[0.011]	[0.010]	[0.009]
USCITIZEN	-0.017	-0.105	-0.047	-0.249**	-0.200*	-5.066*	-3.345	-0.135	-0.019	-0.071	-0.023	-0.102	-0.047
	[0.064]	[0.066]	[0.062]	[0.060]	[0.089]	[2.211]	[2.129]	[0.099]	[0.094]	[0.087]	[0.084]	[0.080]	[0.078]
RESPONSEMODETIME	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	2.390**	2.871**	2.499**			6.886	1.387	2.301**	2.244**	2.468**	2.219**	1.981**	1.959**
	[0.521]	[0.444]	[0.529]			[14.209]	[15.984]	[0.535]	[0.646]	[0.559]	[0.613]	[0.541]	[0.625]
Observations	1,400	1,400	1,400	1,400	1,400	1,400	1,400	715	715	1,071	1,071	1,400	1,400
Chi-square	3510.317	380.009	5047.425	263.006	333.796			297.645	1173.926	264.788	2328.51	343.441	5512.411
df	22	20	24	20	20	20	24	20	24	20	24	23	27

Note: Clustered standard errors in brackets; *=significant at 5%, **=significant at 1%.

Table 5: Preferences for careers in industry

	Full Sample								Industry or Academia most preferred			
	1		2		3		4		5		6	
	a	b	a	b	a	b	a	b	a	b	a	b
	mlogit		mlogit		mlogit		mlogit		mlogit		mlogit	
	Ind=Acad	**Ind>Acad**	**Ind=Acad**	**Ind>Acad**	**Ind=Acad**	**Ind>Acad**	**Ind=Acad**	**Ind>Acad**	**Ind=Acad**	**Ind>Acad**	**Ind=Acad**	**Ind>Acad**
PREF_PUB	-0.838**	-1.270**										
	[0.067]	[0.082]										
PREF_INCOME	0.635**	1.174**										
	[0.112]	[0.146]										
PRICEPUB			-0.024**	-0.058**	-0.024**	-0.058**	-0.022**	-0.057**	-0.022**	-0.066**	-0.018**	-0.061**
			[0.003]	[0.005]	[0.004]	[0.005]	[0.004]	[0.005]	[0.003]	[0.007]	[0.004]	[0.008]
ABILITYSELF					-0.120**	-0.128**	-0.116**	-0.122	-0.122*	-0.176**	-0.118*	-0.174*
					[0.033]	[0.049]	[0.039]	[0.071]	[0.048]	[0.059]	[0.053]	[0.085]
NRC_RANKING					0.002	-0.001	0.003	0.004	0.000	-0.004	0.002	0.005
					[0.004]	[0.004]	[0.004]	[0.004]	[0.005]	[0.005]	[0.005]	[0.004]
STARTPHD_IND=ACAD							1.330**	1.488**			1.635**	1.733**
							[0.169]	[0.179]			[0.192]	[0.227]
STARTPHD_IND>ACAD							1.913**	3.716**			2.057**	4.109**
							[0.199]	[0.260]			[0.268]	[0.354]
JOBAVAIL_ACAD	0.001	-0.004	0.001	-0.004	0.002	-0.004	0.003	-0.001	0.005	0.000	0.007*	0.006
	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]
JOBAVAIL_IND	0.000	-0.001	-0.001	-0.002	-0.001	-0.001	-0.002	-0.003	-0.003	0.000	-0.004	-0.004
	[0.004]	[0.004]	[0.003]	[0.004]	[0.003]	[0.004]	[0.004]	[0.005]	[0.004]	[0.005]	[0.005]	[0.006]
MALE	-0.386**	-0.911**	-0.189	-0.540**	-0.140	-0.488**	-0.076	-0.324	-0.150	-0.653**	-0.053	-0.394
	[0.136]	[0.176]	[0.144]	[0.157]	[0.136]	[0.159]	[0.134]	[0.183]	[0.185]	[0.189]	[0.194]	[0.211]
AGE	0.018	-0.023	0.003	-0.057**	0.002	-0.060**	0.005	-0.055*	0.002	-0.080**	0.004	-0.075*
	[0.018]	[0.022]	[0.018]	[0.022]	[0.018]	[0.022]	[0.021]	[0.026]	[0.020]	[0.027]	[0.025]	[0.032]
USCITIZEN	-0.376	-0.288	-0.307	-0.318	-0.389*	-0.387	-0.257	-0.147	-0.473*	-0.375	-0.309	-0.109
	[0.236]	[0.248]	[0.184]	[0.222]	[0.188]	[0.218]	[0.209]	[0.269]	[0.184]	[0.245]	[0.220]	[0.295]
SUBFIELD	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
RESPONSEMODETIME	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.	incl.
Constant	14.600**	15.421**	14.856**	17.599**	15.611**	18.388**	13.940**	16.176**	15.353**	18.134**	14.027**	16.093**
	[1.229]	[1.430]	[0.803]	[1.375]	[0.826]	[1.437]	[0.884]	[1.614]	[0.825]	[1.498]	[0.936]	[1.223]
Observations	1,400		1,400		1,400		1,400		1,011		1,011	

