



National Center for Science and
Engineering Statistics

Survey

Nonprofit Research Activities (NPRA) Survey | 2021

The NPRA Survey collects information on research and experimental development performed by tax-exempt nonprofit organizations in the United States.

Survey Description

Survey Description

Survey Overview (FY 2021 Survey Cycle)

Purpose

The Nonprofit Research Activities (NPRA) module of the Annual Business Survey collects information on research and experimental development (R&D) performed or funded by nonprofit organizations in the United States.

Data collection authority

Title 13, United States Code, Sections 8(b), 131, and 182; Title 42, United States Code, Section 1861-76 (National Science Foundation Act of 1950, as amended); and Section 505 within the America COMPETES Reauthorization Act of 2010, authorize this collection. Sections 224 and 225 of Title 13 require mandatory response. Office of Management and Budget No. 0607-1004.

Major changes to recent survey cycle

Not applicable.

Key Survey Information

Frequency	Annual.
Initial survey year	FY 2020. (The questions in the FY 2020 survey were first developed and used on the FY 2016 NPRA survey, which was the first national survey of R&D activities in the U.S. nonprofit population since 1997. Due to the differences between the previous surveys and the current module, FY 2020 is considered the initial survey year for the annual series.)
Reference period	FY 2021.
Response unit	Organizations.
Sample or census	Sample.
Population size	40,650 nonfarm businesses filing Internal Revenue Service (IRS) tax form 990 as tax-exempt organizations and with payroll of \$500,000 or more were in scope for the nonprofit R&D module.
Sample size	8,050 organizations.
Key variables	Key variables of interest are listed below. <ul style="list-style-type: none"> • R&D expenditures by source of funds (federal government, state and local government, business, foundations, universities, other nonprofits, internal funds, individual donors, and other) • R&D expenditures by type of R&D (basic research, applied research, and experimental development) • R&D expenditures by type of cost (salaries, wages, and fringe benefits; depreciation and amortization; and all other costs)

- R&D expenditures by field of R&D (agricultural sciences and natural resources and conservation; biological, biomedical, and health sciences; engineering; geosciences, atmospheric sciences, and ocean sciences; mathematics, statistics, and computer and information sciences; physical sciences; psychology and social sciences; humanities; and other fields)
- Headcounts and full-time equivalent counts of R&D personnel (researchers, technicians, and other support personnel)
- Headcounts for contract employees and volunteers working on research activities
- R&D funding provided to others, by type of funding (grants and contracts vs. subawards and subcontracts)

Survey Design

Target population

Included are all nonfarm businesses filing Internal Revenue Service (IRS) tax forms as nonprofit organizations with an annual payroll of \$500,000 or more.

Sampling frame

The sampling frame was constructed from the final 2020 Business Register and the Exempt Organizations Business Master File Extract (EO BMF). The Business Register is the Census Bureau's comprehensive database of U.S. businesses. Organizations were excluded from the frame if they were outside the scope of the survey (e.g., churches, government organizations, educational institutions, or organizations located outside the United States). A financial threshold was also imposed to increase the efficiency of reaching organizations that perform research.

Sample design

The nonprofit R&D frame is stratified by state and primary nonprofit activity (hospitals, other health care, science and technology, and all others), and is systematically sampled within each stratum. A standard type of estimation for stratified systematic sampling is used. Certainty cases have a selection probability of one and a sampling weight of one and represent only themselves. Specifically, firms were selected with certainty based on the following criteria:

- firms with known R&D activity (950 organizations)
- firms larger than stratum-specific payroll thresholds (1,200 organizations).

The nonprofit R&D sample consisted of 8,050 organizations; 2,150 were selected with certainty.

The remaining 5,900 noncertainty cases were selected using the stratified systematic random sample selection described above. The maximum sample weight was 6.9.

Data Collection and Processing

Data collection

The survey was mailed to 8,050 nonprofit organizations in July 2022. Organizations were sent a letter informing them of their requirement to report. The letter also provided instructions on how to access the survey and submit online. There were three mail follow-ups and four separate e-mail follow-ups conducted to increase response. The collection period closed 30 December 2022.

