

Career Dynamics of Doctoral Scientists and Engineers

Natalia Mishagina*
Analysis Group, Inc. †

This version: January 2008‡

Abstract

This paper specifies and estimates a dynamic model of task choices with symmetric learning about one of task-specific abilities and dependence on past performance. The model is applied to the career choice problem of doctoral scientists and engineers to explain the following puzzle: Early in their career, the majority of doctorates are employed in R&D and earn the lowest salaries relative to other tasks. However, as careers progress, they leave R&D for more applied tasks or completely change careers, while those who stay in R&D experience large earnings growth. This paper explains this puzzle with the presence of incomplete information about research ability which can be obtained by engaging in R&D. The model is fit to the rarely used Survey of Doctoral Recipients (1973-2001): a longitudinal data set on the employment histories and earnings of doctorates educated in the US. The parameters of the model are estimated using the Method of Simulated Moments. The predictions of the model are used to evaluate the effects of two counterfactual experiments on the supply of the research skills. First, different learning schemes are compared to the case of full information. Second, the effects of R&D subsidies and changes in the employment options outside sciences and engineering are assessed.

Keywords: *occupational choices; learning; science and technology; human capital; high-skill labor markets.*

JEL Classification Codes: J24, J44, J62, D83.

*I am grateful to Christopher Ferrall and Susumu Imai for their guidance in working on this paper. I appreciate comments from Robert Miller, Sam Kortum, Charles Beach, Michael Waldman, Kala Krishna, Luca Rigotti, Charles Bellemare, Suqin Ge, as well as the participants of the NIEPA poster session (2006) and NASM (2007). I would also like to thank the Division of Science Resources Statistics (NSF) for providing the data and valuable assistance. Finally, I acknowledge Research Fellowship from the NSF/American Statistical Association.

†Analysis Group, Inc. 1080 Beaver Hall, Room 1810 Montreal QC H2Z 1S8 Canada Tel: (514) 394-4460. Email: nmishagina@analysisgroup.com

‡Please do not cite without author's permission. Updated versions of the paper can be found at <http://www.econ.queensu.ca/students/phds/mishagina/Mishagina-JMP.pdf>

1 Introduction

The current state of higher education in sciences and engineering (S&E) in the US has recently become of concern. In particular, policy makers, economists and scientists¹ debate whether the system is adequate to sustain and improve the country's competitiveness in global innovation, Jackson (2003), Freeman et al. (2001), Freeman (2005). More recent re-assessments of the issue suggest that, contrary to some studies, training rates in S&E programs are growing, and labor market data provides no evidence of a shortage but rather of an excess supply of scientists, Butz et al. (2003), Mervis (2003), Teitelbaum (2004).

Ensuring high graduation rates from S&E programs, however, is not sufficient to guarantee retention of high quality specialists in S&E professions. Studies find that a large fraction of S&E baccalaureates work in non-S&E jobs, Lowell and Salzman (2007), Preston (2004). This fact may or may not be of big concern since undergraduate education to a large degree is supported with private loans and savings. At the graduate, and especially doctoral level, attrition from S&E becomes more of a problem since graduate training is predominantly financed with public funds, and, therefore, is costly, Thurgood et al. (2006).

It is found that a large fraction of doctoral research skills are employed outside R&D, Mishagina (2007). In 2001, only 55 percent of S&E doctorates worked in R&D-related tasks, and about 22 percent have never worked in R&D in their life. Employment choices vary throughout a career: Doctorates start their careers in R&D (72 percent) but only 45 percent are still in R&D thirty years later. Finally, about 8 percent works in tasks previously thought unsuitable for S&E doctorates such as financial and other business services, which presents a dramatic career change. These facts suggest that the doctoral skill set including the level of their technological advancement is marketable outside S&E, which facilitates outward mobility when the relative values of employment options change. These changes in career prospects in sciences can be caused by the overproduction of doctorates, the decrease of R&D funding, or falling demand due to outsourcing, Freeman et al. (2001), Lowell and Salzman (2007). Effective policies aimed at the retention of doctorates require understanding the economics behind their career choices.

The observed career dynamics of the S&E doctorates present the following puzzle: The

¹“Is the US Party Over?” by Robert Palazzo (Rensselaer Polytechnic Institute) in “The Scientist”, Volume 25(2), 2008.

first part of the puzzle is that high employment rates in R&D early in a career are accompanied by the lowest starting salaries in R&D relative to other employment options. The second part concerns the attrition from R&D with age. The transition rates out of R&D are non-monotonic and accompanied by the rapidly growing earnings of those who stay. The standard model of self-selection - where ability is known - is unable to explain the observed behavior. The first part of the puzzle cannot be reconciled with the static occupational choice model unless large non-pecuniary benefits are present. Sociology and economic literature concerned with the motives of researchers and the reward structure in sciences traditionally emphasizes that scientists extract high psychological satisfaction from engaging in research, termed as “the joy of puzzle-solving” or “the taste for science”, Stern (1999), Levin and Stephan (1991). The magnitude of these psychological benefits, however, must be enormous to generate the observed employment rates in R&D early in the career considering the value of relative earnings. In the absence of self-selection, they should cover the difference between earnings in R&D and other tasks: \$9,540 a year compared to applied tasks (or 27 percent of annual earnings in R&D) and \$13,000 a year or 35 percent respectively compared to non-S&E. Alternatively, the dynamic occupational choice model would suggest that there is an investment motive of engaging in R&D. For example, R&D can provide training that pays either later on in R&D (Moen (2001)) or in other tasks, thus serving as a “stepping stone job” (Jovanovic and Nyarko (1996)).

Neither of these possible explanations, however, is consistent with the second part of the puzzle, namely high exit rates and large earnings growth for the stayers. First, high non-pecuniary benefits do not explain why individuals leave R&D unless these benefits are decreasing with time. Second, if R&D compensates for early losses in earnings due to training, it is not clear why researchers would leave exactly when they begin to receive a return on their training. If, however, they receive this return performing tasks other than R&D, their earnings in R&D should exhibit the lowest profile compared to the earnings in other tasks. Finally, an alternative explanation to decreasing participation in the profession with age could be related to decreasing productivity with age, which makes the profession “the game of the young”. This case is best illustrated by careers of professional athletes whose performance falls due to the inability of an ageing body to meet the physical demands of

the sports. However, the studies of the productivity of academic scientists found no conclusive evidence that productivity measured in publications or citations falls with age (Stephan (1996)), and the productivity of non-academic scientists has never been studied. In addition, the “game of the young” would manifest itself in the large curvature of the earnings profile and would imply large negative coefficients at the quadratic term in the wage equation². However, average earnings in R&D do not have an early peak but rather grow very rapidly.

This paper provides the following explanation for the observed phenomenon: Productivity in R&D compared to other tasks strongly depends on one’s talent to do research. This talent *ex ante* is not known with certainty but can be learned by engaging in R&D. Therefore, incomplete information about one’s talent drives young researchers into R&D. Later on it forces those with low realized ability to leave the task, while the survivors in R&D enjoy the high growth of earnings. Research talent comes into play in two ways: directly through the ability to pose novel research questions and indirectly through accumulated research capital that allows to complete the research projects. The research capital can be thought of as research skills or developed scientific reputation in the field, which evolves differently for people with different levels of ability: it accumulates faster (depreciates slower) for researchers with higher level of ability. This way even small differences in research ability can generate large differences in productivity over time.

To test the proposed hypothesis and to study the responses of the supply of research skill to changes in the economic environment, the paper develops and estimates a dynamic occupational choice model with symmetric learning about R&D ability and stochastically evolving human capital. Using information on their endowments of task-specific abilities and skill prices, individuals specialize in tasks related versus unrelated to S&E. Tasks within S&E are of two types: a) creation of new knowledge through R&D and b) application of accumulated scientific knowledge through teaching or professional services. The types of tasks differ in their skill requirements and private returns. Ability to produce scientific breakthroughs is incomplete information until the individual engages in research. While in research, ability is revealed with some probability. Research-related human capital evolves

²While many studies report the decreasing performance of athletes with age, an actual decrease in earnings is not well documented.

stochastically: More talented researchers have higher probability to upgrade their research capital while in research. They are also less likely to lose it during research employment interruptions. This can be thought of as a simple way to incorporate dependence on past performance.

Note that this paper defines “career” as a sequence of tasks rather than occupations. This approach is based on the findings that some scientific occupations involve little to no actual involvement in R&D, which is the main interest of policy analysis, Mishagina (2007). Moreover, it was found that a traditional “research career” implies switching away from R&D tasks to other activities. For example, it was found that academic scientists tend to spend less time in R&D and more time teaching or consulting after being tenured, Mishagina (2007). Therefore, if one is interested in studying participation in R&D, conditioning on occupations rather than tasks can be misleading. For example, a switch from industrial to academic research would be recorded as a switch from “scientist” to “post-secondary teacher” and would be mistakenly considered as “leaving research”. Alternatively, promotions would be recorded as a switch from “scientist” to “manager” and would be considered as a “career change”. Conditioning on tasks disentangles mobility due to promotions or sectoral changes as long as an individual is still mainly involved in research on the main job. Finally, mobility described in this paper does not happen due to firm-to-firm mobility, which mainly occurs due to up-or-out contracts in academia. As long as scientists spend most of their time in research, mobility from one university to another is not considered a career switch.

The assumption of unknown research ability early in the career and stochastic learning adopted in the paper are not unintuitive. Scientific projects are the results of the collaboration of multiple researchers. For example, the average number of authors per article in academic biochemistry is 2.90, and in physics 3.6, [Stephan (1996), Preston (2004)]. When observing the final output of a team, it is not straightforward to infer the contribution of each team-member. Individual productivity, and therefore ability, can only be inferred if the person engages in an individual project or becomes a group leader. Therefore, the probability of learning assumed in this paper can be thought of as a probability of receiving a chance to engage in such a project. The latter can be possible if individuals receive a research grant or

access to research facilities early in their career. Therefore, a higher probability of learning can be related to improvement in the access to research funding either because of an overall increase in available resources or a decrease in grant competition.

The dependence of R&D productivity on past performance has been previously emphasized in the S&E literature. Stephan (1996) suggests that the survival of scientists in research depends on their ability to achieve a “critical size within a limited time frame”, similarly to the survival of new firms. Resources in sciences are highly competitive and conditional on a variety of factors, one of which is the novelty of the research project and the scientific reputation of the investigator. Through acquiring resources scientists can develop more ambitious projects and gain scientific reputation in their field. High reputation, in turn, lets them “keep their place in the funding queue” and grow further. This generates cumulative advantage and a growing gap between researchers with different levels of research ability reflected in scientific productivity parameters such as citations, publications, and earnings, Diamond (1986), Neal and Rosen (2000). In addition, it was found that accumulated knowledge depreciates fast, McDowell (1982). Therefore, research career interruptions are costly to doctorates. The speed of research human capital depreciation is assumed to be slower for more talented researchers, which also contributes to the gap in productivity.

The model is fit to the rarely used Survey of Doctorate Recipients (SDR), a longitudinal data set on employment histories and earnings of doctorates educated in the US. The data have been collected biennially by the National Science Foundation since 1973. The survey is unique because of its longitudinal nature and large sample size. Doctorates constitute only 0.5 percent of the labor force, and therefore would be represented by a handful of observations in national surveys. The SDR follows doctorates regardless of their labor force status and employment sector as long as they reside in the US and are under the age of 76. This makes it possible to study their career choices even if they quit S&E. In addition, it contains rich employment information specific to S&E. The parameters of the model are estimated using the Method of Simulated Moments. Empirical moments on career histories, earnings, and transitions of 23,700 doctorates in Life Sciences are matched with the corresponding simulated moments predicted by the model.

The main finding of the paper is that information about research ability plays an important role in explaining the observed career dynamics. Under the full information scenario, participation in R&D falls for young scientists by 44 percent, and lifetime participation in R&D by 20 percent. At the same time the average quality of the supplied skill and retention in R&D dramatically improve. Full information allows individuals to make optimal decisions from the very beginning of their careers and improves their expected discounted lifetime value. The effect of incomplete information intensifies with the rate of information arrival, which in this paper is related to funding possibilities. Faster learning increases participation in R&D, the quality of the retained skill due to the faster attrition of the lower ability researchers and an increase in the discounted lifetime value of the individuals. Next, the predictions of the model are used to understand how the supply and quality of research labor responds to changes in the economic environment and availability of its outside options. This is important for designing and evaluating policies aimed at the retention of talented scientists and engineers in R&D. Using the estimates of the model parameters, I evaluate the effects of R&D subsidies and changes in the employment options outside S&E. I find that the supply of research skill is sensitive to changes in relative skill prices in- and outside R&D. This finding contradicts previous results that the supply of S&E skill is inelastic.

This paper is related, first of all, to the literature on the economics of S&E such as Stephan (1996), Freeman (1975), Ryoo and Rosen (2004), Majumdar and Shimotsu (2005). In particular, it contributes to the studies of the S&E workforce; for example, Ferrall (1997), Biddle and Roberts (1994), Zucker and Darby (2006), and Fallick et al. (2005). The first novelty of this study is its explicit modelling of selection into S&E tasks previously taken as exogenous. The second is its consideration of employment options outside S&E as part of the choice set. The third contribution is the possibility to use the model to analyze the behavioral responses of doctorates to various policy changes. This paper is also related to the general literature on occupational choices such as the static self-selection model by Roy (1951) and Heckman and Sedlacek (1985), dynamic models by Keane and Wolpin (1997), Miller (1984), and their application to specific professions. Examples of the latter are the study of careers of lawyers by Sauer (1998), politicians by Keane and Merlo (2007) and Diermeier et al. (2004), and secondary school teachers by Stinebrickner (2001). The latter

study and Scafidi et al. (2007) are especially closely related to this work because they address similar issues. In particular, both papers assess the timing and reasons for the attrition from teaching professions, including switching to non-teaching occupations which constitutes career changes. The model developed in this paper is not specific to scientists and can be applied to study career decisions in other professions, where dependence on past performance and high ability are the key features, such as performing arts or sports.

