Wrapping it up in a person: Examining employment and earnings outcomes for Ph.D. recipients

Nikolas Zolas,1 Nathan Goldschlag,1 Ron Jarmin,1 Paula Stephan,2,3 Jason Owen-Smith,4 Rebecca F. Rosen,5 Barbara McFadden Allen,6 Bruce A. Weinberg,7,3,8,9,10 Julia I. Lane6,3,8,9,10

In evaluating research investments, it is important to establish whether the expertise gained by researchers in conducting their projects propagates into the broader economy. For eight universities, it was possible to combine data from the UMETRICS project, which provided administrative records on graduate students supported by funded research, with data from the U.S. Census Bureau. The analysis covers 2010–2012 earnings and placement outcomes of people receiving doctorates in 2009–2011. Almost 40% of supported doctorate recipients, both federally and nonfederally funded, entered industry and, when they did, they disproportionately got jobs at large and high-wage establishments in high-tech and professional service industries. Although Ph.D. recipients spread nationally, there was also geographic clustering in employment near the universities that trained and employed the researchers. We also show large differences across fields in placement outcomes.

Ten years ago, Jack Marburger challenged academics to provide scientific evidence about the impact of research investments (1). The United States Congress has been even more insistent: requiring the National Science Foundation to “better articulate the value of grants to the national interest.” The aim of the current study is to investigate the labor market outcomes of doctoral recipients.

Little is known about where research-funded Ph.D.’s go when they graduate and enter the private sector, and even less is known about the characteristics of the businesses that employ them. Thus, it has been difficult to capture the human dimension of the impact of research on the economy. What little evidence there is has been based on an analysis of patent clusters (2, 3); the geographic and industry placement of new Ph.D.’s (4–6); or on bibliometric approaches linking grants, patents, and publications (7, 8). One noteworthy exception has been expensive human dimension of the impact of research investments (4–6); or on bibliometric approaches linking grants, patents, and publications (7, 8). One noteworthy exception has been expensive— the United Kingdom spent more than £34 million (U.S. $51 million) in explicit costs, and much more in implicit costs, to generate almost 7000 case studies. Unfortunately, they lack a common framework or shared standards of evidence and presentation. Consequently, the extent to which this type of information provides rigorous, systematic, aggregate insights into economic value is far from clear (9).

We drew on recent investments to build administrative data that cover researchers supported by both federally and nonfederally funded grants in eight major universities that are members of the Committee on Institutional Cooperation (CIC) (10). Participating universities were Indiana, Iowa, Michigan, Minnesota, Ohio State, Purdue, Penn State, and Wisconsin. Those data, derived from the CIC’s UMETRICS project (11), provided the share of time charged to funded research projects by all people employed under them (from undergraduate and graduate students to postdocs to staff and faculty), as well as purchases for those projects. Graduate students were linked to their dissertations (from ProQuest) and to all subsequent employers and earnings in the United States through matches to Census Bureau data.

We documented the 2010–2012 earnings and placement outcomes of people receiving doctorates in 2009–2011. The universities have provided identifiers that allow the UMETRICS data to be linked to administrative and survey data housed at the U.S. Census Bureau under strict confidentiality protocols. The data are protected by law and are for statistical use only (anonymized unique identifiers are used for match keys), and all results are reviewed to ensure that no identifiable information is disclosed. We performed two distinct, but related, analyses of the jobs obtained by doctoral recipients. The first describes the characteristics of the establishments and firms where people obtained jobs. The second emphasizes the earnings of individuals at those jobs. It should be noted at the start that these analyses are designed to be descriptive and do not demonstrate causation.

