**Data Appendix**

**For**

**Show me the Money:**

**Federal R&D Support for Academic Chemistry, 1990-2009**

**Joshua L. Rosenbloom**

**Donna K. Ginther**

All data used in our analysis are available online. The dataset and documentation have have been deposited in the University of Kansas Digital Repository and can be accessed here: <https://doi.org/10.17161/1808.18234>. This appendix provides a brief explanation of the sources of each of the major data elements in our analysis and how they were linked.

Research & Development Expenditures

 These data are derived from the National Science Foundation’s Survey of Research and Development Expenditures at Universities and Colleges/Higher Education Research and Development Survey (<http://webcaspar.nsf.gov>). Data are available annually since 1973 for total and federally funded R&D expenditures by discipline. They are obtained from survey responses completed by institutions of higher education, which are responsible for classifying all research expenditures by discipline. We computed non-federally funded R&D expenditures as the difference between total and federally funded R&D expenditures.

 Sample institutions were selected from the universe of institutions represented in this data by summing real federally funded R&D expenditures (in prices of 2005) for chemistry and chemical engineering between 1990 and 2009 and then ranking institutions in descending order. We initially selected the top 150 institutions but as described in the text were obliged to drop three of these from the analysis because of inconsistencies in coverage. Before adopting this sampling strategy, we examined several other rankings, using total R&D expenditures and using nominal rather than real expenditures. The lists produced in each case were quite similar. The full list of institutions included in the study in declining order of federally-funded chemistry R&D expenditures is provided in Table 1.

 Institutions report these data for the fiscal year corresponding most closely to the federal fiscal year. In most cases this is likely to run from July of one year to June of the following calendar year. Data are labeled with the calendar year in which the fiscal year ends. Hence data for 2009 most likely cover expenditures from July 2008 through June 2009.

 In addition to the expenditures data, this source also contains information on type of control (private or public) and standardized Carnegie Classifications that we use to categorize university types.

Doctorates Awarded and Postdoctoral Researchers

 These data are derived from the National Science Foundation and National Institutes of Health Survey of Graduate Students and Postdoctorates in Science and Engineering (graduate student survey) which is conducted annually by the National Center for Science and Engineering Statistics. The survey is conducted in the fall semester of each academic year and data are collected at the department level. These data are available from <http://webcaspar.nsf.gov>.

 The level of institutional detail provided in this survey is greater than in the R&D expenditure data. In the latter survey a number of multi-campus state systems report a single aggregated number. To link the data sets, we were obliged to aggregate the data in the student survey to match the level of aggregation of the R&D data.

Publications and Citations

 Publication and citation data were computed by Thomson Reuters, Research Analytics from the data underlying the Web of Science publication and citation database. Thomson Reuters subject area experts categorize journals into subject classes based on detailed analysis of the content and focus of the journals. See http://wokinfo.com/media/essay/journal\_selection\_essay-en.pdf for additional details regarding the selection process used by Thomson Reuters in compiling the Web of Science data. The Web of Science is relatively selective about which journals are included, reflecting subject expert judgment and objective metrics of journal impact. Our research began with the full set of journals that Thomson Reuters categorizes as Chemistry and Chemical Engineering. We also conducted an analysis of all journal titles indexed by Thomson Reuters and added a small number of additional journals that contain significant chemistry content.

 We then worked closely with Thomson Reuters staff to match publications by author affiliation to universities in our sample. In addition to institution name, we considered city, state and zip code information associated with authors to verify the accuracy of article linkages.

 After verifying the full list of publications, Thomson Reuters analyzed them to produce summary statistics describing the number of publications each year produced by each institution, the number of citations that those publications received in 3 and 5 year windows beginning with the publication year, and a variety of other citation related metrics.