Note: IND<ACAD is the omitted category of the dependent variable. Clustered standard errors in brackets; *=significant at 5%, **=significant at 1%. STARTPHD_IND<ACAD is the omitted category in models 4 and 6.

REFERENCES

Agarwal, R. & Ohyama, A. 2010. Industry or academia, basic or applied? Career choices and earnings trajectories of scientists, *Working Paper*.

Aghion, P., Dewatripont, M., & Stein, J. 2008. Academic freedom, private-sector focus, and the process of innovation. *RAND Journal of Economics*, 39(3): 617-635.

Allen, T. J. & Katz, R. 1992. Age, education and the technical ladder. *IEEE Transactions on Engineering Management*, 39(3): 237-245.

Arrow, K. J. 1962. Economic welfare and the allocation of resources for invention. In R. R. Nelson (Ed.), *The Rate and Direction of Inventive Activity*: NBER / Princeton University Press.

Austin, A. 2002. Preparing the next generation of faculty: Graduate School as socialization into the academic career. *Journal of Higher Education*, 73(1): 94-122.

Azoulay, P., Ding, W. W., & Stuart, T. E. 2007. The determinants of faculty patenting behavior: Demographics or opportunities? *Journal of Economic Behavior and Organization*, 63: 599-623.

Baron, J. & Spranca, M. 1997. Protected values. *Organizational Behavior and Human Decision Processes*, 70: 1-16.

Baron, R. M. & Kenny, D. A. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6): 1173-1182.

Beattie, J. & Barlas, S. 2001. Predicting perceived differences in tradeoff difficulty. In E. Weber & J. Baron & G. Loomes (Eds.), *Conflict and tradeoffs in decision making*: Cambridge University Press.

Becker, G. 1962. Investment in human capital: A theoretical analysis. *Journal of Political Economy*, 70(5): 9-49.

Blumenthal, D., Gluck, M., Louis, K., & Wise, D. 1986. University-industry research relationships in biotechnology: Implications for the university. *Science*, 232: 1361-1366.

Boudreau, K. & Lakhani, K. 2011. "Fit": Field experimental evidence on sorting, incentives and creative worker performance, *Working Paper*.

Brown, C. 1980. Equalizing Differences in the Labor Market. *Quarterly Journal of Economics*, 94(1): 113-134.

Camerer, C. & Lovallo, D. 1999. Overconfidence and excess entry: An experimental approach. *American Economic Review*, 89(1): 306-318.

Charles, K. K. & Guryan, J. 2008. Prejudice and wages: an empirical assessment of Becker's The economics of discrimination. *Journal of Political Economy*, 116(5): 773-809.

Chubin, D. E. & Hackett, E. J. 1990. *Peerless Science: Peer Review and U.S. Science Policy*: SUNY Press.

Cockburn, I. M. & Henderson, R. M. 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. *Journal of Industrial Economics*, 46(2): 157-182.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. 2000. Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not), *NBER Working Paper #7552*.

Cyranoski, D., Gilbert, N., Ledford, H., Nayar, A., & Yahia, M. 2011. The PhD factory. *Nature*, 472(7343): 276-279.

Dasgupta, P. & David, P. A. 1994. Toward a New Economics of Science. *Research Policy*, 23(5): 487-521.

Ding, W., Levin, S., Stephan, P., & Winkler, A. 2010. The impact of information technology on academic scientists' productivity and collaboration patterns. *Management Science*, 56(9): 1439.