The remainder of the paper consists of four sections. The model and its numerical solution are outlined in Section 2. It is followed by the description of the data, stylized facts on the career dynamics of doctorates, and the estimation algorithm. Section 4 presents estimation results and several counterfactual experiments. The final section concludes and proposes directions for further research on the topic.

2 Model

In this section I describe the theoretical dynamic model of the occupational choices of S&E doctorates. The second part of the section describes the algorithm for the numerical solution of the model.

2.1 Basics

Time is discrete and indexed by t . A scientist begins to make decisions when he becomes a PhD. The decision to become a PhD and the choice of the field are taken as exogenous. Individuals are also assumed to graduate at the same age and not to time their graduation to the labor market conditions. A scientist is active on the labor market for T periods, which is assumed to be deterministic and the same for everyone. At $T + 1$ everybody exits the market with probability 1 and never returns. At the current stage, I capture retirement as random attrition out of the sample with task-specific probability ϕ_i starting at period 10.

Each individual is endowed with three types of abilities: research-related, denoted as R ; applied, A ; and unrelated to S&E tasks, N . “Research” ability can be thought of as a talent to pose novel research questions and generate interesting ideas. “Applied” ability is a

talent to “sell” or apply existing knowledge to given problems and tasks rather than posing *new* ones. Finally, non-S&E ability can be thought of as people skills or general analytical abilities. These abilities are assumed to be constant over the individual’s career. Denote the vector of person i ’s ability endowment as $\mathbf{x} = (x_R, x_A, x_N)$, with $(\mathbf{x} \sim N(\mu, \Sigma_x))$, where $\mu = (\mu_R, \mu_A, \mu_N)$, and

$$\Sigma_x = \begin{pmatrix} \sigma_R^2 & \sigma_{RA} & \sigma_{RN} \\ & \sigma_A^2 & \sigma_{AN} \\ & & \sigma_N^2 \end{pmatrix}$$

I allow for abilities to be correlated but do not explicitly impose the sign of the correlation coefficient. The parameters of the ability distribution (μ and Σ_x) are model parameters to be estimated. These intrinsic abilities are the only source of *ex ante* heterogeneity in the model³.

Suppose also that there exist three tasks: research (r), application of existing scientific knowledge (a), and some combination of non-S&E tasks (n). Throughout the paper, capital letters are used to indicate the type of ability, and small letters to indicate the corresponding task. Hereafter I assume that the three types of abilities are supplied inelastically, that is, an individual supplies his entire endowment x_m , if he chooses occupation m .

2.2 Learning and the evolution of human capital

It is assumed that x_A and x_N are common knowledge, while x_R is *ex ante* unknown to anyone. It can be symmetrically learned with probability δ if the scientist engages in R&D. For individuals who never worked in R&D x_R can be inferred from the observed realizations of x_A and x_N and the parameters of the skill distribution, which are public knowledge. Denote the perceived level of research ability x_R as \tilde{x}_R , and let τ be an indicator that x_R has been revealed. Then:

$$\tilde{x}_R(\tau) = \tau(\delta)x_R + (1 - \tau(\delta))E[x_R|x_A, x_N] \quad (1)$$

³Due to the limited data access described later, incorporating the characteristics of observed heterogeneity at this stage is impossible.

It is assumed that learning about research ability evolves according to:

$$\Pr(\tau' = 1 | \tau = 0) = \delta I(d = r) \quad (2)$$

$$\Pr(\tau' = 1 | \tau = 1) = 1 \quad (3)$$

That is, conditional on employment in research, individual's research ability is revealed with probability δ . Once research ability has been learned it remains public knowledge forever. It is assumed that research ability gets revealed through an individual research project or the possibility of leading one. It is assumed that this chance is higher for scientists who are believed to be more productive (e.g. research grants that are crucial for individual projects are easier to obtain for scientists with a better research record or higher expected quality of the project and the probability of that project to become a breakthrough *ceteris paribus*). Therefore, the probability of discovering one's true research ability is increasing in one's perceived ability:

$$\delta(\tilde{x}_R) = \frac{\exp(\delta_0 + \delta_1 \tilde{x}_R)}{1 + \exp(\delta_0 + \delta_1 \tilde{x}_R)}$$

Productivity in each task m , is a function of the task-specific ability, x_m , and accumulated human capital, h_m . The former is constant throughout career, while the latter evolves according to the following technology:

$$\Pr(h'_m = h_m + 1) = f^m(\tilde{x}_m) I(d = m), \forall m \quad (4)$$

$$\Pr(h'_m = h_m - 1) = g^m(\tilde{x}_m) I(d \neq m), \forall m \quad (5)$$

In applied and non-S&E tasks the evolution of human capital is standard as in Keane and Wolpin (1997), i.e an additional period of employment these tasks increases the corresponding task-specific human capital by one unit and there is no skill depreciation (i.e. $f^m(\tilde{x}_m) = 1$, $m = a, n$ and $g^m(\tilde{x}_m) = 0$, $m = a, n$).

This process is different for research-specific human capital: It is assumed to change at various rates for individuals with different levels of research ability. Modelling human capital growth conditional on the past performance explicitly would require an introduction of a continuous state variable, which would complicate the solution of the model. At this stage I model the technology for h_r to change stochastically by one period, which makes h_r

a discrete variable as are h_a and h_n .

$$f^r(\tilde{x}_R) = \frac{\exp(\gamma_{f0} + \gamma_{f1}\tilde{x}_R)}{1 + \exp(\gamma_{f0} + \gamma_{f1}\tilde{x}_R)}$$

$$g^r(\tilde{x}_R) = \frac{\exp(\gamma_{g0} + \gamma_{g1}\tilde{x}_R)}{1 + \exp(\gamma_{g0} + \gamma_{g1}\tilde{x}_R)}$$

Thus, human capital appreciates slower for lower ability researchers. Since human capital is assumed to change stochastically only in research, the task superscript is dropped for the remainder of the paper.

2.3 Earnings

Earnings in task m , w_m , is the product of an individual's total skill of type m , s_m , and the rental price of a unit of skills, r_m :

$$w_m = r_m s_m \tag{6}$$

$$= r_m \exp(\tilde{x}_m + \sum_{i=r,a,n} \alpha_{1m} h_i - \sum_{i=r,a,n} \alpha_{2mi} h_i^2 + \varepsilon_m), \tag{7}$$

where $\tilde{x} = (\tilde{x}_R, x_A, x_N)$, and $\varepsilon = (\varepsilon_r, \varepsilon_a, \varepsilon_n)$ is the vector of idiosyncratic serially-uncorrelated productivity shocks. The shocks are assumed to be normally distributed $\varepsilon \sim N(0, \Sigma_\varepsilon)$, where Σ_ε is assumed to be diagonal, that is, shocks are uncorrelated across tasks. Note that the pre-PhD experience is not included in the wage equation. The first reason for that as reported by Thurgood et al. (2006), pre-doctoral experience is very infrequent. The data set used in this study does not provide information on the pre-doctoral experience. The second reason for that is given by several studies who found no significant returns to pre-doctoral experience in the earnings of PhDs.

Individual's state in period t is described by the vector S :

$$S = \{\tau, \tilde{x}, h, pd, t\} \tag{8}$$

The state vector consists of the following components: an indicator that the research ability has been revealed τ , the vector of perceived ability \tilde{x} , the vector of accumulated task-specific human capital $h = (h_r, h_a, h_n)$, last period employment choice pd , and an individual's age t . The next period's state space, S' is given by:

$$S' = \{\tau', \tilde{x}', h', d, t + 1\} \tag{9}$$

where τ' and $h' = \{h'_R, h'_A, h'_N\}$ evolve according to the laws of motion described by (2) and (4) respectively.

Each individual's preferences over task choices d given their state S are represented by a utility function $u(S, d)$:

$$u(S, d) = w(S, d) + b_d + c_{pd,d}I(pd \neq d) \quad (10)$$

where $w(S, d)$ is a wage offer as described above, b_d is a choice-specific non-pecuniary benefit, and $c_{pd,d}$ is a transition-specific switching cost from task pd to task d if $pd \neq d$.

Non-pecuniary benefits b_d are expected to capture non-wage and other benefits as well as personal preferences to perform different tasks. For example, many studies of the organization of science describe research as a puzzle-solving activity where the very process is in itself a reward to its participant, Stephan (1996). Several studies found the evidence of this reward in the life-time earnings of scientists, Stern (1999), Levin and Stephan (1991). The latter study includes the “taste” for puzzle-solving directly into the scientist's utility function in addition to accounting for the monetary benefits research activity brings through the enhancement of the total human capital. The switching costs are expected to account for the adjustments caused by changing tasks. For example, switching from primarily teaching to research would require “catching up” on the recent findings in the area. Alternatively, switching from non-teaching to teaching would require preparing a course outline, assignments and other materials. The non-pecuniary benefits and switching costs are assumed to be time-invariant and the same for all individuals. They are also assumed to be known to all but unobserved by the econometrician.

Every period t individual i observes the vector of state variables S and chooses task $d_i \in D$ to maximize his discounted expected utility. Define a career path as the sequence of task choices each period from some age t until the last period T . Denote it as $\tilde{d}_t = \{d(i)\}_{i=t}^T$, where $d(i) \in D, \forall i$. Denote also the family of all possible career paths as $\mathcal{D} = \{\tilde{d}_t, \forall t\}$. The individual's problem is to choose a career path $\tilde{d} \in \mathcal{D}$ to maximize his discounted expected lifetime utility:

$$\max_{\tilde{d}_t \in \mathcal{D}} E_\epsilon \sum_{i=t}^T \beta^{i-t} u(S, d), \quad (11)$$

where $u(S, d)$ is as described above and $\beta \in (0, 1)$ is a discount factor.

Each period the scientist chooses a task $d \in D$ with value:

$$v(d, S) = u(d, S) + \beta E_\varepsilon V(S') \quad (12)$$

where

$$V(S') = \max\{v(r, S'), v(a, S'), v(n, S')\} \quad (13)$$

The value in R&D depends on whether the information about the research ability has been revealed ($\tau = 1$) or not ($\tau = 0$). If the ability has not been revealed, the value function in R&D is defined by:

$$\begin{aligned} v(r, 0, \tilde{x}, h, pd, t) &= u(r, 0, \tilde{x}, h, pd, t) \\ &+ \beta E_\varepsilon \{ \delta E_{x_R|x_A, x_N} V(1, \tilde{x}', h', r, t+1) \\ &+ (1 - \delta) V(0, \tilde{x}, h', r, t+1) \} \end{aligned}$$

where the right-hand side reflects that with probability δ it will be revealed after one period and $1 - \delta$ it will not. To evaluate $V(1, \tilde{x}', h', r, t+1)$, possible realizations of x_R are drawn from the conditional distribution $f(x_R|x_a, x_n)$ and the expectation of the future values over all x_R is calculated. If the ability is not revealed after one period, the perceived ability \tilde{x}_R remains unknown and the value function is calculated based on its expected value conditional on the known endowments of the other two abilities. In this case the value function in R&D is given by:

$$v(r, 1, \tilde{x}, h, pd, t) = u(r, 1, \tilde{x}, h, pd, t) + \beta E_\varepsilon V(1, \tilde{x}, h', r, t+1)$$

In either case, the value in R&D also reflects whether the research human capital will be augmented or not:

$$\begin{aligned} V(\tau', \tilde{x}', h', r, t+1) &= \pi_{+1} V(\tau', \tilde{x}', h_r + 1, h_a, h_n, r, t+1) \\ &+ (1 - \pi_{+1}) V(\tau', \tilde{x}', h, r, t+1), \quad \tau' = 0, 1 \end{aligned}$$

where $\tilde{h}_r = \{h_r + 1, h_a, h_n\}$.

The values of the two other tasks depend only on whether the R&D capital will depreciate or not. Thus, the value in applied task is given by:

$$\begin{aligned} v(a, \tau, \tilde{x}, h, pd, t) &= u(a, \tau, \tilde{x}, h, pd, t) + \\ &\quad \beta E_{\varepsilon} \{ \pi_{-1} V(\tau, \tilde{x}, h_r - 1, h_a + 1, h_n, a, t + 1) \\ &\quad + (1 - \pi_{-1}) V(\tau, \tilde{x}, h_r, h_a + 1, h_n, a, t + 1) \}, \quad \tau = 0, 1 \end{aligned}$$

and in non-S&E by:

$$\begin{aligned} v(n, \tau, \tilde{x}, h, pd, t) &= u(n, \tau, \tilde{x}, h, pd, t) + \\ &\quad \beta E_{\varepsilon} \{ \pi_{-1} V(\tau, \tilde{x}, h_r - 1, h_a, h_n + 1, n, t + 1) \\ &\quad + (1 - \pi_{-1}) V(\tau, \tilde{x}, h_r, h_a, h_n + 1, n, t + 1) \}, \quad \tau = 0, 1 \end{aligned}$$

2.4 Numerical solution

Since the individual optimization problem is finite it is solved recursively starting at T . The terminal value, v^T , is assumed to be equal to zero for all individuals regardless of their state: $v(d, S) = 0, \forall S : t = T + 1$.

