

Methods for Producing Economic Data: Are Surveys Becoming Less Reliable? Evidence, Risks, and a Quality-Assurance Framework for Nigeria and Beyond

Prof. Williams Aminadokiari Samuel Abomaye-Nimenibo Ph.D

Director of Postgraduate Studies, School of Postgraduate Studies, Obong University,
Obong Ntak, Etim Ekpo LGA, Akwa Ibom State, Nigeria

DOI: <https://doi.org/10.51583/IJLTEMAS.2026.150900051>

Received: 29 August 2026; Accepted: 10 September 2026; Published: 07 October 2026

ABSTRACT

Economic data shape fiscal and monetary policy, social protection, investment decisions, credit ratings, development finance, and the evaluation of public programmes. Yet the reliability of the data used for these purposes cannot be taken for granted. This article critically examines whether survey-based economic data are becoming less reliable and, more importantly, identifies the conditions under which survey estimates remain fit for purpose. The paper is an integrative methodological review that reorganises and extends the original manuscript through the total survey error perspective, official-statistics quality frameworks, and recent evidence from household-survey practice, with particular attention to Nigeria. The analysis distinguishes sampling error from coverage error, nonresponse, measurement error, interviewer and mode effects, processing and imputation errors, and analytical or reporting distortions. It argues that declining response and changing communication technologies create genuine risks, but that low response rates do not mechanically imply high bias, and surveys should not be contrasted with “statistical data” because surveys are themselves a major source of official statistics. Evidence from Nigeria illustrates both the continuing value of well-designed probability surveys and the importance of mode-specific validation: recent experimental work finds meaningful differences between phone and face-to-face responses across common economic and welfare indicators. The article proposes a data-quality strategy based on probability sampling where feasible, mixed-mode and responsive designs, computer-assisted collection, transparent metadata, validation against administrative or transaction records, selective integration of geospatial and digital data, explicit treatment of uncertainty, and reproducible analysis. The central conclusion is that surveys are not obsolete; rather, single-source, weakly documented and poorly validated data-production systems are increasingly difficult to defend. Reliability is best achieved through transparent, multi-source statistical systems in which surveys remain a core but continuously audited component.

Keywords: economic data; survey reliability; total survey error; nonresponse; measurement error; data quality; official statistics; Nigeria

INTRODUCTION

Economic data are part of the infrastructure of modern government and markets. Measures of employment, household income, prices, production, consumption, poverty, agricultural output, enterprise activity and expectations influence public budgets, monetary decisions, social transfers, debt assessments, investment appraisals and academic research. When such data are timely and credible, they reduce uncertainty and improve accountability. When they are biased, outdated or opaque, they can misdirect resources and create a false sense of precision.

The concern motivating this paper is therefore important: can economic data produced through surveys still be trusted to the same degree as in earlier periods? The question has become more pressing as traditional household and establishment surveys face rising costs, lower willingness to participate, rapid urban and occupational mobility, changing telephone ownership, internet-based recruitment, security constraints in fieldwork and greater demand for high-frequency indicators. At the same time, administrative records,

payment systems, satellite imagery, mobile-phone data, online prices and other digital traces have created alternatives that were not available to earlier generations of statisticians.

An adequate answer requires conceptual precision. Survey data and statistical data are not opposing categories. Properly designed surveys are one of the principal means by which official statistics are produced. National accounts, labour-force indicators, household welfare measures and business statistics commonly combine surveys with censuses, administrative records and estimation procedures. The relevant question is therefore not whether “statistical data” are inherently more reliable than “survey data,” but whether a particular data-generating process produces estimates with acceptable error, transparency, comparability and fitness for the intended use.

The modern literature evaluates this issue through the total survey error (TSE) paradigm. In that framework, survey quality depends on the combined effect of errors arising from the frame, sample selection, nonresponse, questionnaire and measurement, interviewers and modes, data processing, imputation and estimation (Biemer, 2010; Groves & Lyberg, 2010). This is a major improvement over judging reliability from one indicator alone, such as sample size or response rate. Groves and Peytcheva (2008), for example, show that nonresponse rates are imperfect predictors of nonresponse bias. The substantive issue is whether respondents and nonrespondents differ in ways that are related to the variables being estimated.

