# Survey Costs: Summary and Commentary from the 2026 Survey Costs Workshop

Kristen Olson, University of Nebraska-Lincoln
James Wagner, University of Michigan
Jill DeMatteis, Fors Marsh
John Eltinge, Retired, U.S. Census Bureau
Daifeng Han, Westat
Chris Jackson, SSRS
Eric Rancourt, Statistics Canada

July 2026

**Acknowledgements and Disclaimer:** The authors thank Ryan Green at AAPOR and the Westat Marketing team for overseeing the logistics of the Survey Costs Workshop. The views expressed in this white paper are those of the authors, and do not represent the opinions or policies of the organizations with which the authors are affiliated and do not necessarily reflect the views of the AAPOR Executive Council. The authors, who retain the copyright to this report, grant AAPOR a non-exclusive perpetual license to the version on the AAPOR website and the right to link to any published versions.

# Table of Contents

## Executive Summary

Survey costs are one of the most important constraints on the survey design options available to survey researchers and practitioners. Decisions on different survey design features, data collection modes, administrative data systems, estimation approaches, and reporting methods all depend on the cost of the options and the available budgets. In a call to action to the survey community, a special topics Workshop on Survey Costs was held at the Westat office space in Bethesda, Maryland on February 9-10, 2026. The workshop was generously sponsored by the American Association for Public Opinion Research, the American Statistical Association, Westat, SSRS, Miravoice, and the International Association of Survey Statisticians. At this workshop, approximately 100 survey research professionals, methodologists, statisticians, field operations specialists, and finance and accounting specialists convened for two days of presentations and conversations. The first day consisted of an opening talk on definitions, followed by a plenary panel, and two sets of presentations to the full group. The second day consisted of two sets of four concurrent presentation sessions, plus four concurrent end-of-workshop summary and brainstorming sessions.

This white paper summarizes the main takeaways from the presentations and discussions at the workshop, along with the Organizing Committee's recommendations for future actions on survey costs. The ten takeaway themes are listed below. The report provides details on each of these themes.

> Takeaway #1: Cost reporting should be transparent about what kinds of costs (e.g., fixed or variable; monetary or nonmonetary) are being reported, the data sources for the costs, the components included in the costs, how inputs for calculating total and relative costs are defined, and what assumptions are being made about the contributions of fixed, infrastructure, and variable costs.
>
> Takeaway #2: Consistent definitions for different cost measures, permitting comparisons across studies and design features, are clearly needed.
>
> Takeaway #3: Survey costs are important factors in all aspects of the survey design and production lifecycle. Survey costs extend well beyond data collection costs, reflecting a full statistical production system. These aspects include planning for new data collection, project budget development, evaluating and adapting historical cost data for future studies, survey design decisions, survey monitoring, and data processing and reporting.
>
> Takeaway #4: Cost monitoring is important and methods to do so vary tremendously across organizations. Cost monitoring is complicated by different practices and administrative systems holding different cost information, costs being recorded at varying

levels of detail, costs being categorized differently across organizations, and some costs being difficult to fully enumerate.

Takeaway #5: Survey design features affect survey costs and survey response rates and other proxy error indicators. How these design features affect cost or error outcomes varies tremendously across studies. Some of this variation arises because of different definitions for costs or different error indicators. Design factors, target populations, and implementation decisions interact with each other and are shaped by environmental and contextual factors (e.g., time of year; country, region, or state; societal, economic, or political influences on study outcomes), all affecting survey costs. How design features affect specific types of costs across studies, populations, and implementation decisions requires more research.

Takeaway #6: Stating specific statistical models linking design features and environmental and contextual factors to both survey costs and survey errors is a necessary next step in research on survey costs.

Takeaway #7: Observed cost data are likely to have measurement errors, which may be biasing errors that consistently over- or underestimate actual costs, or variable errors that vary in the direction of errors across studies or implementations. Understanding the cost data generating process, including the level at which the cost information is recorded and stored, will be fruitful in understanding whether these errors make the cost information fit for its specified use.

Takeaway #8: Although being aware of the potential for measurement error in observed cost data is important, one should not wait for perfectly measured cost information to begin the process of cost evaluation, estimation, or modeling. Cost models may be reexamined as improved cost data become available. In the meantime, sensitivity analyses can be used to evaluate the robustness of findings to plausible assumptions about measurement error.

Takeaway #9: Many challenges exist in understanding and predicting costs, especially across subgroups and data collection stages. There are potential limitations to existing cost data that can limit its utility. Uncertainty should be acknowledged and incorporated into estimates wherever possible. Cost drivers, especially related to statistical modeling expertise, that vary by type of data source, probability or nonprobability samples, or study complexity, are not well understood.

Takeaway #10: The value of surveys or a statistical production system extends beyond the cost of the data collected. The value of survey research should be viewed in terms of both its immediate utility and its broader and longer-term societal benefits.

From these ten themes, the Organizing Committee makes the following 12 recommendations. These recommendations are aimed at researchers within survey organizations to cultivate more internal research and understanding of survey costs. To the extent that this internal research can be shared broadly without disclosing proprietary cost information, these recommendations also provide directions for future research on survey costs.

- Develop and use standardized definitions and measures.
- Connect cost measurement structures with management decisions.
- Justify decisions on components included in evaluation of aggregate costs.
- Justify decisions on quality measures used in evaluation of cost-quality trade-offs.
- Evaluate cost-quality trade-offs through decisions on specific design factors.
- Account for environmental factors.
- Quantify uncertainty in cost analyses.
- Despite imperfections, it is necessary to use available information on costs to design and improve surveys.
- Consider extensions to additional components of the collection, production, dissemination and use of statistical information.
- Develop tools to incentivize managers to examine costs and reduce management burden for examining costs.
- Examine related empirical research and meta-analyses of cost information.
- Engage in transparent and actionable communication with multiple stakeholders for cost analyses.

## 1. Overview and Goals of the White Paper

Survey costs are one of the most important constraints on the survey design options available to survey researchers and practitioners. Decisions on different survey design features, data collection modes, administrative data systems, estimation approaches, and reporting methods all depend on the cost of the options and the available budgets. Yet virtually all existing research on these issues falls under the umbrella of research on survey errors, largely under a Total Survey Error Framework (Groves 1989), and rarely addresses costs. When cost is included, it is as a principle or dimension of quality, or a goal by which to measure effectiveness, but generally without a discussion on the costs themselves (Eurostat 2021). Recent calls for more research into survey costs highlight their importance (Olson 2021; Olson et al. 2020), but also identify myriad issues in our current understanding of costs. In a call to action to the survey community, a special topics Workshop on Survey Costs was held at the Westat office space in Bethesda, Maryland on February 9-10, 2026. The workshop was generously sponsored by the American Association for Public Opinion Research, the American Statistical Association, Westat, SSRS, Miravoice, and the International Association of Survey Statisticians. At this workshop, approximately 100 survey research professionals, methodologists, statisticians, field operations specialists, and finance and accounting specialists convened for two days of presentations and conversations. The first day consisted of an opening talk on definitions, followed by a plenary panel, and two sets of presentations to the full group. The second day consisted of two sets of four concurrent presentation sessions, plus four concurrent end-of-workshop summary and brainstorming sessions.

The primary goal of the workshop was to facilitate conversations on survey costs across organizations and roles, with the longer-term goal of starting a research agenda on survey costs. Secondary goals of these conversations and presentations were to identify common threads across organizations in cost budgeting, monitoring, and evaluation; to understand what drives variation in survey costs across design features; and to present research that links costs and survey errors. While working to understand the conceptual, methodological and empirical takeaways that could be transportable across surveys, researchers, and organizations, the aim was to avoid revealing proprietary cost structures and details.

Taken as a whole, the workshop presentations and related group discussions indicated that different survey components (e.g. modes, interviewers, mailing, instrument programming, management) have cost structures with different levels of development and specificity across studies and organizations. In some cases, presentations focused on broad qualitative descriptions of cost measures, which can lead to consensus regarding certain core concepts, terms and definitions. Other presentations translated those concepts, terms and definitions into realistic methods for measurement of costs, or related approximations. Still other presentations modelled costs as functions of specified design and environmental factors or evaluated various cost-quality trade-offs. Different parts of the workshop offered insights into each of these elements of cost description, measurement, modeling, and evaluation.

This white paper summarizes the main conversations from the two days of the workshop, along with suggestions from the Organizing Committee and workshop participants on next steps for future research on understanding survey costs. This white paper does not review the full existing literature on survey costs, largely focusing on the presentations and discussions at the Survey Costs workshop.

## 2. Terms and Definitions

The Workshop opened with a discussion of cost-related definitions and terms to facilitate conversations across participants and encourage a common vocabulary (Olson 2026). Previous work (Olson et al. 2020) had identified the ***type of cost information*** as either measures that are ***monetary***, measured in units of currency, or measures that are ***nonmonetary***, typically obtained from a measure of field effort (e.g., number of call attempts; number of mailings) or employee effort (e.g., staff full-time equivalent (FTE), interviewer hours; mileage traveled).

The measurement scale and unit also varies across studies. Costs can be measured and reported in aggregate as a ***total overall*** or as a ***relative*** measure. Relative costs can be the costs per sampled unit, cost per complete (also called a cost per respondent or cost per interview), cost relative to a previous year's data collection effort, cost relative to an expected baseline, or the percentage of costs that are attributable to a particular design feature or other component of the study.

The ***data sources*** for cost information also vary. ***Estimated costs,*** costs that come from budgets prior to data collection that are assumed for design purposes, may be different from ***observed costs,*** the costs that are available for analysis from accounting, payroll, paradata, or other recordkeeping systems. Both estimated costs and observed costs may be different from the ***actual costs,*** or the costs that are actually incurred while designing, conducting, and processing a survey. Myriad reasons exist for discrepancies between actual and observed costs. Some organizations may have accounting systems that do not allow employees to disentangle work on different projects (e.g., a field interviewer who contacts multiple housing units in adjacent neighborhoods for different studies on the same day) or different parts of projects (e.g., employees charge total hours to a project, and the accounting format does not distinguish among time spent on different parts of a study design, data collection, or analysis). Some organizations may have employees who regularly work more than a 40-hour work week, but these employees are not permitted to charge more than 40 hours per week to studies. Other organizations may inform their employees about how much time can or should be charged to any individual part of a survey project (e.g., up to 20 hours for programming a web questionnaire) and time above that cannot be allocated to that function, even if the task takes longer than the budgeted time. In some cases, cost information that is collected may not be shared across different parts of the organization, limiting its utility for informing study designs and necessitating the use of estimated costs, even if observed or actual costs are available somewhere in an organization. Estimated, observed, and actual costs may reflect both monetary and nonmonetary measures. For instance, monetary measures come from budgets, accounting, and financial systems whereas nonmonetary measures come from paradata, payroll systems, or travel systems.

Costs also can vary because ***different design features are included***. Costs may reflect a single design feature, which Olson, Wagner, and Anderson refer to as a ***survey component*** (Olson et al. 2020). Survey components include individual features such as printing, postage, data entry, incentives, interviewer hours for travel, project manager or survey methodologist hours on questionnaire design, or sampling statistician hours for sample design. Costs may also reflect an aggregate of design features, combining across multiple survey components. Sometimes researchers report what features go into the aggregate measure (e.g., sum of printing, postage, and incentive costs); others simply report that the costs are for "data collection."

Survey components can be divided into three general categories related to infrastructure costs, fixed costs, and variable costs. The initial introduction drew on past work from the context of self-administered web and mail surveys (Olson et al. 2024), but workshop participants regularly applied these ideas to their own work and contexts. ***Infrastructure costs*** are costs that survey organizations incur to support the organization and study designs and are not study-specific costs. Infrastructure costs, including software and equipment costs, non-project personnel costs, or building maintenance costs, can be amortized across studies or paid for through other means. ***Fixed costs*** are study-specific, but do not (directly) scale with increased data collection efforts or sample size. Fixed costs are typically personnel costs at a survey organization, but can also be software, equipment, or other supplies purchased for a specific study. ***Variable costs*** are study-specific and do scale to some degree with increased data collection efforts and/or with increased sample size. Variable costs can include personnel costs, travel costs, supply costs, and incentive costs, and may be set by the survey organization or by external actors (e.g., postage prices).

All of these measures, from all of these data sources, for both components and totals, as well as infrastructure, fixed, and variable costs were reported during the workshop (see Table 1 below for examples). Over the workshop, participants reported on both monetary costs and, more often, nonmonetary costs. Across the workshop presentations, relative costs were frequently reported, with total costs reported by fewer presenters. Presentations reflected observed, actual, and estimated costs, and sometimes multiple types of costs were reported in the same presentation (e.g., both observed and estimated costs). Workshop participants more often reported on some individual survey components, but sometimes reported aggregated costs. Much of the workshop focused on which type of costs was being reported, with some participants discussing fixed or infrastructure costs, and most discussing variable costs. Notably, Kronschnabl (2026) analyzed contracts across 28 countries in the Survey of Health, Ageing and Retirement in Europe (SHARE) as to what is categorized in fixed versus variable costs, finding wide disagreement across vendors in what components are identified as fixed components versus variable components, even within the same survey system and the same contracting mechanisms. There was also discussion among participants about what proportion of a total survey budget was allocated for variable costs versus fixed or infrastructure costs, although empirical data on this was limited.

**Table 1. Examples of Presentations from Survey Costs Workshop Illustrating Various Cost Measurements**

| Cost Dimension | Types Reported | Examples from Workshop Presentations |
|---|---|---|
| **Cost Information Type** | Monetary costs | (Brenzel 2026; Ferrara et al. 2026; Gentry et al. 2026; Guyer et al. 2026; Olson 2026; Reardon 2026; ZuWallack et al. 2026) |
| | Nonmonetary costs | (Carle 2026; Chardoul, Chang, et al. 2026; Guyer et al. 2026; Lancaster 2026; Lewis et al. 2026; Olson 2026; Owens 2026; Sjoblom et al. 2026; Tolliver 2026; Ubartas and Kuseler 2026; Wagner 2026) |
| **Cost Metric** | Relative costs | (e.g., Amaya et al. 2026; Bertoni 2026; Boulet 2026; DeMatteis and Cook 2026; McRoy 2026; Murphy 2026; Sjoblom et al. 2026) |
| | Total costs | (e.g., Ferrara et al. 2026; Kaye 2026; Murphy 2026; Reardon 2026) |
| **Cost Basis** | Observed, actual, or estimated costs | (e.g., Carle 2026; Chardoul, Chang, et al. 2026; Kaye 2026; Keppenne et al. 2026; Murphy 2026; Reardon 2026; Sjoblom et al. 2026; Stepler 2026; Tolliver 2026; Wagner 2026; ZuWallack et al. 2026) |
| | Multiple cost bases in one presentation | Some presentations reported both observed and estimated costs. It was sometimes difficult to identify these cost sources from the slides alone. |
| **Cost Scope** | Individual survey component costs | (Bertoni 2026; Brenzel 2026; Endres 2026; Gentry et al. 2026; Guyer et al. 2026, 2026; Kaye 2026; Stepler 2026) |
| | Aggregated costs | (e.g., Amaya et al. 2026; Ferrara et al. 2026; Kotch et al. 2026) |
| **Cost Structure** | Fixed or infrastructure costs | (e.g., Kaye 2026; Miladi and Holmes 2026; Wong 2026) |
| | Variable costs | (e.g., Carle 2026; Chardoul, Chang, et al. 2026; Lancaster 2026; Murphy 2026; Sjoblom et al. 2026) |

Recognizing this heterogeneity in cost measures and what was included when calculating costs, participants notably called for some sort of standardization and transparency in reporting. A common comment throughout the workshop by participants was the need to have widely shared definitions of terms, including what constitutes variable, fixed, and infrastructure costs. One easier place to start is for organizations and researchers to clearly define how they are using these terms, including specific details about what is included or excluded from reported costs, whether monetary or nonmonetary. Additionally, participants often commented that fixed costs can make up a large proportion of a survey budget, but most of the research and monitoring systems presented at the workshop examined only variable costs related to data collection. Monitoring and reporting on fixed costs would create the empirical basis for evaluating their importance in a survey budget, as well as how to navigate tradeoffs between fixed and variable costs across studies.