Define the expected term in the value function 12 as Emax:

$$E_{\max}(S) = \int_{\varepsilon} V(S') f(\varepsilon) d\varepsilon \quad (14)$$

Due to the assumption of normality of the productivity shocks, the integral in equation 14 does not admit an analytical solution but can be evaluated numerically. The expected values of the value function is approximated through simulation using a simple frequency simulation method. For computational reasons the distribution of the shocks is discretized with R_{ε} points of support with R_{ε} being fairly large⁴. Let $\varepsilon^l = (\varepsilon_r^l, \varepsilon_a^l, \varepsilon_n^l)$ be a l 'th draw from the trivariate normal distribution with mean zero and a given variance-covariance matrix, where $\varepsilon^l = (\varepsilon_r^l, \varepsilon_a^l, \varepsilon_n^l)$, and let $\varepsilon^l : l = 1, 2, \dots, R_{\varepsilon}$ be the matrix of these draws. For each of these draws I construct simulated approximations for the Emax function. Define $\tilde{V}^l(S) = \max\{v^l(r, S), v^l(a, S), v^l(n, S)\}$, where $v^l(d, S)$ is a choice-specific value function

⁴In this paper I chose 500 points

evaluated at ε^l . Then the $\text{Emax}(S)$ is calculated as:

$$\text{Emax}(S) = \frac{1}{R_\varepsilon} \sum_{l=1}^{R_\varepsilon} \tilde{V}^l(d, S) \quad (15)$$

Since the terminal value is assumed to be equal to zero, in the final period T , the approximation is given by:

$$\text{Emax}(S) = \frac{1}{R_\varepsilon} \sum_{l=1}^{R_\varepsilon} u^l(S, d) \quad (16)$$

In each period prior to T the $\text{Emax}(S')$ function in the next period is already known and can be used in the evaluation of the $\text{Emax}(S)$ in the current period:

$$\text{Emax}(S) = \frac{1}{R_\varepsilon} \sum_{l=1}^{R_\varepsilon} \{u^l(S, d) + \beta \text{Emax}(S')\} \quad (17)$$

The distribution of types was also discretized to have 700 points (types). Each type is represented by a triple describing the individual's endowment of the abilities of three types, x . Each type is therefore a draw from the trivariate normal distribution with the parameters defined by the values of μ and Σ_x . Types were re-drawn every time the distribution parameters changed. For the matrix of the values of abilities x the matrix of Emax was calculated by recursively solving the problem for each feasible state for each type. By discretizing the continuous variables ε and x and assuming stochastic evolution of the human capital in research, the state space becomes a grid of 20,250 points for each type for 15 periods. The grid includes some states that will never be visited by an individual and can therefore be excluded to reduce the dimensionality of the problem. The solution of the model involved searching over the feasible points and did not require polynomial approximation as compared to Keane and Wolpin (1994).

Because the state space is now discrete, the choice probabilities have to be smoothed. The smoothing is done by weighting the value of the actions compared to the value of the state by some parameter ρ after solving the optimization problem:

$$P(d = m) = \frac{\exp(\rho(V(m, S) - \text{Emax}(d, S)))}{\sum_i \exp(\rho(V(i, S) - \text{Emax}(d, S)))} \quad (18)$$

3 Estimation

This section outlines the estimation procedure and the data set used in this paper. It also provides the stylized facts on the career choices of scientists and engineers, occupational transitions and earnings.

3.1 Data

The model is estimated using the Survey of Doctorate Recipients (SDR), collected biennially by the National Science Foundation since 1973. This longitudinal survey includes individuals who graduated from the US doctoral programs in S&E and resided in the US at the time of the survey. They were followed from the moment of their graduation until the age of 76. In the original 1973 survey, the target population included graduates between 1930 and 1973. With every survey, a fraction of new graduates are added, and some of the previous respondents are removed from the survey because they reach the age of 76 or for sampling reasons. Information on these individuals is obtained from the Doctorate Records File (DRF) maintained by the NSF. The primary information for the DRF is the Survey of Earned Doctorates, a one-time census type survey on all doctorates from the US institutions in S&E or health fields collected at the time of their graduation. Since the latter started in 1957, the information source on the graduates prior to 1957 are taken from a registry on highly qualified scientists and engineers. The latter is assembled by the National Academy of Sciences from university catalogues, federal laboratories, and other sources. This way, in 2001 the survey had a sample of 40,000 individuals representing 650,000 doctorates under the age of 76 residing in the US. Due to the information collection techniques (multiple follow-ups and phone interviews), the survey has very high response rates (80 percent). This potential source of error is adjusted in the survey using the weights. This paper bases its analysis on the weighted data. Non-response rates for the employment and personal information used in this paper are low (0-1.2 percent and 0-2.7 percent respectively).

The survey is unique, first of all, because it provides information on the group of professionals that is small relative to the population. Due to this problem, other data sets (e.g. CPS) would have a very small number of individuals with PhD degrees. Secondly, the

data in the SDR is longitudinal, and allows following the individuals as their careers unfold. Finally, the survey asks about academic achievements and other profession-specific information usually unavailable in the general type surveys. There exist several studies that used other data sets, primarily collected by the authors, for example, Mangematin (2000), Robin and Cahuzac (2003), Gaughan and Robin (2004), Oyer (2005), Diamond (2001), Grimes and Register (1997). Those data sets were constructed by the authors using individuals' CVs posted on academic websites. The advantage of this approach is that it allows to construct complete employment histories and to have full productivity parameters not available in the SDR. Unfortunately, I could not follow their example or use one of their data sets because of the small number of observations (at best a few hundred individuals). Another drawback of these data sets is their focus on the academic careers in contrast to the SDR that includes doctorates regardless of the relevance of their employment to the S&E, which provides crucial information for the purposes of this study.

The major interest to this paper were employment records. They contain information on the employer (e.g. geographic region, detailed sector, size, type of industry), occupation and its relevance to the PhD major, primary and secondary activities (i.e. "activities that occupy most time on a typical work week"), information on professional activities (membership in scientific societies, participation in conferences), as well as scientific productivity (publications and patents). Unfortunately, the questions on publishing and patenting activities were asked only in selected years (1983, 1993-2001), and referred only to two years preceding the survey. Due to very low response rates to these questions it was impossible to infer the total number of publications and patents even for a small subsample of individuals. For these reasons, direct scientific productivity variables were not used in this study. The survey was designed to collect information only about employment at the time of the survey.

Because the retrospective information on employment was not available, career histories were constructed using available data for the individuals who responded to more than two consecutive surveys and provided enough information to assign them to one of the three tasks⁵. The resulting subsample consists of 23,700 individuals whose personal and job characteristics are described in the last column of Table 6. Individuals in the subsample are more

⁵The assignment principles are described in more detail in Appendix 1

likely to be white US citizens. They are more likely to be men although the participation of women in Life sciences is the highest compared to other disciplines. Academia employs roughly 60 percent of all individuals, and private sector accounts for another 30 percent. Roughly 45 percent of the academics were tenured and another 17 percent held tenure-track appointments. Finally, 10 percent of all doctorates were employed as postdoctorates and another 23 percent held temporary academic appointments.

Individuals in the sample were assigned to one of three tasks: two tasks in the S&E sector, and the third type includes tasks unrelated to S&E. I distinguish between S&E and non-S&E tasks to pin down the “career changes”. Within S&E all occupations are divided by the primary task: creation of new knowledge, further referred to as “R&D” and the application of accumulated knowledge or “application”. The non-S&E types includes tasks that are not directly related to sciences and engineering, for example, financial- or other non-technical business services. To assign individuals to one of three tasks I used information about their primary activities⁶, and the industry of the employer. The principles of assigning individuals to different tasks are outlined in more details in Appendix 1. The detailed empirical analysis of career choices, their relevance to S&E, and career transitions are presented in the companion paper, Mishagina (2007).

3.2 Stylized facts on career choices, transitions, and earnings

This section provides the stylized facts on career dynamics and the earnings of Life Sciences doctorates. The latter group includes agricultural and food scientists, biologists and biochemists, environmental and health scientists. Table 6 compares individuals in three tasks by observed personal and job characteristics⁷. As can be noted, the main difference comes mostly from average age. Scientists employed in non-S&E tasks are on average older (48.50 versus 45). They are also more likely to be married (0.67 versus 0.47) which is probably related to their being relatively older. Scientists in R&D are slightly more likely to come

⁶Primary activity in the questionnaire is defined as “activity that occupies most time in the typical work week”. Individuals were offered to choose from a list of 14 activities. For more detailed information on this and other variables see Mishagina (2007).

⁷Descriptive statistics are shown to compare individuals in different tasks. However, these characteristics are not used in the estimation due to the data access limitations.

from private Carnegie Research I and II universities. The fraction of graduates from the top schools⁸ is also higher in R&D than in applied jobs but similar to that in non-S&E. Next, there is a difference by citizenship status. The fraction of temporary residents in non-S&E is smaller compared to R&D. This difference may be attributed to age differences as temporary residents receive green cards if they stay in the US. Alternatively, immigration policy may have different policies about hiring foreigners or assisting them with green cards in different tasks. Similarly to the first argument, there are fewer permanent residents in non-S&E and more naturalized citizens. Finally, doctorates in non-S&E tasks are more likely to have pre-PhD degrees in non-S&E fields, which may suggest their preference towards non-S&E tasks or task-specific skills.

R&D tasks are distributed mostly between industry (0.48) and academia (0.40), while applied jobs are more likely to be academic (0.68), which captures teaching activities. All postdoctorates in the sample are concentrated primarily in the R&D tasks. Academic researchers are more likely to be employed in Carnegie Research I/II universities compared to those in applied tasks (0.601 versus 0.337 respectively). In non-S&E, the fraction of academics in Carnegie Research I/II is also high, but they are employed in non-S&E departments (see assignment principles in Appendix 1). Finally, more academics in R&D hold temporary positions other than postdoctoral appointments when compared to other tasks (0.484 versus 0.277 versus 0.36).

Figure 1 show the participation rates of the Life sciences doctorates in each of the three tasks by time since graduation. The first observation is that career choices vary throughout career. Doctorates tend to start their careers in R&D (72 percent) but over time they leave R&D for applied and non-S&E tasks so that after 30 years since graduation only 45 percent of the doctorates are still employed in R&D. Non-S&E tasks account only for 3 percent of the newly minted doctorates, but after roughly 15 years since graduation employment in non-S&E grows to 10 percent. This fraction remains constant until the rest of the career, while the participation in applied tasks continues to grow. In addition to the individuals who start their careers in R&D, another 6 percent join the task later in the career. The difference between the two numbers suggests that late entries into R&D are costly. Finally,

⁸CalTech, UC Berkeley, Stanford, MIT, Harvard, Princeton and Yale.

the remaining 22 percent of scientists never work in R&D in their entire career.

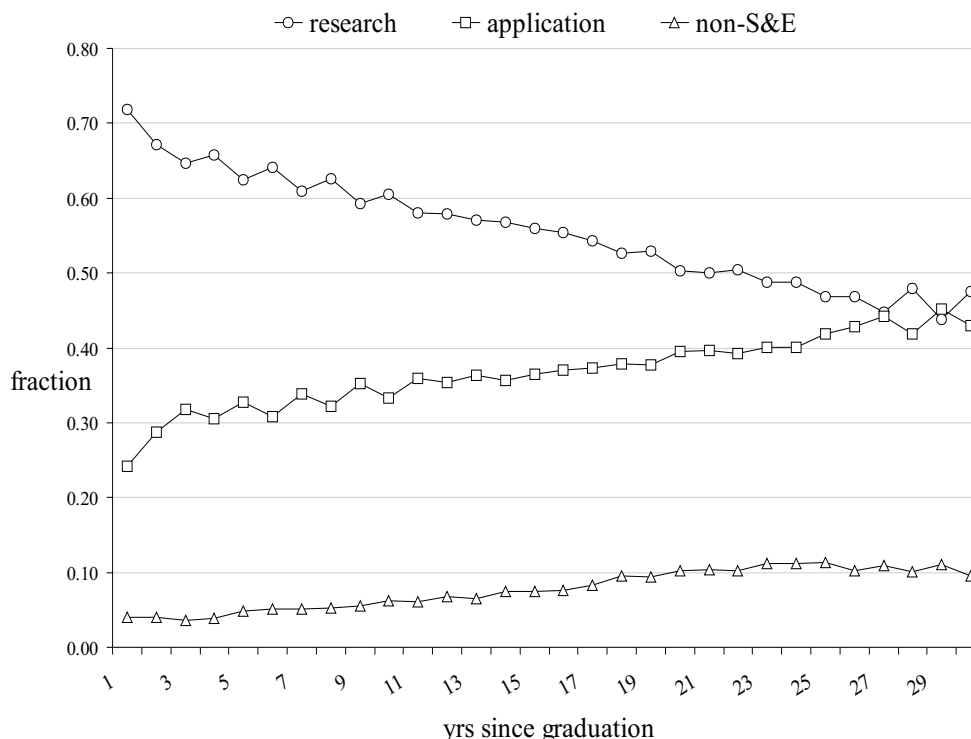


Figure 1: Employment rates of life science doctorates, by task and years after graduation.

Figure 2 presents lifetime earnings profiles by the choice of employment in constant 2001 dollars. It can be seen that individuals employed in R&D early in the career have the lowest starting salaries relative to the other tasks. On average, researchers start with \$35,500 a year which is consistent with the finding that many life scientists start their careers in postdoctoral and other non-faculty appointments with relatively low salaries, Stephan and Ma (2004). These earnings are on average 26 percent and 35 percent lower than those in applied tasks and non-S&E (\$45,000 and \$48,000 respectively). However, R&D salaries grow and catch up with those in applied tasks by the fifth year after graduation, and with those in non-S&E by the thirteenth year. At the end of career, the R&D earnings dominate those in application and non-S&E by 21 percent and 14 percent respectively. Overall, “stayers” in R&D experience a 180 percent growth of earnings by thirty years after graduation compared to 80 percent growth in both applied tasks and non-S&E.

Some studies explain low starting earnings in R&D relative to those later in the career

to be a payment for the on-job-training, Moen (2001). This however does not explain why scientists leave R&D if they expect higher returns on the investment. It seems to be more likely that the high growth of the salaries is caused by the outward selection of those with lower ability and thus lower earnings. In order to better understand the nature of the mobility, the transition rates between tasks are next analyzed.

