Management Science

Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists

Rajshree Agarwal, Atsushi Ohyama,

Please scroll down for article—it is on subsequent pages

INFORMS is the largest professional society in the world for professionals in the fields of operations research, management
science, and analytics.
For more information on INFORMS, its publications, membership, or meetings visit http://www.informs.org

Industry or Academia, Basic or Applied? Career Choices and Earnings Trajectories of Scientists

Rajshree Agarwal

R. H. Smith School of Business, University of Maryland, College Park, Maryland 20742, rajshree@umd.edu

Atsushi Ohyama

Graduate School of Economics and Business Administration, Hokkaido University,
Sapporo, Hokkaido 060-0809, Japan, ohyama@econ.hokudai.ac.jp

We extend life cycle models of human capital investments by incorporating matching theory to examine the sorting pattern of heterogeneous scientists into different career trajectories. We link differences in physical capital investments and complementarities between basic and applied scientists across industry and academic settings to individual differences in scientist ability and preferences to predict an equilibrium matching of scientists to careers and to their earnings evolution. Our empirical analysis, using the National Science Foundation's Scientists and Engineers Statistical Data System database, is consistent with theoretical predictions of (i) sorting by ability into basic versus applied science among academic scientists, but not among industry scientists; and (ii) sorting by higher taste for nonmonetary returns into academia over industry. The evolution of an earnings profile is consistent with these sorting patterns: the earnings trajectories of basic and applied scientists are distinct from each other in academia but are similar in industry.

1. Introduction

The dynamism of the U.S. economy has been attributed to advances made in academic and industry settings in both basic and applied domains. The underlying scientific labor markets can be characterized by two-sided matching: Hiring institutions choose among scientists who differ in their ability and preferences, and scientists—particularly as they embark on their careers—choose among career options regarding *what* to do—basic or applied science—and *where* to do it—academia or industry. The match of job and scientist characteristics may also have path-dependent outcomes and lasting implications on their earnings trajectories.

However, a systematic study that accounts for sorting of scientists across each of the four career options simultaneously is lacking. Although the basic versus applied nature of research is recognized as an important characteristic in both academia and industry (Sauermann and Stephan 2012), scholars have largely examined these career options *two at a time*. For example, studies have focused on the choice between basic and applied research in academic settings (Levin and Stephan 1991, Stuart and Ding 2006, Thursby and Thursby 2002) or in industry settings (Cohen and Levinthal 1990, Rosenberg 1990, Stern 2004), or the choice between industry and academia

(Roach and Sauermann 2010). For the underlying theoretical mechanisms, scholars have relied on life cycle models of human capital investments (Levin and Stephan 1991, Thursby et al. 2007) or on differences in taste for science and nonpecuniary returns (Stern 2004, Roach and Sauermann 2010). These studies have provided valuable insights and generated many empirical findings regarding outcomes of scientific labor markets; nonetheless, we do not have a unified theoretical model that examines the following research question: How do (demand-side) differences in complementarities between basic and applied science across industry and academic settings interact with (supply-side) differences in ability and preferences among scientists to impact their career choices and their resultant earnings evolution?

Our study answers the above research question by investigating sorting patterns of scientists into alternative careers and the earnings trajectories associated with each career option. Our model builds on the premises that (i) academic and industry settings provide different access to physical capital for basic and applied scientists and differ in the complementarity between basic and applied scientists, and (ii) scientists differ in their ability and preferences for nonmonetary "taste for science." We show that competition among institutions and among scientists

for the most desirable partners results in a positive assortative matching: more able scientists are matched with institutions that provide them more physical capital investments, and, when applicable, more able basic scientists are matched with more able applied scientists. Further, the taste for nonmonetary returns separates industrial and academic scientists in equilibrium. For earnings evolution, our model predicts that complementarities between basic and applied scientists in industry relative to academia results in a positive assortative matching between them, implying similar earnings for basic and applied industrial scientists. Academic scientists' earnings are offset because of their taste for nonmonetary returns, and the relative absence of complementarities between basic and applied scientists and differences in access to physical capital result in a divergence in their earnings. We find support for the model implications using the longitudinal Scientists and Engineers Statistical Data System (SESTAT) developed by the National Science Foundation (NSF) for the 1995–2006 period.

We contribute to Stephan's (1996) call for better modeling of the labor markets in science by incorporating matching theory (Becker 1973) into traditional life cycle models of human capital investments (Becker 1962). In doing so, our study provides useful guidance on how possible self- or sample-selection issues in life cycle models may impact empirical findings. Our model builds from the differences in the scientific production functions in academia and industry: basic scientists may work independently of applied scientists in academia, but have to work closely with applied scientists in industry. Because academic institutions make higher per capita investments in basic research relative to applied, our model generates some novel implications that are backed by the empirical analysis: in academia, scientists of higher ability sort into basic rather than applied research, and initial earnings of basic scientists are lower but the slope of their earnings is higher relative to applied scientists. In industry, by contrast, there is no such ability sorting, and the earnings trajectories of basic and applied scientists are similar.

Additionally, by examining all four career options in tandem rather than any two in isolation, our model is able to generate other implications that extend the findings from earlier studies. In particular, we contribute to the nascent literature that examines early career selection of scientists because of ability and preferences (Roach and Sauermann 2010, Stern 2004) by adding the life cycle component. We highlight the fact that observed earnings differentials across the four career options may stem from differences in human and physical capital investments and may not entirely be a result of a taste for science. For example,

consistent with Roach and Sauermann (2010), we find that scientists with a higher nonmonetary "taste for science" are more likely to choose academia than industry, thus foregoing higher earnings. However, even within academia, taking productivity and returns from human capital investments into account, we find that earnings of academic basic scientists at later life stages "catch up" with industry earnings. In his cross-sectional study examining industry job offers of postdoctoral biologists at top-tier research institutions, Stern (2004) found that the "preference effect" overshadowed the "productivity effect," so that scientists seeking jobs in industry are willing to accept lower wage offers to engage in basic research. Our longitudinal study shows that although this may be true for the very initial years (see Figure 2 in §4.2), basic scientists in industry enjoy the same earnings as their applied counterparts for most of their lifetime. We attribute this to the enhanced productivity of industry basic scientists since an assortative matching allows them to access higher levels of complementary physical and human capital. Importantly, by examining the consequences of differences in complementarity of basic and applied scientists across institutional settings, our study also has numerous managerial and policy implications. Chief among them is the need, within universities that train young scientists, to accommodate the requirements for success of the various career paths that they may choose.

2. Theoretical Analysis

2.1. Labor Markets for Science: Received Literature and Stylized Features

2.1.1. Demand for Scientists Across Institutional Settings. In her influential literature review, Stephan (1996) describes salient aspects of the demand for scientists, which stems largely from institutions of higher education (academia) and for-profit businesses (industry), with government and nonprofit organizations being a distinct minority. The two institutional settings—academia and industry—seek scientists for different reasons, are governed by different norms and incentive structures, and—importantly—differ in the relative emphasis on basic and applied science. Consider, for instance, the statistics reported in Table 1 on U.S. 2003 research expenditures and employment of scientists. The total research expenditures in academia and industry are the same, but the proportions spent on basic and applied research are vastly different. Basic research accounts for 74% of academic research expenditures but only 17% of industry expenditures. In contrast, applied research is only 26% of academic research expenditures but 83% in industry. The employment and

Table 1 Research Expenditures, Number of Scientists, and Research Expenditure per Scientist

	Academia		Industry	
	Basic	Applied	Basic	Applied
Research expenditure (millions dollars)	27,956	9,721	6,526	30,883
Number of scientists	204,542	167,865	104,393	310,569
Research expenditure per scientist (millions dollars)	0.14	0.06	0.06	0.10

Sources. National Science Board (2010) (research expenditures data from Appendix Tables 4-4, 4-5, 5-1, and 5-2) and SESTAT 2003 (labor force data, estimated using sample weights).

Note. Numbers are for the year 2003 and expressed as 2000 constant dollars.

research expenditures per scientist also mirror the differential focus within each institutional setting: the academic sector employs more and has higher expenditures per scientist in basic versus applied science, whereas the reverse is true for industry. Nonetheless, more than one-third of the scientists pursue "off-diagonal" careers of applied research in academia and basic research in industry.

Within academia, the primary focus on basic science is fueled by social norms and reward structures that promote nonpecuniary motives such as priority of discovery, recognition of merit awards, and reputation (Merton 1973). Also, although applied science is rooted in the land grant mission of many public universities, the influential Bush (1945) report to President Roosevelt resulted in a shift in focus toward basic science through increased funding and creation of government agencies such as the NSF. Together, these forces resulted in universities placing more emphasis on discoveries of scientific fundamentals than applications of scientific discoveries (Argyres and Liebskind 1998, Goldfarb 2008, Murray and Stern 2007). Recently and after the Bayh-Dohl act, Etzkowitz (1998) highlights the "second revolution" in academia and the increased attention to the commercialization of science. However, the connectivity between basic and applied research or research relevant to Pasteur's quadrant (Stokes 1997) is largely realized only for certain areas of specialization (e.g., biotechnology) and/or follows a "science-push" model, where basic research is conducted prior to the pursuit of potential applications through subsequent technology transfer or entrepreneurship (Aghion et al. 2008, Bercovitz and Feldman 2008, Stuart and Ding 2006, Thursby and Thursby 2002). Tenure norms imply that in the early career years, there are fewer incentives to engage simultaneously in basic and applied research or for basic and applied scientists to work together (Boardman and Ponomariov 2007, Boardman and Bozeman 2007, Braxton et al.

2002, Geisler 1989). Boardman and Ponomariov (2007) and Boardman and Bozeman (2007) report that early career basic and applied scientists tend to work in isolation of each other, not only in traditional academic departments, but even in "multidisciplinary, multipurpose" or "cooperative" university research centers whose primary responsibilities include applications of scientific knowledge. They provide quantitative and qualitative evidence that center-affiliated basic scientists in early career stages (pretenure) are reluctant to invest an effort in working with their applied counterparts and note significant "role strain" and "shirking of responsibilities" in areas where they need to consult with applied and industry partners.

Within industry, the demand for scientists stems from the need to innovate (Stephan 1996), with the goal of transforming scientific knowledge into commercially valuable outputs and appropriating economic value in the form of profits (e.g., Gittelman and Kogut 2003, Aghion et al. 2008). Thus, the demand for applied research is greater, and if firms engage in basic research, it is as a byproduct or coproduct of their applied endeavors (Rosenberg 1990). Fundamental breakthroughs in basic science may occur in industrial labs, given their need as a foundation for applied work (Hounshell and Smith 1988, Rosenberg 1990, Stokes 1997); alternatively, firms may engage in basic science to create absorptive capacity (Cohen and Levinthal 1990) and enhance the productivity of their R&D efforts (Griliches 1986). Anecdotally, Hounshell and Smith (1988) highlight how basic science at DuPont served to search for new fields of chemistry, provided foundations for scientific investigation of DuPont's existing technology, and showcased the company's technical competence and capabilities. Importantly, these authors underscore the need for basic and applied scientists working together, best exemplified by the following statement by DuPont's basic research department director Charles Stine: "Perhaps the most important idea which I have attempted to keep before my assistants in this work is the necessity for continuous intimate contact with the various departments and subsidiaries which we have been attempting to serve" (Hounshell and Smith 1988, p. 137). Thus, relative to academia, where basic and applied scientists can work in parallel or independently of each other, sequencing these activities over time, or for different purposes, there is a more direct link and greater synergies between basic and applied scientists in industry.[1]

---

[1] These distinctions in industry/academia highlight the extreme differences that are salient at the time a scientist embarks on her career. Of course, a continuum of possibilities may occur, particularly in later career stages. As Stephan (1996) notes, academic scientists often engage in privatization of knowledge, and industrial scientists often disclose their knowledge voluntarily.

