

TABLE 1

Features of UMETRICS and the SED

(Responses)

Key features	Administrative data: UMETRICS	Survey data: SED
Data collection purpose	Operational	Mandated to respond to specific requests
Units of analysis	Multiple: individual, research team, networks, and collectivities (e.g., demographic groups, fields, programs, departments)	Individuals and collectivities (e.g., demographic groups, fields)
Coverage	All individuals paid by research grants at participating universities	All graduating doctorate recipients
Temporality	Fully longitudinal, granular to the level of months	Cross-sectional, collected at the point of graduation, retrospective and prospective questions subject to known forms of response bias
Content	Can be constructed from source data	Restricted to questions asked on survey
Major processing challenge	After data collection: data standardization and normalization	Before data collection: design and analysis

AY = academic year; SED = Survey of Earned Doctorates; UMETRICS = Universities: Measuring the Impacts of Research on Innovation, Competitiveness, and Science.

Source(s):

National Center for Science and Engineering Statistics, Survey of Earned Doctorates, AYs 2014–17; Institute for Research on Innovation and Science, UMETRICS, 2020 release data.