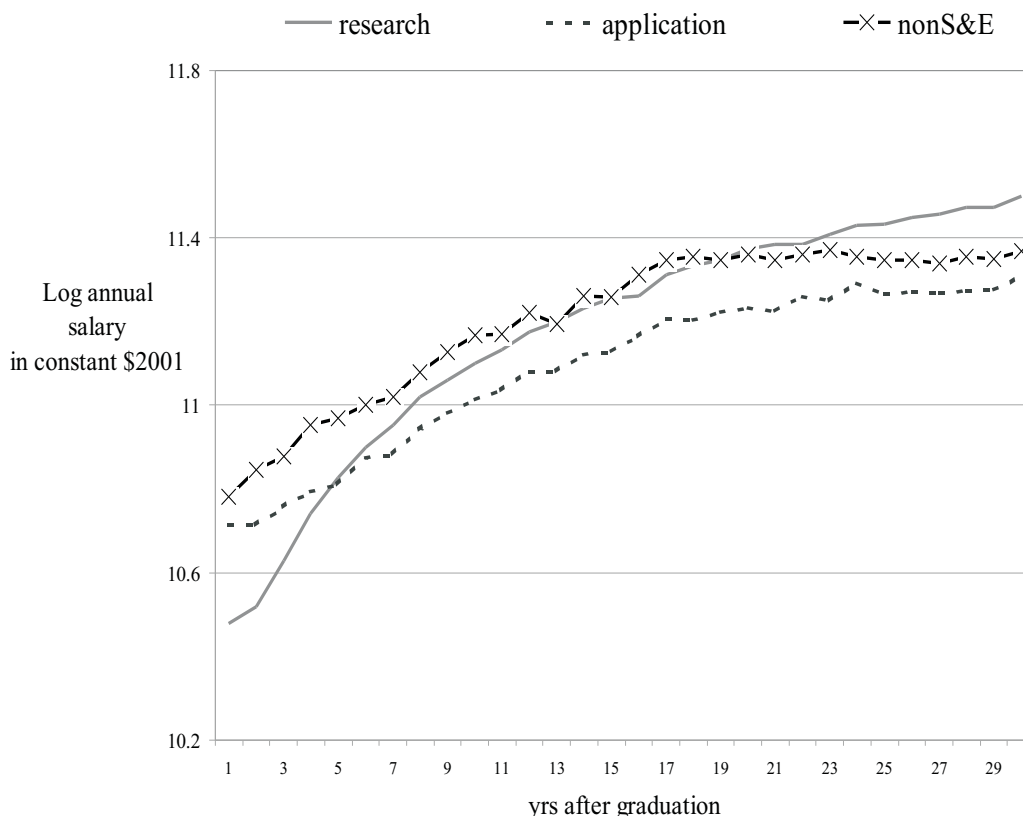


Figure 2: Lifetime earnings profiles (log) in constant 2001 dollars, by task

The empirical transition rates shown on Figure 3. The transition rates are calculated as the number of individuals who switched tasks after t years of task-specific experience over the total number of the individuals still employed in the task of origin by the year t . The first observation is that transition rates vary in their shape and timing by origin and destination. Mobility between S&E-tasks is higher than mobility out of S&E and differs in pattern. The transitions from R&D to application have several pronounced peaks: at 4, 10, and 18 years on the task. Mobility from applied tasks to R&D is monotonically decreasing, which suggests that returns or late entries to research are rare. Transitions into non-S&E

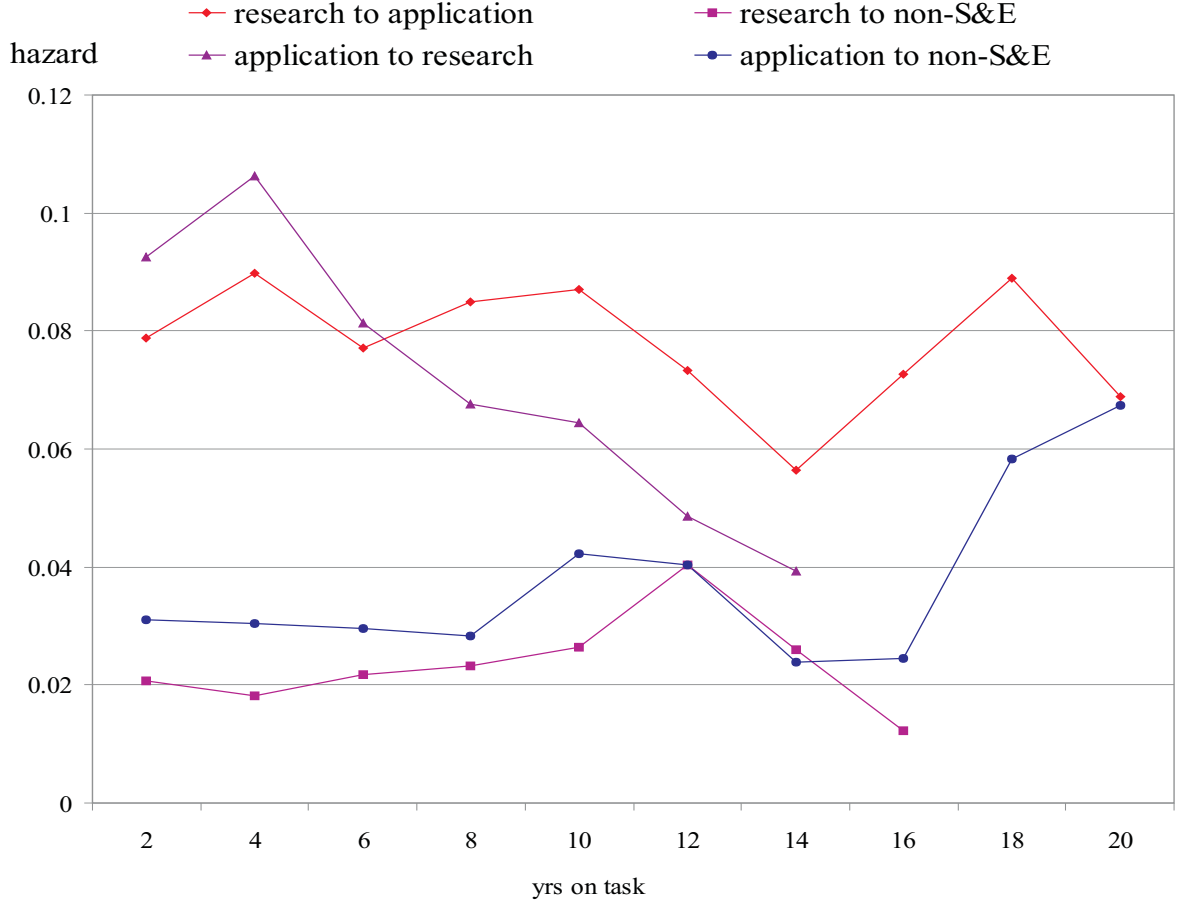


Figure 3: Empirical transition rates within- and out of S&E.

are single-peaked at about 10-12 years on the task and do not substantially vary by origin until later in the career. This observation proposes that non-S&E employers value research- and applied skills and experience in a similar fashion.

The timing of the transition rates out of research early in the career suggests that mobility may be caused by the expiration of tenure probation periods in academia. Figure 4 presents transition rates by sector, which support this hypothesis. However, main transitions to non-S&E happen from the non-academic sectors. These transitions can be correlated with postdoctoral appointments and inability to secure a permanent position. Mishagina (2007) finds that individuals with higher number of postdoctoral appointments have a higher probability of leaving R&D for non-S&E tasks, which may suggest a “discouraged worker” effect.

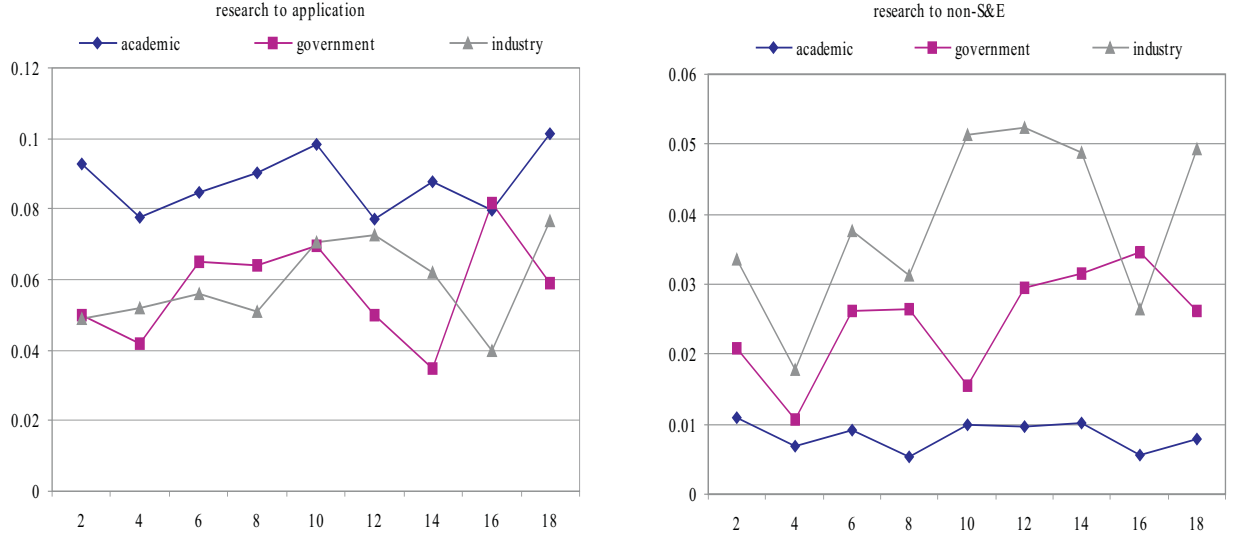


Figure 4: Transitions out of research by sector

Non-monotonic transition rates between tasks were found in the career dynamics of other professionals, for example, junior medical specialists, Van den Berg et al. (2002), or young lawyers, Sauer (1998). The shape of the transitions is traditionally associated in the literature with the presence of incomplete information, Sauer (1998), Nagypál (2007). Notably, the timing of the transitions coincides with the period believed to be crucial for young scientists to establish a foundation to build reputation in R&D. Therefore incomplete information about research ability seems to be a plausible explanation for the observed career dynamics.

The comparison of life sciences with other disciplines shows that these career patterns are not specific to life sciences. Figures 5, 6 show similar patterns for physical sciences and engineering. It is clear that although the magnitudes of retention in R&D slightly differs by discipline, the overall pattern remains the same: Scientists start in R&D but later leave for other tasks. The main transition happens to applied occupations which on average have the lowest earnings compared to those in R&D and non-S&E.