2.1.2. Supply of Scientists and Heterogeneity of Ability and Preferences. On the supply side, research ability and preferences have been highlighted as distinctive characteristics of scientists that affect occupational and activity choices and compensation structure (Dasgupta and David 1994, Stephan 1996, Stern 2004). People differ in their abilities to learn extant knowledge and develop new innovations, either because of innate differences in intelligence or prior investments in knowledge. Also, scientists purposefully raise the level of their human capital and therefore their earnings capacity over time by conducting scientific research (Becker 1962, Ben-Porath 1967, Levin and Stephan 1991, Thursby et al. 2007). The supply-side dynamics of labor markets in science accordingly are often modeled in a life cycle framework wherein human capital investments are undertaken to maximize lifetime utility that includes both monetary and nonmonetary benefits (Levin and Stephan 1991). The inclusion of nonmonetary benefits is consistent with the characterization of the scientist's preferences or "taste for science" independent of ability, which may vary based on the quest for basic research and the desire to apply scientific principles for economic and technological development (Levin and Stephan 1991, Roach and Sauermann 2010, Stephan 1996, Stern 2004). Such preferences reflect attitudes toward open science, freedom of research topics, and intellectual challenge and are informed in the socialization process when receiving scientific education (Stephan 1996). Roach and Sauermann (2010) show that scientists with a higher taste for science prefer careers in academia over industry. Stern (2004) discusses the strong positive correlation between ability and preferences observed among scientists. Although higher-ability scientists command higher compensation, they may also have a greater taste for science, resulting in their willingness to accept a lower wage as a compensating differential.

2.1.3. Implications for Modeling Early Career Choices. In concluding her literature review, Stephan (1996) noted that extant human capital and life cycle models do not capture the complexities of the production of scientific knowledge. In part, this may be because most models do not study the supply and demand side simultaneously. Extending human capital and life cycle models with matching models may result in a more appropriate characterization of scientific labor markets and shed light on factors that influence early career choices. Matching models (Becker 1973, Roth and Sotomayor 1992, Sattinger 1993) are particularly suitable for the study of scientific labor markets because they are better able to accommodate market outcomes in exchanges of heterogeneous and indivisible "goods" (Mindruta 2012).

In developing a model of scientists' early career choices, we incorporate the above salient demand and supply characteristics of scientific labor markets. On the supply side, we assume heterogeneity among scientists in ability and preferences for nonpecuniary returns. On the demand side, the career options represent differences in the physical and human capital that complement a scientist's human capital. Within a multifactor scientific production function context, we define "complementarity" as the positive effect on one input's marginal productivity due to a marginal increase in the other factor. We focus on two types of complementarities. The first is between a scientist and the physical capital provided by the institution to her. We assume, based on Table 1, that the basic scientist has greater access to physical capital than the applied scientist within academia, and the reverse holds in industry. The second type of complementarity is between basic and applied scientists. We employ a simplifying assumption that the focal scientist is *either* a basic *or* an applied scientist but not both, thus ruling out the potential that an individual scientist may undertake basic and applied scientific tasks simultaneously (Mansfield 1995). This assumption is consistent with early career realities that scientists have to focus on one primary activity initially. We later discuss how the model implications may change when allowing for career switches or transitions to Pasteur's quadrant (Stokes 1997) by more senior scientists. At least in the initial (pretenure) years of their career, basic academic scientists are less likely to work with their applied counterparts (Boardman and Ponomariov 2007, Boardman and Bozeman 2007). Within industry, firms that engage in basic research expect basic and applied scientists to work with each other on issues that have potential applications of knowledge for the firm (Hounshell and Smith 1988, Cohen and Levinthal 1990). Accordingly, we assume that basic and applied scientists complement each other in industry knowledge production functions, but not in academia.

2.2. A Brief, Nonmathematical Overview of the Model

This section verbally describes the formal model and propositions in §§2.3 and 2.4 to explicate the underlying logic and intuition. The building blocks, mechanisms, and resultant propositions are also summarized in Figure 1. A scientist's career is defined by a combination of the type of research she conducts (basic or applied) and the type of institution she works in (academia or industry), resulting in four career options: (i) basic scientist in academia, (ii) applied scientist in academia, (iii) basic scientist in industry, and (iv) applied scientist in industry. A scientist supplies her human capital, and the institution she works for supplies the complementary physical and other human capital necessary to conduct scientific research.

Figure 1   Verbal Description of Model Assumptions, Mechanisms, and Predictions

Underlying Assumptions	Mechanisms/Forces in Model	Model Implications/Predictions
Main individual characteristics C1: Scientists differ in ability C2: Scientists differ in their taste for nonpecuniary returns	**Main forces affecting career choices and earnings trajectories** F1: Superior access to physical capital attracts more able scientists F2: Complementarity between basic and applied scientists in industry makes the effective access to physical capital the same for each type of scientists	**Empirical predictions related to career choices** **Proposition 1 for academia**: Given C1, C3, and C4, F1 implies In academia, the average research ability of basic scientists is higher than that of applied scientists. **Proposition 1 for industry**: Given C1, C3, and C4, F2 implies In industry, the average research ability does not depend on whether research type is basic or applied. **Proposition 2**: Given C2 and C3, F3 and F4 imply If nonpecuniary returns are sufficiently large and physical capital availability is relatively modest, those scientists who value nonpecuniary returns are more likely to choose academia over industry.
Main institution characteristics C3: Strong complementarity between basic and applied scientists in industry but not in academia C4: Basic scientists have greater access to physical capital relative to applied scientists in academia; the opposite holds in industry	F3: Complementarity and positive assortative matching between basic and applied scientists in industry results in synergy effects on earnings in equilibrium F4: Academic research environments offer nonmonetary returns F5: Superior access to physical capital makes human capital investments more productive	**Empirical predictions related to earnings trajectories** **Proposition 3**: Given C2 and C3, F3 and F4 imply The average earnings profile of academic scientists is located below those of industrial scientists when the synergy effects in industrial research are sufficiently large. **Proposition 4**: Given C1, C3, and C4, F1 and F5 imply The average initial salary of academic basic scientists is lower than that of academic applied scientists. **Proposition 5 for academia**: Given C1, C3, and C4, F1 and F5 imply The expected value of the slope of the earnings profile is higher for basic scientists than for applied scientists in academia. **Proposition 5 for industry**: Given C1, C3, and C4, F2 implies The expected value of the slope of the earnings profile is similar for basic and applied scientists in industry.

Scientists choose a career option to maximize lifetime utility based on pecuniary and nonpecuniary returns (Roach and Sauermann 2010, Stern 2004) and purposefully invest to raise their human capital and thus earnings over time (Becker 1962, Levin and Stephan 1991, Thursby et al. 2007). Similarly, institutions choose scientists to maximize scientific knowledge production, based on their productivity. As in matching models (Becker 1973), the "match" between a particular scientist and a particular institution is determined as scientists compete with each other to obtain a position that maximizes lifetime utility, and institutions compete with each other to attract highly productive scientists. The sorting of scientists is dependent on the identity of both the scientist and the institution. Because the productivity of human capital investments also depends on the complementary factors available through the institution, more able scientists and institutional settings with access to greater complementary human and physical capital are matched.

Equilibrium outcomes are driven by differences in complementarity across institutional settings and positive assortative matching. By a positive assortative matching, we mean a positive association between pairs of scientists or scientists and institutions with respect to research ability and level of physical capital (Becker 1973). These institutional differences imply that basic academic scientists have access to more physical capital than applied academic scientists, whereas the opposite is true in industry. However, in industry, complementarity between basic and applied scientists makes the effective access to physical capital the same for each scientist type. Further, academia provides greater nonpecuniary returns than industry. In academia, low complementarity between early career basic and applied scientists, combined with relatively lower levels of physical capital in applied science, results in more able scientists sorting into basic rather than applied science (Proposition 1 for academia). In industry, complementarities encourage basic and applied scientists with similar ability to work for the same firm, resulting in a positive assortative matching between basic and applied scientists as well as between scientists and firms. Thus, there is no sorting based on ability across basic and applied domains in industry (Proposition 1 for industry). Further, scientists with higher nonpecuniary preferences are matched to academia relative to industry (Proposition 2).

The above sorting results in scientists' earnings evolution differently within each career option. Differences in research ability and preferences for nonpecuniary returns are associated with the evolution of average earnings for each career path. Given sorting based on nonpecuniary returns and

complementarity among basic and applied scientists in industry, which generates additional monetary gains from synergies realized from positive assortative matching, the earnings profile of academic scientists is lower than industrial scientists (Proposition 3). Further, within academia, basic scientists invest more time initially in building up their human capital than do applied scientists, because access to higher levels of physical capital makes their marginal benefit from human capital investments higher. As a result, initial earnings of basic academic scientists are lower than those of applied academic scientists (Proposition 4). Finally, positive assortative matching among basic and applied industrial scientists results in a similar evolution of their earnings profiles. In contrast, the relative lack of complementarity leads to divergent paths of earnings for basic and applied scientists in academia, with a steeper slope for basic than for applied academic scientists (Proposition 5).

2.3. Formal Theoretical Development

2.3.1. Key Assumptions. The formal model is underpinned by the following assumptions: (i) scientists are perfectly rational; (ii) there is no asymmetric information and no uncertainty; (iii) scientists do not change their career over their life cycle; (iv) within academia, basic scientists have greater access to physical capital than applied scientists; (v) within industry, there are two types of firms—type A firms hire applied scientists only and type B firms hire both applied and basic scientists; (vi) a taste for nonmonetary returns enters the utility function additively and exogenously; and (vii) academia offers nonpecuniary benefits, but industry does not. Some model implications hinge on these assumptions for simplification and tractability of the theoretical analysis. In §5, we discuss how the implications change when certain assumptions are relaxed.

2.3.2. Key Properties of Scientific Knowledge Production Function. We characterize the scientist's research activities to set the stage for analyzing the interaction of market forces and scientists' preferences leading to an equilibrium allocation of scientists into the alternative career paths. Consider a model in which scientists make a human capital investment to increase their knowledge base, and their accumulated knowledge is used as an input to produce final research output. Formally, scientist i's knowledge base evolves as

$$H_{it+1} = H_{it} + R_{it}, \qquad (1)$$

where H_{it} is the level of scientist i's knowledge base at time t, and R_{it} is knowledge added to the stock of her knowledge base. To keep our exposition simple, we assume no depreciation of knowledge base.

Scientist i's knowledge production function, R_{it}, in Equation (1) depends not only on how much time she devotes to human capital investments, but also on where she works (i.e., academia or industry) and what type of research she conducts (i.e., basic or applied). Formally, it takes the form of

$$R_{it} = \gamma_k [\theta_i (l_{it} H_{it})^\eta]^{\varphi_1} [\theta_j (l_{jt} C_{jt})^\eta]^{\varphi_2}, \qquad (2)$$

where θ_i is the ability parameter of scientist i, and l_{it} is i's time spent on building human capital. For scientist i in basic (applied) research, C_{jt} is the complementary human capital provided by scientist j in applied (basic) research. Parameters φ_1 and φ_2 are either 0 or 1 and control the degree of complementarity between basic and applied scientists. Parameter η is exogenously given and takes some value between 0 and $1/2$. The amount of physical capital available to the scientist is denoted by γ_k, which is explained in detail below.

