The (Changing) Knowledge Production Function:
MIT Department of Biology From 1966-2000

Annamaria Conti (annamaria.conti@mgt.gatechedu) and Christopher C. Liu
(chris.liu@rotman.utoronto.ca)

Over the past half-century, the production of science has shifted dramatically. In studies ranging from the size of scientific teams (Jones et al., 2008) and their geographic dispersion (Adams et al., 2005), to the increasing role of technical inputs (Rosenberg, 1982), and the importance of institutions (Furman and Murray, 2008; Jaffe, 1989), numerous scholars have illustrated the changing face of the scientific production process. Yet many aspects of scientific production remain the same. For example, a critical locus of scientific production remains the research university, with professor-led laboratories as the foundational unit (Stephan, 2012). Understanding how these elemental units function, as well as how this function has changed over time, is of critical importance to understanding the role of universities in producing and disseminating innovations (Stephan, 2012; Mansfield et al., 1996)

Broadly, scholars have taken two approaches to the study of laboratories: some have treated laboratories as "black" boxes, linking the variation between laboratories to productive outcomes (e.g., Azoulay et al., 2011). A second stream of research has focused on case studies, trading depth for generalizability (e.g., Jason Owen-Smith, 2001; Shapira, 2009). Despite the central role university laboratories play in the scientific production process, we know strikingly little about their composition, their collaborative structure, and the link between these factors and scientific output. This dearth stems not from a lack of interest, but in the empirical difficulties of collecting micro-level data on university labs.

To examine the internal structure of university laboratories, we have compiled a dataset comprised of laboratories at the MIT Department of Biology, from years 1966 to 2000. Although focused on one university department, this setting has a number of advantages. First, our data source is the department's internal Annual Report, which affords an unparalleled window into laboratory structures over almost four decades (see below). Moreover, mirroring other studies of scientific elites (Zucker et al., 1998; Azoulay et al. 2010), this set of laboratories is important: their scientific discoveries have been critical to the emergence of the biotechnology industry. Lastly, a member of our research team received a doctorate from this department, affording rich insight into technical and organizational aspects of this particular department.

The centerpiece of this dataset is a full list of laboratory members, including their status i.e., graduate student, postdoctoral associate, technician, etc.), and the years when they enter and exit the laboratory. We have supplemented these human capital inputs by compiling all NIH funding within our sample period. Moreover, we have collected all of the publications that arise from these laboratories within a given year.[1] The resulting dataset allows us, for the first time, we believe, to examine a laboratory-level scientific knowledge production function, including both human capital (different classes of laboratory members) as well as financial (NIH grant funding) inputs. Moreover, it allows us to investigate whether this production function has changed over the span of 35 years, a time period that saw the emergence of the biotechnology industry.

Overall, our dataset is comprised of 1,482 laboratory-years, and 20,324 laboratory member-years that span 1966-2000. Within this dataset, there are 119 professors and 6,938 laboratory members, and these unique individuals resulted in 7,553 journal publications. For the

[1] We are in the process of linking the dataset to the NBER patent database.

period 1977-2000, the average laboratory is composed of 5.0 postdocs, 3.3 graduate students 1.4 undergraduates, 2.0 technicians, 0.4 staff scientists, and 0.9 visiting scientists within any given year.[2] Over time, average laboratory size has increased by 41 percent from 1977 to 2000. Much of this increase is driven by an increase in the number of postdocs. An increase in laboratory size is matched by growth in the average number of laboratory publications (rising from 4.3 in 1977 to 5.6 in 2000) and an increase in the average amount of NIH funding, from $267,734 in 1977 to $472,648 in 2000.[3]

Although preliminary, we have chosen to examine the marginal impact of different classes of laboratory members on the publication productivity of these labs. To this scope, we estimate a knowledge production function for laboratory publications using a random effects Poisson model. The inputs we consider are the number of i) postdocs, split into separate categories for those with external (or internal) funding, ii) early graduate students, iii) late-stage graduate students, iv) technicians, v) staff scientists, and vi) grant funding. Our preliminary analysis suggests that adding laboratory members, in descending order of impact, have a positive and statistically significant impact on subsequent paper output: postdocs with external fellowships, internally funded postdocs, graduates students, and grant-funding.

Our next steps involve examination of these trends over time, as well as including additional measures of laboratory output, such as the number of citations per article or the number of journal articles weighted by their impact factor. Additionally, we plan to extend our analysis to the full dataset, incorporating data from the years 1966-1976. Finally, we intend to estimate a dynamic panel model to correct for reverse causality problems.

References

Adams J. D., G. Black, R. Clemmons, and P. Stephan (2005). "Scientific Teams and Institutional Collaborations: Evidence from U.S. Universities, 1981-1999." *Research Policy*, 34: 259-295.

Azoulay P., G. Manso, and J. G. Zivin (2011). "Incentives and Creativity: Evidence from the Academic Life Sciences." *RAND Journal of Economics*, **42**(3): 527-554.

Azoulay P., J. G. Zivin, and J. Wang (2010). "Superstar Extinction." *Quarterly Journal of Economics*, 125(2): 549-589.

J. Furman, F. Murray (2008). "Does Open Access Democratize Innovation? Examining the properties of innovation networks". MIT Sloan Working Paper.

Heinze T., P. Shapira, J. D. Rogers, J. M. Senker (2009). "Organizational and Institutional Influences on Creativity in Scientific Research." *Research Policy*, 38: 610: 623.

Jaffe, A. (1989). "Real Effects of Academic Research." *American Economic Review*, 79(5): 957-970.

Mansfield, E. and J-Y Lee (1996)."The modern university: contributor to industrial innovation and recipient of industrial R&D support." *Research Policy*, 25(7): 1047–1058.

Owen-Smith, J. (2001). "Managing Laboratory Work Through Skepticism: Processes of Evaluation and Control." *American Sociological Review* 66(3): 427-452.

Rosenberg, N. (1982). Inside the Black Box: Technology and Economics. Cambridge University Press, Cambridge, UK.

Stephan, P. (2012). How Economics Shapes Science. Harvard University Press, Cambridge MA.

L.G. Zucker, M.R. Darby, and M.B. Brewer (1998). "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises." *American Economic Review*, 88: 290-306.

[2] We restrict initial analysis to the years 1977-2000 due to changes in personnel titles in the 1976.

[3] Deflated to 1982 US Dollars.

Figure 1: Average Laboratory Composition By Major Categories of Personnel