Placement data were obtained from links to two files derived from administrative data and augmented by survey data: (i) the Business Register (BR), which is the universe of U.S. nonagricultural firms and associated establishments, and (ii) the Longitudinal Business Database (LBD), which contains longitudinally linked data for all firms and associated establishments with paid employees in the United States (12, 13). These files were used to describe the characteristics of the establishments and firms that employ the UMETRICS doctoral recipients covered by the data (an establishment is the physical place where business is conducted and the unit of observation at which industry and geographic location are defined; firms can own one or more establishments). The matched data were used to describe the sector and detailed industry classification of each establishment at which the recipients worked, the geographic location of their place of work, as well as characteristics related to productivity, such as size and payroll per worker (14). Age was determined at the firm level, and establishments owned by R&D performing firms were identified from the Business Research & Development and Innovation Survey.

Table 1. Postgraduation employment of UMETRICS doctoral recipients who were paid by research grants and left the university between 2009 and 2011. The national workforce distribution is calculated from all employment in all establishments covered by the Census’s LBD between 2010 and 2012.

<table>
<thead>
<tr>
<th>Doctoral recipients placed in sector (%)</th>
<th>Industry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R&amp;D firms</td>
</tr>
<tr>
<td>Locale and small</td>
<td></td>
</tr>
<tr>
<td>Placed within sector</td>
<td>170</td>
</tr>
<tr>
<td>National sample (M)</td>
<td>10.8</td>
</tr>
<tr>
<td>Of those in sector, percent placed:</td>
<td></td>
</tr>
<tr>
<td>Within 50 miles</td>
<td>10.1</td>
</tr>
<tr>
<td>Within state</td>
<td>16.6</td>
</tr>
</tbody>
</table>

1U.S. Census Bureau, 4600 Silver Hill Road, Washington, DC 20233, USA. 2Department of Economics, Box 3992, Georgia State University, Atlanta, GA 30302-3992, USA. 3National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138-5398, USA. 4University of Michigan, 500 South State Street, no. 3001, Ann Arbor, MI 48103-1382, USA. 5New York University, 70 Washington Square South, New York, NY 10002, USA. 6Committee on Institutional Cooperation (CIC), 1819 South Neil, Suite D, Champaign, IL 61820, USA. 7Department of Economics, Ohio State University, 1945 North High Street, Columbus, OH 43210, USA. 8Institute for the Study of Labor (IZA), Schumpeter- LippeStrasse 5-9, 53113 Bonn, Germany. 9University of Strasbourg, 61 Avenue de la Forêt Noire, 67 085 Strasbourg, France. 10American Institutes for Research, 1000 Thomas Jefferson Street, Washington, DC 20007, USA.

*Corresponding author. E-mail: weinberg.27@osu.edu
The variables we analyzed, the level at which they were measured, and their source are listed in table S1. Individual earnings data were derived from links to administrative records. Details of the data construction and links are in the supplementary materials (SM), including fig. S1 and table S2, A and B. De-identified versions of the UMERTICS data will be made available to the scholarly research community through a confidentially protected virtual digital enclave being developed at the Institute for Research on Innovation and Science (IRIS) (iris.isr.umich.edu) (10). Data integrated with census data on people and on their employers and their employer’s characteristics will be made available at the Federal Statistical Research Data Centers through a partnership between IRIS and the U.S. Census Bureau (www.census.gov/fsrdc) under strict confidentiality protections.

There were 3,397 graduate students on research payrolls at the sample universities in the period 2009–2011 who received a doctoral degree during that period and who were employed at a different institution in subsequent years. The major sectors into which they flowed in the year after their separation from university employment are summarized in Table 1 and compared with the sectors of employment for the U.S. workforce as a whole, which provides a benchmark (but not a formal control) group. The majority of graduate students (57.1%) went to academia—presumably many to postdoctoral positions. A large percentage (38.7%) found jobs in industry, notably, about 17% in establishments owned by firms that perform research and development (R&D) (for comparison, 10.8% of the U.S. workforce is employed in such establishments) and about 21.7% in establishments owned by firms that do not perform R&D (versus 75.0% of the U.S. workforce). Only a small percentage (4.1%) entered government.