***Takeaway #1: Cost reporting should be transparent about what kinds of costs (e.g., fixed or variable; monetary or nonmonetary) are being reported, how all inputs for calculating total and relative costs are defined, including the data sources for the costs, the components included in the costs, and what assumptions are being made about the contributions of fixed, infrastructure, and variable costs.***

***Takeaway #2: Consistent definitions for different cost measures, permitting comparisons across studies and design features, are clearly needed.***

## 3. Summary from the Workshop Presentations

### 3.1 Uses of Cost Information

Workshop participants described multiple uses for survey cost information. Survey cost information is used in practice to support planning for new data collection systems, budgeting for specific surveys, and making individual survey design decisions, as well as for required billing of survey clients and institutional accounting purposes. Metrics such as total costs, costs per sampled unit, and cost per complete may be used as the basis of cost estimates for these purposes (Olson 2026). Eltinge (2026a) described survey costs as existing within a statistical production ecosystem. In this systemic view of costs, measures that are traditionally used for monitoring data collection (e.g., cost per interview) do not reflect the wide berth of cost drivers across a statistical production lifecycle, from design to production, including both survey data and nonsurvey data. Edwards (2026) made similar points when developing new cost tracking and monitoring systems.

Costs were used when **planning new large-scale data collection systems**. Wong (2026) described planning for a new cohort study using an "open-book accounting" system in which multiple vendors were contracted to collect data for a new birth cohort, and agreed to share monetary (e.g., cost per household) and nonmonetary (e.g., interview length; sample released to the field) cost measures related to data collection with the client and with each other, allowing the client to better manage costs across vendors and across the data collection process. Others described using or building survey infrastructure, such as probability-based panels, to support needs across multiple studies or for federal agencies over time (Baer 2026; Jablonski and van Exel 2026). Here, the aim was to develop cost-effective methods of collecting data from a population over time and across multiple needs.

At the **project budget development** stage, cost data from prior administrations of the survey or from other projects with aspects similar to the project being budgeted were used, adjusting as necessary, to estimate the survey budget. Information for budgets may come from past administrative data, information from project team members, or other sources. For example, Keppenne et al. (2026) described an in-house budgeting tool, developed in Excel, that uses assumptions of costs for supplies, materials, and infrastructure, data collection efforts, and staff costs developed from past studies. This approach allowed any member of the survey organization to have a consistent set of baseline cost assumptions when developing a survey budget. Jackson et al. (2026) described a process where existing cost information and assumptions by different independent teams within a large organization were used to develop an initial overall budget for a survey project. The entire project team then reviewed the project budget holistically, as well as its underlying assumptions, with further discussion leading to possible changes in the budget. This process paired historical information with real-time knowledge from relevant study team members to provide a workable project budget estimate.

Some presentations highlighted **challenges in using historical cost data to budget current and future studies in the presence of uncertainties**, including inflation, increasing rates of nonresponse, and project scope uncertainties. In developing a large new cohort study with multiple data collectors, Wong (2026) described a variety of issues related to budgeting costs for current and future waves of data collection that come with large-scale uncertainties. To address this, the study plans on setting "cost tolerances" for some cost drivers in the form of an interval rather than a single budgeted figure to account for uncertainty in some of the underlying cost assumptions, along with strategic management of contingency funds (Wong 2026). Other uncertainties can arise due to study design, target population, methodological complexity, or allotted time for data collection. Some organizations accounted for this explicitly in budgeting. For instance, Miladi and Holmes (2026) described three factors – a "Technical Complexity Factor" (related to the needs of project management, programming, and reporting staff), a "Field Difficulty Factor" (related to challenges in data collection, and a "Timeline Pressure Factor" (related to how quickly data collection needed to be completed) – that are tied to multipliers in a budget, providing a cost link for uncertainties, capacity constraints, and difficulties when developing a budget for a desired survey project. Kronschnabl (2026) discussed the challenges in developing consistent budgets and outcomes in a multinational longitudinal survey of aging. To reduce variability, the project used a standard form for all potential vendors, allowing the contracting officer to more readily compare why costs varied across contractors, including allocation of costs to different data collection tasks or across fixed or variable costs. It also used (through the eighth wave) financial incentives and disincentives for fieldwork companies to align their practices with the study's design and fielding requirements as a way of managing uncertainty in implementation procedures.

At the **survey design stage**, cost information was used to inform design features including choice of mode(s), frame, interviewer training, sample allocation, adaptive design, and stopping rules (Brown and Jones 2026; Eisenhauer and Han 2026; Guyer et al. 2026; Kotch et al. 2026; McRoy 2026; Reardon 2026; Stepler 2026; Wagner 2026; ZuWallack et al. 2026) (also see Section 3.3 for detailed design feature discussions). Although many presentations on modes of contact and survey administration focused on the variable costs directly associated with data collection (e.g., printing and postage, interviewer pay), others described accounting for fixed costs and infrastructure costs at the planning and design stage (Brown and Jones 2026; Edwards 2026; Eltinge 2026a). These fixed and infrastructure costs include the costs of developing and maintaining survey management systems, increased labor hours for data collection managers, costs associated with recruiting and training interviewers, and the cost of establishing and maintaining contracts with vendors. For instance, Brown and Jones (2026) describe how the costs for field interviewing during a pretest for the NLS-72 extended well beyond interviewing time, and included extensive administrative time, logistics for managing biospecimens and equipment, communicating with longitudinal respondents, training, and monitoring. These lessons on the fixed costs of the planned design led to a substantial reimagining of the design itself, including changing the mode of data collection for some cases to phone instead of face-to-face, allowing some costs to shift from fixed and infrastructure costs to variable costs of data collection. DeMatteis and Cook (2026) compared alternative sample designs with primary sampling units (PSUs) defined at the county versus tract level, where the choice of PSU definition was guided by the cost of using local versus traveling interviewers. Eisenhauer and Han (2026) also used costs to inform sample design, especially when costs vary across important subgroups or strata, including geographic areas and demographics.

Costs were also used at the **monitoring stage**. Edwards (2026) described building a comprehensive cost monitoring system, with inputs from various data sources including paradata, payroll, information from accounting systems, and more. Cost monitoring systems can measure costs at a project level, over time, for a specific sampled case, for areas of the country, for modes of data collection, and for organizational units, for example, but should serve the needs of the people using the system (see discussion of dashboards in Section 3.2 below). The cost monitoring systems described by Edwards are most easily implemented at large organizations, as these systems require substantial IT support to implement. In contrast, Shuttles (2026) described a cost monitoring system that could be implemented by smaller organizations. Shuttles recommended developing a plan for cost monitoring before data collection so that researchers could evaluate costs as they are incurred and potentially mitigate cost-related problems before they happen. This approach, implemented using Excel spreadsheets, compared a small number of monetary costs in specific expenditure categories to budgeted allocations throughout the field period (e.g., 50% of the field period has elapsed, but 60% of the budget has been spent; is this expected?). Cost monitoring often involves the use of paradata. For example, Chardoul et al. (2026) discussed efficiency metrics based on operational cost drivers such as response rate, interview length, household visits per trip, and proportion of hours in travel or training versus real production.

Costs were also discussed as informing the **data processing and reporting stage**. This aspect was highlighted by the opening panel (Chardoul, Coffey, et al. 2026) in their discussion surrounding mixed-mode surveys, which have increased costs for data processing, weighting, and other post-data collection steps. Additionally, Eltinge (2026a) discussed costs related to data cleaning, record linkage, and production of subgroup tabulations (after the production of the primary tabulations) in the context of official statistics as specific costs that are not "data collection" costs but are critical to the production of statistical estimates. Brenzel (2026) provided a comprehensive assessment of costs for immunization surveys in low and middle-income countries, including costs allocated for reporting and technical assistance. Kaye (2026) found that data management, preparation, and processing cost significantly more for nonprobability panel and surveys conducted with an open link rather than a specified list.

The past decade has seen significant change and disruption in survey administration driven by the pandemic, artificial intelligence (AI), and shifting government priorities. These factors have made survey costs at each of these stages more difficult to predict and underscore the importance of explicitly accounting for risk in cost modeling (Edwards 2026). Several presentations emphasized the importance of engaging a broad range of stakeholders in an open and timely manner when using cost information or developing cost models (Eltinge 2026a; Wong 2026).

> ***Takeaway #3: Survey costs are important factors in all aspects of the survey production lifecycle. Survey costs extend well beyond data collection costs, reflecting a full statistical production system. These aspects include planning for new data collection, project budget development, evaluating and adapting historical cost data for future studies, survey design decisions, survey monitoring, and data processing and reporting.***

### 3.2 Cost Budgeting and Accounting Systems and Software

Several presentations focused on **cost accounting systems and software**. Off-the-shelf solutions did not meet all the needs of survey organizations. Rather, many organizations had developed at least some adaptations or extensions of systems. When building in-house cost accounting and monitoring systems, some workshop participants discussed designing a system to meet the needs of most of the organization's projects (Edwards 2026; Keppenne et al. 2026; Miladi and Holmes 2026; Murphy 2026; Shuttles 2026), while others described project-specific approaches (Chardoul, Chang, et al. 2026) or mode-specific approaches (Carle 2026; Sjoblom et al. 2026).

Some organizations showed effective solutions that used readily-available technology (e.g. spreadsheet software such as Excel) with some user-defined functions to enable preparing cost estimates (Keppenne et al. 2026). Others relied on (or were planning for) relatively advanced software solutions that incorporated AI and simulated potential outcomes (Miladi and Holmes 2026; Murphy 2026). A difficult question for organizations to answer was how much time and money to spend on the development of these tools. Workshop participants thought it was difficult to identify an "optimal" solution to how much investment should be put into cost accounting and monitoring systems. Those participants who lacked cost information at their organizations or for individual studies commented that more effort spent in developing systems that yielded readily available cost data could produce dividends. For instance, Stepler (2026) compared two incentive experiments – one where detailed cost information was available and one where cost information was extremely limited, such that decisions based on cost information from a pilot study for the full study were difficult to make.

Other organizations worked on **cost monitoring,** a function that is sometimes integrated into **dashboards**. Dashboard tools can require significant development time and resources, and are more likely to be developed at large survey organizations (Carle 2026; Chardoul, Chang, et al. 2026; Murphy 2026; Sjoblom et al. 2026). These dashboards monitor both nonmonetary and monetary costs at the project level. Dashboards that are targeted at monitoring field interviewers may also allow field managers to monitor costs at an individual interviewer level, a sample line level, or an interviewer-workload level. Some organizations built dashboards using existing statistical software packages (e.g., R Shiny, (Murphy 2026)), whereas others developed homegrown systems that drew on multiple data sources and analysis software (Carle 2026; Chardoul, Chang, et al. 2026; Sjoblom et al. 2026). On the other hand, relatively simple monitoring tools also were built using spreadsheet software (Shuttles 2026), requiring less up-front investment from a survey organization.

Many organizations noted that **cost information was stored in systems that are separated from sample management systems** (Boulet 2026; Carle 2026; Stepler 2026; Tolliver 2026; Ubartas and Kuseler 2026). In order to combine cost (hours and dollars) and production (emails and letters sent, attempts made, surveys completed) data, these two data sources need to be extracted from separate systems and then combined. This disconnect can make it difficult to link cost and effort data. Furthermore, this information is often stored at different levels of analysis (e.g., contact attempt level, interviewer level, interviewer-day level, interviewer-month level, case level) (Tolliver 2026), further complicating analysis and linking disparate data sources. In workshop discussions, several people noted that face-to-face surveys particularly face this issue. Interviewers record hours to be paid in one system related to payroll, but these hours are not tied directly to attempts made on specific units. Linking the two sources of data requires aggregation of one or both sources of data. "Closed" systems where information about costs flowed directly into production monitoring tools appeared to be rare. How these data pipelines are created has a big impact on organizations' abilities to monitor costs and errors. Many organizations instead use measures of levels of effort as nonmonetary proxies of monetary costs

(Chardoul, Chang, et al. 2026; Lancaster 2026; Tolliver 2026; Ubartas and Kuseler 2026; Wagner 2026).

In addition to software systems, some presenters focused on the **organization of teams** for the budget process (Jackson et al. 2026; Miladi and Holmes 2026). These teams included representation from multiple functional units within the organization (account executives, project managers, methods, sampling, operations, data processing, and administration). Involvement of each functional unit aided in communication and enhanced the accuracy of budget estimates and timelines. How to scale this approach across organizations with different organizational structures and to proposals that vary in size was unclear. Nevertheless, one important question for survey organizations that arose from these presentations is how to organize personnel for effective cost estimation. That is, producing accurate cost estimates requires more than good software.

One useful distinction observed early in the workshop was between **fixed costs** and **infrastructure costs**, as reviewed in Section 2. The former might be charged to projects but do not scale with sample size. For example, the fixed cost of hiring a graphic designer to create a study logo is not a function of the sample size. Infrastructure, on the other hand, yields costs that might be supported across multiple studies. For example, software licenses might not be directly charged to projects, but could be amortized across multiple projects that use that software. Although relatively easy to define conceptually, in practice, it may be difficult to assign costs to specific categories like variable, fixed, or infrastructure. For the SHARE studies, vendors proposing to conduct any of the surveys completed a form with line items for key budget components (Kronschnabl 2026). Each vendor was then asked to designate each budget component as "fixed" or "variable." Across 11 components, only one had agreement from all vendors about whether it was fixed or variable.