Ding, W. W. 2011. The impact of founders' professional-education background on the adoption of open science by for-profit biotechnology firms. *Management Science*, 57(2): 257-273.

Edwards, J. 1995. Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Dec. Processes*, 64(3): 307-324.

Elfenbein, D. W., Hamilton, B. H., & Zenger, T. R. 2010. The small firm effect and the entrepreneurial spawning of scientists and engineers. *Management Science*, 56(4): 659-681.

Freeman, R., Weinstein, E., Marincola, E., Rosenbaum, J., & Solomon, F. 2001. Competition and Careers in Biosciences. *Science*, 294(5550): 2293.

Gans, J., Murray, F., & Stern, S. 2010. Contracting Over the Disclosure of Scientific Knowledge: Intellectual Property Protection and Academic Publication, *MIT Working Paper*.

Gans, J. S. & Stern, S. 2010. Is there a market for ideas? *Industrial and Corporate Change*, 19(3): 805-837.

Gibbons, R. S. & Katz, L. F. 1992. Does unmeasured ability explain inter-industry wage differentials? *Review of Economics and Statistics*, 59(3): 515-535.

Gittelman, M. & Kogut, B. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logit of citation patterns. *Management Science*, 49(4): 366-382.

Goddeeris, J. 1988. Compensating differentials and self-selection: An application to lawyers. *Journal of Political Economy*, 96(2): 411-428.

Greenberg, J., Ashton-James, C., & Ashkanasy, N. 2007. Social comparison processes in organizations. *Organizational Behavior and Human Decision Processes*, 102(1): 22-41.

Groves, R. & Peytcheva, E. 2008. The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly*, 72(2): 167-189.

Hackett, E. J. 1990. Science as a vocation in the 1990s. *Journal of Higher Education*, 61(3): 241-279.

Hicks, D. 1995. Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change*, 4: 401-424.

Hwang, H., Reed, W. R., & Hubbard, C. 1992. Compensating wage differentials and unobserved productivity. *Journal of Political Economy*, 100(4): 835-858.

Hwang, H., Mortensen, D. T., & Reed, W. R. 1998. Hedonic wages and labor market search. *Journal of Labor Economics*, 16(4): 815-847.

Killingsworth, M. 1987. Heterogeneous preferences, compensating wage differentials, and comparable worth. *Quarterly Journal of Economics*, 102(4): 727-742.

Konrad, A. & Pfeffer, J. 1990. Do you get what you deserve? Factors affecting the relationship between productivity and pay. *Administrative Science Quarterly*, 35: 258-285.

Kornhauser, W. 1962. *Scientists in industry: Conflict and accommodation*. Berkeley: University of California Press.

Lacetera, N. 2009. Different missions and commitment power in R&D organizations: Theory and evidence on industry-university alliances. *Organization Science*, 20(3): 565-582.

Lacetera, N. & Zirulia, L. 2011. Individual preferences, organization, and competition in a model of R&D incentive provision, *NBER Working Paper #17031*.

Lacetera, N. & Zirulia, L. forthcoming. The economics of scientific misconduct. *Journal of Law, Economics, and Organization*.

Landis, J. & Koch, G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33: 159-174.

Lazear, E. P. 2000. Performance pay and productivity. *American Economic Review*, 90(5): 1346-1361.

Levin, S. G. & Stephan, P. E. 1991. Research Productivity Over the Life Cycle: Evidence for Academic Scientists. *American Economic Review*, 81(1): 114-132.

Lim, K. 2004. The relationship between research and innovation in the semiconductor and pharmaceutical industries (1981-1997). *Research Policy*, 33(2): 287-321.

Luce, M., Payne, J., & Bettman, J. 1999. Emotional trade-off difficulty and choice. *Journal of Marketing Research*, 36(2): 143-159.

Lykken, D. T., Bouchard, T. J., McGue, M., & Tellegen, A. 1993. Heritability of interests: A twin study. *Journal of Applied Psychology*, 78(4): 649-661.

Merton, R. K. 1973. *The sociology of science: Theoretical and empirical investigations*. Chicago: University of Chicago Press.

Merton, R. K. & Barber, E. 2004. *The travels and adventures of serendipity*: Princeton University Press.

Miller, G. A. 1976. Professionals in Bureaucracy: Alienation among Industrial Scientists and Engineers. *American Sociological Review*, 32(5): 755-768.