Evidence that research funding to these universities provides training to a workforce that participates in a national (and likely international) labor market is shown in Table 1. For each of the universities in our sample, more than one in five doctoral recipients stayed in the state in which the university was located, per worker is the average payroll (the total payroll divided by the number of employees) across all employees at the three types of establishments—all U.S. establishments, all U.S. establishments owned by firms that perform R&D, and the establishments that employed UMERTICS doctoral recipients (regardless of whether they are owned by firms that perform R&D). National and R&D establishments are weighted by total establishment employment, whereas doctoral recipient establishments are weighted by the number of doctoral recipients employed. Values for annual payroll per worker are U.S.$1 × 1000.

**Fig. 1. UMERTICS doctoral recipients are placed at establishments that are larger and have higher payrolls per worker.** Medians are dashed inner lines, and means are solid outer lines. The standard deviations in employment at establishments that employed UMERTICS doctoral recipients, at all U.S. establishments owned by R&D performing firms, and all U.S. establishments are 6407, 3661, and 2362, respectively; the standard deviations in annual payroll per worker are $120,199; $56,252; and $44,327, respectively; the differences in employment size and payroll per worker are statistically significant. Annual payroll per worker is the average payroll (the total payroll divided by the number of employees) across all employees at the three types of establishments—all U.S. establishments, all U.S. establishments owned by firms that perform R&D, and the establishments that employed UMERTICS doctoral recipients (regardless of whether they are owned by firms that perform R&D). National and R&D establishments are weighted by total establishment employment, whereas doctoral recipient establishments are weighted by the number of doctoral recipients employed. Values for annual payroll per worker are U.S.$1 × 1000.

**Fig. 2. Annual payroll per worker at establishments that employed UMERTICS doctoral recipients, establishments owned by firms that perform R&D, and all U.S. establishments.** Values for average annual payroll per employee are U.S.$1 × 10,000.
and about 13% stayed within 50 miles of the university. More doctoral recipients stayed in the university's state than moved to any other single state, with the exception of California, which received more students from two of our eight universities. The results are substantially unchanged when the sample was restricted to those supported only by federal research grants (table S3).

The data permit an even deeper examination of the geographic destinations of the doctoral recipients. For example, 19% of the doctoral recipients who left their university's state headed for California, which has only 12% of the U.S. population. We expect that this is partially due to the fact that more R&D is conducted in California than in any other state. The states to which people moved are shown in fig. S2, and a comparison of the locations of Ph.D. recipients to R&D and population by state are shown in table S4. The other major destination states in terms of numbers of graduates among people who leave the state where they trained were Illinois, New York, Texas, North Carolina, Massachusetts, Pennsylvania, and Washington (state), which either have a large share of R&D relative to their populations or are very populous.

It is also possible to examine the specific industries that were most likely to employ the sample of doctoral recipients paid by research grants (table S5); results for doctoral recipients supported by federally funded projects are reported in table S6. The employers were much more likely to be in industries such as engineering, or high-tech and professional service fields (including medicine) than U.S. employers at large: For example, the shares of doctoral recipients employed in pharmaceutical and medicine manufacturing, semiconductors, and computer systems design were between 4 and 19 times the U.S. average. Uncommon employment destinations for doctoral recipients were restaurants and eating places and grocery stores.

The establishment-level data also permitted an examination of the differences between the employers of the UMERTICS doctoral recipients being studied, the typical employer across the U.S. as a whole, and the typical U.S. establishment owned by a firm performing R&D (Fig. 1). Because the distribution of establishments is highly skewed, it is informative to provide information on both the means and the medians. The median establishment that employed the doctoral recipients in our sample has more employees (at 1084) than the national median (at 73) or even than establishments owned by firms performing R&D (341) (Fig. 1). We used payroll per worker (constructed by dividing total payroll by the number of employees) to measure the average earnings at establishments. The median establishment that employed UMERTICS doctoral recipients also tended to have a higher payroll per worker (over $90,000) than the median U.S. establishment, which has a payroll per worker of just over $33,000, or the median establishment owned by a firm performing R&D, which has a payroll per worker of just less than $61,000. The same results hold when we used the mean as the measure of central tendency.