The paper has five objectives. First, it clarifies the meaning and sources of economic data. Second, it explains the principal threats to survey reliability. Third, it assesses whether contemporary evidence supports the claim that surveys are becoming less reliable. Fourth, it examines alternative and complementary methods of producing economic data, including administrative, digital and geospatial sources. Fifth, it develops a practical quality-assurance framework, with particular attention to Nigeria and other data-constrained economies.

The contribution is methodological rather than a new primary-data experiment. The original manuscript assembled a broad discussion of primary and secondary research, interviews, observation, survey methods, field studies and data analysis. The present article consolidates those materials around a single research question and updates them with contemporary survey-science and official-statistics literature. In doing so, it also replaces categorical assertions about survey manipulation with a more defensible distinction between error inherent in data collection and misconduct or analytical discretion occurring after data have been collected.

METHODOLOGICAL APPROACH

This study adopts a critical integrative review and methodological synthesis. It does not claim to estimate a new population parameter or a causal effect from an original sample. Instead, it evaluates the reliability of economic-data production by combining (i) established survey-methodology concepts, (ii) international principles for official statistics and data quality, (iii) peer-reviewed evidence on nonresponse and sampling, and (iv) contemporary examples from Nigerian household-survey practice.

The analytical lens is fitness for use. Reliability is treated as multidimensional: an estimate can be precise but biased, unbiased on average but too noisy for policy, accurate but too late to be useful, or methodologically sound but insufficiently documented for replication. The IMF Data Quality Assessment Framework explicitly recognises this multidimensionality through integrity, methodological soundness, accuracy and reliability, serviceability and accessibility, supported by institutional prerequisites (International Monetary Fund [IMF], 2012). The United Nations Fundamental Principles of Official Statistics similarly emphasise relevance, impartiality, professional standards, transparency of methods and confidentiality (United Nations, 2014).

The review therefore assesses data-production methods against six questions: Does the source adequately cover the target population? Is selection into the observed data understood? Are the variables measured consistently and with limited respondent or instrument error? Are processing and adjustment procedures documented? Can the results be benchmarked or validated? Are users given sufficient metadata to interpret

uncertainty and limitations? These questions form the basis of the recommendations developed later in the paper.

Economic Data and the Architecture of Data Production

Primary and Secondary Economic Data

Primary data are collected directly for a defined statistical or research purpose. In economics, they may arise from household or enterprise questionnaires, direct observation, experiments, price collection, farm measurement, diaries, focus groups, interviews or sensor-assisted measurement. Secondary data are data already collected or compiled by another entity and subsequently reused. They include published national accounts, labour statistics, tax records, customs data, financial statements, administrative registers, previous surveys, censuses and research datasets.

The distinction is about provenance rather than quality. Primary data can be poorly designed, while secondary data can be exceptionally reliable if they originate in strong statistical systems. Conversely, administrative datasets may have near-complete coverage of people who interact with a programme but omit those outside the programme and may reflect rules designed for administration rather than statistical concepts. Data users should therefore document who generated the data, for what purpose, from which population, with what definitions, and through which transformations.

Surveys as Statistical Instruments

A survey is a structured method for obtaining information from a population or a sample. Probability sample surveys assign known non-zero selection probabilities to units in the target population and support design-based inference when the frame and implementation are adequate. Nonprobability surveys recruit units through mechanisms whose selection probabilities are unknown, so population inference depends more heavily on modelling assumptions and post-survey adjustment.

Probability sampling remains the clearest statistical basis for generalisation, although it is not immune to coverage and nonresponse problems. Reviews comparing probability and nonprobability approaches generally find better population accuracy for probability samples, even in an era of declining response (Cornesse et al., 2020). The policy implication is not that every survey must use a single classical mode, but that the path from target population to final estimate should be explicit enough to evaluate selection and uncertainty.