Moorman, R. H. & Podsakoff, P. M. 1992. A metaanalytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behavior research. *Journal of Occupational and Organizational Psychology*, 65: 131-149.

Murray, F. & O'Mahony, S. 2007. Exploring the foundations of cumulative innovation: Implications for Organization Science. *Organization Science*, 18(6): 1006-1021.

National Research Council. 2010. *A data-based assessment of research-doctorate programs in the United States*. Washington, DC: National Academies Press.

National Science Board. 2010. *Science and Engineering Indicators 2010*. Arlington,VA: National Science Foundation.

National Science Foundation. 2008. Survey of Earned Doctorates.

Nelson, R. 2004. The market economy, and the scientific commons. *Research Policy*, 33(3): 455-471.

Nunnally, J. C. 1978. *Psychometric Theory* (2nd ed.): McGraw-Hill.

Parchomovsky, G. 1999. Publish or Perish. *Michigan Law Review*, 98: 926-952.

Penin, J. 2007. Open knowledge disclosure: An overview of the evidence an economic motivations. *Journal of Economic Surveys*, 21(2): 326-348.

Podsakoff, P. M., MacKenzie, J. Y., & Lee, J. Y. 2003. Common methods bias in behavioral research: A critical review and recommended remedies. *Journal of Applied Psychology*, 88(5): 879-903.

Roach, M. & Sauermann, H. 2010. A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. *Research Policy*, 39(3): 422-434.

Rogelberg, S. & Stanton, J. 2007. Understanding and dealing with organizational survey nonresponse. *Organizational Research Methods*, 10: 195-209.

Rosen, S. 1986. The theory of equalizing differences. In O. Ashenfelter & R. Layard (Eds.), *Handbook of Labor Economics*: 641-692: North-Holland.

Santos Silva, J. & Tenreyro, S. 2006. The log of gravity. *Review of Economics and Statistics*, 88(4): 641-658.

Sauermann, H. & Cohen, W. 2010. What makes them tick? Employee motives and industrial innovation. *Management Science*, 56(12): 2134-2153.

Sauermann, H. & Stephan, P. E. 2010. Twins or Strangers? Differences and Similarities between Academic and Industrial Science, *NBER Working Paper #16113*.

Shepard, H. A. 1956. Nine dilemmas in industrial research. *Administrative Science Quarterly*, 1(3): 295-309.

Simeth, M. & Raffo, J. D. S. 2011. What makes companies pursue an Open Science strategy?, *Working Paper*.

Slaughter, S. & Rhoades, G. 2004. *Academic Capitalism and the New Economy: Markets, State, and Higher Education*. Baltimore: Johns Hopkins University Press.

Sorenson, O. & Fleming, L. 2004. Science and the diffusion of knowledge. *Research Policy*, 33(10): 1615-1634.

Stephan, P. forthcoming. *How Economics Shapes Science*: Harvard University Press.

Stephan, P. E. 2004. Robert K. Merton's perspective on priority and the provision of the public good knowledge. *Scientometrics*, 60(1): 81-87.

Stern, S. 2004. Do scientists pay to be scientists? *Management Science*, 50(6): 835-853.

Stuart, T. & Liu, C. 2010. Boundary spanning in a for-profit research lab: An exploration of the interface between commerce and academe, *Working Paper*.

Tuckman, H. & Leahey, J. 1975. What is an article worth? *Journal of Political Economy*, 83(5): 951-967.

Wooldridge, J. 1997. Quasi-Likelihood methods for count data. In M. Pesaran & P. Schmidt (Eds.), *Handbook of Applied Econometrics*, Vol. 2: 352-406. Oxford: Blackwell.

ONLINE APPENDIX

Appendix A: Why do scientists care about publishing? An analysis of open-ended responses

While the Mertonian view suggests that publications serve scientists primarily as a tool to obtain peer recognition (Merton, 1973), we are interested in a deeper understanding of the various reasons scientists value publishing. To our knowledge, large-scale empirical work on this question is absent. In addition to using the three rating measures of reasons to publish employed in the empirical analysis, we also asked respondents the following open ended question after eliciting their reservation wages for jobs with and without publishing: "If you are willing to accept lower pay for a job that allows you to publish, why is publishing so important to you?" We first provide some illustrative responses and then analyze more general patterns.