The data were rich enough to go beyond summary statistics and to characterize the distribution of the average measure for all establishments employing the workers under study. In particular, we calculated the payroll per worker for each establishment (regardless of R&D status) employing individuals in our sample, all U.S. establishments, and all U.S. establishments owned by firms performing R&D (regardless of whether they employ the doctoral recipients in our sample) and report the distribution of each in Fig. 2. Although only 8.3% of the U.S. workforce and 24% of workers at establishments owned by firms performing R&D worked at establishments with payrolls per worker in excess of $100,000, more than half (51%) of the sample of doctoral recipients do. The results are consistent with the hypothesis that doctoral recipients are placed in establishments with characteristics usually associated with greater productivity.

It is possible to examine individual earnings outcomes, as well as the placement outcomes described above. The data also permit the statistical analysis of the relation between a researcher’s field of study (based on ProQuest dissertation data) and subsequent placement and individual earnings. The results of the analysis of individual earnings outcomes in the year after Ph.D. receipt are reported in (Fig. 3). Descriptive statistics and regression results are shown in tables S7 and S8. Although earnings are an imperfect measure of the value of skills, especially for people still investing in their human capital, the labor economics literature finds a strong correlation between skills and earnings (16). The two fields with the highest earnings are mathematics and/or computer sciences and engineering, with mean earnings in excess of $65,000. Although mean earnings for doctoral recipients are low in biology at $36,000, this may be due to many taking jobs as postdoctoral researchers in life sciences. If the sample is subset to only include industry earnings, the average earnings increase by one quarter (although the gap varies by field), with the highest earnings in mathematics and/or computer sciences (almost $90,000) and engineering (almost $80,000).

Placements are reported in the four right panels of Fig. 3. Doctoral recipients with degrees in engineering are most likely to go to work in industry, followed by those with degrees in math and/or computer science. New degree holders in engineering are by far the most likely to go to establishments that are owned by firms performing R&D. In all fields, doctoral recipients going into industry are likely to go into establishments with high payroll per worker (above the median for the establishments within their six-digit industry code (see SM)). The recent

Fig. 3. The annual earnings and placement of doctoral recipients supported by grants vary by field. Young firms are defined to be those <5 years old. High-payroll per worker establishments are defined as those with a payroll per worker above the median for the establishments within their six-digit industry. Means and standard errors (error bars) for each variable.
economics literature has emphasized the importance of young firms in contributing to economic growth (17, 18). We examine placements and find that engineers, physicists, and computer scientists are most likely to go to establishments of young firms.

Just as it was possible to show the distribution of average payroll per worker for employers, it is also possible to describe the distribution of individual earnings. One year after leaving the university, doctoral recipients placed in industry had considerably higher earnings than those who went to government or academia, and considerably more doctoral recipients placed in academia had earnings below $50,000 per year (Fig. 4). There is suggestive evidence that at least part of that may be due to doctoral recipients taking postdoctoral research positions; there is a noticeable heaping of the earnings distribution at just under $50,000 for those in biology, chemistry, and health disciplines. However, the earnings growth (albeit from a lower base) was still robust in all sectors 2 years after doctoral recipients left the university, as shown by the distribution of earnings growth.

This work takes a first step toward describing the links between research funding and the economy by tracing the flows of doctoral recipients employed by research grants subsequent to their separation from the university that employed them. The analysis shows that many doctoral recipients who were employed by funded research projects moved into the nonacademic sector and that, when they do, they disproportionately get jobs at establishments with high payroll per worker and in high-tech and professional service industries. Although the results are descriptive and not causal, the findings are consistent with sociological research regarding knowledge flows. A major way in which knowledge is transmitted from research institutions to the economic marketplace is through the placement of people at businesses that draw on that knowledge (10). As research has shown, and as Oppenheimer pointed out, the best way to send knowledge is to wrap it up in a person (19). Higher earnings and placement in large establishments with high payroll per worker, and in establishments owned by firms performing R&D, all characteristics correlated with higher productivity, are consistent with that view (20).