To model the knowledge production function in academia, we set φ_1 equal to 1 and φ_2 to 0 for an academic setting, which allows for the lack of complementarity between basic and applied scientists. Thus, for scientist i conducting type k research in an academic setting, the knowledge production function is

$$R_{it} = \gamma_k \theta_i (l_{it} H_{it})^\eta. \qquad (3)$$

For industry, we model two types of firms: Type A firms conduct only applied research and have the knowledge production function as given by (3). Type B firms conduct both applied and basic research. For scientist i conducting type k research in type B firms, the knowledge production function is

$$R_{it} = \gamma_k \theta_i (l_{it} H_{it})^\eta \theta_j (l_{jt} C_{jt})^\eta. \qquad (4)$$

Implications of relaxing the above assumptions, which allow us to understand essential forces leading to an equilibrium-matching pattern but at the cost of some loss of generality, are discussed in §2.4.

Scientists are differentiated by ability θ and a taste for nonpecuniary returns δ. We assume that θ is distributed according to a Pareto distribution with a density function

$$f(\theta) = a\theta^{-(a+1)} \quad \text{for } \theta > 1, \qquad (5)$$

where $a \geq 2$ to ensure existence of at least the second moment. The variable δ takes either 0 or 1, indicating that a scientist has a taste for nonpecuniary returns if δ equals 1. Because the taste for nonmonetary returns is better accommodated in academia rather than in industry (Sauermann and Stephan 2012), we incorporate this factor in the model by assuming that scientists with a taste for nonpecuniary returns gain

additional nonmonetary benefits x when they work in academia. The variables θ and δ are independently distributed.

In addition to human capital, physical capital is an important input of research production. Because academia places relatively more emphasis on basic than on applied science and the investments in physical capital are commensurate with these differences (see Table 1), we incorporate this in our model by using the concept of the first-order stochastic dominance. Physical capital per scientist, γ_k, is assumed to be Pareto distributed with a density function

$$g(\gamma_k) = bc_k^b \gamma_k^{-(b+1)} \quad \text{for } \gamma_k > c_k \qquad (6)$$

with

$$b \geq 2, \quad c_{BA} > c_{AA} > 0 \quad \text{and} \quad c_{AI} \geq c_{BI} > 0, \qquad (7)$$

where subscripts BA, AA, BI, and AI indicate "basic research in academia," "applied research in academia," "basic research in industry," and "applied research in industry," respectively. An essential property of the specification by (6) and (7) is the first-order stochastic dominance relation. Under this specification, which makes the model analytically tractable, the fraction of basic academic scientists who enjoy access to large amounts of physical capital is greater than the fraction of applied academic scientists.

2.3.3. Scientists' Predilections over Career Paths. We first derive a scientist's optimal path of human capital investment for each career choice and resultant lifetime earnings. Scientists seek to maximize the discounted present value of their lifetime earnings by allocating their time between building up human capital, l_{it}, and converting accumulated human capital into observable knowledge outputs (e.g., patents, publications, new products and processes, etc.) for earnings, $1 - l_{it}$. We assume that scientists are not directly compensated for the time invested in building human capital but are paid for the conversion of human capital into new knowledge (Ben-Porath 1967, Levin and Stephan 1991, Thursby et al. 2007). Because scientists are paid according to the current level of their human capital (Thursby et al. 2007), their current earnings, Y_{it}, are

$$Y_{it} = w_k(1 - l_{it})H_{it}, \qquad (8)$$

where w_k is the rental rate of human capital when scientist i chooses career path k.[2]

---

[2] Scientists' pay is assumed to be linear to the stock of human capital to avoid computational complexity, but nonlinear forms can be incorporated to reflect that human capital is an input for final research outcomes.

Time flows discretely and T is the terminal period. In her dynamic optimization problem, a scientist optimally divides her time between building human capital, l_t, and converting human capital to earn, $1 - l_{it}$, so as to maximize her lifetime earnings, $\sum_{t=0}^{T} \beta^t Y_{it}$, where β is a discount factor. The Bellman equation for this maximization problem is given by

$$V_t(H_{it}) = \max_{l_{it}} [w_k(1 - l_{it})H_{it} + \beta V_{t+1}(H_{it+1})] \qquad (9)$$

subject to (1) and (3) if a scientist works in academia or in a type A firm and to (1) and (4) if she works in a type B firm. By using the terminal condition $V_{T+1}(H_{T+1}) = 0$ and solving (9) recursively, we obtain the optimal path of l_{it}. All the details of the derivation in this section are placed in the online appendix, which is available from the authors upon request. Given the evolution of the complementary human capital, the evolution of scientist i's human capital and earnings are pinned down by Equations (1) and (8), respectively. These are given by

$$Y_t(H_{it})$$
$$= w_k \left[H_{it} - (\gamma_k \theta_i \eta A_{jt})^{1/(1-\eta)} \left(\frac{\beta - \beta^{T-t+1}}{1-\beta} \right)^{1/(1-\eta)} \right] \qquad (10)$$

and

$$H_{it+1} = H_{it} + (\gamma_k \theta_i A_{jt})^{1/(1-\eta)} \eta^{\eta/(1-\eta)} \left(\frac{\beta - \beta^{T-t+1}}{1-\beta} \right)^{\eta/(1-\eta)}, \qquad (11)$$

where $A_{jt} = \theta_j (l_{jt} C_{jt})^{\eta}$ if she works in a type B firm and $A_{jt} = 1$ otherwise.

Scientists are assumed to be risk neutral and their lifetime utility depends on both monetary and nonmonetary rewards. Formally, for academic and type A firm scientist i conducting type k research, this is

$$U(\theta_i, \gamma_k) = w_k[(\theta_i \gamma_k)^{1/(1-\eta)} \xi_A + \mu] + \delta_i x_i, \qquad (12)$$

where $\xi_A \equiv \eta^{1/(1-\eta)}((1-\eta)/\eta) \sum_{s=1}^{T-1} ((\beta - \beta^{T-s+1})/(1-\beta))^{1/(1-\eta)}$ and $\mu \equiv ((1-\beta^T)/(1-\beta))H_{i1}$. For industrial scientists in type A firms, δ_i is set to zero. The first term in (12) is the optimal lifetime earnings of scientist i when she conducts type k research in academia. If scientist i conducts type k research in type B firms, her lifetime utility is

$$U(\theta_i, \gamma_k) = w_k[(\theta_i \gamma_k)^{1/(1-\eta)} \xi_I(\theta_j) + \mu], \qquad (13)$$

where $\xi_I(\theta_j) \equiv \eta^{1/(1-\eta)}((1-\eta)/\eta) \sum_{s=1}^{T-1} A_{js}((\beta - \beta^{T-s+1})/(1-\beta))^{1/(1-\eta)}$.

Scientists have perfect foresight regarding their stream of future earnings when they choose a career path because there is no uncertainty about human capital investment. Therefore, scientists use optimal

lifetime earnings to anticipate monetary rewards generated from each career choice, and Equations (12) and (13) completely characterize scientist i's predilections over the four alternative careers. Furthermore, the absence of uncertainty implies that scientists do not change their career path in the middle of their career (We discuss relaxation of this assumption §5.1). Scientist i's predilections over the four alternative career paths depend on her ability, θ_i, a taste for nonpecuniary returns, δ_i as well as on what type research she conducts and where, γ_k. Most importantly, other things being equal, a scientist's predilection increases with parameter γ_k of research productivity associated with where she works and what type of research she is conducting.

2.3.4. Institutions' Predilections over Scientists. We now describe predilections of institutions (university in academia and firm in industry) when they hire scientists. An institution's net gain at time t from hiring scientist i in terms of monetary value is given by

$$\pi_t = p[(1 - l_{it})H_{it}]^\phi - w_k(1 - l_{it})H_{it}.$$

To ease the computational burden, we assume that $\phi = 1$ so that cumulative research outputs are transformed linearly into commercially or academically valuable outputs. Under this assumption, a net gain for a university or a type A firm from hiring scientist i throughout her entire career is

$$\Pi = \sum_{t=1}^{T} \beta^{t-1}\pi_t = (p - w_k)[(\theta_i\gamma_k)^{1/(1-\eta)}\xi_A + \mu] \quad (14)$$

and the corresponding gain for a type B firm is

$$\Pi = (p - w_k)[(\theta_i\gamma_k)^{1/(1-\eta)}\xi_I(\theta_j) + \mu]. \quad (15)$$

Thus, given that a firm or university hires scientists to maximize scientific knowledge output, both types of employers have a predilection for hiring scientists with high ability θ.

2.4. Implications of the Theoretical Model

2.4.1. An Equilibrium Sorting Pattern. Based on stable matching (Becker 1973, Roth and Sotomayor 1992, Sattinger 1993), we examine an equilibrium sorting pattern of scientists. Given limited availability of specific positions, scientists compete with each other for a desirable career option. Similarly, the scarcity of able scientists leads research institutions to compete with each other to attract desirable scientists. For the sake of simplicity, we assume that a firm and a university both hire a pair of scientists. A matching outcome is a set of disjoint triplets (scientist i, institution j, type of research k) that indicates that scientist i conducts research type k (applied or basic)

in institution j (academia or industry). A match is stable if there is no other triplet that makes both a scientist and an institution better off (Roth and Sotomayor 1992). In the preceding section, we showed that predilections of both scientists and research institutions strictly increase with research productivity. In particular, the complementarity in knowledge production function (3) or (4) results in scientists' and research institutions' predilections (12)–(15). Standard matching theory thus implies that a stable match is a positive sorting with respect to these characteristics. In both academia and industry, complementarity between a scientist's ability θ and a university's or firm's physical capital γ encourages highly able scientists to be matched with institutions that provide them greater access to physical capital.

Lemma 1. *In both academia and industry, a positive sorting with respect to a scientist's ability θ and an institution's physical capital γ takes place in equilibrium.*

All the formal proofs of the lemma and propositions in this section are in the online appendix. Lemma 1 has implications for differences in the average quality of scientists across different career categories. To see this, Lemma 1 implies that demand and supply of scientific ability in academia satisfy

$$c_{AA}^b \int_\gamma^\infty bx^{-(b+1)}\,dx + c_{BA}^b \int_\gamma^\infty bx^{-(b+1)}\,dx = 2\int_\theta^\infty ax^{-(a+1)}\,dx,$$

which can be simplified to

$$\theta(\gamma) = d_A\gamma^{b/a}, \quad (16)$$

where $d_A \equiv (2/(c_{AA} + c_{BA}))^{1/a} < 1$.

In conjunction with (6) and (7), which state first-order stochastic dominance of the distribution of physical capital available to basic scientists relative to applied scientists in academia, (16) implies that the average value of θ for academic basic scientists is higher than that for academic applied scientists: more able scientists choose basic science given complementarity between human and physical capital.