Respondents with a very high price of publishing

- "I prefer to do science for science sake. At this time the peer review process is the best measure for the caliber of your work."
- "Reading publications enables us to come up with new results; it isn't right not to publish them. Also it will restrict one's career opportunities in academia."
- "It means contribution to knowledge, status, accomplishment. It's the currency in science, it's esteem, etc."
- "It's not science until you communicate to the public what you have found."
- "It is the reason we are scientists. Discover and share to increase our knowledge. By the way, I would not take a job that didn't allow me to publish research I had conducted regardless of salary."

Respondents with PRICEPUB=0

- "I expect to be paid a decent wage whether or not I am required to publish."
- "I honestly find publications to be a distraction. That said, if you can find a job where I don't have to publish either papers, or reports, or reviews, or summaries, or some other periodic distraction from doing science then you are better at this game than I am."
- "In my field, if you choose to go into industry and are not allowed to publish, I do not think it will matter as industry experience would now be the measure of your success and often private sector does not publish." [respondent in cell/molecular biology]

Other interesting responses

- "While a firm has the right to make money from its discoveries, scientists have a moral obligation to better the world at large. Yes, that company is funding my current research, but public funds educated

me and paid for my training. Therefore, I have an obligation to that public. Knowledge is the thing I produce, so I'm obligated to give some of that knowledge (through publication) back to the communities that helped me learn how to obtain it."

- "I believe research should be available to as great of a population as possible. If I have to take a job in which that standard is compromised I assume it is for monetary interest of the company and I would want to share in that."
- "I realized that [publishing] is not that important to me... Keeping data private is important for pharma/biotech to a degree, and I can understand that."
- "I feel like if you are unable to publish in any given job, it will make it difficult to find subsequent employment. Therefore the initial compensation needs to be higher to compensate for lost potential at future jobs."

These statements illustrate the different degrees to which scientists value publishing and the wide range of reasons for which they value the opportunity to publish. It is also notable how some respondents quite clearly expressed the trade-offs between the various benefits of publishing and higher current pay. To provide a more representative picture, we identified a set of 10 categories that capture the breadth of reasons mentioned by our respondents. We then asked three research assistants to independently code the roughly 800 open ended answers into these categories. The inter-rater reliability coefficient (kappa) of 0.71 indicates substantial agreement among the raters (Landis & Koch, 1977), and existing disagreements were resolved by one of the authors. Table A1 provides a description of the categories and indicates the percentage of respondents in each. In cases where respondents stated multiple reasons, the coding reflects the first reason mentioned (presumably the most important reason). We find that 31% of respondents stated reasons related to the value of publications from a career perspective. Roughly 41% of respondents mentioned reasons broadly related to the function of publications as a mechanism that facilitates the cumulative nature of science. Ten percent of respondents mentioned recognition without being more specific as to why they cared about recognition (those who made the link to career opportunities are included in the above counts). The remaining respondents stated a wide range of reasons including, for example, publications as an objective measure of their research accomplishments, a personal sense of achievement and worth derived from publishing, the need for publications to obtain a green card (international students), the central role of publications in obtaining funding for future research, or that publishing implies closure of a particular research project.

Table A1: Reasons for publishing

	Reason (coded)	% of respondents
Science	Publishing allows me to contribute to advancement of science/knowledge; allows others to build on my results/avoid duplicate research; allows me to have an impact, etc.	25%
	Publishing allows me to participate in the scientific community; publishing is essential for being part of the community; important communication mechanism	5%
	I believe that knowledge should be openly shared / there is no point in doing science otherwise [strong normative aspect]	9%
	Publications allow the critical evaluation of results by others; I might get feedback	2%
Career	Publications are useful to find better jobs, higher pay; facilitate career moves; build resume	24%
	Publications are useful to transition to academia	7%
Recognition and Measurement	Publications give me recognition/reputation/respect/credibility/being known	10%
	Publications are an objective measure of my accomplishments, record of achievements, credit for my work	6%
Personal Achievement	Publications give me a personal sense of achievement and worth	3%
Other	Other	9%
Total		100%

Note: Based on 794 open-ended responses, inter-rater reliability kappa=0.71.