One cost mentioned by multiple participants as particularly difficult to enumerate and categorize as either fixed or variable was the **cost of developing and implementing proposals and contracts** (Chardoul, Coffey, et al. 2026; Eltinge 2026a; Stepler 2026). Developing proposals and contracts requires technical expertise across multiple units in an organization, yet is not covered by direct project costs. Legal and technical support for contract development related to procurement of study supplies (e.g., incentives) or administrative data could be a monetary cost to a study or organization more broadly. Many participants also mentioned a cost of contract negotiation as the time to finalize the contract, often taking months or years to negotiate across organizations. Contract development costs also were mentioned by those who had to develop requests for proposals and/or contracts with data collection firms. Despite the difficulty in distinguishing between fixed and infrastructure costs, these concepts were considered to be very useful for thinking about survey budgets.

***Takeaway #4: Cost monitoring is important and methods to do so vary tremendously across organizations. Cost monitoring is complicated by different practices and administrative systems holding different cost information, costs being recorded at varying levels of detail,***

***costs being categorized differently across organizations, and some costs being difficult to fully enumerate.***

### 3.3 Research on Costs and Survey Design Features

Many important cost drivers in surveys are linked to specific design choices. Making these connections more explicit can help integrate cost considerations earlier in the design process and support more informed tradeoffs. Participants examined how various design features influence survey costs, including the interactions among these features and the complexities involved in measuring and modeling their effects. We describe below the different survey design features that were discussed at the workshop as affecting costs. Direct cost measures were sometimes reported (typically as relative costs) but often discussed in broader terms (e.g., mode A is more expensive than mode B). These discussions of design features sometimes linked to error indicators such as response rates, sample composition and representation, or design effects. How design features affect survey costs and survey errors across surveys and across organizations is a much-needed area for research.

**Sample design decisions** can have important cost consequences; as described in well-established cost models used for sample design, some variable costs of data collection can scale linearly with larger samples (Cochran 1977). Presentations at the workshop illustrated survey costs beyond these simple sampling models. For instance, oversampling certain subgroups may improve analytic precision but can increase costs if those groups are more difficult to contact or require greater effort to recruit (Eisenhauer and Han 2026). Similarly, geographic clustering can reduce operational costs in interviewer-administered surveys while introducing tradeoffs with statistical efficiency. DeMatteis and Cook (2026) discussed the cost and precision trade-offs for two alternative primary sampling unit options (i.e., counties vs census tracts) in multi-stage sample designs in light of different interviewer staffing models. Owens (2026) combined probability RDD phone and text messaging samples and nonprobability web samples to manage costs of state-level political polls, subject to university procurement rules. These presentations illustrated that cost effects cannot be considered in isolation from statistical or operational factors; rather, they interact in complex ways that must be accounted for in both planning and analysis.

**Mode of data collection** is another central design feature with direct cost implications. Multiple presentations examined mode as an important design feature driving costs. Lower-cost modes such as web surveys may reduce per-case variable costs of data collection but may not reach all segments of the population effectively, potentially requiring follow-up in more expensive modes. Mixed-mode or sequential designs are often used to balance costs and survey errors, yet they introduce additional complexity in both implementation and cost estimation. Understanding these interactions remains a key research challenge. Boulet (2026) examined whether targeting additional Computer-Assisted Telephone Interviewing (CATI) follow-up to cases with low predicted response propensity can improve survey outcomes. The goal was to reduce nonresponse bias, increase responses in underrepresented groups, and improve precision (measured as effective sample size) while using CATI effort more efficiently. Reardon (2026) demonstrated the gains in participation and cost efficiency using a mail mode compared to push-to-web mode for the New York City Health Panel, examining whether including a mailed invitation letter increased response rates overall and for subgroups over SMS and email invitations alone, as well as the variable costs related to printing and postage. Costs per complete, the costs to obtain an equivalent number of respondents across experimental conditions, and an "efficiency index," reporting the number of extra completes obtained for an additional $100, were the cost measures reported, along with the effects of each experimental treatment on response rates and sample composition. Lewis et al. (2026) reported on a decline in the use of paper questionnaires in a study in Chicago in a sequential web-to-mail design, yielding few differences from a web-only design on response rates and sample composition, noting that the web-only design saved on printing and postage costs. Endres (2026) discussed how different types of postage in mail surveys affect delivery dates and response rates. Brown and Jones (2026) incorporated telephone, web, and paper questionnaires into a planned in-person data collection to reduce anticipated costs after a face-to-face pretest. Boulet (2026) and McRoy (2026) both reported on differential relative costs for modes of data collection.

**Combinations of survey frame and mode** also have substantial cost implications. For instance, ZuWallack (2026) compared costs and errors for using an address-based frame to recruit individuals with no telephone number to a web survey, phone numbers from an RDD frame that were not linked to an address to a telephone survey, and a combined web+phone approach for those units that had an available phone number and address available. The error indicators included demographic characteristics of age, race, and education, a measure of the design effect due to weighting, and health-based survey estimates across these frames and alternative simulated designs. Cost indicators included the estimated variable costs for data collection per complete and per "effective sample size." Similarly, Amaya et al. (2026) compared data collected on three frames with different modes–a web-based nonprobability sample; a web and phone probability-based web panel; and a mixed-mode web, mail, and phone address-based sample. The cost indicator was the relative variable cost related to data collection across these three packages of frames and modes. The error indicators were differences in survey estimates related to political, social, and media engagement across the modes/data sources overall and for selected subgroups, where one of the mode/frame combinations was directionally assumed to be "truth," and the highest cost option did not always yield the "best" data. Owens (2026) reported on multiple state-level studies that used a combination of telephone, nonprobability web panels, and text messages, with error indicators of design effects, demographic composition and distribution of party ID, showing great heterogeneity across studies and across time in both cost and the error indicators. Kaye (2026) examined fixed and variable costs across combinations of modes and frames in web and mail surveys, including probability and nonprobability methods. Although the nonprobability panels and opt-in open link methods yielded no difference in costs for data collection from the other methods, on average, these studies required substantially more time for data management and processing than probability methods or listserv-based samples. In all of these studies, costs and (where reported) errors varied over the combination of mode and frame.

**Fieldwork protocols** are major drivers of survey costs. Increasing the number of contact attempts or extending field periods may improve response rates, but often with diminishing returns. Yet how fieldwork decisions for adding or stopping contact attempts are related to survey costs are difficult to understand, as discussed by two workshop presentations. Evaluating the marginal cost of additional effort relative to its contribution to data quality is challenging (Tolliver 2026), as is evaluating the cost savings from reducing additional effort (Wagner 2026). Understanding the costs associated with different types of field outcomes during the data collection period requires models. Tolliver (2026) used regression models to predict the hours spent on a case, the miles charged for traveling, and the number of contact attempts over the course of a field period using paradata (e.g., indicators of prior reluctance, prior outcome code), characteristics of a housing unit (e.g., the address has a yard), and characteristics of the area (e.g., percent Black or Hispanic). Although model specification was informed by knowledge about the survey process, the model predictions had large and variable residuals and failed to predict the cases that needed the highest number of contact attempts. Similarly, Wagner (Wagner 2019, 2026) used paradata to predict the number of hours worked by an interviewer according to different call outcomes (e.g., number of interviews, number of noncontacts) and measures of workload (number of segments worked, number of days worked). Both modeling efforts aimed to estimate cost savings that might arise from stopping field effort, extrapolating from the models to estimate what would happen should data collection stop at a particular point (e.g., after 3 more contact attempts). From the workshop discussion, this remains an area in need of further research.

**Staffing structures for interviewer-administered surveys** also affect survey costs. Larger interviewing pools require more supervisors, which adds to survey data collection costs in nonlinear ways (Brown and Jones 2026). Traveling interviewers or other field data collectors overnight into areas may or may not save costs (DeMatteis and Cook 2026; Kotch et al. 2026), depending on the scale, the study, and the tasks at hand. DeMatteis and Cook (2026) found that traveling interviewers for a large-scale national survey yielded cost savings with a large national interviewing work pool, whereas Kotch et al. (2026) found that travelers were more expensive in a small local study requiring phlebotomists. Other costs related to staffing were mentioned as difficult to measure through data collection-oriented paradata, but extremely important to consider for the costs of a project. These include the costs of hiring, recruitment, training, and supervising interviewers and accounting for interviewer attrition (Brown and Jones 2026; Chardoul, Chang, et al. 2026).

**Incentive structures** affect survey costs. While higher prepaid incentives increase total expenditures for a fixed sample size, they sometimes reduce field effort, improve response among harder-to-reach groups, or allow researchers to reduce an initial sample size in anticipation of higher response rates. The overall cost impact is context-dependent and nonlinear, making it difficult to predict outcomes reliably. Experimental or adaptive approaches to incentive design were examined in presentations. Bertoni (2026) compared $1 cash to a $5 virtual gift card for recruitment into a probability-based panel, showing differences in incentive costs across these two incentives types, both overall and for important subgroups. McRoy (2026) described experiments on various types of mode inducement incentives in a sequential mixed-mode design, attempting to increase the proportion of web or telephone interviews over mailed questionnaires among later respondents. These incentives successfully pushed respondents to participate in the induced modes and affected cost structures, sometimes increasing costs (when encouraging telephone participation) and sometimes decreasing costs (when encouraging web participation or discouraging paper participation). Stepler (2026), in addition to examining costs directly, noted that administering cash incentives would add some operational costs that were often not accounted for in existing research. The presentation provided an example of estimating labor hours for mail assembly that could be used to inform mailout operations on future survey iterations. Gentry et al. (2026) examined alternative prepaid and promised incentive levels over time for different subgroups of interest on return rates for a media diary. Guyer et al. (2026) examined different combinations of prepaid and postpaid incentives on response rates and number of days to respond, as well as the mailing costs per complete and average costs per complete.

**Questionnaire design** affects costs through its influence on interview length and respondent burden (Chardoul, Chang, et al. 2026; Gillen et al. 2026; Lancaster 2026). Longer or more complex instruments increase interviewer time (by definition) and potentially increase respondent burden (Yan et al. 2020; Yan and Williams 2022); longer questionnaires may also increase the risk for nonresponse or breakoff (although this was not empirically examined at the workshop) (Bogen 1996). This complexity is exacerbated when **biomeasures** are collected (Brown and Jones 2026; Kotch et al. 2026) or multiple languages are needed (Gillen et al. 2026). Despite longer instruments requiring more interviewer and respondent time, systematic methods for linking instrument design to specific cost outcomes (beyond a measure of interview length on its own) are underdeveloped, highlighting a gap in current research on survey costs.

**Adaptive or responsive design features** attempt to integrate cost considerations directly into ongoing data collection. By using interim data to guide where effort is allocated, these approaches aim to improve efficiency. Wagner (2026) showed a model-based approach for estimating cost savings in a face-to-face survey by forecasting future call attempts using hours as a proxy of monetary costs. Lancaster (2026) used multiple measures to assess case-level effort, including number of contact attempts, how many times a respondent had accessed a web survey, or whether the case had ever refused, to track field outcomes and guide follow-up attempts. Some of these data also were used to estimate a screening eligibility propensity model, instructing interviewers where to visit in upcoming contact attempts. Boulet (2026) used response propensity models to guide mode allocation decisions (web versus phone) in an adaptive design. However, adaptive designs require additional infrastructure elements including systems to capture paradata, monitor field progress, and support to develop prediction models and statistical decision rules, as well as technical staff to evaluate these models and decisions.

Adaptive designs also require relevant historical data or previous experimentation in order to identify the optimal assignment of data collection protocols to subgroups in the population. The need of adaptive and responsive designs to have systems and data available may limit their application to organizations that can develop this infrastructure or projects that can absorb some of these costs. Further, these designs create tradeoffs such as design and operational efforts before and during collection that are not yet well understood.

**Cross-cultural, multinational, and multilingual surveys** also pose unique cost challenges. In large-scale multinational surveys, data collection agencies implement a common set of design features and common questionnaire. Yet each country/data collection organization pair implements the study in its own way with its own cost structure, even when common cost categories are used or encouraged. This heterogeneity in cost drivers and cost structures for multinational and cross-cultural studies was emphasized in surveys in Europe (Kronschnabl 2026) and in low- and middle-income countries (Brenzel 2026). Employment arrangements, benefits, cost of living, typical hiring practices, paired with design differences, multiple languages, variable internet coverage, and different views of what constitutes fixed or variable costs complicate cost comparisons across countries. Even within one country, it can be difficult to develop good cost models for multiple languages. Gillen et al. (2026) describe a study in which an assumed interview length underestimated the total length of the interview overall, but especially underestimated for interviews conducted in languages other than English.

Across these design features, a recurring theme in discussion was that cost effects are often **nonlinear, context-dependent, and shaped by interactions among multiple factors**. Despite these nonlinearities and need for context, few presentations examined these contextual factors (outside of country). A notable exception to this was Gillen, et al. (2026), who examined how factors ranging from holidays, the US election cycle, language of interview, state or region of the target population, and sponsor affected different cost outcomes and response rates. Although there was some examination of the cost of incentives across harder-to-recruit groups (Bertoni 2026; Gentry et al. 2026) and use of probability panel information to target mode decisions to lower propensity groups (Boulet 2026) or to recruit hard-to-recruit groups overall (Baer 2026), few presentations reported cost variation across subgroups, sample frame information, or other heterogeneity in costs. Some presentations did examine variation in implementations of the same study over time, simultaneously reporting response rates and costs per complete and per sampled case. For instance, Ferrara et al. (2026) examined variation in study design and implementation decisions, response rates, and total observed (including fixed and variable) costs per complete and per sampled case for six years of a statewide RDD survey. In any cost study over time, changes in costs and response rates can be partially attributed to design feature variation over the years, although disentangling these from other global trends directly is difficult.

Regardless of the cause, **interactions among design features, shifts in effectiveness of design features over time and other contextual information affecting outcomes complicate cost modeling**. These complications highlight the difficulty of making predictions that can be generalized about cost outcomes for researchers who may make different decisions. Although pretesting offers an opportunity to better understand cost drivers before full-scale implementation (Baer 2026; Brown and Jones 2026; DeMatteis and Cook 2026; Eisenhauer and Han 2026; Stepler 2026), including refining assumptions about field effort, testing alternative design features, and quantifying uncertainty in cost models, pretests come with their own costs and assumptions.