In addition to the complementarity between scientists and physical capital, there is complementarity between basic and applied scientists in type B firms. Our model allows scientists to choose a side of the match between basic and applied scientists. As Kremer and Maskin (1996) extensively analyze, complementarity does not suffice for a positive assortative matching between agents when agents can choose a side. In particular, they point out that a positive assortative matching between agents may not occur when the amount of outputs from joint production depends on who in a matched pair does which

task. Although complementarity encourages the formation of a match of the same traits, this asymmetric force discourages it and creates a gap in traits of matched individuals. In our knowledge production function (4), agents' productivities θ_i and θ_j enter in a perfectly symmetric way so that, for a given level of physical capital, the asymmetric force is totally absent. In other words, we assume that the total amount of knowledge created is unchanged if roles of basic and applied scientists are exchanged. Furthermore, a scientist has access to her partner's physical capital indirectly through a matching in equilibrium. This implies that effective access to physical capital does not depend on whether she conducts basic or applied research. Therefore, a basic (applied) scientist with a given level of research ability is matched with an applied (basic) scientist with that same level of research ability. Under this assumption, the following proposition holds:

PROPOSITION 1. *The average research ability of basic scientists in academia is higher than that of applied scientists in academia. In industry, the average research ability does not depend on whether the research type is basic or applied.*

A scientist's choice between academia and industry depends on nonmonetary returns, x, as well as gains from the complementarity. Because of Lemma 1, the difference in lifetime utility between working in industry and academia is given by

$$\Delta = (\lambda - 1)wz\gamma^{(a+b)/(a(1-\eta))} - \delta x, \qquad (17)$$

where $\lambda \equiv (1/2)^{(a\eta - (1-2\eta)(1-\eta))/(a(1-2\eta)(1-\eta))} \eta^{\eta/((1-2\eta)(1-\eta))}$. $\gamma^{(b+a\eta)/(a(1-2\eta)(1-\eta))}((c_{BI} + c_{AI})/(c_{BA} + c_{AA}))^{1/(1-\eta)}$, $z \equiv d_A^{1/(1-\eta)}\xi_A$ and $w_k = w$ for any k.

The first term is the monetary gain from the complementarity from working for type B firms; the second term is the nonmonetary return to working in academia. For some scientists, λ can be less than 1, but λ increases with γ_k. In other words, the synergy from positive assortative matching between basic and applied research in type B firms is realized for pairs of highly able scientists and institutions, but such a synergy is negligible or absent if the level of human capital of scientists and of physical capital of an institution is low. In our model, in addition to the effect of taste for science, complementarities between basic and applied scientists also create lifetime earnings differentials between academic and industrial scientists.

PROPOSITION 2. *If nonpecuniary returns are sufficiently large and physical capital availability is relatively modest, those scientists who value nonpecuniary returns are more likely to choose academia over industry.*

2.4.2. Earnings Evolution. This section characterizes the evolution of the earnings profile along with an equilibrium path. Equations (10) and (16) imply

that the earnings profile of scientists in academia and type A industrial firms evolves in equilibrium according to

$$Y_t = w_k[H_t - \gamma_k^{b/(a(1-\eta))}d_A^{1/(1-\eta)}\xi_t], \qquad (18)$$

and the earnings profile of industrial scientists in type B firms evolves in equilibrium according to

$$Y_t = w_k[H_t - \lambda\gamma_k^{b/(a(1-\eta))}d_A^{1/(1-\eta)}\xi_t], \qquad (19)$$

where $\xi_t = \eta^{\eta/(1-\eta)}((\beta - \beta^{T-t+1})/(1-\beta))^{\eta/(1-\eta)}$.

The above two equations, coupled with (17), suggest several implications for differences in earnings profiles among the four career options. First, we examine differentials across institutional settings. As can be seen from (17), beyond the sorting of scientists based on nonpecuniary returns, the model predicts that the synergy effect in industry shifts the earnings profile of industrial scientists up. The magnitude of the synergy effects is positively correlated with the quality of the pool of scientists. In other words, the synergy effect is more likely to generate monetary gains when the distribution of ability has a fatter right tail. Sufficiently large synergy effects relative to nonpecuniary returns create a positive earnings differential between industrial and academic scientists at any given point in time.

PROPOSITION 3. *The average earnings profile of academic scientists is located below those of industrial scientists when the synergy effects in industrial research are sufficiently large.*

Next we turn to expected differentials between basic and applied scientists within each institutional setting. Within academia, basic scientists invest more in building up their human capital than do applied scientists, because access to a higher level of physical capital makes their marginal benefit from human capital investments higher. Because the initial level of human capital is assumed to be the same for all scientists, the average initial earnings of academic basic scientists are lower than those of academic applied scientists.

PROPOSITION 4. *The average initial salary of academic basic scientists is lower than that of academic applied scientists.*

The positive sorting also affects the earnings growth of academic scientists. Under positive sorting, more able scientists are on average drawn to basic science within academia. For a given level of physical capital, therefore, basic scientists are more productive in human capital investment than applied scientists. Additionally, on average, basic scientists have access to a higher level of physical capital than applied scientists. Both factors lead basic scientists in academia to invest more in human capital at early stages of their

life cycle, making the earnings profile of basic scientists in academia steeper than that of applied scientists in academia.

In contrast, a different mechanism is at work in industry. For type B firms, basic and applied scientists are paired in research activities to exploit possible synergy effects (see Equation (17)). Thus, the synergy effect shifts up the earnings profile of industrial scientists in type B firms, relative to the earnings profile of applied scientists in type A firms. This force works to direct more able scientists toward type B firms. However, inequality (7) implies that applied scientists in type A firms have more favorable access to physical capital. This is because the effective physical capital for a scientist in type B firms is the average of physical capital associated with a pair of basic and applied scientists. This force shifts up the average earnings profiles of type A industrial scientists. In equilibrium, these two opposite forces cancel out to make scientists with a given ability indifferent between choosing these two types of firms. The earnings profile of basic scientists in type B firms coincides with that of applied scientists in type B firms, but may be different from the earnings profile of applied scientists in type A firms. The average earnings profile of industrial applied scientists is a weighted average of the earnings profiles of applied scientists in types A and B firms and is pinned down by the distribution of types A and B firms. The average earnings profiles of basic and applied scientists in industry can differ from one another, but the difference is smaller than the corresponding difference in academia. In an extreme case where there are no type A firms, the earnings profile of basic scientists is the same as that of applied scientists in industry. Accordingly,

Proposition 5. *The expected value of the slope of the earnings profile is* (a) *higher for basic scientists than for applied scientists in academia and* (b) *similar for basic and applied scientists in industry.*

3. Data and Methodology

The empirical analysis uses the NSF restricted files of the Survey of Doctorate Recipients (SDR) for the years 1995, 1997, 1999, 2001, 2003, and 2006. SDR, a part of SESTAT, contains information about the employment, education, and demographic characteristics of scientists and engineers in the United States.[3] In using the SDR database, we focus on *doctorate degree* recipients from U.S. institutions, given that a doctorate is usually a necessary qualification for both basic and academic research occupations in science and engineering fields. Further, the complementary

relationship in knowledge production between basic and applied scientists in industry is much more likely to exist for highly specialized scientists such as doctoral degree holders than for graduates who hold only a bachelor's or master's degree. The inclusion of bachelor's and master's degree holders from the integrated SESTAT database does not change our empirical results significantly, but it blurs the main points of our analyses. Unless otherwise noted, we focus our empirical analysis on scientists and engineers who are employed in the four career options of interest, are aged 65 or younger, report nonzero annualized basic salary, and work full time (working weekly for at least 30 hours and annually for at least 48 weeks).[4] Postdocs are included in our empirical analysis, but doctoral recipients working in teaching and consulting positions are excluded. See Table I in the online appendix for a description of criteria based on which individuals are included/excluded from the sample. As seen in Table 2, the total number of individuals in the sample from 1995 to 2006 is 33,776. These longitudinal data are the main data for our empirical examinations, but repeated cross-sectional and panel data are also used to ensure that our results are robust to cohort effects.

3.1. Variable Definitions and Descriptive Statistics

3.1.1. Career Options. Our career choice variable is obtained using the responses on two questions in the SDRs. The first question relates to the type of principal employment, and the second relates to the type of job activity on which the majority of time is spent. For the first question, respondents who reported that their employer was a four-year college or university, medical school, or university-affiliated research institute are identified as working in academia, whereas respondents who reported that their employer was private for profit are identified as working in industry. For the second question, respondents who reported basic (applied) research as their primary job activity are identified as such. Combining these two pieces of information permits the sorting of scientists in the data into each of the four career options. We recognize the lack of consensus on the definition and measurement of the "basic" or "applied" nature of research across various prior studies (Murray and Stern 2007, Sauermann and Stephan 2012) and that research may be within a continuum. The reliance on questions that

---

[3] For the details of SESTAT, see http://www.nsf.gov/statistics/sestat/.

[4] Academic contracts may consist of either 12-month appointments or nine-month appended with summer contracts ranging from one to three months. Seventy percent of all academics in the database report working at least 48 weeks (approximately 11 months). The results are robust to analysis where we included the additional 20% of academic scientists that work between 36 and 47 weeks (corresponding to a minimum of a nine-month contract).

Table 2   Descriptive Statistics

	Academia		Industry		
	Basic	Applied	Basic	Applied	All
Demographics					
Age	41.44	43.19	40.52	42.05	42.04
Gender (male)	0.67	0.67	0.75	0.80	0.72
White	0.52	0.52	0.50	0.51	0.51
Married	0.75	0.76	0.74	0.80	0.77
U.S. citizen	0.85	0.86	0.80	0.85	0.85
Earnings					
Real annual earnings	55,285	59,236	74,365	78,632	65,288
Scientist ability					
Years to bachelor degree	4.07	4.40	4.35	4.11	4.17
School ranking	3.35	3.12	3.37	3.21	3.25
Grant recipient	0.59	0.42	0.45	0.41	0.47
Father's education— 4-year college or higher	0.53	0.44	0.53	0.45	0.48
Mother's education— 4-year college or higher	0.37	0.29	0.34	0.28	0.31
Job importance					
Salary	3.39	3.38	3.50	3.45	3.43
Benefits	3.52	3.49	3.50	3.50	3.50
Job security	3.51	3.51	3.40	3.42	3.46
Job location	3.47	3.43	3.40	3.40	3.43
Opportunity for advancement	3.50	3.50	3.43	3.41	3.47
Intellectual challenge	3.86	3.87	3.80	3.77	3.83
Level of responsibility	3.44	3.45	3.40	3.38	3.42
Degree of independence	3.76	3.75	3.63	3.60	3.70
Contribution to society	3.56	3.53	3.31	3.32	3.45
Number of observations	12,370	8,084	1,194	12,128	33,776

Source. Authors' estimation using restricted use SDR data for 1995–2006.
Notes. A subsample from 1995, 1997, and 1999 SDRs is used for statistics on scientist ability. For job importance, 2001 and 2003 SDRs are used to produce descriptive statistics.

ask scientists to self-report their *primary* job activity as one still permits them to have aspects of the other while nonetheless categorizing themselves in a primary activity, and alleviates some of the problems of the other measures, as described by Sauermann and Stephan (2012).

3.1.2.   Earnings Profile. Survey respondents report their annualized salary, excluding bonuses, overtime or other additional compensation. We deflate these annualized earnings by the consumer price index (base year is 1995). As seen in Table 2, average earnings (averaged over individuals and time) of applied scientists in industry are the highest among the four groups, and the average earnings of basic scientists in academia are the lowest.[5] Notably, there is an approximately $23,000 annual earnings differential between these two groups.