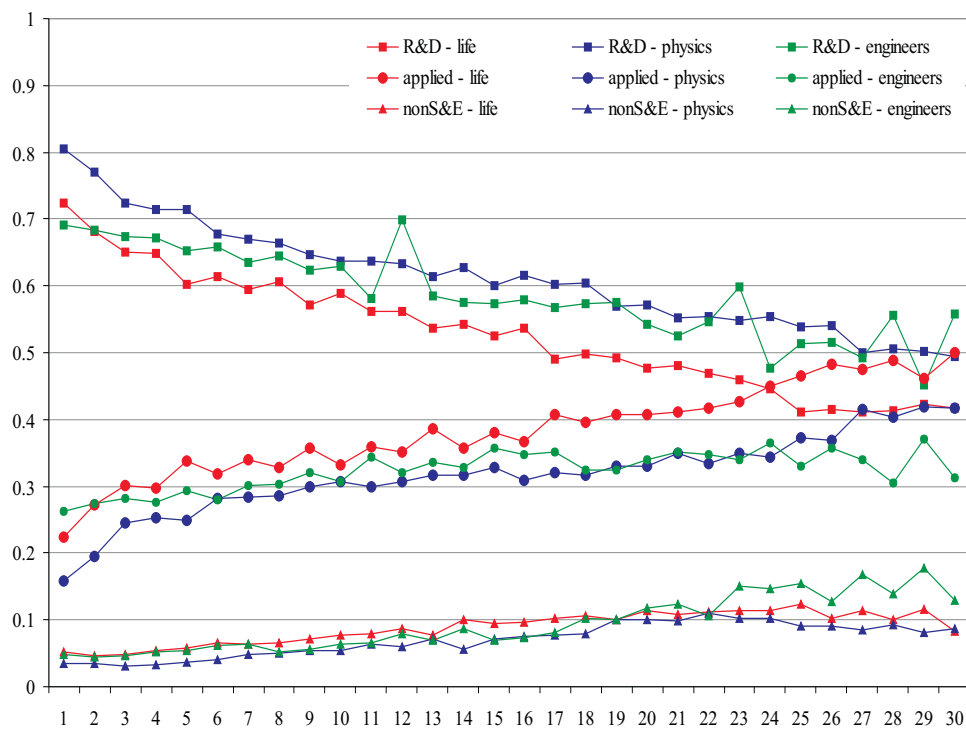


Figure 5: Comparison by discipline: participation rates

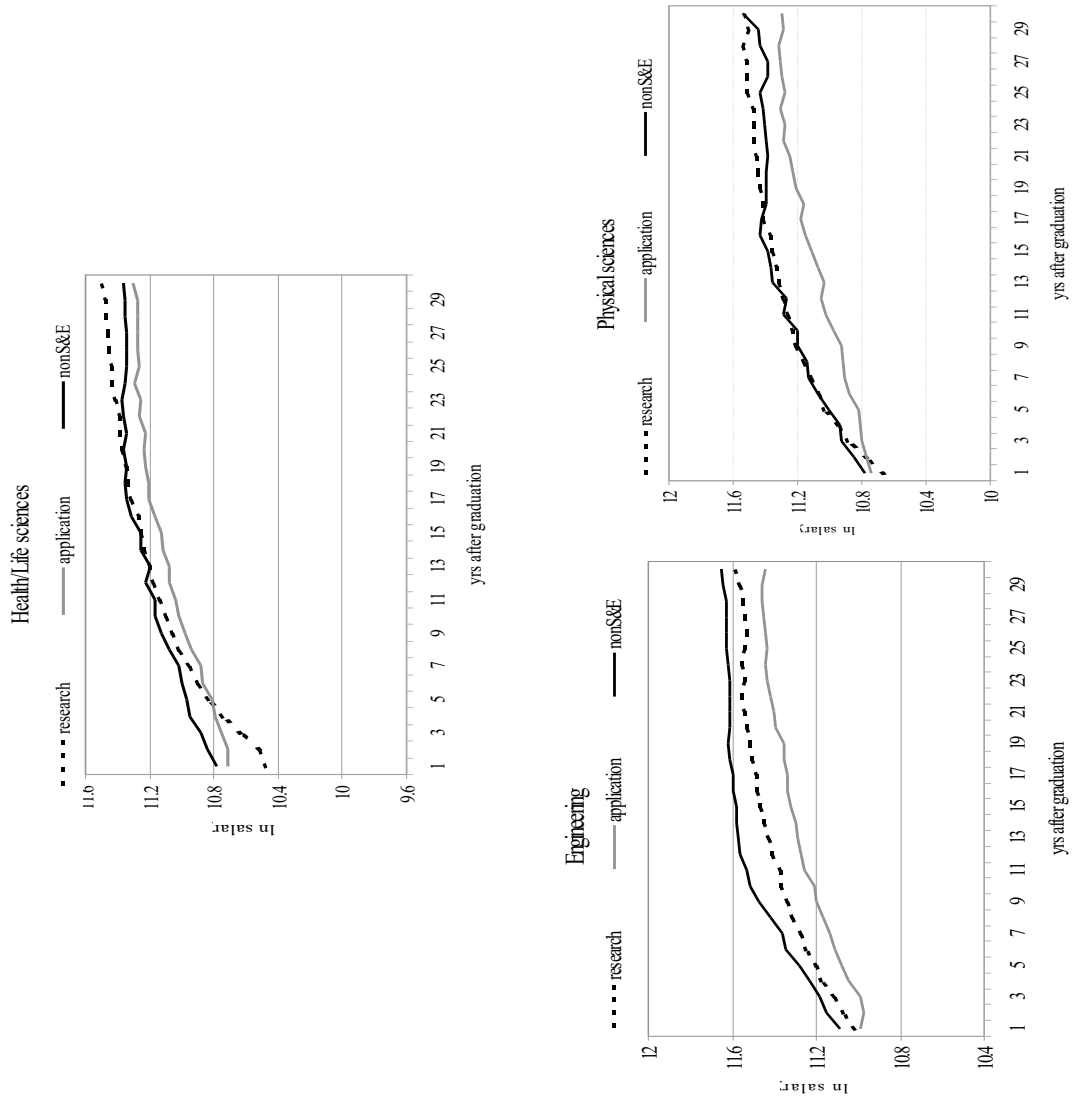


Figure 6: Comparison by discipline: earnings

3.3 Estimation procedure

Denote the vector of parameters to be estimated:

$$\theta = (\alpha^1, \alpha^2, b, c, \mu, \Sigma_x, \Sigma_\varepsilon, \delta_0, \delta_1, \gamma, \phi),$$

where α^m is a vector of the wage equation coefficients with a typical element $\{\alpha_{ij}^m\}$ for $m = 1, 2$ and $i, j = \{r, a, n\}$; $b = (b_r, b_a, b_n)$ is a vector of task-specific non-pecuniary benefits; c is a vector of task-switching costs, $\{c_{ij}\}_{i \neq j}$, $i, j = \{r, a, n\}$; μ and Σ_x are parameters of the ability distribution; Σ_ε is a correlation matrix of the distribution of shocks; $\delta_{0,1}$ are parameters of the learning technology; $\gamma = (\gamma_{f0}, \gamma_{f1}, \gamma_{g0}, \gamma_{g1})$ is a vector of the parameters of the human capital evolution technology in research; ϕ 's are task-specific sample attrition probabilities⁹.

Due to the limited access to the SDR data files, it is impossible to estimate the model by any method that requires direct utilization of the individual observations. This problem is overcome by using the Method of Simulated Moments. For the given parameter values θ the individual optimization problem is solved and choice probabilities are computed. Next, I simulate individual 30,000 career paths. For each simulation, a type is randomly chosen out of the matrix of pre-drawn ability vectors and an initial state vector is formed. A career path is generated starting from the first period and moving forward to period T : a random value is drawn and using the computed probability vector the optimal choice and wage are recorded. The vector of state variables is updated according to the laws of motion of the state variables described in (9). The simulation is repeated for 15 periods, and an individual career history is recorded. I use these career histories to compute a set of 400 moments to be matched with the corresponding moments in the data. The average difference between the two sets of moments is weighted by the inverse of Λ , a diagonal matrix with a typical element $\{\lambda_{ii}\} = \sigma_i$, where σ_i is the standard deviation of the i^{th} moment. This weighted distance between the simulated and actual moments gives the criterion function, which is minimized by the choice of the model parameters θ :

$$\hat{\theta} = \arg \min_{\theta} (\bar{m}(\theta) - \bar{m})' \Lambda^{-1} (\bar{m}(\theta) - \bar{m}), \quad (19)$$

⁹The discount factor β is not estimated and set to 0.95 a year.

where \bar{m} is a vector of empirical moments obtained from N observations, with a typical j^{th} element $\bar{m}_j = \frac{1}{N} \sum_{i=1}^N m_{ij}$, and $\bar{m}(\theta)$ is a vector of the moments obtained from \mathcal{S} simulated careers at parameter values θ with a typical j^{th} element $\bar{m}(\theta)_j = \frac{1}{\mathcal{S}} \sum_{i=1}^{\mathcal{S}} m_{ij}(\theta)$. The criterion function is minimized using the Nelder-Mead optimization method.

Both the model solution and moment simulation were executed in parallel using 10 processors. Since individuals and types are uncorrelated, both procedures were evaluated separately for 10 subgroups of 70 types and 1,000 simulations each. Simulated moments obtained for each subgroup were then collected and merged by the main server which used them to calculate the value for the criterion function and update the vector of parameters. This way the actual computing time was reduced by 10 and could be even further reduced by shrinking the size of the subgroup each processor works with and increasing the number of processors utilized.

3.4 Sources of identification

It is assumed that a scientist works for 30 years, which would correspond to retirement at 62 given the median age at graduation in S&E doctoral programs of 32, Thurgood et al. (2006). The available data on the labor force status of doctorates suggest that participation drops from 92 percent in 55-59 age group to 80 percent in 60-64 age group and further to 40 percent in the 65-75 age group, Kang (2003). More detailed information for the latter age group is not available. The life span is divided into 15 periods of 2 years to reflect the biennial nature of the data.

The set of moments is described in Table 7. The moments can be divided into three groups: a) moments pertaining to career choices and earnings, b) moments related to lifetime participation in R&D, c) moments describing transitions between tasks. Each set of moments contains data on choice or change in choices, corresponding mean salaries, and salary standard deviations. Moments on the lifetime participation in research include individuals who ever tried R&D versus individuals who never tried R&D. Moments on earnings and their standard deviations help identifying the parameters of the ability distribution and the parameters of the wage equation. The parameters of the human capital evolution tech-

nology are identified jointly by the participation rates and salaries. The former also help pinning down non-wage benefits. Moments on the lifetime participation in R&D and relevant earnings and moments related to the transitions identify learning rates and transition costs. The costs of leaving non-S&E are especially difficult to identify since the corresponding transition rates are not available. This is why it is important to have information on lifetime participation in research and transition-specific salaries.

4 Results

Estimation results summarized in Table 1. For identification, non-salary benefits in non-S&E are normalized to zero, therefore the estimated values for benefits in other two tasks reflect the benefits relative to the non-S&E. The estimated coefficients b_a is essentially null while b_r is approximately \$4,770. Recall that the standard occupational choice model with non-pecuniary benefits described earlier would suggest high benefits in R&D compared to other tasks (\$9,500 and \$13,000 respectively).

The model predicts high switching costs from non-S&E sector to both R&D and application: \$17,260 and \$17,002 respectively. These results support the common perception that late entries or returns to S&E are difficult. The transitions between applied and R&D are virtually free. Returns to R&D from the applied tasks are more expensive (\$1,380) but not as expensive as returns from non-S&E tasks (\$1,940). These low estimates of the transition costs within S&E produce higher transition rates within the sector than predicted by the data. Estimated values for the technology of human capital evolution correspond to the average probabilities of the human capital growth of 0.80 and depreciation of 0.43.

Parameter	research	application	non-S&E
<i>Wage equation</i>			
α_{ri}^1	0.2171	0.0010	0.0062
α_{ai}^2	0.0978	0.1232	0.0068
α_{ni}^1	0.0312	0.1149	0.1010
α_{ri}^2	-0.0147	-0.0001	-0.0001
α_{ai}^1	-0.0009	-0.0036	-0.0001
α_{ni}^2	-0.0001	-0.0005	-0.0100
<i>Human capital evolution</i>			
γ_{f0}	-5.7852		
γ_{f1}	0.7146		
γ_{g0}	-5.6295		
γ_{g1}	-0.0745		
<i>Learning probability</i>			
δ_0	-8.3104		
δ_1	0.4758		
<i>Utility function</i>			
b_i	110.50	925.70	0.00
c_{ri}	-	7.67	9,970.17
c_{ai}	1,380.43	-	1,940.41
c_{ni}	17,260.35	17,002.13	-
<i>Ability distribution</i>			
μ_i	10.3630	10.5093	9.2254
σ_x^2	0.0643	0.1135	0.1499
ρ_{ra}	0.6738		
ρ_{rn}	0.5584		
ρ_{an}	0.7825		
<i>Shocks distribution</i>			
$\sigma_{\varepsilon_i}^2$	0.9764	0.0882	0.6350

Table 1: Estimated coefficients.

Another result is positive correlation coefficients between the abilities, which suggest hierarchical sorting rather than sorting based on comparative advantage. The strongest correlation of 0.78 is found for R&D and non-S&E abilities. Correlation coefficient between R&D and non-S&E, and application and non-S&E are 0.67 and 0.56 respectively. This finding suggest that depending on the variance in the distribution of the skills in the population and relative variance among the skills mobility from task to task is easier than under the comparative advantage sorting scenario. The estimated variance of abilities is the largest in non-S&E followed by R&D ability (0.51 versus 0.36) capturing the high variance of the observed earnings in non-S&E due to the heterogeneous nature of tasks by the construction of the category. Productivity in applied tasks is also the least subjected to shocks, whose variance is estimated to be 0.08 compared to 0.64 in non-S&E and 0.98 in R&D. Mean ability in non-S&E is the lowest (9.58). Therefore, individuals choosing non-S&E since right after graduation have very large realized abilities in this task. Finally, the coefficient in the learning technology function predicts the probability of learning after 1 period to be approximately 0.3. More discussion on the effect of the various schedules of information arrival follows after the discussion of the model fit.

4.1 Model Fit

This part of the section demonstrates how well the model fits the observed career dynamics and earnings. Model predictions are plotted against the empirical moments on Figures 7, 8, and 9. The model fits major moments such as participation rates and earnings well. In particular it captures distribution across the tasks at all ages as well as the distribution and the growth rates of the earnings. It was challenging to fit the transition rates. The model suggests high transition rates between the S&E tasks relative to the data and underpredicts exits into the non-S&E.

Modelling career choices allows to pin down some variables of interest that are not available in the data such as research ability. Figures 10 and 11 demonstrate model predictions regarding the choices and earning of individuals by the quartile of their actual research ability.