***Takeaway #5: Survey design features affect survey costs and survey response rates and other proxy error indicators. How these design features affect cost or error outcomes varies tremendously across studies. Some of this variation arises because of different definitions for costs or different error indicators. Design factors, target populations, and implementation decisions interact with each other and are shaped by environmental and contextual factors (e.g. time of year; country, region, or state; societal, economic, or political influences on study outcomes), all affecting survey costs. How design features affect specific types of costs across studies, populations, and implementation decisions requires more research.***

### 3.4 A General Statistical Model for Survey Costs and Survey Errors

Surveys and other data collection systems are designed with the goal of estimating statistics from the survey. Any estimate will be subject to some set of survey errors. Whether certain types of costs are related to survey errors is of longstanding interest to the statistical and survey community (Groves 1989). Researchers using cost models for design purposes explore costs to inform survey design decisions through pretesting or through adaptive designs. This approach reframes cost models as reflecting inputs rather than outputs: instead of asking “what did this design cost?,” the focus shifts to “what design is likely to achieve objectives within a given cost structure?” Minimizing survey errors, or improving survey quality more broadly, is one objective that researchers (may) attempt to design for within a certain budget. However, explicit connections between survey errors and survey costs are rare.

To connect survey costs with survey quality, it is necessary to define a statistical model for costs that can be linked with a statistical model for survey quality outcomes. A statistical model related to both survey costs and survey quality should reflect decisions related to production steps (including, but not limited to, data collection) and to more general environmental factors. That is, both survey costs and survey quality – measured as statistical errors (e.g., accuracy, reliability) or any other measure of survey quality, such as relevance, comparability, accessibility, timeliness, and credibility (Federal Committee on Statistical Methodology 2020) – are functions of design features, defined broadly below, and environmental factors. In particular, following Eltinge (2026a), we define “design” broadly to include all input elements needed for design decisions (denoted $X$) related to the following five areas:

*Market:* The principal stakeholders applicable to the statistical production system. This may include the intended principal data users; groups of secondary data users (sometimes not well defined *a priori*); intermediaries important in the dissemination and use of statistical information products; survey respondents, providers of administrative and commercial data, and others affected by the data collection process; and related funding and oversight groups. In many ways, the market – or stakeholders – for the data determine its use and whether the data meet a "fitness for use" criterion on both costs and quality. The market, in this definition, also reflects the target population for the statistical estimates. We will refer to this as $X_{Market}$.

*Data Sources:* The full set of data sets and sources used to produce statistical estimates. This may include sample surveys (including sample frames and paradata), nonsurvey data sources like administrative and commercial records, and auxiliary data collected either as part of the survey or as additional or auxiliary information, such as sensors, biomarkers, and environmental samples. We will call this $X_{Source}$.

*Methodology:* All decisions related to the design of the survey or data sources themselves. These include, but are not limited to, methods for sample design; questionnaire development and testing; modes of data collection, recruitment and field procedures; weighting, editing and imputation; and related modeling, estimation, and privacy protection. We will call this $X_{Methodology}$.

*Technology:* The full body of technology employed in the process of statistical information production, dissemination and use. This category includes, but is not limited to, technology used for data collection as well as by survey management staff, data dissemination, statistical analysis, and reporting. We will call this $X_{Technology}$.

*Administration:* All aspects of the administrative and managerial processes used in the production, dissemination, and use of statistical information. Examples include management of the survey fieldwork; negotiations with prospective providers of administrative or commercial data; negotiations with respondents and other data sources regarding allocation of privacy-loss budgets; design and management of accounting systems that may provide important cost-structure information; allocation of funds to training of personnel; and investments in the development, testing, implementation and maintenance of systems; and investments in multiple forms of intangible capital. We will call this $X_{Admin}$.

Of these categories, most of the workshop presentations focused on the Methodology and Administration groups, with some discussion of Technology and Data Sources. There was less explicit discussion of the Market category.

### *3.4.1 Notation*

Rather than referring to these individually, we can formally describe these design features through a vector:

$$X = \left(X_{Market}, X_{Source}, X_{Methodology}, X_{Technology}, X_{Admin}\right)$$

We also define a separate vector $Z$ that includes *environmental factors* that are important for the production, dissemination and use of statistical information. Examples include societal, economic, or political factors that can have an important influence on either survey costs or survey errors. For instance, these may include general environmental factors that affect survey unit contact, initial response, and willingness to cooperate with the full survey process, possibly including repeated cooperation for panel surveys; similar factors that may affect the willingness of governmental or private-sector organizations to provide administrative or commercial data, under specified privacy protections; economic and social conditions that may be relevant to initial participation in, and retention for, recruited survey panels; and labor market conditions that may affect availability of personnel for survey fieldwork, and also affect related labor costs. These environmental factors may also affect how accurately a respondent answers survey questions overall or for particular survey topics, as well as how transportable survey constructs are across subpopulations. These environmental factors may be as large as the country in which the study is taking place (Brenzel 2026; Kronschnabl 2026), time of year or state of residence (Gillen et al. 2026), and can vary across data collection providers (Wong 2026). The *environmental factors* vector, $Z$, is analogous to the "essential survey conditions" (Groves and Lyberg 2010; Hansen et al. 1961)

### *3.4.2 Conceptual Statistical Models for Costs and Quality*

We are interested in survey costs, which we will denote with $C$; and survey quality, denoted with $Q$. Because costs and quality are multidimensional, we will represent them as a collection of costs and quality indicators, statistically represented as a vector. In particular, let $C$ represent a vector of *actual* costs incurred in the statistical information process. Different elements of $C$ may include monetary or non-monetary costs that can be attributed to different parts of the full process of producing estimates in a statistical information system; and may be expressed in either absolute or relative terms.

In addition, let $Q$ represent a vector of quality measurements considered important for a specified suite of statistical information products; this is where survey errors come into the model, although quality is larger than simple accuracy and reliability. The elements of $Q$ may include multiple measures of quality for each of several specified estimands. Notable quality measures include accuracy (often including point estimator bias and variance, and confidence interval coverage), relevance, comparability, interpretability and punctuality. In keeping with

(Eltinge 2023, 2026b, 2026a), we thus define a multivariate performance profile $P = (Q, C)$, where the vectors $Q$and $C$ are these multivariate measures of quality and cost, respectively.

Costs, *C*, are a function of both the inputs into the design of the study (which we denoted by *X* above) and the environment, *Z*. Quality is also a function of the study design *X* and environment *Z*. We can use this to specify general conceptual statistical models for cost and quality, respectively:

$$C = g_C(X,\ Z;\ \gamma_C) +\ d_C \tag{C.1}$$

$$Q = g_Q\big(X,\ Z; \gamma_Q\big) +\ d_Q \tag{Q.1}$$

where $g_C(\cdot,\ \cdot;\cdot\ )$ and $g_Q(\cdot,\ \cdot;\cdot\ )$ are parametric functions of specified form for the cost and quality, models, respectively; and $\gamma_C$ and $\gamma_Q$ are associated parameters those models. In addition, $d_C$ and $d_Q$ are "equation error terms" assumed to have mean zero, conditional on specified $(X,\ Z)$.

Variants on models (C.1) and (Q.1) are common in the sampling literature. For example, for stratified random sampling, Cochran (1977, eq. 5.17) considers the linear cost function,

$$C = c_0 +\ \sum_{h=1}^{L} c_h n_h \tag{C.2}$$

where we have $L$ strata labeled $h = 1, \ldots, L$; stratum-level sample sizes $n_h$; stratum-level per-unit costs $c_h$; and a fixed cost term $c_0$. In this equation (C.2), $g_C(\cdot,\ \cdot;\cdot\ )$ is a relatively simple linear model in which total cost is a linear function of sample size within a stratum ($n_h$) and the survey costs per stratum ($c_h$). That is, in this cost model, the only design features *X* considered explicitly are the number of strata and the stratum-level sample sizes. In addition, the simplified cost model is essentially conditioning on other design features (e.g., costs of fieldwork, as well as production system development) and environmental conditions (e.g., the degree of respondent cooperation or reluctance), which are implicitly absorbed in fixed cost term $c_0$ and the stratum-level cost terms $c_h$. This model excludes the equation error term $d_C$. In other words, model (C.2) implicitly assumes that the difference between the true cost and this formula is small enough to ignore.

Survey sampling literature, and survey methodology literature more broadly, contains many examples of the quality function (Q.1). For example, Cochran (1977, eq. 5.18), pairs the cost function for a stratified random sample (C.2) with the variance of the sample mean ($\bar{y}$) as a measure of quality:

$$V(\bar{y}) = \sum_{h=1}^{L} \left(\frac{W_h^2 S_h^2}{n_h}\right)(1 - f_h) \tag{Q.2}$$

where $N$ is the overall population size, the stratum size in the population $W_h =\ N_h/N$ , the sampling fraction within each stratum $f_h =\ n_h/N_h$, and $S_h^2$ is the finite population variance for stratum $h$. Note that the terms $S_h^2$ in expression (Q.2) may depend on features of the general design $X$ (e.g., decisions on stratum boundaries) and environmental conditions $Z$ (e.g., societal or economic conditions that may affect dispersion of a survey variable $Y$ within specified strata). Indeed, a "good" stratification scheme attempts to make $S_h^2$ small within strata. In addition, note that expression (Q.2) excludes the equation error term $d_Q$, reflecting an implicit assumption of approximation error in the estimator variance that is small enough to ignore. In a report on a 2006 workshop on survey costs, Karr and Last (2006) discuss related issues for cases involving quality measures beyond sampling error variances.

### 3.4.3 *Trade-Offs Between Cost and Quality: Partial Control via Changes in Design*

The survey sampling literature has developed a wide range of standard formulas for tradeoffs between survey quality (largely as measured by the variance of an estimator) and survey costs for a specific set of conditions, as well as "optimizing" designs taking cost and quality into account (typically, the smallest variance for a given cost model). In particular, these models generally assume that the error terms $d_C$ and $d_Q$ are small enough to ignore, environmental conditions $Z$ are relatively stable, and the cost and quality model coefficients are known (see Cochran 1977, Theorem 5.7).

However, most of the work presented at the workshop moved beyond these assumptions. Here, we need a more general approach to cost-quality trade-offs for different design features. For example, assume for the moment that we have specified environmental conditions $Z$; and consider two distinct design specifications $X_0$ and $X_0^*$, respectively. For example, $X_0$ and $X_0^*$ might represent different specifications for initial sample size; different screening designs; different levels of interviewer training and field supervision; different incentive levels; different data collection modes; or different strategies for initial unit contact and nonresponse follow-up.

The associated differences in expected cost and quality between design $X_0$ and $X_0^*$ are, respectively:

$$g_C(X_0^*, Z; \gamma_C) - g_C(X_0, Z; \gamma_C) \tag{D.1}$$

and

$$g_Q(X_0^*, Z; \gamma_Q) - g_Q(X_0, Z; \gamma_Q) \tag{D.2}$$

We can ask whether those changes are large or small, relative to our model error terms, and relative to thresholds for changes in cost and quality that are considered to be of practical importance. In addition, we can ask whether the coefficients in our models are stable – over time, and across different survey settings. If they are stable, we can hope to borrow some cost information across different surveys. However, if the coefficients are not stable, then we may need to develop our empirical cost information separately for each survey or project separately, e.g., through preliminary pilot studies.

### 3.4.4 *Regression Estimation of Model Coefficients $\gamma_C$ and $\gamma_Q$*

These general models are only useful if they can be estimated. Thus, we now turn to a conceptual model for the (general) steps for obtaining estimates of parameters in these models. Using the terminology from Section 2, let $C_i$ represent the vector of true costs incurred under a set of design and environmental conditions $(X_i, Z_i)$; and let $\hat{C}_i$ represent the corresponding observed costs. We assume that the observed costs may have some errors in its measurement $e_{Ci}$, as discussed in section 2 and 3.5. Similarly, let $\hat{Q}_i$ represent estimators of quality achieved under the conditions $(X_i, Z_i)$, which may also have measurement errors, $e_{Qi}$. We can then estimate the model coefficient vectors $\gamma_C$ and $\gamma_Q$ and the variances for the equation errors $d_C$ and $d_Q$. These models generally require:

(a) Specified forms for the functions $g_C(\cdot,\ \cdot;\cdot\ )$ and $g_Q(\cdot,\ \cdot;\cdot\ )$

(b) The availability of empirical measures of cost and quality under specified design and environmental conditions $i$:

$$\hat{C}_i = C_i + e_{Ci} \text{ and } \hat{Q}_i = Q_i + e_{Qi} \tag{E.1}$$

(c) Assumptions or empirical information on the expectations and variances of the measurement error terms $e_{Ci}$ and $e_{Qi}$ in (E.1). Note that the error terms $e_{Ci}$ may arise from the limitations of the accounting systems and reporting constraints noted in Section 2. We may not have direct empirical information, but can discuss the limitations of the measures that are available.

(d) Customary statistical estimation methods (generally parametric linear or nonlinear regression)

To illustrate estimating cost models using a regression approach, we draw on two presentations. During the workshop, Tolliver (2026) and Wagner (2026) described using multilevel linear regression models (where $g_C$ is a linear or log-linear link function) to predict measures of survey costs. Regression coefficients reported by Wagner (2026) provide an example of sample-based estimates of the $\gamma_C$ terms in the models for a limited set of X values; residual plots by Tolliver (2026) reflect sample-based estimates of $d_C$.

Explicit estimation of error models during the workshop was largely based on sample allocation models, in concert with costs. These sampling error models were discussed when deciding how to allocate sample across strata and how to decide on the size or type of clusters (DeMatteis and Cook 2026; Eisenhauer and Han 2026), in which costs increase with a larger sample size. For instance, DeMatteis and Cook (2026) reported on differences in costs and sampling precision for using a county-based versus tract-based PSU sample design.

When considering survey costs and survey errors, most of the presentations examined packages of design features (*X*), typically implemented in an experiment or with planned variation across sampled units (*X**) within a single study. In these studies, cost indicators

($\hat{C}_i$) typically focus on variable costs related to data collection, although what exactly is included in these costs varies from study to study. Furthermore, sometimes $\hat{C}_i$ was an observed cost derived from accounting systems, but sometimes was an estimated cost, using assumed costs for materials, supplies, and labor. Quality indicators ($\hat{Q}_i$) typically included response rates and measures of sample representation. Sample representation was measured through comparing demographic characteristics to a population distribution or to an assumed "best" method of data collection as a measure of accuracy. Others used a measure of the design effect due to weighting or a measure of effective sample size as a measure of variability. Table 2 below summarizes the presentations that explicitly mentioned both cost and quality indicators along with design features.

Given the heterogeneity across the studies presented at the workshop in design features, measures of costs, and error indicators, it is difficult to draw conclusions about the association between survey costs and survey errors across these studies. The error indicators from these studies focus almost exclusively on participation rates overall, sample composition, and design effects. Whether this is because these outcomes are measurable, required to be disclosed, or the most important error indicators available for the design features tested is not clear. The costs examined here are almost all at the end of data collection, and most focused on some subset of variable costs, either estimated or observed. How the cost model describing those costs relates to the design features or error indicators in question is not necessarily specified. Yet reporting costs in any metric and any kind of error indicator is an important step to begin to develop an understanding of the relationship between survey errors and survey costs.