3.1.3.   Scientist Characteristics. For characteristics of scientists, key variables in our theoretical model relate to their research ability θ, a taste for nonpecuniary returns δ, and accumulation of human capital over time. Because there are no universal empirical measures for these unobservables, we utilize several pieces of information from the survey to proxy for these variables. To proxy for differences in research ability (time invariant θ), we rely on information related to (i) time to complete first baccalaureate degree, (ii) Ph.D. program ranking by field of science, (iii) whether a scientist received a grant during her doctoral program, and (iv) parental educational levels.[6] We obtain data on Ph.D. program ranking from the National Research Council's evaluation of Ph.D. program quality (Golderberger et al. 1995), and other ability measures are within SDR. These variables represent variations in quality and capture the scientist's ability prior to entering in the labor market; thus, they should correlate highly with their research ability. To gauge preferences and determine proxies for a taste for nonpecuniary returns δ, we utilize survey responses to questions asking scientists to rate the relative importance they placed on the following job characteristics: (i) opportunity for advancement, (ii) benefits, (iii) intellectual challenge, (iv) independence, (v) location, (vi) responsibility, (vii) salary, (viii) security, and (ix) contribution to society. The answers were coded on a four-point scale, one being not important at all to four being very important.

We use a variable for labor market experience to estimate how the earnings profile in each career evolves over time (Propositions 3–5). This variable is measured as age minus years of schooling minus six years, and we include both linear and quadratic terms in the earnings equation. Instead of labor experience, age is included as a control in the cross-sectional regressions testing scientists' career choices (Propositions 1 and 2). Additionally, we use several demographic control variables such as dummies for gender, race, marital status, and U.S. citizenship, with one denoting male, white, married, and U.S. citizen, respectively. Table 2 provides descriptive statistics for the data used in the analysis.

3.2.   Methodology
Propositions 1 and 2 (career choices) are tested using a probit model and are robust to logit, multinomial logit, or linear probability specifications. For

---

[5] Given the 48 weeks threshold for full-time workers noted above, these salary figures include at least two months of "summer" compensation for academics on nine-month contracts.

[6] The relationship between parents' and children's ability, particularly as measured by educational levels, has been well documented in prior studies (cf. review by Haveman and Wolfe 1995), and attributed to both biological (i.e., genetic) and environmental (e.g., home investments) factors (Leibowitz 1974, Sacerdote 2000).

Proposition 1, the following model is tested separately for academia/industry:

$$\Pr(D_{ji} = 1) = \Phi(\beta_{0j} + \beta_{1j}\, ability\, measures_{ji} + \beta_{2j}X_{ji}), \quad (20)$$

where the dummy dependent variable takes the value of 1 if a scientist i conducts basic research and 0 if applied research within each institutional setting j. We use information from the 1995, 1997, and 1999 SDR files, given lack of relevant information for the other years. In addition to the reported empirical measures of an individual's research ability, we included the demographic variables listed above and the scientist's educational field as controls, although these coefficients are not reported. For Proposition 2, we use the 2001 and 2003 SDR files, because the job importance questions were only asked in these two survey years. The dependent variable in (20) is modified to indicate whether a scientist works in academia or industry, and ability measures are replaced by job importance variables. The analysis reported in Tables 3 and 4 for Proposition 1 and 2 is conducted by pooling the cross-section data across the relevant survey years, with one observation per individual to ensure independence of observations.

Propositions 3, 4, and 5 related to earnings profile of scientists are tested using all SDR files and several

regression techniques. Specifically, we first use pooled cross-sectional data to estimate

$$\ln(y_{gi}) = \beta_{0g} + \beta_{1g}experience_{gi} + \beta_{2g}(experience_{gi})^2$$
$$+ \beta_{3g}X_{gi} + \varepsilon_{gi}, \quad (21)$$

where y_{gi} is self-reported annualized salary for individual i in career choice g. Here, experience relates to labor market experience, and X is a vector of control variables described above. In Tables 5(a) and 5(b), we report the analysis using the longitudinal data, clustering standard errors for individuals who appear multiple times to account for nonindependence of observations.[7] Further, given the potential of confounding of labor experience and cohort or period effects (e.g., Glenn 2005), we need to ensure that the results are not being driven by differences in vintages of cohorts (Stephan 1996). Although labor market experience and cohort effects cannot be completely separated out empirically with nonexperimental data, we can nonetheless examine the extent to which cohort effects affect our main findings. In Tables 6 and 7, we use panel data for the subsample of repeated observations of scientists. We define seven cohorts of scientists and engineers based on doctoral degree completion during a given five-year period, starting from the 1970–1974 cohort and ending with the 2000–2006 cohort. Further, we create labor experience categories of five-year spans (starting with less than 5 years and up to greater than 25 years). We then perform ordinary least squares (OLS) regressions for each labor experience category, with the dependent variable as the log of real annualized salary and independent variables as cohort dummies and demographic control variables used above. We also estimate a slope of an earnings equation—the effect of labor experience—separately for each cohort, and report the analysis for the middle cohorts of the 1985–1989 and 1990–1994 graduates. Labor market experience in the former cohort varies 5–24 years and in the latter cohort from 0 to 19 years. Although they are not reported because of space constraints, we obtained similar results for other cohorts.

Table 3   Test for Proposition 1 (Choice of Basic/Applied Research Within Each Institutional Setting)

	Academia	Industry
Years to completion of first bachelor	0.009*	−0.006**
	(0.005)	(0.003)
Grant	0.13***	0.01
	(0.02)	(0.01)
Ph.D. program ranking	0.05***	0.009
	(0.01)	(0.005)
Father's education—4-year college or higher	0.05***	0.02
	(0.02)	(0.01)
Mother's education—4-year college or higher	0.03	0.01
	(0.02)	(0.01)
Number of observations	3,707	3,002
Pseudo R-squared	0.12	0.04
Log-likelihood	−2,218.50	−804.90

Source. Authors' estimation using SDR samples of restricted SESTAT data for 1995, 1997, and 1999.

Notes. The dependent variable is a dummy variable that takes 1 if a scientist is a basic scientist and 0 if he or she is an applied scientist. Years to complete first bachelor degree is defined as negative of years spent in completing first bachelor degree (e.g., 4-year is coded as −4). A marginal increase in this variable means less time of completion of 1st bachelor degree. A marginal increase in the variable of Ph.D. program ranking means a higher ranking of a Ph.D. program. Estimated marginal effects evaluated at sample mean values are reported in the table. Numbers in the parentheses are standard errors. Dummies of demographic characteristics and a field of study are included in the regression, though these estimates are not reported.

***, ** and * indicate that coefficients are significant at the 1%, 5%, and 10% levels, respectively.

4.   Results

4.1.   Propositions Related to Career Choices of Scientists

The estimated marginal effects for Proposition 1 are reported in Table 3. We test whether proxy variables for research ability result in sorting of basic or applied research among academic and industrial scientists,

---

[7] The results are robust to specifications where we use only one observation per individual.

Table 4   Test for Proposition 2 (Choice of Academia vs. Industry Based on Pecuniary and Nonpecuniary Returns)

	Panel (a)		Panel (b)		Panel (c)	
	(I)	(II)	(I)	(II)	(I)	(II)
Salary	−0.12***	−0.09***				
	(0.01)	(0.01)				
Benefits	0.01	0.01				
	(0.01)	(0.01)				
Job security	0.08***	0.06***				
	(0.01)	(0.01)				
Job location	−0.005	0.001				
	(0.01)	(0.01)				
Opportunity for advancement	0.04***	0.04***				
	(0.01)	(0.01)				
Intellectual challenge	0.01	0.04**				
	(0.02)	(0.02)				
Level of responsibility	−0.06***	−0.04***				
	(0.01)	(0.01)				
Degree of independence	0.10***	0.12***				
	(0.01)	(0.01)				
Contribution to society	0.11***	0.11***				
	(0.01)	(0.01)				
Nonpecuniary factors			0.13***	0.20***	0.07***	0.09***
			(0.02)	(0.02)	(0.006)	(0.006)
Pecuniary factors			−0.06***	−0.04***	−0.01*	−0.003
			(0.01)	(0.01)	(0.006)	(0.006)
Real annual salary ($1,000)		−0.006***		−0.007***		−0.007***
		(<0.001)		(<0.001)		(<0.0001)
No. of observations	7,466	7,466	7,466	7,466	7,466	7,466
Pseudo *R*-squared	0.06	0.16	0.03	0.14	0.03	0.15
Log-likelihood	−4,769.08	−4,264.31	−4,932.72	−4,381.60	−4,903.30	−4,341.52
Inclusion of real salary	No	Yes	No	Yes	No	Yes

Source. Authors' estimation using restricted SDR data for 2001 and 2003.
Notes. The dependent variable is a dummy variable that takes 1 if a scientist is an academic scientist and 0 if he or she is an industrial scientist. In panel (a), each variable of job importance is four-point scale, coded as 4 if very important, 3 if somewhat important, 2 if unimportant, and 1 if not important at all. In panel (b), the variable of nonpecuniary factors is constructed from taking the average of the original variables of challenge, responsibility, and independence. Similarly, the variable of pecuniary factors is constructed from taking average of the original variables of salary and benefit. In panel (c), the variables of nonpecuniary and pecuniary factors are predicted values from the factor analysis by Bartlett scoring method. The nonpecuniary factor is highly correlated with challenge, responsibility, and independence. Similarly, the pecuniary factor is highly correlated with salary and benefit. Estimated marginal effects evaluated are reported in the table. Numbers in the parentheses are standard errors.
    ***, ** and * indicate that coefficients are significant at the 1%, 5%, and 10% levels, respectively.

respectively. Most demographic variables are not statistically significant. The estimated results show support for Proposition 1: for academic scientists, the estimated marginal effects of grant reception, Ph.D. program ranking, and father's education are positive and statistically significant at the 1% significance level, although this is not the case for industrial scientists. For academia, the probability of being basic over applied is likely to increase when scientists spent less time to complete the first baccalaureate degree, received a grant during doctoral study, and had higher parental education levels. By contrast, in industry, estimated effects of a grant reception and parents' education are statistically insignificant, and time to complete a bachelor's degree has an estimated negative impact on the career option of a basic scientist.

To test Proposition 2 regarding importance of nonpecuniary benefits, we use survey respondents'

evaluation of the importance of nine attributes of their job. Because no universal way to classify a job attribute to either pecuniary or nonpecuniary benefits exists, we use three related measures to test Proposition 2. Table 4, panel (a) reports the marginal effects on the likelihood of having an academic career (basic or applied). First, we use all nine original variables of job importance independently. According to model (I) in panel (a) the probability of being an academic rather than an industrial scientist decreases when salary or responsibility is rated as an important job aspect. In contrast, independence, job security, and contribution to society have a positive effect (similar in magnitude to salary) on the likelihood of choosing academia. To address the concern that receiving a high salary correlates with the likelihood of rating it as an important attribute, we add real annual salary in model (II) of Table 4, panel (a) (and also

Table 5(a)   Tests for Propositions 3, 4, and 5 (Earnings Profile for Each Career Option) OLS

	Academia		Industry	
	Basic	Applied	Basic	Applied
Labor experience	0.06***	0.04***	0.04***	0.03***
	(0.002)	(0.003)	(0.007)	(0.003)
Labor experience squared	−0.001***	−0.001***	−0.001***	−0.001***
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
Gender dummy	0.17***	0.14***	0.11***	0.05***
	(0.01)	(0.02)	(0.05)	(0.02)
Marriage dummy	0.08***	0.09***	0.02	0.05***
	(0.01)	(0.02)	(0.04)	(0.02)
White dummy	0.01	0.05***	−0.03	0.07***
	(0.01)	(0.01)	(0.03)	(0.01)
Citizenship dummy	0.18***	0.23***	0.04	0.05***
	(0.02)	(0.02)	(0.04)	(0.02)
Constant	9.77***	9.97***	10.61***	10.71***
	(0.02)	(0.03)	(0.06)	(0.03)
Number of observations	12,370	8,084	1,194	12,128
R-squared	0.31	0.16	0.10	0.05

Source. Authors' estimation using restricted SDR data for 1995–2006.
Notes. The dependent variable is logarithm of annualized earnings. All regressions use pooled data of restricted SDR data for 1995, 1997, 1999, 2001, 2003, and 2006. Numbers in the parentheses are robust standard errors clustered at individual level.