Data processing

Prior to tabulating the data, response data were reviewed and edited to correct reporting errors. R&D data were tabulated for records reporting \$50,000 or more in R&D expenditures. Survey analysts reviewed the R&D reported by the survey respondents. Research was done by evaluating the reported R&D to expenses ratio, and organization website information.

Additional data errors were detected and corrected using an automated data edit system designed to review the data for reasonableness and consistency. The editing process interactively performed corrections by using standard procedures to fix detectable errors. Quality control techniques were used to verify that operating procedures were carried out as specified.

Estimation techniques

Where possible, missing data were imputed using previous survey data or other publicly available documents, such as annual reports and financial statements. Weights were used to compensate for unequal probabilities of selection; unit nonresponse; and to calibrate sample estimates of expenses to match total expenses on the frame. Measures of sampling variability were estimated using the delete-a-group jackknife variance estimator.

Survey Quality Measures

Sampling error

Sampling error is the difference between estimates obtained from the sample and results theoretically obtainable from a comparable complete enumeration of the sampling frame. This error results because only a subset of the sampling frame is measured in a sample survey. For published estimates from NPRA, standard errors are produced for estimated percentages, while relative standard errors (RSEs) are produced for all other estimates. Tables of the estimated measures of sampling variability corresponding to each data table are available upon request.

Coverage error

Coverage error occurs when the frame fails to completely enumerate the population of interest. There can be both undercoverage error, where units are not included in the frame, and overcoverage error, where units included in the frame are out of scope for the population of interest. The NPRA module uses the prior year Business Register to construct the frame so any changes in businesses that would change the inclusion or exclusion of the business to the survey scope could be sources of coverage error. Prior to tabulation, survey unit information is updated with the most recent available Business Register data to mitigate this source of error.

Nonresponse error

Unit nonresponse is treated by adjusting weighted reported and imputed data by multiplying each organization's sampling weight by a nonresponse adjustment factor. Detailed descriptions of the adjustments for nonresponse are available in the Technical Notes.

Measurement error

The most common source of measurement error was reporting in different monetary units (for example, reporting whole dollars rather than thousands of dollars). This was corrected during data processing. Another source of error involved incorrect inclusion of organizations already represented in other R&D data collections. The R&D of these respondents was set to 0 where it was determined their R&D was already represented in other survey responses. These cases included nonprofit organizations managing federal laboratories and some university-affiliated hospitals.

Data Availability and Comparability

Data availability

Data are available at <https://nces.nsf.gov/surveys/nonprofit-research-activities/>.

Data comparability

The questions in the FY 2020 module were first developed and used on the FY 2016 Nonprofit Research Activities Survey—which was the first national survey of R&D activities in the U.S. nonprofit population since 1997. Due to the differences between the previous surveys and the current module, data are not comparable for trend analysis.

Data Products

Publications

NPRA data will be published in NCSES InfoBriefs and data tables available at <https://nces.nsf.gov/surveys/nonprofit-research-activities/>.

Electronic access

The NPRA module contains confidential data that are protected under Title 13 and Title 26 of the United States Code. Two types of data are currently available: public-use tabular statistics and restricted microdata. Public-use tabular statistics can be obtained on the NCSES website (<https://nces.nsf.gov/>) and by contacting NCSES. Restricted microdata will be available at any of the 15 secure Research Data Centers administered by the Center for Economic Studies (CES) at the Census Bureau. Researchers interested in accessing microdata can apply for a restricted-use license by submitting a proposal to the CES, which evaluates proposals based on their benefit to the Census Bureau, scientific merit, feasibility, and risk of disclosure. To learn more about the Research Data Centers and how to apply, please visit the [CES page on research with restricted-use data](#). For additional information about the application process, including how to initiate a project, please contact the administrator at the primary site where the research will be conducted. Per the Federal Cybersecurity Enhancement Act of 2015, the data are protected from cybersecurity risks through screening of the systems that transmit the data.