***Takeaway #6: Stating specific statistical models linking design features and environmental factors to both survey costs and survey errors is a necessary next step in research on survey costs.***

Table 2. Summary of design features, cost indicators, and quality indicators among Workshop presentations that reported on all three features

| Study | Design feature ($X$) | Cost indicator ($\hat{C}_i$) | Quality indicator ($\hat{Q}_i$) |
|---|---|---|---|
| (Amaya et al. 2026) | • Sample source + mode | • Costs relative to probability panel<br>• Sample purchase, Nonprobability sample "cost per complete";<br>• Programming, CATI interviewers, Contact attempts, Incentives, Panel management; Vendor management | • Survey estimates in three sample source/mode combinations |
| (Bertoni 2026) | • Incentives | • Incentive cost<br>• Number of redeemed virtual incentives | • Response rates in two subgroups<br>• Sample composition |
| (Boulet 2026) | • Mode | • % of telephone surveys; Relative cost per complete; relative budget per mode | • Relative effective sample size; response rates across subgroups |
| (DeMatteis and Cook 2026) | • Type of PSU<br>• Number of PSUs<br>• Local vs. traveling interviewers | • Ratio to overall cost for average cost per screener, average cost per extended interview | • Design effect |
| (Eisenhauer and Han 2026) | • Sample allocation | • Assumed relative cost | • Minimum detectable effect size<br>• Statistical power |
| (Endres 2026) | • Postage and mailing type<br>• Mail tracking | • Postage cost<br>• Time in the field<br>• Mail tracking software cost<br>• Printing cost<br>• Labor for mailing | • Response rates |
| (Ferrara et al. 2026) | • Variation over time<br>• Mode<br>• Prenotifications<br>• Incentives<br>• Survey topic | • Total cost and cost per complete, including project management, programming, interviewer time, sample purchase, equipment, phone service, printing, postage, incentives<br>• Sample size<br>• Number of call attempts<br>• Number of completes | • Response rates |

| Study | Design feature ($X$) | Cost indicator ($\hat{C}_i$) | Quality indicator ($\hat{Q}_i$) |
|---|---|---|---|
| (Gentry et al. 2026) | • Incentives<br>• Variation over time<br>• Postcards | • Costs including sample purchase, printing, postage, incentives, survey processing | • Response rates in six subgroups |
| (Guyer et al. 2026) | • Incentive<br>• Phone reminders<br>• Mode | • Days to complete<br>• Number of mailings<br>• Number of emails<br>• Cost of mailings<br>• Staff effort | • Response rates |
| (Kotch et al. 2026) | • Mode | • Cost per interview<br>• Invitation cost, scheduling costs, travel cost, downtime for interviewer; interview cost; travel paperwork; post-interview administrative tasks; airfare; hiring cost; training cost | • Retention rates |
| (Lewis et al. 2026) | • Mode + incentives | • Relative printing and postage cost<br>• Relative incentive cost | • Response rates<br>• Sample composition<br>• Key survey estimates |
| (McRoy 2026) | • Incentives<br>• Mode | • % of web, telephone, paper completes<br>• Relative costs per complete, including processing, inbound mail, interviewers, incentives | • Conditional response rates<br>• Participation in optional module |
| (Owens 2026) | • Mode<br>• Sample source | • Completed interviews per $15,000<br>• Number of telephone surveys<br>• % of telephone surveys in total sample | • Design effect<br>• Sample composition |
| (Reardon 2026) | • Mode | • Cost per complete; Cost per 2000 completes; Extra completes per $100<br>• Printing, postage, labor for mailing | • Response rates<br>• Sample composition |
| (Stepler 2026) | • Nonmonetary incentive: lanyard vs. pen<br>• Digital incentive<br>• Mode | • Total cost, cost per sampled unit; cost per complete; relative costs for components out of total cost<br>• Costs for printing letters; envelopes; questionnaires; postage; incentive; labor for incentives; labor for mailing package; incentive processing fees | • Response rates |
| (Ubartas and Kuseler 2026) | • Mode | • Average collection time per case<br>• Number of contact attempts | • Response rates<br>• Contact rates |

| Study | Design feature (*X*) | Cost indicator ($\hat{C}_i$) | Quality indicator ($\hat{Q}_i$) |
|---|---|---|---|
| (ZuWallack et al. 2026) | • Mode + frame | • Estimated cost per complete and cost per effective complete, including printing, postage, incentives, interviewer labor, telephone cost, texting labor, texting charges | • Response rates<br>• Sample composition<br>• Key survey outcomes<br>• Sample yield<br>• Design effect<br>• Confidence interval width |

### 3.5 Measurement Errors in Cost Data

Although the statistical models described in section 3.4 include measurement errors for both cost and quality indicators, the mechanisms for measurement errors in cost data are multidimensional and vary across organizations and researchers. Per expression (E.1), the quality of available empirical observed cost information $\hat{C}_i$ is reflected in the mean and variance of the error terms $e_{Ci}$. **Measurement errors arise in survey cost data due to challenges with the accuracy, completeness, and interpretability of the cost data**. These limitations may have implications for analysis and decision-making from cost information. Depending on the nature of the measurement error, the resulting cost estimates may be biased (resulting in systematic under- or over-estimation of the true cost) or may have relatively high variance (resulting in inaccurate estimation of the true cost and attenuation of relationships). Both of these types of errors could lead to faulty decision-making. However, Eltinge (2026a) noted that the limitations should be considered in the context of fitness-for-use, focusing on the degree of accuracy that is sufficient to inform design choices and tradeoff analyses. For example, in the analysis of costs of alternative treatments in methodological experiments, it may not be important that the cost data used in the analysis do not include all costs, provided the excluded costs may be assumed to be the same for all treatments; or are small relative to other cost elements. Statistically, in some cases, the accounting systems that produce the direct cost estimates $\hat{C}_i$ may involve aggregation effects, outdated information sources, or definitional issues that produce biases, i.e., nonzero expectations of the error terms $e_{Ci}$. In other cases, the direct cost estimates $\hat{C}_i$ may be available only for a subset of design and environmental conditions $i$.

**Procedures used for timekeeping and billing of costs may result in inaccurate allocation of costs to specific tasks or projects** (Stepler 2026; Tolliver 2026). For example, in presentations and discussions, interviewer timekeeping was specifically mentioned as an area at risk of having measurement errors. When interviewers work on multiple projects concurrently, accurate reporting of actual time spent on each project may be challenging (Chardoul, Coffey, et al. 2026). Furthermore, it may be challenging or impossible to appropriately allocate costs associated with face-to-face interviewer travel for particular projects if an interviewer is working on multiple projects simultaneously. Staff time was also discussed as an area in which costs may have measurement errors. Some organizations allow salaried staff to only bill a fixed amount of time overall (e.g., 40 hours per week) or a budgeted number of hours per project (e.g., no more than 20 hours total) or per project task (e.g., only 5 hours on programming), thus failing to account for costs associated with hours worked beyond these constraints. Some measurement errors in cost data may be structural. Costs may only be available for analysis on a delay (e.g., monthly payroll data is available after the month has elapsed). Some organizations may require staff to record time spent on a project overall, but not for specific tasks on that project. Still other organizations may record costs related to supplies and materials for projects in the number of units used (e.g., the number of envelopes from a bulk purchase order), using a cost model to allocate those costs to projects specifically.

In almost every presentation, cost data reported by organizations often **do not include certain types of costs**, such as costs associated with shared resources (e.g., computing costs, accounting costs), indirect costs, the cost of infrastructure investments, and post-data collection processing tasks (Chardoul, Coffey, et al. 2026). Furthermore, researchers using experimental or quasi-experimental designs, for instance, often **make assumptions that costs for certain tasks are uniform** across the design features of interest, without empirically evaluating whether this assumption holds. Empirical evaluations of these assumptions may yield hidden, surprising, or unanticipated differences (Kaye 2026).

Accounting systems used by survey research organizations are designed to capture costs for billing and administrative purposes and are typically not designed or structured to capture costs in a manner suited for methodological purposes, a point discussed frequently at the workshop. As a result, **costs may be captured at multiple levels for analysis, including the project, case, day, interviewer, or attempt level**. These varying levels of analysis can make using these data jointly difficult, as well as require assumptions related to aggregating data to appropriate levels. For cost information to be useful for budgeting, design, and methodological evaluations, cost accounting systems need to be designed to capture costs at the level(s) suited for the particular analysis. Alternatively, analysts must make decisions about how to aggregate or disaggregate the information. To answer this, a cost model is particularly useful to address questions such as: Are the design and environmental factors expected to affect costs at a particular level of analysis? Or are the costs measured at a particular level likely to affect survey errors? Clear specification of these models could clarify when costs do not replicate across time, studies, or designs.

It is not clear whether accounting systems necessarily capture related information that is crucial for survey cost modeling. For example, if case-level accounting records exist, can they be linked with other data for the same case (e.g., sample member or respondent characteristics, frame information, reluctant-respondent indicators, other interview-process paradata, and interview-specific consistency checks)? Also, in most cases, we have only rough proxies for true case-specific costs, e.g., measures of hours worked by an interviewer or a project manager, instead of the full cost of working with a given sample unit, requiring models to extrapolate to monetary costs per case (e.g., Tolliver 2026). Evaluating the extent to which these issues, or other accounting-system issues, may limit the value of the available data $\hat{C}_i$ in fitting our cost models and in applying the modeling results to planning and evaluation of trade-offs is important.

***Takeaway #7: Observed cost data are likely to have measurement errors, which may be biasing errors that consistently over- or underestimate actual costs, or variable errors that vary in the direction of errors across studies or implementations. Understanding the cost data generating process, including the level at which the cost information is recorded and stored, will be***

***fruitful in understanding whether these errors make the cost information fit for its specified use.***

***Takeaway #8: Although being aware of the potential for measurement error in observed cost data is important, one should not wait for perfectly measured cost information to begin the process of cost evaluation, estimation, or modeling. Cost models may be reexamined as improved cost data become available. In the meantime, sensitivity analyses can be used to evaluate the robustness of findings to plausible assumptions about measurement error.***

## 4. Challenges in Cost Measurement, Analysis, and Research

Workshop participants discussed a range of challenges associated with incorporating cost considerations into survey design, drawing on insights shared through presentations, breakout sessions, and informal discussions during breaks. Despite growing interest in this area, they noted persistent practical and methodological challenges, including inherent uncertainty of survey costs, the institutional structures that govern cost management, and limitations in available data and conceptual frameworks (Boulet 2026; Tolliver 2026; Wagner 2026). More broadly, these issues were seen as extending beyond measurement and modeling to the broader contexts in which surveys were planned and conducted (Eltinge 2026a). This section explores some of these issues, highlighting key topics such as cost variability, case-level differences, data limitations, contractual influences, and the link between cost and survey quality.

### 4.1 Uncertainty in Survey Costs

A foundational challenge in survey cost analysis is the inherent uncertainty in estimating costs. Costs are inherently variable, particularly in complex or longitudinal studies, yet planning and estimation often treat them as fixed. Key drivers—such as response rates and the effort required to reach different segments of the sample—are difficult to predict in advance. Small errors in these assumptions, especially for harder-to-reach cases, can lead to substantial deviations from projected costs. Despite these dynamics, cost models and methodological discussions do not always fully capture such variability, limiting their ability to reflect real-world implementation conditions.

One way to consider this issue is through the statistical models presented in section 3.4 ((C.1) and (Q.1)). The relationship between design features and costs and quality (represented by estimates of the model coefficient vectors $\gamma_C$ and $\gamma_Q$ from section 3.4 above) as well as the model fit (represented by the variances for the equation errors $d_C$ and $d_Q$) may vary over studies, time, and organizations. Some of this variation may depend on environmental factors (*Z)* that are out of the control of the survey organization or researcher. Building solid information about the variability in these kinds of models (or relationships between design features, costs and quality) is important. If these models (or relationships) are stable over time, and across different prospective applications, then information is less uncertain and more transportable across studies. Furthermore, if costs systematically vary across subgroups, geographic areas, urbanicity, combinations of design features, or time in the field period (for instance), then better understanding this variation could lead to better prediction of costs and potentially errors.

Consequently, it is important to evaluate, communicate, and provide appropriate justification for the types of design and environmental conditions under consideration when developing and evaluating cost and quality models. In many applications, survey management will appropriately choose to focus primarily on the span of design conditions that they can reasonably hope to implement and the span of environmental conditions that they consider realistic. The resulting analyses may be useful within the specified span, but one should not overstate the extent to which those analytic results can be extrapolated to other settings.

### 4.2 Cost Variation Across Cases

Beyond aggregate uncertainty, costs vary substantially across individual cases. Average measures, such as cost per complete, often mask these differences. Some cases are resolved quickly ("easy" completes), while others require repeated contact attempts or more intensive follow-up ("hard" completes). Case difficulty is not binary (easy versus hard cases), but rather varies across multiple dimensions, creating substantial variation in costs.

This variation poses challenges for cost analysis. Without stratification by sample unit or respondent characteristics, paradata, or fieldwork phases, it can be difficult to identify where costs are concentrated or which cases drive overall expenditures. Measures such as cost per complete provide only a coarse summary, and alternative approaches—like cost per base weight—remain complex and not widely applied, limiting their usefulness in practice.

The uneven distribution of effort (as discussed in Section 2) also complicates planning and resource allocation. High-cost cases may be unpredictable and identifying them in advance is often difficult. These patterns illustrate the persistent uncertainty inherent in survey costs and the limitations of current approaches for capturing and analyzing variation across cases.

From a predictive cost modeling perspective, in some cases, the cost and quality models (C.1) and (Q.1) may have strong model fits, as reflected in, e.g., values of $R^2$ that are high and standard deviations of $d_C$ and $d_Q$ that are relatively small. For those cases, the cost model may have considerable practical value in budget planning, and in discussion of cost-quality trade-offs. Conversely, relatively weak model fits may limit the value of cost models in those discussions. Here, further model specification that reflects these types of individual case- or group-level variation in costs is important. One might strive to develop predictive models that allow high quality case-level or small-group level prediction of costs (and errors) to more fully understand what kinds of cases contribute substantially to costs (and errors).

### 4.3 Quantifying and Communicating Cost Uncertainty

Estimating costs is only part of the challenge; quantifying and communicating that uncertainty to survey designers and managers within an organization is equally critical. Cost projections are often reported as single point estimates, which can obscure the full range of plausible outcomes. Factors such as differential response rates, fieldwork disruptions, and unanticipated operational issues can introduce substantial variation. That is, the environmental factors may be as important as the design factors when describing costs. This suggests a need to move toward ranges, scenarios, or probabilistic representations of cost, yet practical and widely accepted methods for doing so remain underdeveloped. It may be possible to incorporate this uncertainty into the survey design and associated budget (Burger et al. 2017). For example, if cost estimates associated with recruiting a subgroup are highly variable, then the design may be set such that there is a high probability of remaining within budget. Alternatively, the initial design can be altered based on the results from a pretest, with improved cost estimates (e.g., Brown and Jones 2026).