***, ** and * indicate that coefficients are significant at the 1%, 5%, and 10% levels, respectively.

Table 5(b)   Tests for Propositions 3, 4, and 5 (Equality of Estimated Coefficients in Table 5(a))

	Within institution		Between institution	
	Basic vs. Applied		Academia vs. Industry	
	Academia	Industry	Basic	Applied
(i) Constant	< 0.001	0.15	< 0.001	< 0.001
(ii) Labor experience	< 0.001	0.15	0.02	0.001
(iii) Labor experience squared	0.35	0.35	0.90	0.24
(ii) and (iii)	< 0.001	0.14	< 0.001	< 0.001
(i), (ii), and (iii)	< 0.001	0.17	< 0.001	< 0.001

Source. Authors' estimation using restricted SDR data for 1995, 1997, 1999, 2001, 2003, and 2006.
Note. Numbers in the tables are *p*-values for testing the hypothesis that the coefficients are equal across two comparison groups.

in Table 4, panels (b) and (c)). The results remain largely unchanged and are strengthened by the finding that actual salary received results in a lower likelihood of choosing an academic career. Second, we use factor analysis of the nine attributes to cluster and account for underlying factors that may systematically influence respondents' answers to survey questions about job attributes. Using the customary 0.7 of factor loading as a cutoff, the factor analysis suggests that salary and benefits may be bundled; accordingly, we use the average value of these two variables to construct a new variable that captures the importance of pecuniary benefits. Similarly, the factor analysis suggests that intellectual challenge, responsibility, and independence can be bundled together, and we interpret the average value of these variables as indicating the importance of nonpecuniary benefits. Regression results using these new variables are reported in Table 4, panel (b). Finally, we predict values of the two factors by using the Bartlett scoring method and report results from these variables in section of the table. All models show that the more important nonpecuniary benefits are relative to pecuniary returns, the more likely scientists are to sort into academia than into industry. We note also that our classification of variables accords with extant literature. For example, Heneman and Schwab (1985) classify "benefits" as a pecuniary return. Our literature review also suggests that "opportunity for advancement" is a gray area: on one hand, promotions come with monetary returns (Lazear and Rosen 1981); on the other hand, advancement allows for nonpecuniary returns such as greater control/authority (Rynes et al. 1983). In general, "job security" is nonpecuniary (e.g., Hertzberg et al. 1959). People are willing to sacrifice some monetary returns to secure a job. Taking the literature review and regression results into account, we conclude that scientists who value nonpecuniary benefits highly are drawn to academic research. Thus, Proposition 2 is supported by the data.

4.2. Propositions Related to Evolution of Earnings

Tests for Propositions 3–5 using longitudinal data are reported in Tables 5(a) and 5(b). Table 5(a) provides the OLS estimation results for earnings based on labor market experience for each career option. Table 5(b) provides the statistical tests for differences in coefficient values across career options for the estimated earnings profile, with the null hypothesis that each or a combination of estimated coefficients (constant, labor experience, and labor experience squared) is the same between any relevant two options. In Figure 2, we provide the estimated earnings trajectories, based on the coefficients of the linear and quadratic terms of labor experience and at sample mean values of all other independent variables.

The results of the control variables are consistent with past labor economics studies: regardless of the chosen career option, married, white, U.S. citizen, and male scientists on average earn more than their counterparts, and the estimated earnings profiles are concave with respect to labor market experience (Stephan 1996). The earnings trajectories are consistent with Proposition 3; the average earnings profile of both basic and applied academic scientists are lower than for industrial scientists (see Figure 2 and tests of equality in Table 5(b)). Consistent with Proposition 4,

Figure 2   Estimated Earnings Profiles for Each Career Option
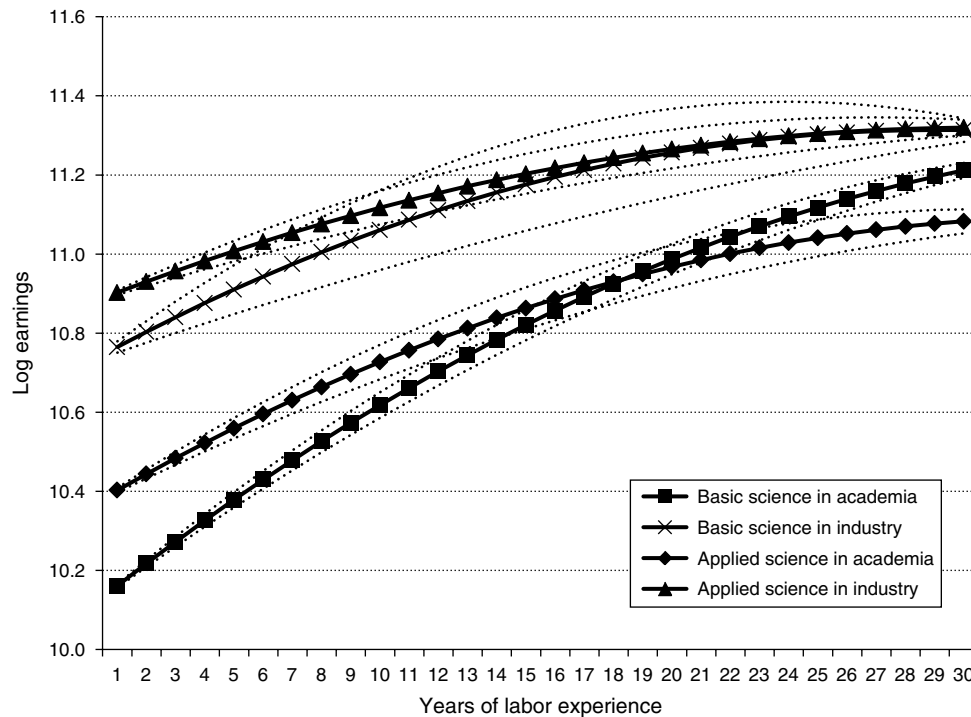
Source. Authors' estimation using restricted SDR data for 1995, 1997, 1999, 2001, 2003, and 2006.
Note. Dotted lines represent 95% confidence intervals.

the initial earnings of academic basic scientists are lower than the initial earnings of academic applied scientists. Indeed, academic basic scientists make the lowest initial earnings of all other scientists, as is evident from Figure 2, tests reported in Table 5(b), and alternative analysis in the online appendix, limiting the sample to observations for labor experience of less than or equal to two years. However, consistent with Proposition 5(a), academic basic scientists also have the steepest slope, such that the earnings of academic basic scientists "catch up" to the earnings of the industrial scientists toward the very end of the career. The coefficient values for academic basic scientists are significantly different from the coefficient values for academic applied scientists. Consistent with Proposition 5(b), there is no such statistically significant relationship among industrial scientists for basic and applied research; the null hypotheses for equality of coefficients cannot be rejected for any combination of the constant, linear, and quadratic terms of labor experience.

We also test the estimations of differences in career options using longitudinal data, particularly to investigate whether the results are driven by cohort effects. Given space constraints, we limit our analysis of Proposition 5 to academia, because our theory predicts differences in basic and academic science only

for academic scientists, and not for industrial scientists. Table 6 reports the OLS regressions for basic and applied research, respectively. The dependent variable is the log of real annualized salary within each labor experience category, and the independent variables include cohort dummies and demographic control variables used for cross-sectional regression analyses. The cohort effects in Table 6 are statistically insignificant at conventional significance levels in most experience categories, regardless of basic or applied research. We also estimate the relationship between earnings and labor experience separately for the 1985–1989 and 1990–1994 cohorts. Consistent with Proposition 5, in both cohorts, the estimated slope of an earnings equation in Table 7 is steeper for academic basic scientists than for academic applied scientists for the relevant range of labor experience. Further, in both cohorts, consistent with Propositions 3 and 4, the estimated coefficient of the intercept term for industrial scientists is higher than for academic scientists; within academia, it is higher for applied than for basic scientists (statistical tests available on request). Thus, taken together, cohort effects reveal minimal impact on the observed pattern of earnings evolution (Table 6); further, tracking specific cohorts (Table 7) yields a similar pattern to the longitudinal analysis.

Table 6　Tests for Cohort Effects in Academic Basic and Applied Research

	Academic basic research					Academic applied research				
	Labor experience categories					Labor experience categories				
	< 5 years	5–9 years	10–14 years	15–19 years	20–24 years	< 5 years	5–9 years	10–14 years	15–19 years	20–24 years
Cohort 1 (1970–1974 graduates)					Base					Base
Cohort 2 (1975–1979 graduates)				Base	−0.04 (0.06)				Base	−0.06 (0.07)
Cohort 3 (1980–1984 graduates)			Base	0.008 (0.04)	0.013 (0.07)			Base	−0.01 (0.04)	−0.20 (0.20)
Cohort 4 (1985–1989 graduates)		Base	−0.03 (0.04)	−0.009 (0.07)	−0.16 (0.17)		Base	−0.01 (0.04)	0.08 (0.05)	0.05 (0.20)
Cohort 5 (1990–1994 graduates)	Base	−0.08*** (0.03)	−0.05 (0.05)	−0.03 (0.11)		Base	−0.16*** (0.04)	−0.02 (0.05)	−0.15 (0.11)	
Cohort 6 (1995–1999 graduates)	−0.007 (0.03)	−0.05 (0.06)	−0.26** (0.11)			−0.03 (0.04)	−0.04 (0.05)	−0.25 (0.17)		
Cohort 7 (2000–2006 graduates)	0.05 (0.03)	−0.27** (0.11)				−0.07 (0.08)	−0.25 (0.16)			
Constant	10.22*** (0.04)	10.49*** (0.05)	10.78*** (0.09)	10.60*** (0.17)	10.44*** (0.67)	10.24*** (0.06)	10.75*** (0.10)	10.74*** (0.10)	11.02*** (0.09)	11.26*** (0.21)
No. of observations	1,567	1,135	910	793	600	693	581	449	418	322
p-value	< 0.001	< 0.001	< 0.001	0.001	0.09	0.005	< 0.001	< 0.001	< 0.001	0.33
R-squared	0.02	0.06	0.05	0.03	0.06	0.02	0.04	0.05	0.05	0.03

Source. Authors' estimation using restricted use SDR data for 1995–2006.
Notes. The dependent variable is logarithm of annualized earnings. Numbers in the parentheses are standard errors. The *p*-values are for a joint significance test.
***, ** and * indicate that coefficients are significant at the 1%, 5%, and 10% levels, respectively.