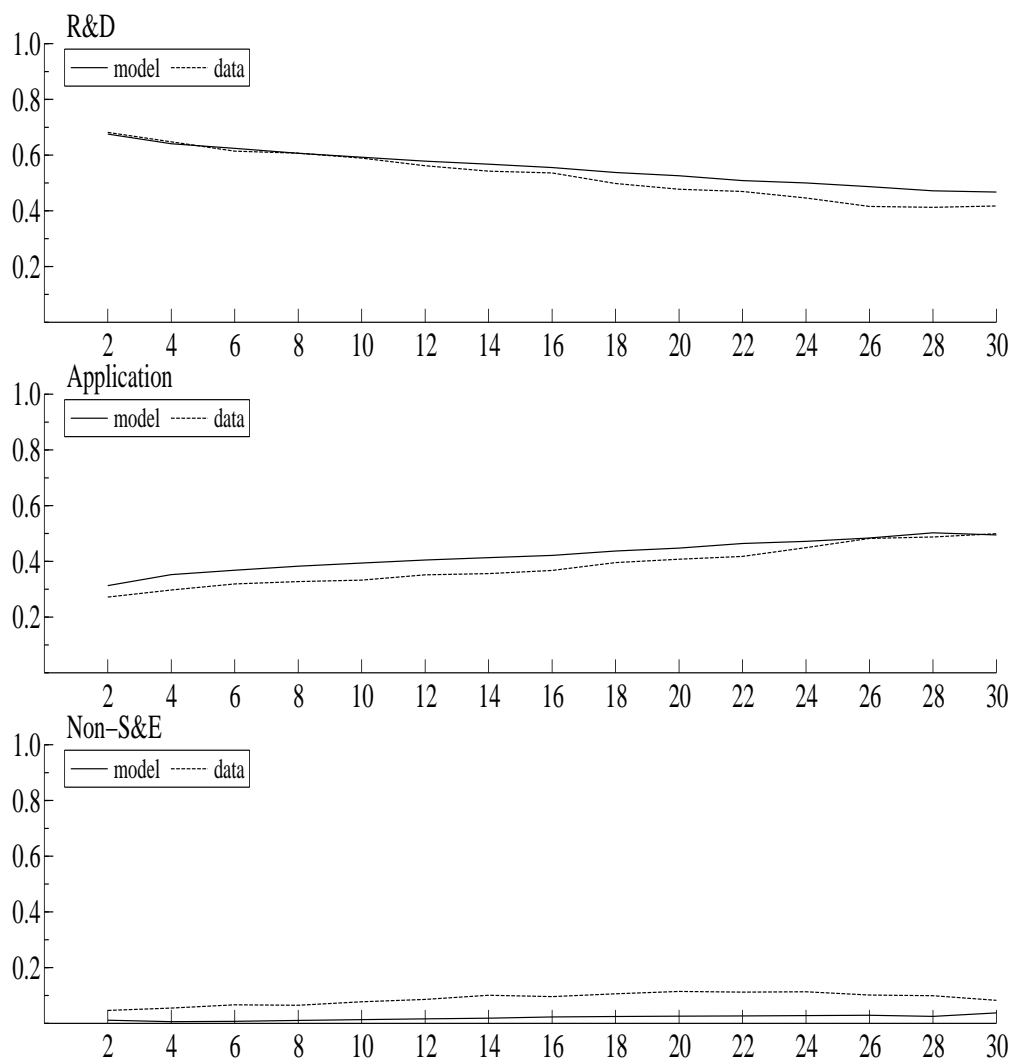


Figure 7: Employment rates, actual versus predicted

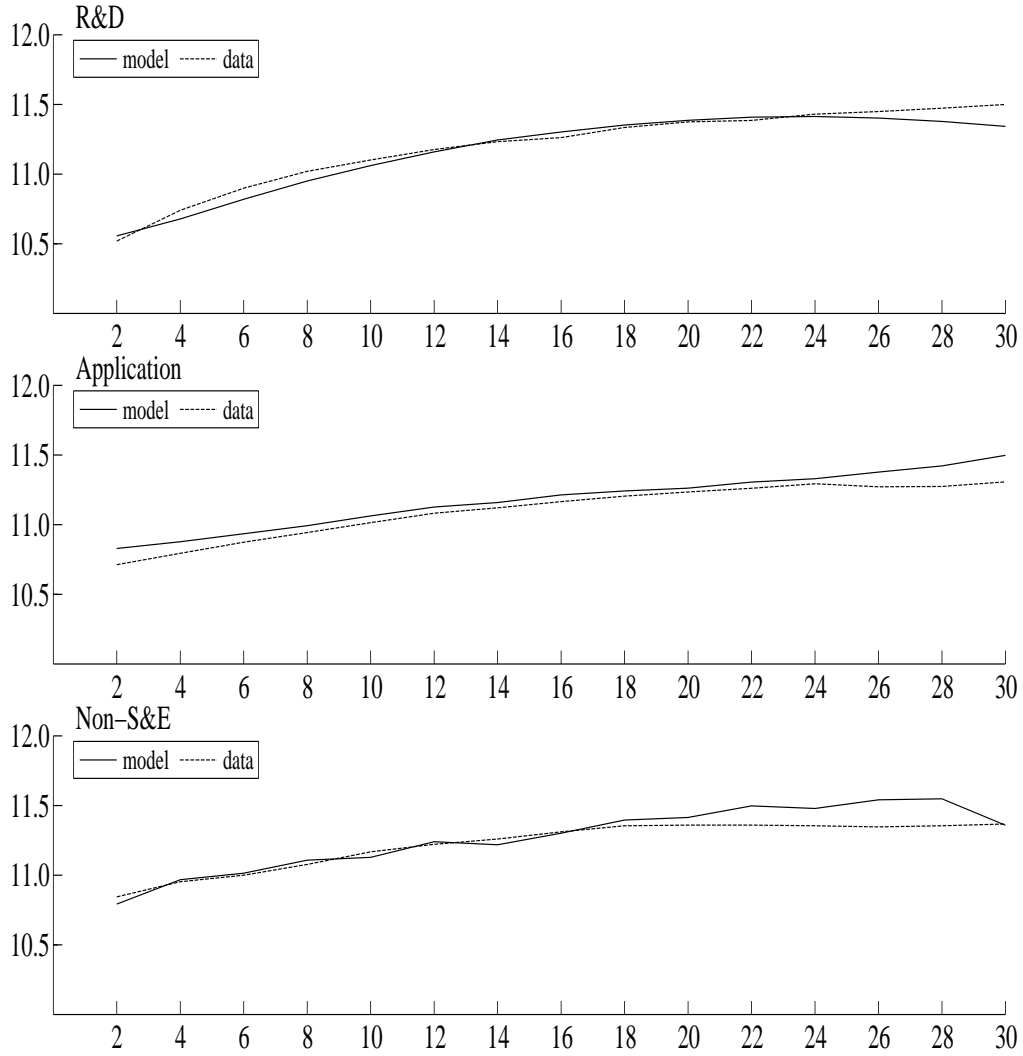


Figure 8: Average salaries, actual versus predicted.

4.2 Effect of information about ability

To understand how incomplete information about research ability affects career choices, the predictions of the baseline model are compared to the case of full information. If the R&D ability is public knowledge, R&D earnings and individual decisions are based on the actual rather than expected x_R . Predicted participation rates under full information case are plotted against the baseline model on Figure 12. The main result is that if individuals had known their R&D ability *ex ante*, they would have chosen the optimal task early in the career and stayed there until retirement. Incomplete information as predicted increases participation in

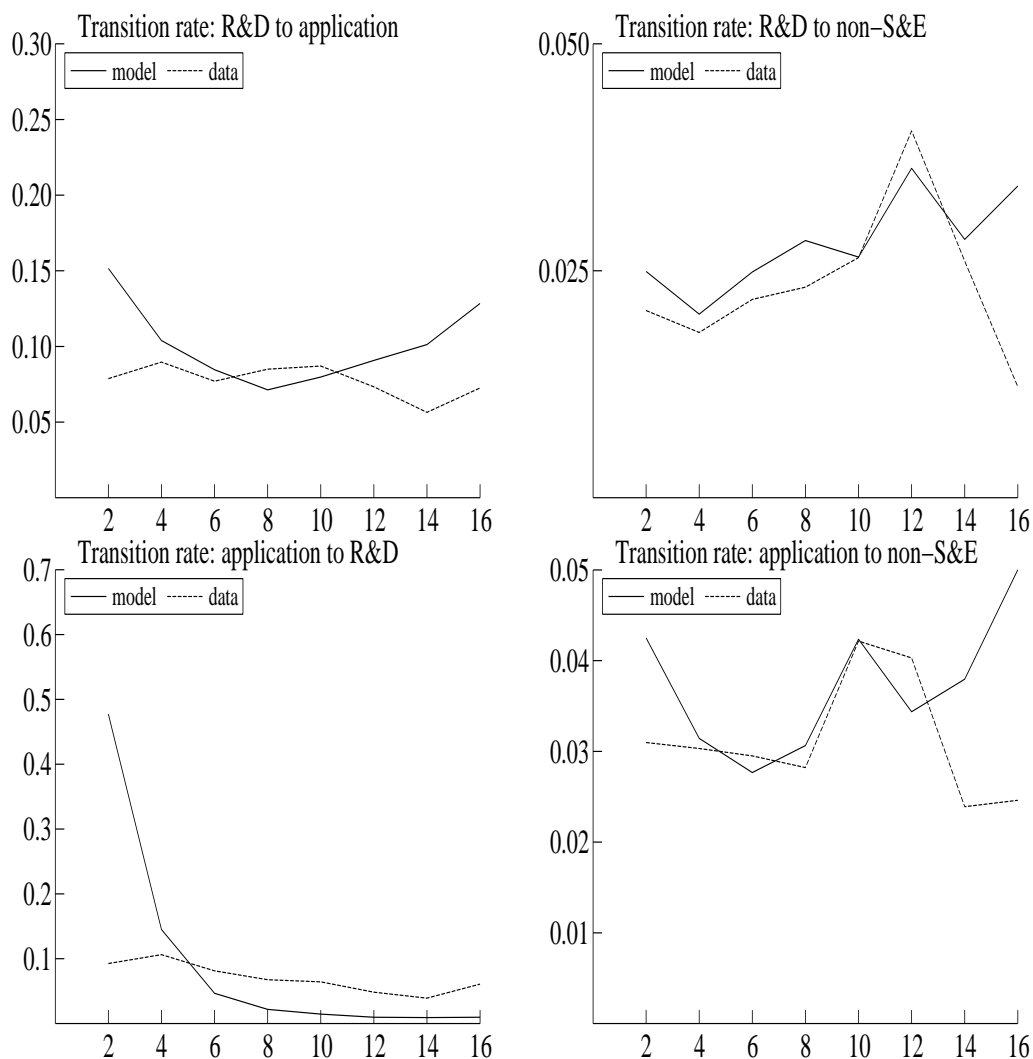


Figure 9: Transition rates, actual versus predicted.

R&D at all ages. The difference may include individuals with relatively low research talent who would have chosen research if they have not known their true ability *ex ante*. The difference between the predicted rates is especially high early in the career: in the first year after graduation there are 45 percent more researchers in R&D if ability is unknown. The latter is also due to the high predicted costs of entering R&D later in the career and finite time horizon. The gap between the participation rates gradually narrows down as low ability researchers leave the task. This is consistent with the data where major exits happen early in the career. Table 2 presents the results of the full information versus standard human

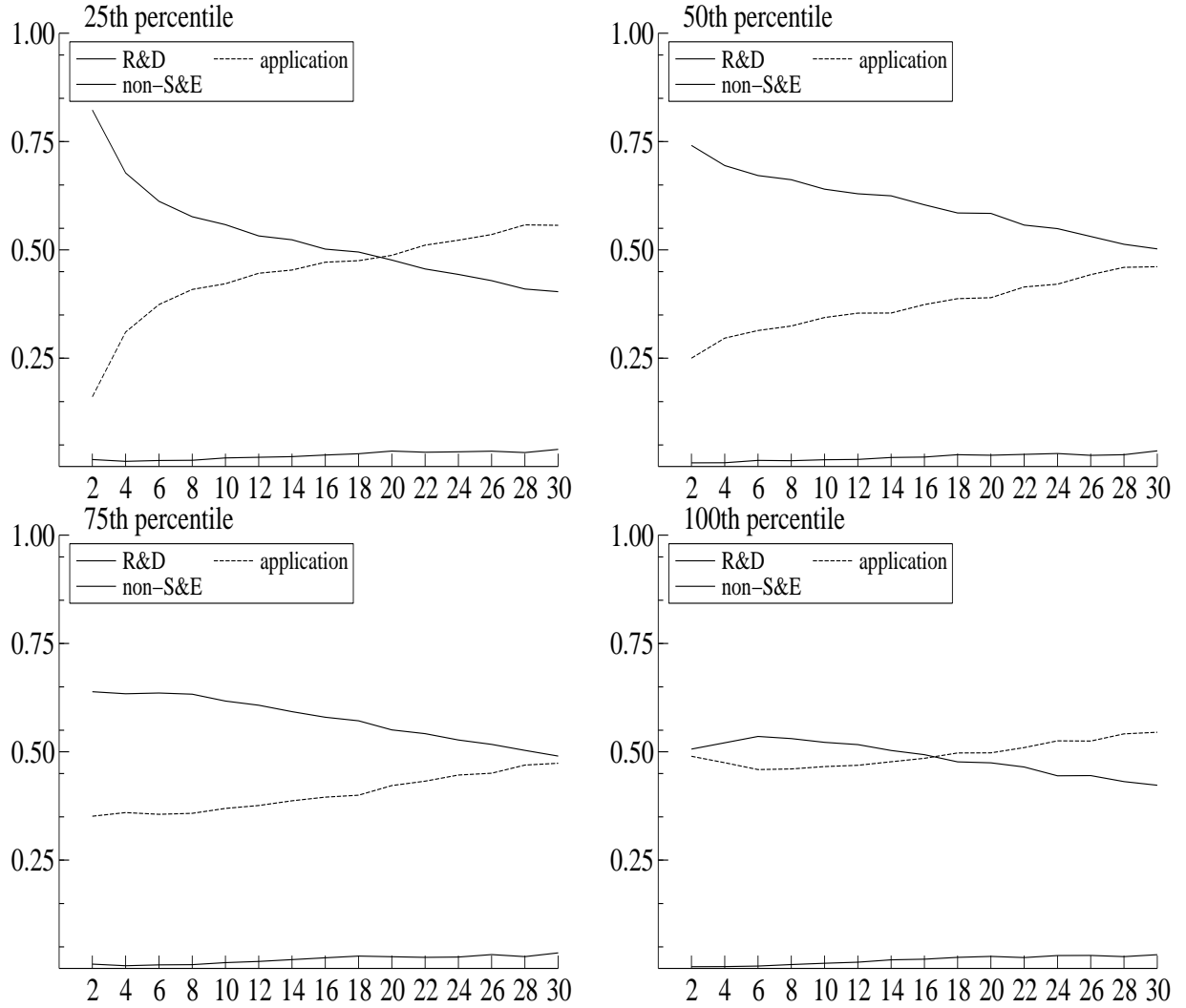


Figure 10: Predicted choices by percentile of research ability.

capital accumulation scenarios. Overall under the full information scenario the fraction of doctorates who ever tried R&D is lower by 14.73 percent. The average quality of researchers who choose to do R&D when ability is unknown increases by 17.6 percent because lower ability researchers prefer other employment options. Retention in R&D also increases by 17.4 percent. These new optimal career increase the discounted expected value of a doctoral degree by 4.6 percent.