Table B1: Correlations

		1	2	3	4	5	6	7	8	9
1	WAGE_PUB	1								
2	WAGE_NOPUB	0.7367*	1							
3	PRICEPUB	0.1487*	0.7782*	1						
4	PRICEPUB%	-0.1176*	0.5538*	0.9192*	1					
5	PREF_INCOME	0.0901*	-0.0580*	-0.1686*	-0.2021*	1				
6	PREF_PUB	0.0178	0.2668*	0.3737*	0.3787*	-0.1133*	1			
7	REASON_PAY	-0.0209	-0.0552*	-0.0613*	-0.0658*	0.3924*	0.1170*	1		
8	REASON_CONTR	0.0188	0.1691*	0.2299*	0.2131*	-0.1035*	0.3982*	0.0983*	1	
9	REASON_RECOG	0.0189	0.0762*	0.0939*	0.0997*	0.1029*	0.3044*	0.3508*	0.2913*	1
10	ABILITYSELF	0.1901*	0.1785*	0.0846*	0.0274	0.0628*	0.1704*	0.1021*	0.1561*	0.1064*
11	NRC_RANKING	0.0483	0.0792*	0.0710*	0.0650*	-0.0778*	0.0149	-0.0569*	0.0509	-0.0012
12	IND<ACAD	0.0946*	0.3015*	0.3530*	0.3280*	-0.2709*	0.3846*	-0.1250*	0.1814*	0.0710*
13	IND>ACAD	-0.0810*	-0.2609*	-0.3064*	-0.2972*	0.2368*	-0.3601*	0.1004*	-0.1931*	-0.1086*
14	JOBAVAIL_ACAD	0.002	0.0599*	0.0857*	0.0813*	0.0209	0.1257*	0.023	0.0794*	0.0346
15	JOBAVAIL_IND	0.1517*	0.1165*	0.0295	-0.0227	0.0016	0.0241	-0.0233	0.0654*	-0.0023
16	MALE	0.1056*	0.1110*	0.0642*	0.0223	0.0623*	-0.0159	0.0315	0.0187	-0.0153
17	AGE	-0.0133	0.0217	0.044	0.0567*	-0.0416	0.0978*	-0.0132	0.0642*	-0.0835*
18	USCITIZEN	-0.0255	-0.0499	-0.0493	-0.0399	-0.0864*	-0.1009*	-0.0419	-0.0472	-0.1382*
19	STARTPHD_IND<ACAD	0.0363	0.1965*	0.2536*	0.2476*	-0.2289*	0.2260*	-0.0999*	0.1260*	0.0705*
20	STARTPHD_IND>ACAD	-0.0554*	-0.1530*	-0.1723*	-0.1719*	0.1698*	-0.1939*	0.0577*	-0.1213*	-0.0837*

		10	11	12	13	14	15	16	17	18
10	ABILITYSELF	1								
11	NRC_RANKING	-0.0382	1							
12	IND<ACAD	0.0977*	0.0662*	1						
13	IND>ACAD	-0.0701*	-0.0758*	-0.5814*	1					
14	JOBAVAIL_ACAD	0.0785*	-0.0609*	0.0655*	-0.0967*	1				
15	JOBAVAIL_IND	0.0929*	0.013	0.0002	-0.0103	0.2681*	1			
16	MALE	0.1542*	-0.0641*	0.0783*	-0.0998*	0.009	0.0336	1		
17	AGE	-0.0013	-0.0541*	0.0585*	-0.0932*	-0.0135	-0.0585*	0.1162*	1	
18	USCITIZEN	-0.1582*	0.1658*	0.0399	0.0049	-0.0563*	0.0563*	-0.0813*	-0.1097*	1
19	STARTPHD_IND<ACAD	0.0626*	0.0911*	0.4818*	-0.4121*	0.0813*	0.0143	0.0710*	0.0441	0.0653*
20	STARTPHD_IND>ACAD	-0.0419	-0.1035*	-0.3874*	0.4873*	-0.0689*	0.0213	-0.0841*	-0.0442	-0.0443

Note: *=significant at 5%

Appendix C: Accounting for pre-existing career preferences

We follow the prior literature in conceptualizing preferences for job attributes such as pay or publishing as antecedents of career and job choices. As discussed in section 5.1., however, one may be concerned that individuals who have always been interested in a particular career adjust their preferences to match those desired careers. For example, a scientist who has always been interested in becoming an academic may report a strong preference for publishing because publishing is essential for a successful academic career and may also report a lower preference for money because academia pays relatively little. In this section, we briefly address this issue both conceptually and empirically.