### 4.4 Limitations of Available Cost Data

Practical analysis of survey costs is often constrained by significant limitations in available data. As noted above, cost information may not be captured at the level of detail needed for meaningful analysis, and linking costs to specific cases, activities, or outcomes can be difficult. Many organizations rely on complex or fragmented tracking systems. In addition, some cost data are proprietary, further limiting access. Also, definitions of cost components often lack consistency, and key elements may be partially measured or inconsistently allocated. These limitations complicate interpretation and restrict the ability to identify cost drivers, although should not prevent analysis of the cost data that are available.

### 4.5 Contractual and Risk Considerations

Cost management is also shaped by contractual and institutional structures. Different contract types—such as fixed-price versus cost-reimbursable agreements—create varying incentives for managing effort, documenting costs, and adapting to changing conditions. These arrangements influence who bears the financial risk: under fixed-price contracts, contractors may limit additional effort to contain overruns. Conversely, sponsors may absorb cost overruns under cost-reimbursable contracts. Contract structures can also limit the transparency and granularity of cost data, making it difficult to conduct a thorough or accurate analysis.

### 4.6 Linking Cost and Survey Quality

A key conceptual challenge is connecting cost to survey quality. While both cost and quality are central to design decisions, they are often measured and analyzed separately, making it difficult to understand how resource allocation affects errors, coverage, or other quality dimensions. Costs evolve dynamically throughout data collection as real-time decisions—such as adjusting contact strategies or reallocating effort—alter both costs and outcomes. On the other hand, quality often tends to be seen as a fixed target but should be considered in a context of fitness for use. Hence, there is a need for approaches that jointly consider cost, quality, and operational dynamics to better understand these interactions and support informed decision-making in survey design and management. Variants on the models outlined in Section 3 may help to address these issues.

**4.7 Costs with Alternative Data Sources, including Non-Probability Surveys**

As discussed in section 3.3 above, one design feature examined by some of the presentations at the workshop involved comparisons of nonprobability samples with probability samples on costs and errors. In the existing literature, discussions of nonprobability sampling often cite the potential for substantial cost savings due to a reduction in expenditures for nonresponse follow-up and some other features of customary probability sample surveys. However, other studies of no-probability sampling have noted the potential for substantial quality problems due to issues with biases arising from undercoverage of some subpopulations. Other bias issues can arise from measurement errors attributed to inattentiveness and other forms of disengagement by some participants in a series of nonprobability sample surveys. Due to these concerns, numerous articles have developed methods for evaluation and mitigation of some components of these biases, under specified (sometimes restrictive) conditions (Ang et al. 2025; Chen et al. 2024; Cornesse et al. 2020; Hammon and Zinn 2025; Kim 2025; Liao and Biemer 2025; Liu and Valliant 2023; Marella 2023; Pfeffermann and Sverchkov 2025; Unangst et al. 2020; West et al. 2021; Wu 2022).

One area that was not discussed by many presenters at the workshop with nonprobability samples, but remains a challenge for those making population inference, especially in the context of official statistics, has to do with the data processing and analysis costs to implement these sophisticated statistical methods for evaluation and implementation. This is relevant because both new data and new analysis methods are needed. That is, in addition to data from the nonprobability sample itself, these methods generally require substantial amounts of data from other supplementary data sources, e.g., administrative records or probability sample surveys; and require substantial analytic work in development and testing of associated models.

Realistic evaluation of cost-quality trade-offs for nonprobability surveys varies from one use case to the next. Some researchers use nonprobability data with no adjustment or little adjustment, without addressing bias concerns in depth; others use weights provided (perhaps) by the nonprobability sample provider; still others use complex statistical models to blend or link nonprobability and probability data or administrative data. A complete comparison (especially for those using nonprobability data in concert with probability data; such as (Owens 2026)) would evaluate the costs of the non-probability data collection as such, plus the costs of the above-mentioned supplementary data sources; costs for related work with record linkage; costs for model development and monitoring; costs for related analyses required to implement the methods in the literature cited here; and costs for development, testing, operation and maintenance of related systems. At the workshop, Kaye (2026) reported large costs for data processing for nonprobability panel or opt-in data collection. Also, evaluation of the "quality" dimension of cost-quality trade-offs would require application of model checks, diagnostics and sensitivity analyses anchored in that literature. The empirical results of those analyses may vary substantially across different estimands. The costs of nonprobability samples may be beyond the technical and methodological costs, but also in the institutional costs arising from potential degradation of the quality of information provided (see section 5 below). The challenge here comes with clearly articulating all the cost drivers beyond the nonprobability sample acquisition, and in making the data and statistical analysis from it "fit for use."

This same kind of evaluation applies when combining probability sample data with administrative data. Purchasing administrative data can be extremely expensive, especially for those outside a federal statistical agency (Chardoul, Coffey, et al. 2026). Processing this data requires sophisticated knowledge to use on its own, or to link with probability data. Virtually all of the aspects of costs and quality evaluation described above for nonprobability sample data can be applied to administrative data.

***Takeaway #9: Many challenges exist in understanding and predicting costs, especially across subgroups and data collection stages. There are potential limitations to existing cost data that can limit its utility. Uncertainty should be acknowledged and incorporated into estimates wherever possible. Cost drivers, especially related to statistical modeling expertise, that vary by type of data source, probability or nonprobability samples, or study complexity, are not well understood.***

## 5. Measuring the "Value" of Data

When organizations commission survey research, the conversation almost always arrives at cost. How much will this cost? Can the study be conducted for less money? Is there a more affordable option that yields "close enough" estimates? Budgets are real constraints, guiding what methods can be used to answer a research question. But the framing of cost as the primary decision variable misses something important: the value of survey research is not just what it tells the client or user. The value of survey research should be viewed in terms of both its immediate utility and its broader and longer-term societal benefits. Additionally, survey errors alone (in terms of accuracy and reliability) fail to reflect all dimensions of quality (the quality vector $Q$ in section 3.4 above). Thus, we now examine other domains of survey quality, some of which were discussed at the workshop, especially by the opening plenary panel (Chardoul, Coffey, et al. 2026), and others that build off of frameworks for data quality.

### 5.1 Connecting Costs with Value: Use Value, Option Value and Related Measures of Quality

Previous sections discuss connections between survey cost structures and the quality of information, with emphasis on cost-quality trade-offs. This included multiple dimensions of quality, including accuracy, relevance, comparability, temporal and cross-sectional granularity, punctuality, interpretability, accessibility, scientific integrity, credibility, and confidentiality (Federal Committee on Statistical Methodology 2020). To some degree, these quality measures serve as indicators of the value delivered to stakeholders by a given suite of statistical information products and services. However, stakeholder perceptions of value also can depend on many factors specific to a given stakeholder and their operating environment. To explore this further, we draw on the distinction in the economics literature between two notions of "value." Arrow and Fisher (1974) and the literature on environmental economics provide some general background on the concepts of, and measurement methods for, *use value* and *option value*.

For the current discussion, we define "use value" as the value conveyed to a given stakeholder (data user) by the use of a specific item (suite of statistical information products or services) under specific circumstances. An example is the value conveyed to a steel producer by its use of certain entries in the Bureau of Labor Statistics (BLS) Producer Price Index to adjust the prices paid under a previously negotiated contract for a particular type of steel. The use value has to do with current and actual uses for these data.

In addition, we define "option value" as the current value conveyed to a given stakeholder by the prospective future use of a specific item. An example involving statistical information is the value that a policymaker assigns to a statistical information series on public health that they have never used, but might use in the future during a public health crisis. A related example is the value that a policymaker assigns to the ability of a statistical organization to create and publish a new statistical information product (e.g., on employment in a not-yet-defined occupational category) quickly and with a specified level of quality. The option value has to do with the future and/or potential uses of these data.

For statistical information products and services, assessment of their value, and of related cost-value trade-offs, generally will involve the cost-quality trade-offs reviewed in previous sections. However, those assessments also may require consideration of other factors, including important intangibles like the credibility of the statistical organization, trust by a given set of stakeholders, and the scientific integrity of the methods that are used. The remainder of this section explores these expanded dimensions, and difficult-to-quantify measures of survey quality, drawing on the framework for data quality from the Federal Committee on Statistical Methodology (Federal Committee on Statistical Methodology 2020).

### 5.2 Balancing Priorities: Cost Management Conditional on Quality Criteria and Perceived Stakeholder Value

The Federal Committee on Statistical Methodology defines two aspects of data quality as scientific integrity and credibility (Federal Committee on Statistical Methodology 2020). These aspects of data quality reflect what users can trust, what survey organizations can defend, and what it means to the study's target population.

Higher-quality survey research — often characterized by the use of probability-based sampling, well-tested questionnaire design, large sample sizes, fielding across modes and languages, extensive follow-up attempts, careful weighting, and validation — is more expensive than its alternatives. Whether large probability-based samples with other markers of higher quality research costs more than studies without these design features is generally not a useful question. Rather, the important question is whether higher cost data yield smaller errors of representation and measurement (e.g., higher quality surveys as measured by criteria and accuracy) and also provide more utility in terms of specific decisions made today (use value), future decisions and uses (option value), and mitigating risk. Utility may come when the estimates produced meet needed quality domains including accuracy, relevance, comparability, temporal and cross-sectional granularity, punctuality, interpretability, and accessibility, each outcome of which may come with different study design needs.

### 5.3 Credibility as Value: Confidence in the Data Used to Make Decisions

The most immediate value for data is also the most practical: data collected are only as useful as a user's confidence in its utility for decisions and understanding a population. Credibility is often gained through widespread confidence in assessments of accuracy. Decision-makers in government, philanthropy, healthcare, media, and the private sector routinely use survey findings to allocate resources, design programs, craft communications, and set strategy. When the underlying data is unreliable or inaccurate, those decisions are built on a foundation that cannot hold weight – that is, their use value is low.

The use of lower-cost survey data may be justified in certain settings with well-understood implications. For example, methods for integrating nonprobability data with probability data are an active area of research, with estimates compared to (probability survey-based) benchmarks. However, there may be long-term implications that are not well understood when sample sources change over time, especially as AI-generated "silicon samples" based only on model predictions and not including real humans become part of a lower-cost product list sold to researchers. The gap between higher-cost and lower-cost survey methods arises in multiple domains. One area where higher cost surveys attempt to be accurate is in having high coverage of the population. That is, probability sample methods have a known sampling frame that is designed to represent the full population of interest; higher cost surveys will use methods that attempt to bring in sample members who are reluctant to participate or are difficult to reach. As such, lower cost nonprobability or convenience samples can systematically exclude people who are harder to reach — those without reliable internet access, non-English speakers, rural residents, older adults, people with lower incomes. Opt-in panels, by definition, consist of people who have chosen to participate in surveys, which means they are not randomly selected and may differ from the broader population in ways that matter enormously for a funder's or researcher's research questions (e.g., Baker et al. 2010; Kennedy et al. 2016). A probability-based sample, by contrast, gives every member of the target population a known chance of being included, providing the statistical connection between the sample and the population of interest (Cochran 1977). For instance, Baer (2026) described the possible value of building a probability-based panel to help the Internal Revenue Service make decisions quickly and also maintain quality estimates for this need.

For decision-makers, the practical implication is straightforward: less expensive research may feel like savings initially, but may introduce uncertainty that could lead to costly mistakes downstream. The risk of making decisions based on bad data could exceed the savings realized on the research budget; even when the decision is "empirically based," inaccurate or unreliable data may move decisions in the wrong direction or provide an incorrect snapshot of a population. For example, organizational risks may increase for healthcare organizations that make program decisions based on a flawed survey of patient attitudes, foundations that allocate grants based on inaccurate data about community needs, or companies that launch a product based on inaccurate consumer sentiment research. That is, the use value from lower quality or less accurate data may be quite limited, and the risk to the credibility of the organization itself may be great.

### 5.4 Scientific Integrity as Value: The Risk of Publishing Data One Cannot Stand Behind

For organizations and researchers that publish or publicly release their research findings, the stakes of releasing statistical information with poor data quality are even higher. Published survey data is cited by journalists, referenced by policymakers, shared on social media, and picked up by other researchers. If that data is later found to be methodologically flawed — if the sampling approach was not defensible, if the weighting was inadequate, if the response rates were too low to draw reliable conclusions — the reputational consequences can be significant and lasting. In this way, scientific integrity underlies organizational credibility. The scientific integrity domain of survey quality is relevant here (Federal Committee on Statistical Methodology 2020).

This risk is not hypothetical. High-profile polling failures in recent election cycles have prompted widespread public skepticism about survey research generally (e.g., Kennedy et al. 2018), and organizations associated with poor-quality data have faced serious credibility damage. In an environment where trust in institutions and information sources is already fragile, this threat can occur even when the quality of the data may not have changed, but public perception of the quality of the data changes. For instance, a recent analysis determined that the effect of threats to official statistics through the firing of the BLS Commissioner is reduced trust in official statistics, costing the US economy about $20 billion (Bloom et al. 2026). Lower trust in the quality of estimates, even when no evidence exists that such quality has changed, can yield real economic impact.

Investing in scientifically defendable methods is, among other things, a form of maintaining the scientific integrity of the survey organization or researcher. When a journalist or critic examines a study's methodology, transparency about methods is critical (see Hackett 2026 for examples of opt-in surveys without clear methodological statements). Beyond this transparency, having methods that have yielded higher-quality statistics (e.g., more accurate, more reliable) over repeated studies (e.g., a probability-based sample, a transparent weighting approach, a published methods statement, and a fielding process that can withstand scrutiny) has its own value in articulating the quality of the work. These conditions allow published findings to be defended with known scientific and statistical frameworks underlying them. The value of methodological credibility is real, as is the cost of absence of transparency.

### 5.5 The Ethical Responsibility to Respondents and Other Data Subjects as People

Survey researchers abide by codes of ethics across multiple disciplines. An important part of conducting research on humans is the researchers' ethical responsibilities to those respondents, including maintaining confidentiality and privacy, another measure of survey quality (Federal Committee on Statistical Methodology 2020).

A dimension that is less often discussed in procurement conversations but is extremely important is the ethical value of ensuring that a survey actually represents the people it claims to represent. Survey research, at its best, is an act of listening at scale. It is a mechanism for giving voice to the experiences, attitudes, and needs of populations who might not otherwise be heard in the rooms where decisions are made. For instance, health systems surveying patients about their care experiences, foundations measuring community attitudes toward civic institutions, and media organizations polling the public on a pressing policy issue all want data that provide insights about real people. Surveys that contain real respondents and represent a target population carry weight beyond their statistical quality, providing relevance and comparability to other studies. Protecting the identities of those individuals is critical because surveys collect information from real people, families, businesses, and organizations.