As with any initial examination of new data, there are a number of issues and future extensions to note. There is a need to develop a statistical framework to make causal inferences, which involves identifying appropriate counterfactuals or quasi-experimental variation. The Census Bureau links will make it possible to study both longer-term career trajectories and the characteristics of businesses started by researchers and to compare those with the careers of different types of workers and other business startups in the United States. In addition, the type of analysis we do here for Ph.D. recipients can be extended to study the career outcomes of other groups involved in research—particularly postdoctoral researchers, graduate students who do not get doctoral degrees, undergraduate students, research staff, and people employed by nonfederal projects. As the database expands, researchers can begin to study how outcomes relate to whether funding is federal or nonfederal.

There are important caveats, however. Although extensible to all research universities, the institutions we studied should not be viewed as representative of all academic research institutions. They are large public institutions in the Midwest, and many have large engineering programs and medical schools. Also, the use of U.S.-based administrative and survey data limits the ability to track students who leave the United States. The analysis is explicitly descriptive in nature and is not intended to make any causal assertions. However, because the data will be available at IRIS and the Federal Statistical Research Data Centers, the research community can build on this infrastructure to advance the science of science and to provide policy-makers with research-backed tools to assess the effects of investments in science.

Fig. 4. Annual earnings (U.S.$1 × 1000) and earnings growth of UMERTICS doctoral recipients by sector and discipline. The figure plots the smoothed share of UMERTICS doctoral recipients (the probability density estimated using a Gaussian kernel model) at each level of earnings or earnings growth (with bandwidths of $10,000; $10,000; and 25%, respectively). Individual earnings data are derived from a match to W-2 earnings data.
SOCIAL BEHAVIOR

Sexual fidelity trade-offs promote regulatory variation in the prairie vole brain

Mariam Okhovat, Alejandro Berrio, Gerard Wallace, Alexander G. Ophir,* Steven M. Phelps*

Individual variation in social behavior seems ubiquitous, but we know little about how it relates to brain diversity. Among monogamous prairie voles, levels of vasopressin receptor (encoded by the gene avpr1a) in brain regions related to spatial memory predict male space use and sexual fidelity in the field. We find that trade-offs between the benefits of male fidelity and infidelity are reflected in patterns of territorial intrusion, offspring paternity, avpr1a expression, and the evolutionary fitness of alternative avpr1a alleles. DNA variation at the avpr1a locus includes polymorphisms that reliably predict the epistatic status and neural expression of avpr1a, and patterns of DNA diversity demonstrate that avpr1a regulatory variation has been favored by selection. In prairie voles, trade-offs in the fitness consequences of social behaviors seem to promote neuronal and molecular diversity.

Social behavior emerges from the complex, dynamic, and often strategic interactions of individuals—a complexity that places it among the most challenging and interesting behaviors to study. Neuroscience has elucidated many mechanisms of social behavior (1, 2). In parallel, evolutionary biology has outlined how social interaction can promote variation within a species (3–5). Frequency- or density-dependent selection, for example, maintains individual differences in the parental care of sunfish (3), the territorial defense of lizards (4), and the canibalistic behavior of tadpoles (5). Among humans, similar forces have been proposed to explain differences in personality, resilience, and psychiatric risk (6–8). Given that social diversity is central to behavioral ecology, social psychology, and mental health, it is surprising that we know so little about natural variation in the social brain, how it emerges from the interaction of genetic and epigenetic processes, or how it has been sculpted by evolutionary forces.

We explored individual differences in neuronal gene expression in the monogamous prairie vole, Microtus ochrogaster, a small North American rodent whose males and females form pair bonds and share parental care (9). Prairie vole pair-bonding is governed by multiple modulators and brain regions (2, 10, 11). Of these genes, the vasopressin 1a receptor (V1aR, encoded by avpr1a) is particularly well studied (2, 11–15). V1aR expression can vary profoundly across individual prairie voles (12), and its abundance in a spatial-memory circuit predicts sexual fidelity in males (13, 14) but not females (supplementary materials), a finding consistent with male-specific vasopressin effects in other contexts (15). We used the relationship between avpr1a expression and male fidelity to examine how social forces contribute to brain diversity. Specifically, we asked whether the fitness consequences of male sexual fidelity promote genetic and epigenetic variation in avpr1a.