Other Sources of Economic Data

Administrative records arise from the routine operation of government or organisations: taxation, social protection, education, health, pensions, licensing, customs, payroll and regulatory systems. Their strengths include continuity, lower marginal collection cost and, in some domains, extensive coverage. Their limitations include concept mismatch, policy-induced breaks, missing populations, duplicate identities, weak interoperability and access restrictions.

Digital and transaction data include electronic payments, scanner data, online prices, platform activity, call-detail records and other digitally generated observations. Geospatial and remote-sensing data can measure land cover, rainfall, night-time lights, crop conditions and accessibility. These sources can improve timeliness and geographic granularity, but they are not automatically representative of people or firms. Their statistical use requires careful linkage, calibration, privacy protection and validation.

Accordingly, the strongest contemporary statistical systems increasingly use data integration rather than source substitution. Surveys can provide rich socioeconomic concepts and variables absent from administrative systems; administrative records can benchmark survey responses and reduce respondent burden; geospatial data can measure environmental exposures; and digital data can increase frequency. The objective is a coherent statistical architecture in which the limitations of one source are partly compensated by the strengths of another.

Table 1. Main economic-data sources: strengths, risks and appropriate uses

Source	Principal strengths	Principal risks	Best role in an integrated system
Probability surveys	Rich concepts; known sampling design; direct questions; supports subgroup analysis	Coverage error; nonresponse; recall and measurement error; cost	Core source for household/firm characteristics, calibrated to frames and benchmarks
Censuses	Broad coverage; small-area detail; benchmark frames	High cost; infrequent; limited questionnaire depth	Population benchmark and sampling-frame infrastructure
Administrative records	Frequent; low marginal collection cost; may cover complete programme populations	Concept mismatch; policy changes; exclusion of nonparticipants; linkage/privacy issues	Benchmarking, imputation, longitudinal follow-up and burden reduction
Transactions and digital traces	High frequency; behavioural rather than recalled data; large volume	Platform selection; ownership bias; changing algorithms; commercial opacity	Nowcasting and specialised indicators after calibration
Remote sensing/geospatial data	Spatial coverage; repeated measurement; independent environmental observation	Measurement-model error; resolution limits; linkage problems	Environmental, agricultural and infrastructure measurement linked to survey units
Qualitative and observational data	Context, mechanisms and interpretation; useful for instrument development	Limited population generalisability; observer/researcher effects	Exploration, questionnaire design and interpretation of quantitative patterns

Source: Author’s synthesis from the survey-methodology and official-statistics literature.

Why Survey Reliability Can Deteriorate

Total Survey Error and Mean Squared Error

The reliability of an estimate cannot be inferred from a questionnaire alone. Survey error accumulates across the entire production chain. A useful statistical summary is mean squared error (MSE):

$$\text{MSE}(\theta) = \text{Var}(\theta) + \text{Bias}(\theta)^2$$

This expression distinguishes variance, which reflects instability across repeated samples, from bias, which reflects systematic deviation from the population quantity of interest. In practice, different sources of error can interact, so quality management requires more than adding mechanically separable components. Nevertheless, the TSE framework disciplines analysis by asking where error may enter and how each source can be detected or reduced (Biemer, 2010).

Coverage and Frame Error

Coverage error occurs when the sampling frame does not correspond adequately to the target population. Undercoverage may arise where informal settlements, newly created communities, migrants, nomadic populations, unregistered enterprises or households without particular technologies are absent from the frame. Overcoverage occurs when duplicate or ineligible units remain. Rapid demographic change can make a once-strong frame progressively less representative.

Technology-specific surveys create additional coverage problems. A telephone survey may exclude households without reliable phone access; an online survey may exclude those without internet access or digital literacy. Weighting can reduce some disparities when suitable auxiliary variables exist, but it cannot reliably correct selection on unobserved characteristics that strongly predict the outcome.

Sampling Error and Design Effects

Even a perfectly implemented probability survey observes only a sample. Sampling error is therefore inherent, but it is quantifiable. Stratification can improve precision when strata are internally homogeneous, while clustering usually increases variance relative to simple random sampling when sampled units within clusters are similar. Survey weights, unequal selection probabilities and finite-population corrections also affect uncertainty.