Low-quality research methods undermine that claim in ways that matter beyond methodology. An opt-in panel that over-represents highly engaged, internet-savvy respondents does not capture the full range of human experience in the population of interest, but instead amplifies some voices while omitting others. These people may be those who are historically underrepresented in data, policy, and public life. Increasingly, opt-in survey panel data are filled with "bogus respondents" (Kennedy et al. 2024) who provide answers with the goal of screening into surveys, largely to make money, bots, or AI-generated responses. These studies distort the perspective that is shared, failing to be grounded in real experiences.

Probability-based research, by contrast, is grounded in a commitment to representativeness that is fundamentally democratic. When every person in a population has a known, positive chance of being included in a study, the research treats all members of that population as worthy of being heard. This inclusion of all voices allows survey research to be done responsibly.

### 5.6 Thinking About the Value of Survey Cost Information

Contemplating these issues of integrity as a measure of quality does not mean that every research question requires the most expensive possible approach, or that budget considerations should be set aside entirely. Research designs involve genuine tradeoffs, and experienced survey designers can help researchers identify the methods that best match their needs, timeline, and resources. But the value of data is not defined by "what is the cheapest way to get an answer," but rather what we will need to be done with the data. That is, costs should be evaluated in line with fitness for use (Biemer and Lyberg 2003).

When considering costs relative to fitness for use, researchers could ask questions such as: How confident do we need to be in this data before we act on it? What are the consequences — for an organization's decisions, reputation, and the people or communities that are served — if the data are wrong? And does the research we commission reflect the respect and care that ethical responsibilities require to the people whose voices the study is trying to understand?

Survey research is an investment in knowing something about the world, hopefully yielding insights that could come from no other method. The cost of knowing something about a population, with confidence, in a way that has methods that are transportable and reproducible, and in a way that respects the people behind the data, has value beyond the cost of the data themselves.

***Takeaway #10: The value of surveys or a statistical production system extends beyond the cost of the data collected. The value of survey research should be viewed in terms of both its immediate utility and its broader and longer-term societal benefits.***

## 6. Recommendations

We now turn to some summary recommendations to advance our collective understanding of survey costs.

*Develop and use standardized definitions and measures.* Implementation of the ideas in this white paper across researchers and organization will benefit from the use of standardized definitions and measures for cost components. Although this white paper has presented some definitions to address this goal, these general definitions provide little specific guidance to researchers on what to do. This task is larger than this white paper can accomplish. A first step could be to compare definitions and measures across multiple empirical studies. An additional step could compare definitions and measures used for surveys across organizations in academia, government and the private sector, and evaluate how these align with standard accounting terms. These information gathering steps could be the basis of a longer term project to develop a standard set of definitions for cost measurement, as there are standard definitions for response rates (American Association for Public Opinion Research 2023).

*Connect cost measurement structures with management decisions.* In evaluation of survey cost structures, specify some concrete high-priority management decisions that require cost-structure information. Some management decisions require cost-structure information at an aggregate level. For instance, this might be the relative costs of surveys at an organization conducted in mode A versus mode B to inform the mode selection of a future survey. Other management decisions require detailed and low-level information that will shed light on specific cost-quality trade-offs. For instance, these decisions may be how much of a budget is used to “close out” a lingering case after 10 contact attempts in a telephone survey. Identifying what information would be useful, and whether that information is available, is an important first step.

*Justify decisions on components included in evaluation of aggregate costs.* When evaluating aggregate costs, include practical justification of which cost components were included or excluded. This includes both the infrastructure and fixed costs described in Section 2; and the various activities, materials, supplies, and labor costs attributed to variable cost factors. Also, some organizations encounter costs for overhead, contract development, unfunded mandates, or “enterprise services” that are variously classified in accounting systems as infrastructure, fixed or variable costs. Describing when these costs are included (or not) in reported cost estimates will help researchers and project managers understand how and why costs vary across studies and, if reported publicly, organizations.

*Justify decisions on quality measures used in evaluation of cost-quality trade-offs.* When the survey management decisions center on cost-quality trade-offs, it is important to demonstrate a clear connection between a specific class of survey management decisions and relevant measures of quality. For example, some current assessments of cost-quality trade-offs tend to use quality measures based primarily on response rates or related quantities like R-indicators. Cost-quality trade-offs could examine a wider range of quality measures, including assessment of biases attributable to issues with frame coverage; imperfect record linkage; wave, unit and item nonresponse; and measurement error (e.g., Boonstra et al. 2021; Little et al. 2019; Meng 2018; Schroeder and West 2025). Similar comments apply to cost-quality trade-offs related to instrument design and measurement errors. For example, in the measurement of income, expenditures, dietary intakes or health conditions, one may may need to decide between the use of high-quality measurement methods that are expensive (e.g., diaries; questions that require respondents to use records), but provide data that align closely with the underlying target concepts; or the use of rough proxy measures that are less expensive (e.g., single-item questions), but provide data that are aligned less well with the underlying concepts.

*Evaluate cost-quality trade-offs through decisions on specific design factors.* In many survey settings, decisions on a given set of design factors can affect both cost and quality. Consequently, evaluations of cost-quality trade-offs depend inherently on which specific design factors can feasibly be changed. In addition, relationships between design factors and the resulting cost and quality profiles will generally involve some uncertainty, as reflected in the error terms in expressions (C.1) and (Q.1). Thus, practical discussion of prospective cost-quality trade-offs need to occur in the context of changes in specific design factors and need to account for uncertainties in the associated models.

*Account for environmental factors.* In evaluation of cost-quality trade-offs, identify practical ways in which one can account for environmental factors, as well as customary design factors. This can be of special interest when some environmental factors are considered important for some dimensions of cost and quality, the resulting environmental effects may interact with design factors, and those environmental conditions may not be constant across studies. Examining costs over months of a year, for instance, may be informative. Although aggregate social, economic and political factors have been examined in the context of response rates (Harris-Kojetin and Tucker 1999), less attention has been placed on how these factors affect costs.

*Quantify uncertainty in cost analyses.* In evaluating available empirical information on costs, it may be useful to extract cost data at a level that permits quantitative measures of uncertainty or variation in costs. When uncertainty is measured, survey managers and budget officers can develop methods that use that information about uncertainty to make practical contributions to the development of budgets, to real-time management of the costs of survey operations, and to system development and maintenance that permits regular evaluation of – and possibly improved control over – uncertainties in cost. A high degree of caution is warranted in interpreting cost-structure estimates that have indications of a high degree of uncertainty. For instance, standard errors of reported totals (aggregate costs), means or quantiles (cost-per-case), and regression coefficients (in modeling of survey cost structures) may be able to be calculated if the cost information is extracted in a way that allows for variance measurement or if multiple similar studies are aggregated to form a "sample" of similar studies over which costs may vary. Other relevant indicators of uncertainty include $R^2$ or other measures of how well the independent variables predict the outcome in regression models for costs. This information may be especially important when a researcher seeks to predict and control costs at fine levels of aggregation, e.g., for personal-visit interviews within geographically dispersed primary sample units. Incorporating this uncertainty into survey designs may allow for better control of costs – either by specifying probability statements about potential over-runs or by using risk mitigation strategies, such as planning for design changes based on interim results. Further, making this uncertainty explicit may encourage the development of more precise methods, i.e. better measurement and prediction of costs.

*Despite imperfections, it is necessary to use available information on costs to design and improve surveys.* Methods and approaches exist to measure costs and should be used. Ideally, attempts should be made at differentiating infrastructure, fixed and variable costs and standardizing these across surveys.

*Consider extensions to additional components of the collection, production, dissemination and use of statistical information.* A very limited number of survey components are reported in some cost calculations (e.g., only incentive costs; only printing and postage costs). Broad consideration of the full suite of components can contribute to improved efficiency of each component, and also can help in setting priorities for improvements across the range of components.

*Develop tools to incentivize managers to examine costs and reduce management burden for examining costs.* (Note that this topic is distinct from the empirical study of costs related to incentives offered to survey respondents or sampled units.) During the workshop, some participants noted that within a given survey organization, managers encounter many requests for information, including cost information. Survey managers may not have detailed information readily available and may have little reason to examine cost information that cannot be retrieved from the existing accounting systems, dashboards, or spreadsheets. Ironically, the assembly of cost information may itself be very costly. Consequently, survey managers generally will respond to the requests that are most directly connected with standard management incentives, and that impose relatively low burden on the survey organization. Thus, one may consider concrete ways in which to incentivize managers to explore cost information in ways that will contribute the most to subsequent practical improvement of cost-quality trade-offs. Funders, for example, may include requirements to provide detailed empirical cost information in bids and in subsequent contract reports. Within an organization, upper-level management may ask project managers what decisions they are making using data versus impressions or “experience” as the guide.

In parallel with these management incentives, it may also be necessary to reduce the burden associated with the capture and analysis of survey cost data through the creation of new infrastructure. For example, in keeping with the discussion in Section 3.2, one may seek to develop relatively standardized software (including adaptation of customary spreadsheet tools) for the collection and analysis of survey cost data that can be transported across most studies. In addition, for cases that are novel or involve complex contingencies that are not readily captured through standard accounting and paradata methods, one could ask for initial structured insights from field managers about cost structures, especially in initial pilot work or repeated survey administrations, akin to Bayesian prior-elicitation methods (see Coffey et al. 2020 for an example of this approach for response propensity models). Also, for high-priority cases that require extensive empirical work with survey costs, one may consider different experimental approaches. These experiments may involve a package of large changes (perhaps administered using fractional factorial designs) in which a few dominant design features are considered important to evaluate for their effects on cost and quality (e.g., shifts in mode and frame altogether; changes to computerization). Experiments may also test relatively minor, but important changes in design (e.g., different incentive levels). Experimental tests of relatively small changes may be of special interest for cases in which only incremental changes are feasible in production-scale testing.

*Examine related additional empirical research and meta-analyses of cost information.* All of these recommendations will require additional empirical research and testing. Many current empirical studies of survey costs are specific to a particular study. If enough cost information accumulates using relatively similar measures, meta-analyses of currently available empirical information could be conducted. An initial step for such a meta-analyses could be an in-depth side-by-side comparison of reported results if cost measures are similar enough to be combined in this way. That comparison could be somewhat analogous to the Miller et al. (2020) side-by-side comparison of empirical analyses of nonresponse bias.

*Engage in transparent and actionable communication with multiple stakeholders for cost analyses.* Discussion materials (including graphics, when informative and feasible) designed to report results of cost analyses to different groups of stakeholders would be useful. These groups could be focused on design and implementation of surveys (and other statistical information production) at various levels of scale; survey field operations; general survey management; related work with system design; and funding and oversight decisions. Sharing cost information to feed discussion and enable research among various survey designers and development areas in organizations should be happening. In preparation of those discussion materials, researchers and survey managers should focus especially on communicating in ways that stakeholders will consider transparent and actionable.

## 7. References


Amaya, Ashley, Scott Keeter, Elisa Shearer, and Colleen McClain. 2026. "What Does More Money Buy You in Bias Reduction of Engagement Measures?" Paper presented at Survey Costs Workshop. February 9.

American Association for Public Opinion Research. 2023. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. 10th ed. American Association for Public Opinion Research. https://aapor.org/wp-content/uploads/2023/05/Standards-Definitions-10th-edition.pdf.

Ang, Lyndon, Robert Clark, Bronwyn Loong, and Anders Holmberg. 2025. "An Empirical Comparison of Methods to Produce Business Statistics Using Non-Probability Data." *Journal of Official Statistics* 41 (1): 3–34. https://doi.org/10.1177/0282423X241298243.

Arrow, Kenneth J., and Anthony C. Fisher. 1974. "Environmental Preservation, Uncertainty, and Irreversibility*." *The Quarterly Journal of Economics* 88 (2): 312–19. https://doi.org/10.2307/1883074.

Baer, Justin. 2026. "How Probability-Based Panels Can Transform Federal Evidence Building." Paper presented at Survey Costs Workshop. February 9.

Baker, Reg, Stephen J. Blumberg, J. Michael Brick, et al. 2010. "Research Synthesis: AAPOR Report on Online Panels." *Public Opinion Quarterly* 74 (4): 711–81. https://doi.org/10.1093/poq/nfq048.

Bertoni, Nick. 2026. "Cost Evaluation of a $5 Visible Gift Card versus a Visible $1 Bill in Panel Recruitment Mailings." Paper presented at Survey Costs Workshop. February 9.

Biemer, Paul P., and Lars E. Lyberg. 2003. *Introduction to Survey Quality*. John Wiley & Sons, Inc.

Bloom, Nicholas, Erica L. Groshen, Duncan Hobbs, and Michael R. Strain. 2026. "The Value of Reliable Statistics." Working Paper No. 35135. Working Paper Series. National Bureau of Economic Research, April. https://doi.org/10.3386/w35135.

Bogen, Karen. 1996. "The Effect of Questionnaire Length on Response Rates - A Review of the Literature." *Proceedings of the Survey Research Methods Section, American Statistical Association*, 1020–25.

Boonstra, Philip S., Roderick J. A. Little, Brady T. West, Rebecca R. Andridge, and Fernanda Alvarado-Leiton. 2021. "A Simulation Study of Diagnostics for Selection Bias." *Journal of Official Statistics* 37 (3): 751–69. https://doi.org/10.2478/jos-2021-0033.

Boulet, Cilanne. 2026. "Adjusting and Optimizing Social Survey Designs According to Cost by Mode at Statistics Canada." Paper presented at Survey Costs Workshop. February 9.

Brenzel, Logan. 2026. “Costs of Immunization Survey Methods Costs in LMICs.” Paper presented at Survey Costs Workshop. February 9.

Brown, Shana, and Justin Jones. 2026. “What Varies and What Stays the Same: Cost Modelling for a National Mixed-Mode Study.” Paper presented at Survey Costs Workshop. February 9.

Burger, Joep, Koen Perryck, and Barry Schouten. 2017. “Robustness of Adaptive Survey Designs to Inaccuracy of Design Parameters.” *Journal of Official Statistics* 33 (3): 687. https://doi.org/10.1515/jos-2017-0032.

Carle, Jill. 2026. “Westat Survey Cost Metrics and Tracking.” Paper presented at Survey Costs Worksho. February 9.

Chardoul, Stephanie, Wen Chang, and Margaret Hudson. 2026. “How Do We Think About and Monitor Survey Costs? Dashboards and Reporting.” Paper presented at Survey Costs Workshop. February 9.

Chardoul, Stephanie, Stephanie M. Coffey, Ashley Kirzinger, and Jordon Peugh. 2026. “Plenary Panel: ‘Why Discuss Costs?’” Paper presented at Survey Costs Workshop. February 9.

Chen, Yilin, Pengfei Li, and Changbao Wu. 2024. “Dealing with Undercoverage for Non-Probability Survey Samples.” *Survey Methodology* 49 (2). https://www150.statcan.gc.ca/n1/pub/12-001-x/2023002/article/00005-eng.htm.