5.　Model Extensions, Alternative Explanations, Robustness Checks, and Limitations

5.1.　Implications of Relaxation of Assumptions and Limitations of the Theoretical Model

Several simplifying assumptions enabled us to focus attention on the salient features of sorting of scientists and the resultant earnings evolution. Chief among them are the assumptions of no uncertainty and no change across careers to make the model analytically tractable, which prevented us from studying scientists' sequential choices of careers. The transition matrices in the SDR database reveal that 97% do not switch careers between industry and academic settings, and 86% do not switch between basic and applied science (Agarwal and Sonka 2010). Further, there is no significant difference in the fraction of junior scientists in the sample (within two years of graduation) doing applied versus basic research within each institutional setting: 16.8% versus 18.5% in academia and 11.1% versus 14.9% in industry. Nonetheless, this simplifying assumption precludes the possibility that the complementarity between basic and applied scientists in academia increases as they become more senior and that scientists make career choices strategically to maximize their lifetime utility (Dasgupta and David 1994). Relaxation of this assumption leads to a revised model that predicts a "cash out" story: basic academic scientists, particularly those with higher ability, will switch to applied science at later life cycle stages. In addition, the model suggests that only able scientists can actually pursue strategic career switches, because such options are of limited availability and an assortative matching likely prevails. This is consistent with the literature on star scientists being more likely to engage in applications of their scientific discoveries (e.g., Stuart and Ding 2006). Because our model does not allow scientists to make such career changes, this results in underevaluating the earnings evolution of able scientists.

Further, the assumptions that scientists work in either basic or applied science but not both and that taste for science is exogenous abstract away from within-individual complementarity of the two activities (Mansfield 1995) and the potential that taste for science may be endogenous to time allocation of a scientist toward basic and applied scientific activities (Levin and Stephan 1991, Thursby et al. 2007). Thus, our approach is unable to address important issues such as effects of within-individual complementarity in basic and applied science and endogenous taste for science on optimal time allocation, evolution of human capital along basic and applied dimensions, and reputation-building strategies and outcomes. Incorporation of these elements is beyond

Table 7   Earnings Equation by Cohort

	Academia		Industry	
	Basic	Applied	Basic	Applied
Cohort of 1985–1989 graduates				
Labor experience	0.05***	0.02**	0.02	0.04***
	(0.01)	(0.009)	(0.02)	(0.009)
Labor experience squared	−0.001***	−0.001**	−0.001***	−0.001***
	(0.0003)	(0.0002)	(0.0006)	(0.0003)
Constant	10.16***	10.50***	11.14***	10.87***
	(0.09)	(0.09)	(0.27)	(0.07)
Number of observations	1,866	1,243	210	2 2,259
p-value for joint significance	< 0.001	< 0.001	0.664	< 0.001
Cohort of 1990–1994 graduates				
Labor experience	0.06***	0.05***	0.06***	0.03***
	(0.005)	(0.007)	(0.02)	(0.006)
Labor experience squared	−0.001***	−0.001**	−0.001**	−0.001***
	(0.0002)	(0.0002)	(0.0007)	(0.0002)
Constant	9.88***	10.00***	10.55***	10.81***
	(0.04)	(0.07)	(0.13)	(0.03)
Number of observations	2,765	1,829	279	2,733
p-value for joint significance	< 0.001	< 0.001	< 0.001	0.004

Source. Authors' estimation using restricted use SDR data for 1995–2006.
Notes. The dependent variable is logarithm of annualized earnings. The regression specification is the same as the one in Table 5(a), though we only report constant, labor experience, and labor experience, squared. Numbers in the parentheses are standard errors. For a joint significance test, the null hypothesis is that all coefficients are zero.

***, ** and * indicate that coefficients are significant at the 1%, 5%, and 10% levels, respectively.

the scope of this paper, given additional complexity and model intractability. Thus, our analysis focuses on where scientists conduct research (similar to Aghion et al. 2008 and Stern 2004, who also treat taste for science as exogenous), assuming that there are no switches or gradual shifts over time between basic and applied research.

Another simplifying assumption relates to the ability parameters of basic and applied scientists entering symmetrically into the knowledge production function for type B firms in industry. Relaxing this assumption increases the complexity of the model significantly and obscures the model's prediction regarding similarity of earnings for industry basic and applied scientists. The prediction hinges on the fact that the asymmetric force tends to create a gap in earnings profiles of basic and applied scientists, whereas the synergy and complementary forces work to narrow that gap. However, for a reasonable range of parameters, we can still obtain the model prediction. Further, the assumption of stable matching (and related assumption of full information) helped derive an equilibrium matching pattern. As in Shimer (2005), when a coordination friction is introduced at

the expense of analytical complexity, the assortative matching becomes imperfect. Some less able scientists are paired with more capable research institutions or scientists, but the average characteristic of a sorting pattern and earnings evolution remains the same qualitatively.

In the context of the earnings profile, relaxing the assumption that scientists' pay is not linked to current investment in human capital but only to accumulated human capital permits the possibility that current human capital investments also create immediate returns and will result in more time devoted to human capital investment, provided that scientists' other activities are assumed away. This makes the earnings profiles steeper than currently predicted. Alternatively, if one models the scientists' other activities as requiring both time and current research output, human capital accumulation slows down, resulting in a flatter earnings profile, though because of shift up because of income from other activities.

5.2. Robustness Checks, Alternative Explanations, and Limitations for the Empirical Findings

Our empirical analysis used a large sample of scientists and therefore provided general empirical insights regarding the model's predictions. However, several alternative explanations need to be addressed, particularly as they relate to underlying heterogeneity of jobs and characteristics of scientists, resulting in similar predictions. We first examine whether the differences in earnings profile between academic and industry scientists may be a result of gender composition: robustness checks reported in the online appendix reveal that the propositions are supported for both male and female subsamples. A second concern is whether mean estimates of earnings profiles from OLS regressions are consistent with the data at different points of the earnings distribution— i.e., it is possible that opposite effects of high earners and low earners cancel each other out to generate the average earnings profiles estimated from OLS regressions. Quantile regressions at the first, second, and third quartiles show that the model's predictions regarding earnings profiles are supported (see the online appendix). Furthermore, this result indicates that the assortative sorting story is consistent with the data and strengthens our empirical analysis. Third, we examine whether the propositions are robust to within-science field comparisons for life and physical sciences, particularly because the delineation of "basic" versus "applied" may differ across them. We could not conduct this analysis for other science and engineering areas, because of the lack of sufficient observations for meeting NSF disclosure requirements. As seen in the online appendix, with one exception, all the propositions are largely supported within the subsamples. The only proposition

for which we lacked support was Proposition 4 in the physical sciences; we did not find support for differences in initial earnings between academic basic and applied scientists. We also note that in the life sciences, we received partial support for Proposition 5: in addition to the predicted steepness of the earnings slope for basic scientists relative to applied scientists in academia, we also find similar results in industry (as opposed to the predicted no significant difference in slopes in industry). Consistent with extant studies that use biotechnology as an exemplar of an industry that is based on basic science, the results indicate that basic scientists, regardless of whether they are in academia or industry, enjoy higher returns relative to applied scientists. In unreported regressions, we also confirmed that our analysis is robust to the exclusion of post doctoral scientists and of nontenure track faculty in academia.

Among the empirical limitations of our study, chief is our inability to rule out some competing explanations. Nonetheless, we are able to examine the explanatory power of alternative models relative to ours by using our model's unique implications as guidance. In particular, we believe that although some of the results may be explained piecemeal by alternative theories, these theories cannot simultaneously account for all the empirical findings. For example, earnings differentials between academia and industry could be attributed to a relatively small supply of academic positions, but this simple supply-and-demand story cannot explain similarities and/or differences in the evolution of earnings profiles between basic and applied scientists. High fixed costs in applied science may induce research institutions to hire able scientists and lead to higher earnings of applied scientists, but they cannot explain differences in the slopes observed across institutional settings, particularly when compared with the earnings slopes of basic scientists. Similarly, differences in preferences for nonpecuniary benefits between basic science and applied science within each research institution may result in earnings differentials between the two, but would not explain why the differential exists within academia and not industry. Alternatively, as Nickerson and Zenger (2008) argue, peer envy may result in similar compensation for basic and applied scientists in industry, but this explanation leaves unanswered why academic settings are free from such peer envy. More importantly, it is hard to explain the evidence of this study that salary is a more important factor for industrial scientists than academic scientists.

Although the empirical results of this study are largely consistent with the theoretical implications, we acknowledge that our empirical analysis is descriptive in the sense that it explores empirical patterns of sorting and earnings evolution rather than establishing a causal relationship. Accordingly, some empirical findings in this paper need to be interpreted with caution. Further, we note that our main interest is not to estimate an earnings profile of each career path for a scientist randomly chosen from the population; thus, we do not attempt to control for potential selection effects in the empirical analysis. Nonetheless, our theoretical model regarding the sorting of scientists guides us to infer a direction of the potential selection effects in our empirical results. Because this and other studies report that academic scientists have a taste for science, an estimated earnings profile of academic scientists in our study would be downward biased from random selection. Ability or learning ability sorting may result in an upward bias from random assignment for the estimated slope of basic scientists in academia and a downward bias for applied scientists in academia. An estimated earnings profile of industrial scientists is likely upwardly biased from random selection, given the positive sorting between basic and applied scientists. Finally, we are also cognizant of the fact that additional useful insights may be obtained by empirically examining a causal or structural relationship between sorting of scientists and earnings evolution. Unfortunately, a simple self-selection model such as a Roy model (Roy 1951, Heckman and Honoré 1990) does not resolve a self-selection problem because of the nature of two-sided matching. The problem is further exacerbated by the fact that we do not have data from a random experiment. Accordingly, we leave it for future research to firmly establish a causal relationship by developing a full structural estimation model with a two-sided matching estimation.

6. Discussion and Conclusion

Our paper integrates matching theory with life cycle models of human capital investments to analyze labor markets for scientists, where heterogeneous characteristics of scientists and institutions result in an optimal match between specific individuals and the alternative career options. Specifically, we model how scientists who differ across two dimensions—ability and taste for nonpecuniary returns—choose their careers. The sorting patterns for early career scientists, along with differences in complementarities between basic and applied scientists and in investments in human and physical capital within academia and industry also have implications for the earnings trajectory for scientists within each career option. We test the model implications using comprehensive and rich data compiled by the NSF's SESTAT program and find support for its predictions.

Our model's predictions for the sorting of scientists and the resultant earnings evolution contribute to the literature in the economics of science. By integrating insights from matching theory (Becker 1973)

into traditional models of scientific labor markets that build on life cycles in human capital investments (Becker 1962), we address Stephan's (1996) call for better modeling of the economics of science. For example, some of our model implications are very consistent with the implications generated by life cycle models (e.g., Levin and Stephan 1991, Thursby et al. 2007), but incorporating the two-sided matching element permits us to generate new implications: our study models how competition within scientists for preferred jobs and within institutions for preferred scientists results in matching based on ability, preferences, and complementarity in multifactor scientific production functions. The matching outcomes in turn influence the evolution of human capital and earnings over time. Our model thus generates some novel insights that are backed by the empirical analysis: in academia, scientists of higher ability sort into basic rather than applied research, and initial earnings of basic scientists are lower but the slope of earnings is higher relative to applied scientists. In industry, by contrast, there is no such ability sorting, and the earnings trajectories of basic and applied scientists are similar to each other. Thus, we hope that our study sheds light on how possible self- or sample-selection biases may influence applications of the life cycle model to specific empirical contexts.