First, recall that we designed our trade-off question such that all respondents stated their price of publishing when employed in a firm. Thus, all respondents evaluated publishing in the context of industrial science, even those who would prefer a career in academia. Of course, scientist with a strong desire to work in academia may assign a high price to publishing in industry because they would like to have the option to switch from industry to academia at a later point. While this mechanism may explain *why* some scientists assign a larger price to publishing in industry than others, it does not change the fact that they do assign a high value to publishing in industry, with the associated implications for scientific labor markets and industrial employers.

To gauge the extent to which our regression results regarding determinants of PRICEPUB change when we control for respondents' initial career preferences, we re-estimate key models in table C1. The measures of initial career preferences are STARPHD_IND=ACAD and STARTPHD_IND>ACAD (STARTPHD_IND<ACAD is omitted) and are described in more detail in section 5.1. As expected, we find that a pre-existing preference for industrial research is associated with a lower PRICEPUB. More importantly, the coefficients of our featured variables change only little once initial career preferences are included and all results remain significant. Thus, while initial career preferences may shape preferences for particular job attributes, this mechanism is not the primary driver of our results.

Table C1: Regressions including career preferences at start of PhD program

	1	2	3	4	5	6
	poisson	poisson	poisson	poisson	poisson	poisson
	PRICEPUB	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**	**PRICEPUB**
PREF_PUB	0.419**	0.393**	0.410**		0.422**	0.413**
	[0.023]	[0.029]	[0.024]		[0.023]	[0.024]
PREF_INCOME	-0.152**	-0.106**	-0.125**		-0.150**	-0.124**
	[0.026]	[0.031]	[0.029]		[0.027]	[0.029]
REASON_PAY		-0.073**				
		[0.027]				
REASON_CONTR		0.180**				
		[0.047]				
REASON_RECOG		-0.014				
		[0.036]				
REASON_PAY=REASON_CONTR			-0.142*			-0.136*
			[0.056]			[0.056]
REASON_PAY>REASON_CONTR			-0.294*			-0.290*
			[0.127]			[0.126]
ABILITYSELF				0.033*	0.002	0.003
				[0.014]	[0.015]	[0.015]
NRC_RANKING				0.003*	0.003*	0.003*
				[0.001]	[0.001]	[0.001]
STARTPHD_IND=ACAD	-0.304**	-0.286**	-0.291**	-0.460**	-0.301**	-0.288**
	[0.062]	[0.063]	[0.061]	[0.055]	[0.063]	[0.062]
STARTPHD_IND>ACAD	-0.328**	-0.310**	-0.314**	-0.579**	-0.311**	-0.298**
	[0.064]	[0.066]	[0.064]	[0.064]	[0.064]	[0.064]
SUBFIELD	incl.	incl.	incl.	incl.	incl.	incl.
JOBAVAIL_ACAD	0.001	0.001	0.001	0.002	0.002	0.002
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
JOBAVAIL_IND	0.001	0.000	0.000	0.001	0.000	0.000
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
MALE	0.131**	0.132**	0.133**	0.080	0.132**	0.134**
	[0.049]	[0.049]	[0.049]	[0.054]	[0.050]	[0.050]
AGE	-0.003	-0.004	-0.003	0.005	-0.001	-0.002
	[0.009]	[0.009]	[0.010]	[0.011]	[0.009]	[0.009]
USCITIZEN	-0.067	-0.061	-0.065	-0.151	-0.083	-0.080
	[0.082]	[0.071]	[0.078]	[0.079]	[0.083]	[0.079]
RESPONSEMODETIME	incl.	incl.	incl.	incl.	incl.	incl.
Constant	2.319**	1.897**	2.370**	3.088**	2.289**	2.333**
	[0.577]	[0.546]	[0.582]	[0.559]	[0.584]	[0.588]
Observations	1,400	1,400	1,400	1,400	1,400	1,400
Chi-square	3626.736	5071.195	5327.612	1572.098	5805.058	7560.785
df	22	25	24	22	24	26

Note: Standard errors in brackets; *=significant at 5%, **=significant at 1%.