A recurrent mistake in applied research is to report ordinary standard errors as if observations were independently and identically distributed. Reliable economic inference requires use of the actual design: strata, clusters and weights. Large nominal sample size does not compensate for an inefficient design, severe weight variability or a frame that excludes relevant parts of the population.

Nonresponse and Attrition

Unit nonresponse occurs when sampled households, firms or persons do not participate; item nonresponse occurs when participants omit particular questions. Panel attrition is nonresponse across waves. These problems have become more visible as survey cooperation declines in many settings. However, response rate is not a direct measure of bias. A survey with a lower response rate can still produce good estimates if response propensities are weakly related to key outcomes after adjustment, while a higher-response survey can be biased if a small but distinctive group is systematically missing (Groves & Peytcheva, 2008).

What matters is the relationship among response propensity, auxiliary variables and the statistic being estimated. Nonresponse adjustment should therefore use available frame variables, paradata, previous-wave data or administrative information, and agencies should publish diagnostics rather than only a headline response percentage. The National Research Council (2013) similarly recommends a broader research and operational agenda encompassing responsive design, incentives, paradata, mixed modes and administrative records.

Measurement, Recall and Social-Desirability Error

Questions can fail even when the correct person is successfully interviewed. Respondents may not understand concepts, may not know the answer, may forget events, may round amounts, or may deliberately withhold sensitive information. Recall periods affect accuracy: a long period may increase forgetting, while a very short period may capture atypical behaviour. Complex household expenditure and income modules are especially vulnerable because economic flows are frequent, irregular and sometimes shared among household members.

Question wording, ordering, response categories and translation also matter. Tourangeau, Rips and Rasinski (2000) describe survey responding as a cognitive process involving comprehension, retrieval, judgement and response. Error can arise at every stage. Good practice therefore requires cognitive testing, piloting, clear definitions, appropriate recall windows and interviewer instructions that do not lead the respondent.

Interviewer and Mode Effects

Face-to-face, telephone, web and self-administered modes can produce different answers even when questions are nominally identical. Interviewers can improve comprehension and completion but can also introduce social-presence and interviewer effects. Self-administered modes reduce some social-desirability pressures but increase the need for literacy and careful interface design. Mixed-mode approaches can improve coverage and control costs, although mode comparability must be tested (de Leeuw, 2005; Dillman et al., 2014).

This issue is especially salient for low- and middle-income countries that expanded phone surveys during and after the COVID-19 period. Recent experimental evidence from rural Nigeria isolates survey-mode effects by comparing phone and in-person measurement across common outcomes. Markhof et al. (2026) report systematic differences across health, labour, shocks, wellbeing and food-security measures, with a median difference of about 17–18 percent between modes. The implication is not that phone surveys should be abandoned; rather, mode itself must be treated as a potential measurement mechanism requiring calibration, instrument adaptation and validation.

Processing, Coding, Imputation and Editing Error

Errors can be introduced after fieldwork through data entry, coding, unit conversions, duplicate handling, inconsistent skip patterns, outlier treatment, imputation and merging. Computer-assisted personal interviewing (CAPI) reduces some traditional entry errors by enforcing ranges, skips and consistency checks at the point of interview. It also permits timestamps, GPS information and paradata that can support quality monitoring. But software does not eliminate conceptual error, and poorly designed validation rules can themselves distort data.

Imputation is often necessary, but users need to know which values are observed and which are imputed, the methods used, and whether uncertainty from imputation is carried into estimates. Meyer, Mok and Sullivan (2015) show that survey-quality concerns can include nonresponse, rising imputation and measurement error simultaneously. Focusing only on field response therefore understates the broader problem.

Researcher Discretion, Selective Analysis and Manipulation

The original concern that researchers can manipulate survey evidence to fit prior beliefs should be separated from ordinary survey error. Fabrication, falsification, selective deletion, undisclosed recoding, outcome switching and selective reporting are research-integrity problems. They are not intrinsic properties of surveys; administrative and digital datasets can also be manipulated or selectively interpreted.