Although prairie voles are socially monogamous, they are not sexually exclusive (16). Approximately 25% of young are conceived outside a pair bond (termed extra-pair fertilizations, or EPFs). Male fidelity is often thought to depend on spatial strategies that balance the demands of mating-guarding against the value of mating multiply (17, 18). To examine the relationship between space use and sexual fidelity among male prairie voles, we estimated the intensity of a male's space use by fitting kernel density estimates to animal positions measured over several weeks by radiotelemetry (Fig. 1, A and B, and fig. S1). By overlaying these maps of space-use intensity, we could estimate how often males encounter other individuals either at home or in neighboring territories. We found that the spatial behavior of EPF males differs from that of males who sire young only with a partner (intra-pair fertilizations, IPF). EPF males have larger home ranges (P < 0.05; Fig. 1C), and they more frequently encounter extra-pair females (P < 0.0001; Fig. 1D), intrude on territories (P < 0.01; Fig. 1E), and are intruded upon (P < 0.01; Fig. 1F). The rate at which a male intrudes on a neighbor's territory is correlated with the rate at which he encounters extra-pair females [Pearson's correlation coefficient (r) = 0.69, P < 0.0001], but also with the rate at which he is intruded upon by other males (r = 0.83, P < 0.0001; Fig. 1G). Overall, the data suggest that venturing away from a male's core home range increases encounters with both extra-pair females and their aggressive mates; these intrusions may offer the opportunity for extra-pair paternity, but they also increase the rates at

*Department of Integrative Biology, University of Texas at Austin, 1 University Station, Campus Code C0930, Austin, TX 78712, USA. †Department of Psychology, Cornell University, 224 Uris Hall, Ithaca, NY 14853, USA.
*Corresponding author. E-mail: sphelps@mail.utexas.edu

**REFERENCES AND NOTES**


**ACKNOWLEDGMENTS**

This research was supported by NSF SfCSP Awards 1064220 and 1262447; NSF Education and Human Resources Division of Graduate Education Awards 1348691, 1547507, 1348701, 1535399, and 1535370; NSF National Center for Science and Engineering Statistics award 1423706; NIH P01AG039347; 1535399, and 1535370; NSF National Center for Science and Engineering Statistics award 1423706; NIH P01AG039347; and the Ewing Marion Kaufman and Alfred P. Sloan Foundations. Data were generously provided by the Committee on Institutional Cooperation and its member institutions. We thank W. Cheng, C. Connell, C. Jones, E. Kloochman, and J. Toke for research support; G. Carr, M. Harrison, D. Mayo, M. Sweet, J. Van Horn, and S. Willis for help with data issues; and J. Walsh, R. Weiss, and C. Whittacre for their continuing support. The research agenda draws on work with many coauthors, but particularly work with J. Mairesse and M. Pezzoni. Opinions expressed are those of the authors to do not necessarily represent the views of the Census Bureau.

**SUPPLEMENTARY MATERIALS**

www.sciencemag.org/content/350/6266/1367/suppl/DC1

Materials and Methods  
Figs. S1 and S2  
Tables S1 to S8  
References (21–32)  
18 May 2015; accepted 4 November 2015  
10.1126/science.aac5949
Tracking the knowledge economy

Although the U.S investment in scientific research can be documented readily, its output is harder to track. Zolas et al. combined data obtained from eight universities on their doctorate recipients with data from business registries and the U.S. Census Bureau. This allowed them to link Ph.D. recipients to all their subsequent employers. Doctoral recipients tended to stay in academia or join large companies with high salaries. Roughly 20% stayed in the state in which they received their degree. In the year after receiving a Ph.D., mathematicians and computer scientists received the highest salaries, and biologists received the lowest.

*Science*, this issue p. 1367