Cochran, William G. 1977. *Sampling Techniques, Third Edition*. Third Edition. John Wiley and Sons.

Coffey, Stephanie, Brady T. West, James Wagner, and Michael R. Elliott. 2020. “What Do You Think? Using Expert Opinion to Improve Predictions of Response Propensity Under a Bayesian Framework.” *Methods, Data, Analyses* 14 (2): 159–94. https://doi.org/10.12758/mda.2020.05.

Cornesse, Carina, Annelies G. Blom, David Dutwin, et al. 2020. “A Review of Conceptual Approaches and Empirical Evidence on Probability and Nonprobability Sample Survey Research.” *Journal of Survey Statistics and Methodology* 8 (1): 4–36. https://doi.org/10.1093/jssam/smz041.

DeMatteis, Jill M., and Tammy Cook. 2026. “Modernizing Multistage Sampling: Cost-Efficient Sample Design for Face-to-Face Household Studies.” Paper presented at Survey Costs Workshop. February 9.

Edwards, Brad. 2026. “Developing Systems for Collecting, Tracking, and Predicting Survey Costs.” Paper presented at Survey Costs Workshop. February 9.

Eisenhauer, Elizabeth Willow, and Daifeng Han. 2026. “Sample Size Determination in Survey Design: Balancing Cost, Precision, and Bias.” Paper presented at Survey Costs Workshop. February 9.

Eltinge, John L. 2023. “Competing Effects of Scale, Scope and Complexity in the Production, Dissemination and Use of Official Statistics.” *Proceedings 64th ISI World Statistics Congress* (Ottawa, Canada). https://www.isi-next.org/media/abstract_files/ottawa-2023_72968dbf2f837793e77245287efc21dc_isi-2023-eltinge-scale-scope_PUOeVq8.pdf.

Eltinge, John L. 2026a. “Cost Structures for the Design and Operation of Statistical Information Systems.” Paper presented at Survey Costs Workshop. February 9.

Eltinge, John L. 2026b. “Ethical Issues in the Dynamic Environment for Public-Stewardship Statistical Information.” *Data Science in Science* 5 (1): 2605771. https://doi.org/10.1080/26941899.2025.2605771.

Endres, Kyle. 2026. “Approaches to Offset Rising Postage Costs?” Paper presented at Survey Costs Workshop. February 9.

Eurostat. 2021. *European Statistical System (ESS) Handbook for Quality and Metadata Reports — Re-Edition 2021*. Publications Office of the European Union. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-021.

Federal Committee on Statistical Methodology. 2020. *A Framework for Data Quality*. FCSM 20-04. Federal Committee on Statistical Methodology.

Ferrara, Amanda M., Gillian E. H. Russell, Valerie Keppenne, Renee A. Kotch, Joshua G. Rosenberger, and Glenn E. Sterner III. 2026. “Changes in Methodology, Costs per Complete, and Response Rates Over a 5-Year Statewide Dual-Frame Random Digit Dial Project.” Paper presented at Survey Costs Workshop. February 9.

Gentry, Robin, Ally Kochenour, and Kim Hawkins. 2026. “The Great Incentive Swap: Why Less Upfront Might Get You More in the End.” Paper presented at Survey Costs Workshop. February 9.

Gillen, Summer, Kristen Conrad, and Chris Jackson. 2026. “Don’t Forget the Weather Report: The Impact of Exogenous Factors on Research Costs.” Paper presented at Survey Costs Workshop. February 9.

Groves, Robert M. 1989. *Survey Errors and Survey Costs*. John Wiley & Sons, Inc.

Groves, Robert M., and Lars Lyberg. 2010. “Total Survey Error: Past, Present, and Future.” *Public Opinion Quarterly* 74 (5): 849–79. https://doi.org/10.1093/poq/nfq065.

Guyer, Heidi, Ruby Johnson, Carlos Macuada, et al. 2026. “How Much Does It Take? A Cost Assessment of Experiments and Recruitment Strategies in the National Dementia Workforce Study.” Paper presented at Survey Costs Workshop. February 9.

Hackett, Conrad. 2026. “Has There Been a Christian Revival among Young Adults in the U.K.? Recent Surveys May Be Misleading.” Beliefs & Practices. *Pew Research Center*, January

23. https://www.pewresearch.org/short-reads/2026/01/23/has-there-been-a-christian-revival-among-young-adults-in-the-uk-recent-surveys-may-be-misleading/.

Hammon, Angelina, and Sabine Zinn. 2025. "Validating an Index of Selection Bias for Proportions in Non-Probability Samples." *International Statistical Review* 93 (3): 499–516. https://doi.org/10.1111/insr.12590.

Hansen, Morris H., William N. Hurwitz, and Max A. Bershad. 1961. "Measurement Errors in Censuses and Surveys." *Bulletin of the International Statistical Institute* 38 (2): 351–74.

Harris-Kojetin, Brian, and Clyde Tucker. 1999. "Exploring the Relation of Economic and Political Conditions with Refusal Rates to a Government Survey." *Journal of Official Statistics* 15 (2): 167–84.

Jablonski, Wojtek, and Job van Exel. 2026. "Improving Survey Quality While Cutting Costs: ESHPM's Panel Initiative." Paper presented at Survey Costs Workshop. February 9.

Jackson, Chris, Summer Gillen, and Kristen Conrad. 2026. "It Takes a Village: A Case Study in Multidisciplinary Interaction in Cost Estimations for Complex Programs." Paper presented at Survey Costs Workshop. February 9.

Karr, Alan F., and Michael Last. 2006. *Survey Costs: Workshop Report and White Paper*. (Research Triangle Park, NC), Technical Report Number 161.

Kaye, Miranda P. 2026. "Hidden Costs Associated with Design Decisions." Paper presented at Survey Costs Workshop. February 9.

Kennedy, Courtney, Mark Blumenthal, Scott Clement, et al. 2018. "An Evaluation of the 2016 Election Polls in the United States." *Public Opinion Quarterly* 82 (1): 1–33. https://doi.org/10.1093/poq/nfx047.

Kennedy, Courtney, Andrew Mercer, Scott Keeter, Nick Hatley, Kyley McGeeney, and Alejandra Gimenez. 2016. "Evaluating Online Nonprobability Surveys." Pew Research Center, May 2. http://www.pewresearch.org/methods/2016/05/02/evaluating-online-nonprobability-surveys/.

Kennedy, Courtney, Andrew W. Mercer, and Arnold Lau. 2024. *Exploring the Assumption That Commercial Online Nonprobability Survey Respondents Are Answering in Good Faith*. 50 (1): 3–21.

Keppenne, Valerie, Amanda M. Ferrara, Renee A. Kotch, and Joshua G. Rosenberger. 2026. "Simple, Yet Effective: Estimating Survey Costs Using Excel." Paper presented at Survey Costs Workshop. February 9.

Kim, Jae Kwang. 2025. "Calibration Weighting for Analyzing Non-Probability Samples." *Journal of Official Statistics* 41 (3): 881–88. https://doi.org/10.1177/0282423X251318104.

Kotch, Renee A., Amanda M. Ferrara, Valerie Keppenne, and Joshua G. Rosenberger. 2026. "Balancing Cost and Data Quality in Field Studies: Strategies, Pitfalls, and Real-World Lessons." Paper presented at Survey Costs Workshop. February 9.

Kronschnabl, Judith. 2026. "A European Perspective on Harmonising Survey Cost Structures: Insights from the Survey of Health, Ageing and Retirement in Europe (SHARE)." Paper presented at Survey Costs Workshop. February 9.

Lancaster. 2026. "Case-Level Effort as a Measure of Survey Cost in Multi-Mode Surveys." Paper presented at Survey Costs Workshop. February 9.

Lewis, Taylor, Nicole Lee, Lauren Scott, Dain Palmer, Tessa Day, and Nik Prachand. 2026. "Are Paper Surveys Worth the Cost? Comparing a Sequential Mixed-Mode Web and Paper Survey Data Collection Protocol with a Web-Only Protocol." Paper presented at Survey Costs Workshop. February 9.

Liao, Dan, and Paul P. Biemer. 2025. "Blending Administrative and Nonprobability Survey Data to Enhance National and Subnational Estimates." *Journal of Official Statistics* 41 (3): 889–94. https://doi.org/10.1177/0282423X251342183.

Little, Roderick J. A., Brady T. West, Philip S. Boonstra, and Jingwei Hu. 2019. "Measures of the Degree of Departure from Ignorable Sample Selection." *Journal of Survey Statistics and Methodology* 8 (5): 932–64. https://doi.org/10.1093/jssam/smz023.

Liu, Zhan, and Richard Valliant. 2023. "Investigating an Alternative for Estimation from a Nonprobability Sample: Matching plus Calibration." *Journal of Official Statistics* 39 (1): 45–78. https://doi.org/10.2478/jos-2023-0003.

Marella, Daniela. 2023. "Adjusting for Selection Bias in Nonprobability Samples by Empirical Likelihood Approach." *Journal of Official Statistics* 39 (2): 151–72. https://doi.org/10.2478/jos-2023-0008.

McRoy, Martha. 2026. "Incentivized Mode Choice Experiments in a Multimode Survey: Results from the Healthy Illinois Study." Paper presented at Survey Costs Workshop. February 9.

Meng, Xiao-Li. 2018. "Statistical Paradises and Paradoxes in Big Data (I): Law of Large Populations, Big Data Paradox, and the 2016 US Presidential Election." *The Annals of Applied Statistics* 12 (2): 685–726. https://doi.org/10.1214/18-AOAS1161SF.

Miladi, Anis, and John Lee Holmes. 2026. "Survey Budget Governance Framework." Paper presented at Survey Costs Workshop. February 9.

Miller, Peter, Tala H. Fakhouri, Morgan Earp, et al. 2020. *A Systematic Review of Nonresponse Bias Studies in Federally Sponsored Surveys*. FCSM-20-02.

Murphy, Joe. 2026. "Monitoring Survey Costs with the Adaptive Management Platform (AMP)." Paper presented at Survey Costs Workshop, Bethesda, MD. February 10.

Olson, Kristen. 2021. “Unpacking the Black Box of Survey Costs.” *Research in Social and Administrative Pharmacy* 17 (7): 1342–46. https://doi.org/10.1016/j.sapharm.2020.08.014.

Olson, Kristen. 2026. “Survey Costs: Gaining a Common Language and Setting a Research Agenda.” Paper presented at Survey Costs Workshop. February 9.

Olson, Kristen, John Stevenson, Nadia Assad, et al. 2024. “Examining Variation in Survey Costs Across Surveys.” *Sociological Methods & Research* 0 (0): 00491241241298914. https://doi.org/10.1177/00491241241298914.

Olson, Kristen, James Wagner, and Raeda Anderson. 2020. “Survey Costs: Where Are We and What Is the Way Forward?” *Journal of Survey Statistics and Methodology* 9 (5): 921–42. https://doi.org/10.1093/jssam/smaa014.

Owens, Mark. 2026. “Ideal Blends of Phone and Online Samples: Polls from Single States.” Paper presented at Survey Costs Workshop. February 9.

Pfeffermann, Danny, and Michael Sverchkov. 2025. “Use of Nonprobability Samples for Official Statistics, State of the Art.” *Survey Methodology* 51 (1): 169–96.

Reardon, Leigh. 2026. “Assessing the Efficacy and Cost-Efficiency of Mailed Invitations for Survey Participants.” Paper presented at Survey Costs Workshop. February 9.

Schroeder, Heather M., and Brady T. West. 2025. “Analyzing Potential Non-Ignorable Selection Bias in an Off-Wave Mail Survey Implemented in a Long-Standing Panel Study.” *Journal of Survey Statistics and Methodology* 13 (1): 100–127. https://doi.org/10.1093/jssam/smae039.

Shuttles, Chuck. 2026. “Habit #1 of Highly Effective Survey Cost Managers: Be Proactive [in Cost & Schedule Monitoring].” Paper presented at Survey Costs Workshop. February 9.

Sjoblom, Micah, Ting Yan, and Ned English. 2026. “Survey Cost Intelligence: Monitoring Field Data Collection Costs through Reports and Key Indicators.” Paper presented at Survey Costs Workshop. February 9.

Stepler, Renee. 2026. “Assessing the Cost Effectiveness of Survey Incentives: Lessons from Two Case Studies.” Paper presented at Survey Costs Workshop. February 9.

Tolliver, Kevin. 2026. “Decomposing Survey Costs: Insights from Empirical Modeling of SIPP Operations.” Paper presented at Survey Costs Workshop. February 9.

Ubartas, Cindy, and Wade Kuseler. 2026. “Challenges in Evaluating Costs by Mode for Social Surveys at Statistics Canada.” Paper presented at Survey Costs Workshop. February 9.

Unangst, Jennifer, Ashley E. Amaya, Herschel L. Sanders, et al. 2020. “A Process for Decomposing Total Survey Error in Probability and Nonprobability Surveys: A Case

Study Comparing Health Statistics in US Internet Panels." *Journal of Survey Statistics and Methodology* 8 (1): 62–88. https://doi.org/10.1093/jssam/smz040.

Wagner, James. 2019. "Estimation of Survey Cost Parameters Using Paradata." *Survey Practice* 12 (1): no pp. https://doi.org/10.29115/SP-2018-0036.

Wagner, James. 2026. "Estimating Cost Savings from Applying a Stopping Rule in a Face-to-Face Survey." Paper presented at Survey Costs Workshop. February 9.

West, Brady T., Roderick J. Little, Rebecca R. Andridge, et al. 2021. "Assessing Selection Bias in Regression Coefficients Estimated from Nonprobability Samples with Applications to Genetics and Demographic Surveys." *The Annals of Applied Statistics* 15 (3): 1556–81. https://doi.org/10.1214/21-AOAS1453.

Wong, Erica. 2026. "An Open Book Approach for a New Birth Cohort Study in the UK." Paper presented at Survey Costs Workshop. February 9.

Wu, Changbao. 2022. "Statistical Inference with Non-Probability Survey Samples." *Survey Methodology* 48 (2): 283–311.

Yan, Ting, Scott Fricker, and Shirley Tsai. 2020. "Response Burden: What Is It and What Predicts It?" In *Advances in Questionnaire Design, Development, Evaluation and Testing*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119263685.ch8.

Yan, Ting, and Douglas Williams. 2022. "Response Burden – Review and Conceptual Framework." *Journal of Official Statistics* 38 (4): 939–61. https://doi.org/10.2478/jos-2022-0041.

ZuWallack, Randy, Thomas Brassell, Adam Kaderabek, James Dayton, and Olivia Saucier. 2026. "Leveraging Complementary Strengths of ABS and RDD in the IM$^2$ Survey Design." Paper presented at Survey Costs Workshop. February 9.