Standard human capital accumulation increases participation in R&D tasks by almost 6 percent and the average quality of stayers by 5.22 percent. At the same time, retention

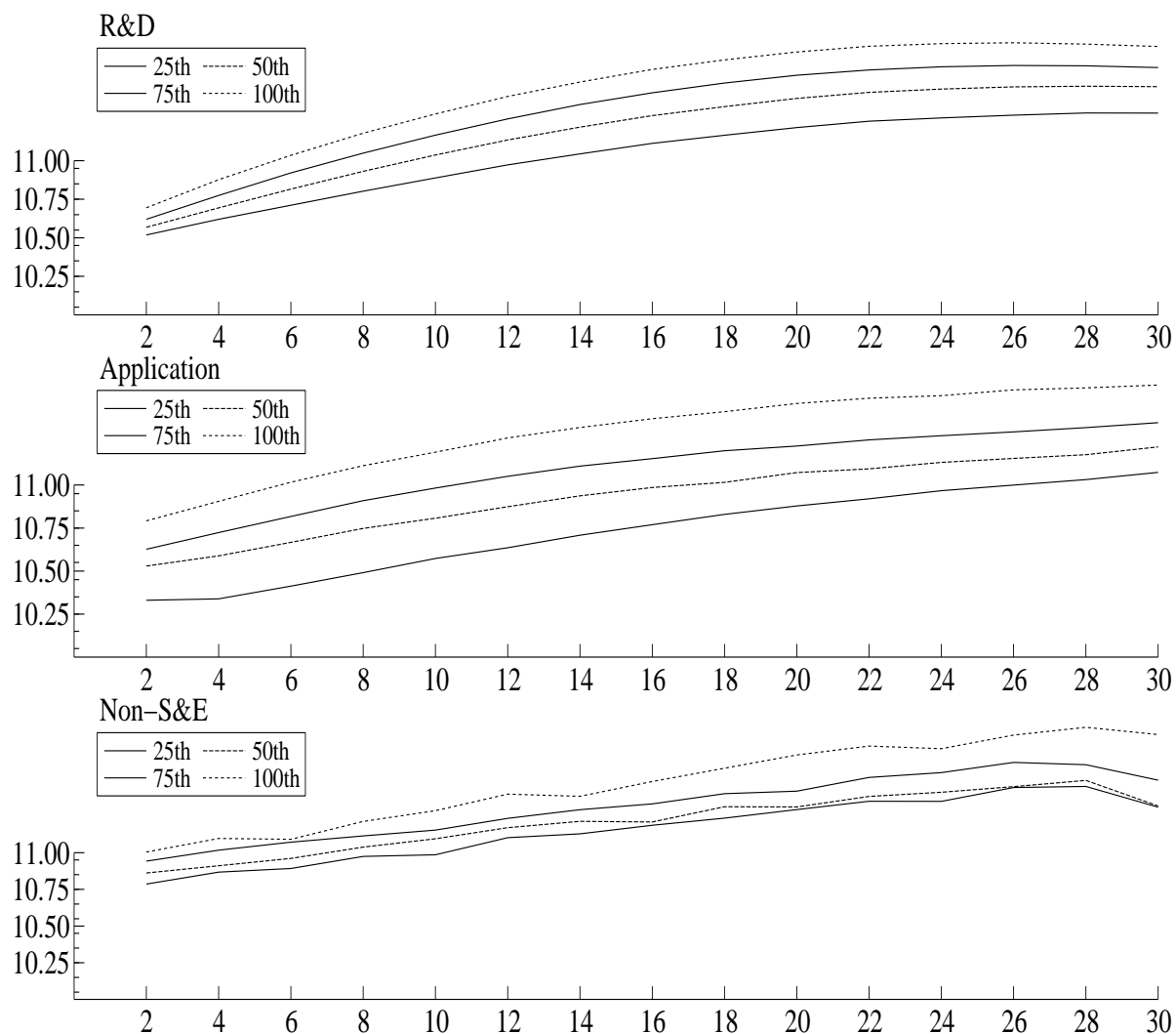


Figure 11: Predicted earnings by percentile of research ability.

in R&D drops by 65 percent. These findings suggest that if human capital accumulation was independent of ability, more people would be attracted to R&D partially because of the perspectives of faster salary growth in R&D and partially because of better perspectives in other options due to the skill transferability. The latter explains lower retention rates in the R&D because more scientists accumulate experience and it happens faster than under the baseline scenario which improves their outside options and facilitates the move. This increased outward mobility occurs at the cost of the relatively lower ability scientists and results in an improvement of the average ability of the stayers.

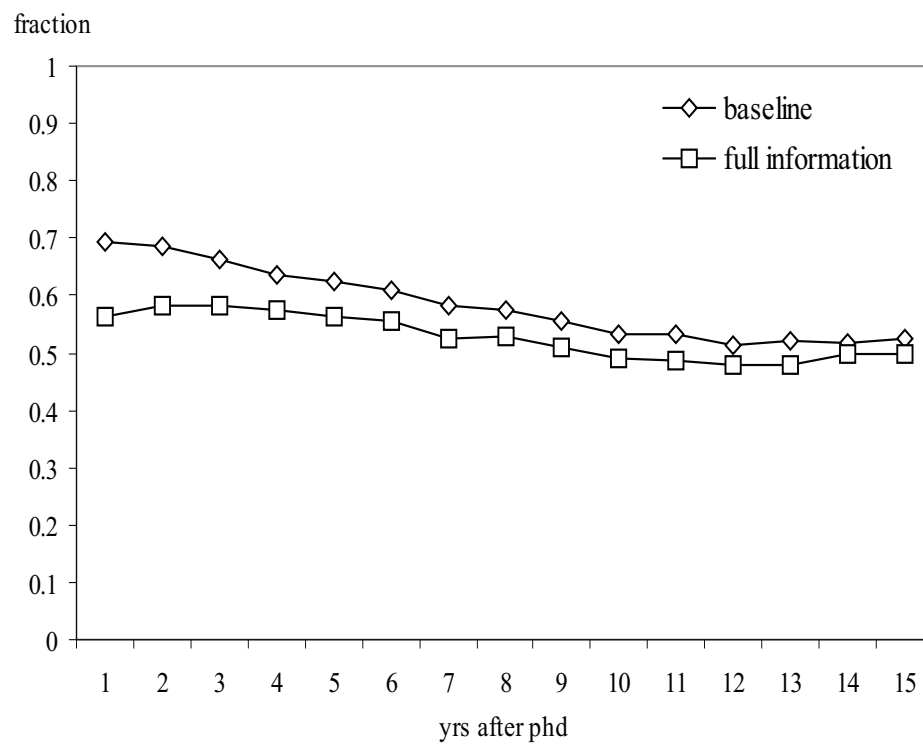


Figure 12: Participation in R&D: full information vs baseline model

Explanation	Full information	Standard human capital
1. Participation in R&D	-14.73%	5.83%
2. Retention in R&D	17.41%	-65.55%
3. Average quality in R&D	17.59%	5.22%
4. Change in the value	4.13%	1.56%

Table 2: Full information and standard human capital scenarios.

Explanation	Participation	Retention
1. Full information	-14.73%	-8.52%
2. No skill transferability	-7.13%	4.25%
3. No non-pecuniary benefits	-5.21%	-5.46%

Table 3: Accounting.

4.3 Alternative explanations and model predictions

As was mentioned earlier, there exist several alternative explanations of the observed career dynamics, that is, high non-pecuniary benefits and skill transferability. The model allows to evaluate how each of the stories contribute to the inter-task mobility and career choices. Table 3 presents simple accounting for the role of each competing explanation. To obtain each of them, first the coefficient at the non-pecuniary benefits were set to zero keeping everything else constant. Then, skill transferability coefficients were restricted to be zero and new predictions were obtained.

The results suggest that imperfect information about ability accounts for about 15 percent of participation in and 8.5 percent of mobility out of R&D tasks. Skill transferability accounts for another 7 percent and 4.25 percent respectively. Finally, non-pecuniary benefits increase participation in R&D by 5.2 percent and increase mobility by 5.46 percent.

5 Counterfactual Experiments

The predictions of the model are used to evaluate the effect of several counterfactual experiments on the supply of research skill and the value of doctoral degree. Two types of experiments were conducted: First I compare different learning schemes to quantify the effect of incomplete information about ability on scientific careers. Secondly, I evaluate how supply and the quality of the research skill reacts to changes in relative skill prices caused by a) R&D subsidies and b) the improvement of the employment options outside S&E.

5.1 Learning Schemes

The speed of information arrival is expected to change the relative value of employment in R&D tasks. To quantify this effect I evaluate three cases: a) slow learning $\delta = 0.05$, b) sure

learning $\delta = 1$, and c) $\delta = 0.75$. The probability of learning in the model is equivalent to having a chance to engage in an individual project, which is possible if a scientist receives a research grant or access to specific equipment or materials. Therefore, higher probability of learning can be associated with improvement in the availability of grants or decreasing competition for them. The results are compared to the cases of full information and the baseline model. The latter predicts that information about research ability arrives gradually over time: the baseline probability of learning δ is 0.30. The results of these experiments are summarized in Table 4.

Probability of learning	$\delta = 0.05$	$\delta = 0.30$	$\delta = 0.75$	$\delta = 1$
1. Participation in R&D:	-4.65%	0	3.20%	3.88%
2. Quality of researchers:	-10.64%	0	6.85%	9.55%
3. Change in value	-0.58%	0	0.45%	0.60%

Table 4: Effect of different learning schemes on the supplied research skill.

The results suggest that incomplete information about ability intensifies with the rate of the information arrival. As probability to learn increases, more scientists choose R&D options: in the “sure learning” case lifetime participation in R&D grows by 2.26 percent compared to the baseline case, while when the probability of learning is low (0.05) participation falls by 3.51 percent. As the probability of learning grows lower ability researchers are “sorted out” earlier resulting in the higher average ability of those who stay (-23.51 percent vs 10.20 percent in the extreme cases). The next result is that the “slow learning” case decreases the expected discounted value by almost 6 percent, while when learning is faster the value increases by 2.19 percent and 2.66 percent depending on the case.

The results of these experiments show that the effect of information about ability on the retention and quality of research skill intensifies with the speed of information arrival.

5.2 Changes in relative prices

The next experiment evaluates how changes in relative skill prices affect the career dynamics of doctorates. First, I evaluate the results of subsidies in R&D sector that would result in an increase in the relative skill prices by 50 percent.

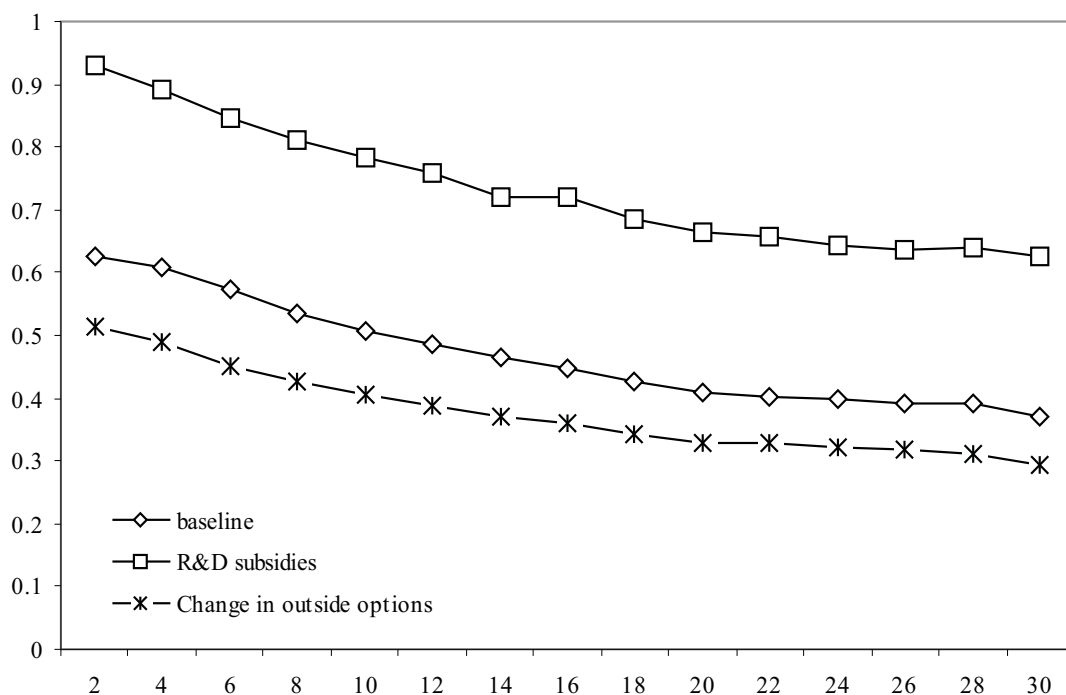


Figure 13: Participation in R&D and changes in relative skill prices

Low salaries of postdoctorates have been emphasized many times as an important source of the disadvantage of scientific careers relative to other options, Freeman et al. (2001). Other studies suggest that raising salaries in R&D is ineffective because scientific labor supply is inelastic due to long and costly training, Goolsbee (1998). However, as shown on Figure 13, the supply of the R&D skill is responsive to changes in skill prices: 50 percent raise of R&D skill prices causes participation in R&D increase by nearly 11 percent throughout a career, and almost doubles retention rates. Increased participation comes at the cost of decreasing quality in R&D. The average quality of R&D labor drops by almost 16.5 percent, which happens because relatively lower ability researchers have higher value of participation in research. Without modelling production of R&D, however, it is hard to tell whether the overall production of new knowledge falls due to the decreased quality of the participants or gains due to increased participation because of the cumulative nature of knowledge production.

The next experiment evaluates how improvement in outside options affects total skill

	R&D up 50 percent	non-S&E up 50 percent	baseline
1. Participation in R&D	25.81%	-5.27%	0
2. Retention in R&D	84.94%	-9.69%	0
3. Average quality of researchers			0
- if ever worked in R&D	8.14%	-0.80%	0
- if stayed in R&D	0.73%	0.31%	0
4. Change in value	42.44%	2.34%	0

Table 5: Effect of changing relative prices in R&D and non-S&E.

supplied to the R&D sector. In particular, I consider the 50 percent raise of skill prices in non-S&E. In this case, participation in R&D falls by 17 percent, although retention does not improve (see Figure 13 and Table 5). The average quality of labor in R&D in this scenario increases by 3 percent.