Our study also contributes new insights by building on extant work related to the "taste for science" (Roach and Sauermann 2010, Stern 2004). We predict and find that a higher taste for nonpecuniary returns sorts scientists in academia over industry. Stern (2004) focuses primarily on industry and shows that scientists with higher "taste for science" are willing to accept a compensating wage differential to work for science oriented firms, and we use a similar logic to show how such preferences may result in these scientists preferring academia over industry. Further, by incorporating life cycle–related trade-offs between current and future opportunities, we generate additional important insights regarding the relation between a taste for nonpecuniary returns and earnings trajectories. Our study reveals that the initial earnings of basic scientists are lower than applied scientists within academia, in spite of the fact that science norms and nonmonetary benefits are very similar in these research domains. We posit and show that basic academic scientists may sacrifice current earnings for steeper growth over their life cycle, relative to applied academic scientists. The empirical analyses support our model prediction that basic and applied scientists have similar earnings in industry, given synergies between the two. Thus, our model implications indicate that the observed earnings differential between academia and industry stem from both differences in a taste for science and

the presence (absence) of complementarities between basic and applied scientists in industry (academia). This result holds even after controlling for observable and unobservable characteristics of scientists.

Further, extant work on careers of scientists typically examines either industry or academic scientists or examines the institutional differences in narrowly defined fields such as life sciences (Bercovitz and Feldman 2008, Roach and Sauermann 2010, Sauermann and Stephan 2012, Stuart and Ding 2006). Using a broad, rich data set allows us to build on extant work by comparing industry and academic scientists across a broad, generalizable range of fields. Our model and empirical findings are consistent with the survey-based study of Roach and Sauermann (2010), who show that scientists self-select between academia and industry based on their "taste for science," and with Sauermann and Stephan (2012), who explore similarities and differences between academic and industrial science. Consistent with Roach and Sauermann (2010) and our study's Propositions 2 and 3, Sauermann and Stephan (2012) find that academic scientists are relatively more satisfied with nonpecuniary benefits than industrial scientists, thus underscoring the importance of nonmonetary benefits in sorting scientists between academia and industry. Further, our longitudinal findings related to Propositions 4 and 5 regarding earnings profiles within and across institutional settings provide a more nuanced relationship that complements their cross-sectional findings. They find that academic scientists earn less than industrial scientists; we find a striking similarity of earnings profiles between basic and applied scientists within industry, but divergence of earnings profiles between basic and applied scientists in academia. Additionally, we find that at later career stages, basic academic scientists' earnings are not significantly different from those of industrial scientists.

Our study provides managerial and policy implications for each institutional setting that we examine. Within industry, policy makers and innovation managers need to recognize that high complementarities between basic and applied scientists permit them to attract equally able scientists in either research domain, and thus these complementarities deserve to be emphasized by creating cultures that recognize the importance of basic science to technological applications. Within academia, our study offers some insights to policy makers and administrators, particularly in the hiring and retention policies for scientists in basic and applied domains. For instance, our model shows that a sorting of ability within basic and applied academic research is driven by the differential investments in complementary physical capital. Thus, to the extent that national science and innovation policy

may encourage universities to engage in more technology transfer, there is a need to develop stronger complementarities between basic and applied scientists. Existing models underscore the value of time allocation of an individual scientist between basic and applied activities, but our study highlights that increasing complementarities across individuals that specialize in basic or applied science may also have beneficial outcomes.

Our study also offers young scientists seeking to embark on alternative careers an understanding of the sorting patterns and average earnings, so that they can make informed decisions based on an assessment of their own ability and preferences. This is even more important because all young scientists are trained in academia, where there is often a pejorative assessment placed on either applied research or working in industry settings (Stephan 1996). Given that only 26% of scientists are employed in basic academic research (Agarwal and Sonka 2010), for the United States to remain competitive in an increasingly knowledge-based economy, there is also a need for renewed assessment of our science education policy, and to examine if our universities provide the other 74% of their graduate student population with the necessary knowledge, skills, and attitudes required for success. Such an assessment may lead to the more widespread development of innovative programs within universities that will provide young scientists with the literacy and experiential learning of the economic, business, and legal issues they may encounter in their careers (Agarwal and Sonka 2010). Similarly, national innovation policies may also want to encourage focused programs developed for lifelong learning skills for scientists and provide for the development of continued education programs that permit scientists in the workforce to efficiently learn the complementary management and entrepreneurship skills that they need to be more successful in leveraging their scientific accomplishments.

In summary, our theoretical model and the empirical findings identify the impact on institutional differences on the microlevel choices made by economic agents (firms and universities on the demand side, and scientists with heterogeneous ability and preferences on the supply side) and is an important first step toward identifying the extent to which market forces guide the allocation of scientists. Our theoretical study demonstrates this point by using insights from two-sided matching theory. Based on the premise that there may be higher complementarities between basic and applied scientists in industry production functions relative to academia, our model derives distinctive implications regarding career choices and resultant earning trajectories of scientists. As a result, our study can be regarded as

a minimal test (i.e., a necessary condition test) for the existence of the complementarity and contributes to the literature that has previously documented the synergies in basic and applied research conducted in industry settings.

Acknowledgments

References

Agarwal R, Sonka S (2010) Different strokes for different folks: University programs that enable diverse career choices of young scientists. Libecap G, Thursby GM, Hoskinson S, eds. *Advances in the Study of Entrepreneurship, Innovation and Economic Growth*, Spanning Boundaries and Disciplines: University Technology Commercialization in the Idea Age, Vol. 21 (Emerald Publishing, Cambridge, MA), 139–164.

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

Argyres N, Liebskind J (1998) Privatizing the intellectual commons: Universities and the commercialization of biotechnology research. *J. Econom. Behav. Organ.* 35(4):427–454.

Becker GS (1962) Investment in human capital: A theoretical approach. *J. Political Econom.* 79(5):9–49.

Becker GS (1973) A theory of marriage: Part I. *J. Political Econom.* 81(4):813–846.

Ben-Porath Y (1967) The production of human capital and the life cycle of earnings. *J. Political Econom.* 75(4):352–365.

Bercovitz J, Feldman M (2008) Academic entrepreneurs: Organizational change at the individual level. *Organ. Sci.* 19(1):69–89.

Boardman PC, Bozeman B (2007) Role strain in university research centers. *J. Higher Ed.* 78(4):430–463.

Boardman PC, Ponomariov BL (2007) Reward systems and NSF university research centers: The impact of tenure on university scientists' valuation of applied and commercially relevant research. *J. Higher Ed.* 78(1):51–70.

Braxton JM, Luckey W, Helland P (2002) *Institutionalizing a Broader View of Scholarship Through Boyer's Four Domains* (ASHE-ERIC Higher Education Report). (Association for the Study of Higher Education, Washington, DC).

Bush V (1945) *Science the Endless Frontier: A Report to the President on a Program for Postwar Scientific Research* (U.S. Government Printing Office, Washington, DC).

Cohen WM, Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation. *Admin. Quart. Sci.* 35(1):128–152.

Dasgupta P, David P (1994) Toward a new economics of science and technology. *Res. Policy* 23(5):487–521.

Etzkowitz H (1998) The norms of entrepreneurial science: Cognitive effects of the new university-industry linkages. *Res. Policy* 27(8):823–834.

Geisler E (1989) University-industry relations: A review of major issues. Link AN, Tassey G, eds. *Cooperative Research and Development: The Industry-University-Government Relationship* (Kluwer, Boston), 43–64.

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

Glenn N (2005) *Cohort Analysis* (Sage Publications, Thousand Oaks, CA).

Golderberger ML, Maher BA, Flattau PE (1995) *Research Doctorate Programs in the United States; Continuity and Change* (National Research Council, Washington, DC).

Goldfarb B (2008) The effect of government contracting on academic research: Does the source of funding affect scientific output? *Res. Policy* 37(1):41–58.

Griliches Z (1986) Productivity, R&D, and basic research at the firm level in the 1970s. *Amer. Econom. Rev.* 76(1):141–154.

Haveman R, Wolfe B (1995) The determinants of children's attainments: A review of methods and findings. *J. Econom. Literature* 33(4):1829–1878.

Heckman JJ, Honoré BE (1990) The empirical content of the Roy model. *Econometrica* 58(5):1121–1149.

Heneman HG III, Schwab DP (1985) Pay satisfaction: Its multidimensional nature and measurement. *Internat. J. Psych.* 20:129–141.

Hertzberg F, Mausner B, Snyderman BB (1959) *The Motivation to Work* (John Wiley & Sons, New York).

Hounshell DA, Smith JK (1988) *Science and Corporate Strategy* (Cambridge University Press, New York).

Kremer M, Maskin E (1996) Wage inequality and segregation by skill. NBER Working Paper w5718, National Bureau of Economic Research, Cambridge, MA.

Lazear EP, Rosen S (1981) Rank-order tournaments as optimum labor contracts. *J. Political Econom.* 89(5):841–864.

Levin SG, Stephan PE (1991) Research productivity over the life cycle: Evidence for academic scientists. *Amer. Econom. Rev.* 81(1):114–132.

Leibowitz A (1974) Home investments in children. *J. Political Econom.* 82(2):S111–S131.

Mansfield E (1995) Academic research underlying industrial innovations: Courses, characteristics, and financing. *Rev. Econom. Statist.* 77(1):55–65.

Merton R (1973) *The Sociology of Science: Theoretical and Empirical Investigation* (University of Chicago Press, Chicago).

Mindruta D (2012) Value creation in university-industry relationships: A matching approach. *Strategic Management J.* Forthcoming.

Murray F, Stern S (2007) Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *J. Econom. Behav. Organ.* 63(4):648–687.

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

Nickerson JA, Zenger TR (2008) Envy, comparison costs, and the economic theory of the firm. *Strategic Management J.* 29(13):1429–1449.

Roach M, Sauermann H (2010) A taste for science? Ph.D. scientists' academic orientation and self-selection into research career in industry. *Res. Policy* 39(3):422–434.

Rosenberg N (1990) Why do firms do basic research (with their own money)? *Res. Policy* 19(2):165–174.

Roth AE, Sotomayor M (1992) *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis* (Cambridge University Press, Cambridge, MA).

Roy A (1951) Some thoughts on the distribution of earnings. *Oxford Econom. Papers* 3(2):135–146.

Rynes S, Schwab DP, Heneman HG (1983) The role of pay and market pay variability in job application decisions. *Organ. Behav. Human Performances* 31(3):353–364.

Sacerdote B (2000) The nature and nurture of economic outcomes. NBER Working Paper 7949, National Bureau of Economic Research, Cambridge, MA.

Sattinger M (1993) Assignment models of the distribution of earnings. *J. Econom. Literature* 31:831–880.

Sauermann H, Stephan P (2012) Conflicting logics? A multidimensional view of industrial and academic science. *Organ. Sci.*, ePub ahead of print September 10, http://dx.doi.org/10.1287/orsc.1120.0769.

Shimer S (2005) The assignment of workers to jobs in an economy with coordination frictions. *J. Political Econom.* 113(5): 996–1025.

Stephan PE (1996) The economics of science. *J. Econom. Literature* 34(3):1199–1235.

Stern S (2004) Do scientists pay to be scientists? *Management Sci.* 50(6):835–853.

Stokes DE (1997) *Pasteur's Quadrant* (Brookings Institution Press, Washington, DC).

Stuart T, Ding W (2006) When do scientists become entrepreneurs? The social structural antecedents of commercial activity in the academic life sciences. *Amer. J. Sociol.* 112(1):97–144.

Thursby JG, Thursby MC (2002) Who is selling the ivory tower? Sources of growth in university licensing. *Management Sci.* 48(1):90–104.

Thursby M, Thursby JG, Gupta-Mukherjee S (2007) Are there real effects of licensing on academic research? A life cycle view. *J. Econom. Behav. Organ.* 63(4):577–598.