The appropriate safeguards are transparency and reproducibility: preregistration where appropriate, documented sampling and instruments, preserved raw data, audit trails for cleaning, published codebooks, explicit treatment of missing values, reproducible statistical code, disclosure of robustness checks and clear separation between exploratory and confirmatory analyses. These practices reduce the scope for undisclosed analytical discretion and allow reviewers to distinguish genuine data limitations from avoidable research choices.

Table 2. Total survey error: diagnostic questions and mitigation

Error domain	Diagnostic question	Illustrative mitigation
Coverage	Who cannot enter the frame or mode?	Update frames; dual-frame designs; area-based supplementation; calibration to benchmarks
Sampling	Are selection probabilities known and design effects incorporated?	Probability selection; stratification; adequate cluster numbers; survey-design variance

		estimation
Nonresponse	Do respondents differ from nonrespondents on outcome-related characteristics?	Responsive follow-up; incentives; nonresponse weights; paradata; administrative auxiliary variables
Measurement	Can respondents understand, recall and report the concept accurately?	Cognitive testing; shorter modules; diaries/records; validation studies; appropriate recall periods
Mode/interviewer	Would the same person answer differently by phone, web or face-to-face?	Mode experiments; harmonised scripts; interviewer training; bridge samples and calibration
Processing/imputation	How are edits, missing values, outliers and derived variables handled?	Automated checks; version control; flags for edits/imputations; multiple imputation where appropriate
Analysis/reporting	Are uncertainty, weights, exclusions and alternative specifications transparent?	Reproducible code; design-based SEs; sensitivity analysis; complete metadata and disclosure

Source: Author’s synthesis based on Biemer (2010), Groves and Lyberg (2010), and related literature.

Are Surveys Becoming Less Reliable?

The most defensible answer is conditional. Certain indicators of survey health have deteriorated in many settings: cooperation can be harder to secure, interviewer contact more difficult, traditional landline frames less useful, respondent burden greater and measurement of complex income or programme receipt imperfect. Meyer et al. (2015) describe serious deterioration in some household-survey measures, particularly when compared with administrative benchmarks. The National Research Council (2013) likewise treats declining response as a significant challenge to the social-science data infrastructure.

Yet the conclusion that surveys as a class have become unreliable is too broad. First, not all surveys face the same populations, frames, modes or topics. Second, response rate alone is a weak proxy for bias. Third, new technologies can improve some dimensions of quality even as they create new ones. CAPI reduces data-entry delay and permits real-time validation; geolocation and paradata support supervision; administrative linkage can benchmark responses; and mixed-mode contact strategies can recover otherwise hard-to-reach units.

Fourth, alternatives are not error free. Administrative records can be highly accurate for programme transactions yet poor representations of people outside programmes. Digital traces can be abundant but selected. Online opt-in samples can achieve large sample sizes while lacking a transparent probability mechanism. Cornesse et al. (2020) conclude that probability samples generally retain an accuracy advantage over nonprobability alternatives, even under contemporary response challenges.

The correct conclusion is therefore that the error profile of surveys is changing. Traditional assumptions—such as high cooperation, stable frames and equivalence across interview modes—can no longer be taken for granted. Reliability increasingly depends on active quality management, validation, transparent documentation and integration with external information. Surveys remain central because many economic concepts—expectations, informal activity, unpaid work, subjective wellbeing, coping responses, household decision-making and many dimensions of consumption—are not directly observable in administrative or digital systems.

Nigeria as an Illustrative Case

Nigeria demonstrates both the necessity of survey data and the opportunities for modernisation. A large informal economy, diverse livelihoods, regional heterogeneity and uneven administrative coverage make household and enterprise surveys indispensable. At the same time, field costs, security conditions, mobility and infrastructure can challenge conventional collection.

