

International survey projects—such as the World Values Survey, European Social Survey, International Social Survey Programme, Eurobarometer together with its worldwide renditions, and many others—constitute the primary empirical foundation for understanding political attitudes and behaviors across countries and over time. These projects generate the datasets that scholars routinely rely upon to study trust in government, satisfaction with democracy, political participation, and a wide range of related topics. However, although these datasets are widely used in social science research, they contain substantial and systematic differences in data quality. These differences arise from multiple sources, including variation in sampling procedures, translation and adaptation of questionnaires, the extent and rigor of fieldwork controls, and the integrity of computer data files. In practice, surveys conducted in different countries—and even within the same project across different years—may vary significantly in how well they adhere to methodological standards.

In recent years, these problems have intensified. The methodological disruptions associated with the COVID-19 pandemic, including sudden shifts to mixed-mode and remote interviewing, staffing shortages, and reduced capacity for fieldwork monitoring, have further increased cross-national variability in survey quality. Despite the apparent importance of these issues, the effects of data quality problems on substantive research findings remain insufficiently understood across different years and may vary widely. One key reason is that the relevant metadata—information about how surveys were conducted and how reliable they are—is typically scattered, incomplete, or difficult for researchers to access. Another reason is the absence of well-established analytical procedures for incorporating survey quality information into statistical models. As a result, most cross-national research proceeds under the implicit assumption that all surveys are equally reliable, an assumption that is rarely justified.

This project addresses this critical gap by developing a comprehensive methodological framework for measuring, documenting, and integrating survey quality into analyses of political attitudes and behaviors. Drawing on the methodology tradition (Total Survey Error) and survey harmonization (Survey Data Recycling), the project will construct an extensive set of indicators that capture three core dimensions of survey quality: (1) the design and execution of the study, including sampling procedures and fieldwork controls; (2) the consistency between survey documentation and the data actually recorded; and (3) the internal quality of computer data files, including the presence of duplicates, missing values, and other anomalies. These indicators will be applied to two large harmonized datasets: Data 1 (2013–2017), created in earlier work, and Data 2 (2022–2025), which will be constructed within the project following comparable methodological principles.

The project will test a series of hypotheses about the structure and evolution of survey quality, cross-national and cross-project differences, and the extent to which data quality affects key political outcomes. Statistical analyses will combine multilevel modeling with innovative procedures—such as partialling out the influence of survey errors and weighting surveys by their quality—to evaluate how data quality shapes empirical conclusions.

The expected outcomes include a publicly accessible database of survey quality indicators, a new harmonized dataset released in the Harvard Dataverse, a user guide and interactive interface to support researchers, high-impact publications, and an international workshop. By demonstrating when, how, and to what extent survey quality influences analyses of political attitudes and behaviors, the project will significantly enhance the reliability of cross-national research and establish new standards for transparent and informed use of survey data.