6 Conclusion

This paper explains the puzzle in the observed career dynamics of S&E doctorates that is hard to reconcile with the standard occupational choice model unless large non-salary benefits in R&D are assumed. The puzzle consists of a) predominance of R&D employment early in the career accompanied by the lowest relative earnings and b) decreasing participation in R&D as careers develop in combination with high growth rates of the stayers. This paper suggests that the dependence of productivity in R&D on ability to produce new research ideas intensified through the dependence on past performance and the lack of *a priori* information about ability account for a large fraction of the non-salary benefits in R&D. The paper develops a dynamic model of occupational choices with symmetric learning about research ability and the stochastic evolution of human capital to capture the proposed mechanisms and evaluate the predicted behavior. The model is estimated on the rarely utilized Survey of Doctorate Recipients for 1973-2001 using the Method of Simulated Moments.

The paper finds that incomplete information about ability to do research plays an important role in career choices. Incomplete information causes higher participation in R&D and reduces the average quality of the supplied research skill. Information about ability accounts for approximately 44 percent of the non-salary benefits in R&D. To better un-

derstand the effect of incomplete information and its realization, several different learning schemes were evaluated and compared to the full information case. The results of these counterfactual experiments suggest that the effect of learning intensifies with the speed of information arrival.

The predictions of the model were used to evaluate the effect of changes in the relative skill prices in R&D and non-S&E. Both experiments suggest that supply of research skill previously thought of as inelastic is sensitive to changes in relative prices and the value of outside options.

References

- Biddle, J. and Roberts, K. (1994). Private sector scientists and engineers and the transition to management. *The Journal of Human Resources*, 29(1):82–107.
- Butz, W., Bloom, G., Gross, M., Kelly, T., Kofner, A., and Rippen, H. (2003). Is there a shortage of scientists and engineers? How would we know? *RAND Corporation*, IP-241-OSTP.
- Diamond, A. (1986). What is a citation worth? *Journal of Human Resources*, 21(2):200–215.
- Diamond, A. (2001). Scientists’ salaries and the implicit contracts theory of labour markets. *Int.J. Technology Management*, 22(7/8):688–697.
- Diermeier, D., Keane, M., and Merlo, A. (2004). A political economy model of congressional careers. *PIER*, Working Paper(04-037).
- Fallick, B., Fleischmann, C., and Rebitzer, J. (2005). Job hopping in Silicone Valley: Some evidence concerning the micro-foundations of a high technology cluster. *NBER*, Working Paper(11710).
- Ferrall, C. (1997). Empirical analysis of occupational hierarchies. *The Journal of Human Resources*, 32(1):1–34.
- Freeman, R. (1975). Supply and salary adjustments to the changing science manpower market: Physics, 1948-1973. *The American Economic Review*, 65(1):27–39.
- Freeman, R. (2005). Does globalization of the scientific/engineering workforce threaten US economic leadership? *NBER*, Working paper(11457).
- Freeman, R., Weinstein, E., Marincola, E., Rosenbaum, J., and Solomon, F. (2001). Competition and careers in biosciences. *Science*, 294(5550):2293–2294.
- Gaughan, M. and Robin, S. (2004). National science training policy and early scientific careers in France and the United States. *Research Policy*, 33:569–581.

- Goolsbee, A. (1998). Does government R&D policy mainly benefit scientists and engineers? *American Economic Review*, 88(2):298–302.
- Grimes, P. and Register, C. (1997). Career publications and academic job rank: evidence from the class of 1968. *The Journal of Economic Education*, 28(1):82–92.
- Heckman, J. and Sedlacek, G. (1985). Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy*, 93(6):1077–1125.
- Jackson, S. (2003). Envisioning a 21st century science and engineering workforce for the united states: Tasks for university, industry, and government. Technical report, Government-University-Industry Research Roundtable of National Academies, Washington, D.C.
- Jovanovic, B. and Nyarko, Y. (1996). Stepping stone mobility. *NBER*, Working Paper(5651).
- Kang, K. (2003). Characteristics of doctoral scientists and engineers in the United States: 2001. Technical Report NSF 03- 310, National Science Foundation, Arlington, VA.
- Keane, M. and Merlo, A. (2007). Money, political ambition, and the career decisions of politicians. *PIER*, Working Paper(07-016).
- Keane, M. and Wolpin, K. (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *The Review of Economics and Statistics*, 76(4):648–672.
- Keane, M. and Wolpin, K. (1997). The career decisions of young men. *The Journal of Political Economy*, 105(3):473–522.
- Levin, S. and Stephan, P. (1991). Research productivity over the life cycle: Evidence for academic scientists. *The American Economic Review*, 81(1):114–132.
- Lowell, B. and Salzman, H. (2007). Into the eye of the storm: Assessing the evidence on science and engineering education, quality, and workforce demand. *The Urban Institute*, Research Report.

- Majumdar, S. and Shimotsu, K. (2005). Enrollment responses to labour market conditions: A study of the canadian market for scientists and engineers.
- Mangematin, V. (2000). PhD job market: professional trajectories and incentives during the PhD. *Research Policy*, 29:741–756.
- McDowell, J. (1982). Obsolescence of knowledge and career publication profiles: Some evidence of differences among fields in costs of interrupted careers. *American Economic Review*, 72(4):752–768.
- Mervis, J. (2003). Scientific workforce: Down for the count? *Science magazine*.
- Miller, R. (1984). Job matching and occupational choice. *Journal of Political Economy*, 92(6):1086–1120.
- Mishagina, N. (2007). Empirical analysis of career transitions of sciences and engineering doctorates in the US. *Queen’s University*, Working Paper(1137).
- Moen, J. (2001). Is mobility of technical personnel a source of R&D spillovers? *NBER*, Working Paper(7834).
- Nagypál, E. (2007). Learning-by-doing versus learning about match quality: Can we tell them apart? *Review of Economic Studies*, 74(2):537–566.
- Neal, D. and Rosen, S. (2000). *Handbook of Income Distribution*, chapter 7, pages 379–427. Elsevier.
- Oyer, P. (2005). The macro-foundations of microeconomics: Initial labor market conditions and long-term outcomes for economists and MBAs. *Working Paper, Stanford Business School*.
- Preston, A. (2004). *Leaving Science: Occupational Exit from Scientific Careers*. Russell Sage Foundation.
- Robin, S. and Cahuzac, E. (2003). Knocking on academia’s doors: An inquiry into the early careers of doctors in life sciences. *CEIS*, 17(1):1–23.

- Roy, A. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3:135–146.
- Ryoo, J. and Rosen, S. (2004). The engineering labor market. *Journal of Political Economy*, 112(1, Pt.2):S110–S140.
- Sauer, R. (1998). Job mobility and the market for lawyers. *The Journal of Political Economy*, 106(1):147–171.
- Scafidi, B., Sjoquist, D., and Stinebrickner, T. (2007). Do teachers really leave for higher paying jobs in alternative occupations? *Advances in Economic Policy and Analysis*, forthcoming.
- Stephan, P. (1996). Economics of science. *Journal of Economic Literature*, 34(3):1199–1235.
- Stephan, P. and Ma, J. (2004). The increased frequency and duration of the postdoctorate career stage. *Georgia State University*, Working paper.
- Stern, S. (1999). Do scientists pay to be scientists. *NBER*, Working Paper(7410).
- Stinebrickner, T. (2001). A dynamic model of teacher labor supply. *Journal of Labor Economics*, 19(1):196–230.
- Teitelbaum, M. (2004). Do we need more scientists? In Kelly, T., Butz, W., Carroll, S., Adamson, D., and Bloom, G., editors, *The US Scientific and Technical Workforce: Improving Data for Decision Making*, CF-194-OSTP/SF, pages 11–20, Arlington, VA. RAND Science and Technology, RAND Corporation.
- Thurgood, L., Golladay, M., and Hill, S. (2006). US doctorates in the 20th century. special report NSF 06-319, National Science Foundation (SRS/SBES), Arlington, VA.
- Van den Berg, G., Holm, A., and van Ours, J. (2002). Do stepping-stone jobs exist? Early career paths in the medical profession. *Population Economics*, 15:647–665.
- Zucker, L. and Darby, M. (2006). Movement of star scientists and engineers and high-tech firm entry. *NBER*, Working Paper(12172).

Appendix 1

6.1 Types of occupations and assignment principles

This section outlines the basic principles of assigning individuals and their employment into one of three types of tasks using information on their primary and secondary tasks, occupation, and sector of employment. The survey questions regarding occupations and activities changed substantially in 1993. These changes, however, did not substantially affect the classification because the suggested responses were detailed enough to determine what type of job the person held. One major difference in the questionnaire is that before 1993 the list of possible activities distinguished between the “management of R&D”, “management of education”, and “management in non-R&D”, which made assignment to different tasks easy. The post-1993 questionnaire aggregated these three activities into one called “management and administration”. For this period I used information on the secondary activity to determine the relevance of each case to R&D. A comparison between occupational assignments for pre- and post-1993 data shows no discrepancies or inconsistencies.

I consider three types of tasks: R&D-related, applied, and non-S&E. A responder is considered to be in the R&D task in a given period of time if the following conditions hold: a) the reported primary activity was either research, development, design, or management of R&D; b) indicated their occupation as scientists (e.g. physicist), engineers (e.g. mechanical engineer), “postsecondary teacher”, or “manager”; c) the reported employer was academic, government, or industry that can be considered as S&E-related. Using information on occupation and sector allows to exclude individuals who performed research tasks in non-S&E sectors.

The second type includes jobs that require mostly application of accumulated scientific or technical knowledge rather than development of new knowledge. Respondents were considered employed in this type of tasks reported: a) activities such as teaching in S&E fields in both secondary and post-secondary institutions, professional services in S&E (e.g. technical consulting and assessment, counselling, surveying, etc.), software development, and managerial activities in these areas, b) the employer’s industry was S&E-related.

The last category, non-S&E tasks, included tasks unrelated to S&E. In order to distinguish the relevance of a job to S&E, I used a taxonomy different from that adopted by the NSF. For example, the NSF taxonomy considers all managerial occupations as unrelated to S&E. This way, the head of a university department or a director of a research laboratory would be considered as someone who changed his career. For the purposes of my analysis, such a classification would give misleading results because mobility from S&E to non-S&E would include both career advancements (exits due to promotions) and career changes. Distinguishing between the two is possible only when managerial occupations are separated by their relevance to S&E. One example of non-S&E tasks would be teaching non-S&E subjects (e.g. in humanities, business¹⁰, law, or arts). Another example would be employment in the areas of legislation, business services, such as finance, accounting, non-technical consulting, or marketing and sales of products and services in non-S&E industries (e.g. tourism and hospitality, entertainment, and media). Some might argue that business consulting or legislation in high-tech industries or manufacturing requires technical knowledge. I agree with this argument but believe that technical education for these professions does not require to be at the level of a research doctorate.

This classification of occupations can seem too general as it does not differ by discipline. However, for the purposes of this study it captures the major features of jobs that are similar across different disciplines. For example, it is true that mathematicians and biochemists investigate very different problems using different methods, and their products have different life-cycle and commercial value. However, both types of scientists use analytical and modelling skills, need to demonstrate high ability to generate original ideas to create scientific breakthroughs, and aim at being independent researchers. Second argument is that the current classification allows for interdisciplinary projects that are very common in industrial research and development, when the difference across disciplines is less defined and scientists from different fields are considered as substitutes. The same argument applies for grouping research occupations from different sectors. It is true that academic and industrial research differ considerably in requirements, earnings, and work environments, however when it comes

¹⁰For certain social science majors, especially for economists, business would not be an unrelated area. However, social science doctorates were excluded from the estimation sample.

to research it is expected to impose similar requirements to research ability, implicitly via project funding queue or explicitly via tenure contracts.

Appendix 2

	R&D jobs	applied jobs	non-S&E	Overall
Men	0.671	0.513	0.512	0.645
White	0.801	0.851	0.812	0.819
Citizen, native	0.811	0.841	0.820	0.813
Married	0.471	0.452	0.670	0.470
Graduate of Research I and II	0.867	0.849	0.854	0.860
Top-school graduate	0.158	0.134	0.155	0.149
Fraction with non-S&E degrees	0.049	0.089	0.128	0.069
Academic sector	0.404	0.658	0.178	0.478
Industry/business sector	0.472	0.278	0.712	0.420
Postdoctorate	0.116	0.020	0.006	0.104
Tenure-track if acad.	0.155	0.187	0.106	0.169
Tenured if acad.	0.361	0.536	0.534	0.451

Table 6: Descriptive statistics by task

Description of the moment	Number of moments in the group
Participation rates	T periods \times 3 tasks
Log(salaries)	T periods \times 3 tasks
St.dev. of log(salaries)	T periods \times 3 tasks
Fraction never employed in research	1
Log(salaries) if ever/never worked in research ^a	T periods \times 2 cases
St.dev. of log(salaries) if ever/never worked in research	T periods \times 2 cases
Lifetime participation in research	T periods \times 2 cases
Log(salaries) by participation in research ^b	T periods \times 2 cases
St.dev. of log(salaries) by participation in research	T periods \times 2 cases
Transition rates ^c	10×4 types of transitions
Log(salaries) conditional on transition	10×4 types of transitions
Log(salaries) of “stayers” in R&D	10

Table 7: Set of empirical moments.

^aEver- versus never employed in R&D

^bExperience in research by certain age: yes or no

^cIn the data the transitions are recorded for every 2nd year of employment. This way 10 records correspond to 20 years on the task