The Nigeria General Household Survey-Panel (GHS-Panel) provides a useful example of a modern probability-based longitudinal platform. The National Bureau of Statistics (NBS), in collaboration with the World Bank's Living Standards Measurement Study, has implemented five waves since 2010. Wave 5 covered 2023–2024 and interviewed 4,715 households across 518 enumeration areas in two visits, post-planting and post-harvest (National Bureau of Statistics [NBS], 2024). The panel structure allows analysts to follow households over time rather than infer all change from separate cross-sections.

Nigerian survey practice has also adopted computer-assisted systems. The 2018/19 Nigeria Living Standards Survey used the World Bank Survey Solutions CAPI platform, allowing completed interviews to be transmitted and reviewed quickly and facilitating consistency checks during fieldwork. Such systems do not guarantee validity, but they materially strengthen supervision, reduce transcription error and shorten the interval between collection and quality review.

The 2026 World Bank experiment on survey mode in rural Nigeria is particularly relevant to the present question. By isolating phone versus in-person modes, Markhof et al. (2026) show that the collection channel can alter reported values across economically important outcomes. This is a caution against assuming that cheaper remote collection is a perfect substitute for face-to-face measurement. It also illustrates the value of methodological experiments: rather than debating mode quality abstractly, statistical agencies can estimate mode effects and design bridge samples, calibration models or topic-specific protocols.

For Nigeria, the broader implication is that strengthening economic statistics should not mean choosing between surveys and administrative data. It should mean building interoperable systems. Household and labour surveys can be linked, under appropriate confidentiality safeguards, to geographic, price, school, health, social-protection, tax or business-register information. Administrative sources can help construct frames and validate survey reports; surveys can measure populations and concepts administrative systems do not capture. Such integration can improve both efficiency and credibility.

A Quality-Assurance Framework for Producing Economic Data

Begin with the decision, not the instrument

Data production should begin with a clearly specified use: national policy, programme targeting, macroeconomic monitoring, academic inference, market analysis or local planning. The required accuracy, frequency, geography and subgroup detail differ across uses. A quarterly macro indicator may prioritise timeliness and revision policy; a poverty assessment may prioritise distributional accuracy and comparability; a small-area allocation formula may require frame completeness and geographic precision.

Once the decision problem is defined, the statistical producer can determine whether a survey is necessary, whether existing administrative sources are sufficient, or whether a combined design is preferable. This “data-needs first” approach helps prevent routine collection of variables that impose burden without clear analytical value.

Prefer probability mechanisms for population inference where feasible

Where the objective is population inference, probability sampling remains the strongest default because inclusion probabilities are known by design. When nonprobability sources are used, the assumptions supporting generalisation should be explicit, and external benchmarks should be used to assess selection bias. Large sample size should never be presented as a substitute for representativeness.

Use mixed-mode and responsive designs strategically

No single contact mode is universally optimal. Face-to-face interviewing can reach populations with limited digital access but is expensive. Telephone collection is rapid and cheaper but can suffer coverage and mode effects. Web collection is efficient for connected populations but can exclude others. Mixed-mode systems should therefore allocate modes strategically and monitor whether estimates shift when mode changes.

Responsive survey design uses information accumulated during fieldwork—contact history, response propensities, costs, interviewer performance and subgroup response—to redirect effort toward cases that matter most for reducing error. This is more efficient than treating every outstanding case identically.

Build validation into data collection

Validation should be planned before fieldwork. Selected survey responses can be compared with administrative records, receipts, direct measurements, GPS-derived areas, transaction histories or re-interviews, subject to consent and legal safeguards. Validation subsamples can identify which variables are most vulnerable to misreporting and can support adjustment models.

For agricultural statistics, for example, self-reported land area can be compared with GPS-based measurement; for programme participation, survey responses can be checked against registries; for prices, enumerator collection can be compared with scanner or online sources where available. The objective is not to declare one source universally “true,” but to understand systematic discrepancies.

Adopt computer-assisted collection with audit trails

CAPI, CATI and web instruments should implement range checks, skips, consistency checks, timestamps and supervisor dashboards. Enumerators should be trained not merely to operate devices but to understand concepts and neutral interviewing. Paradata should be used to identify unusual duration, repeated answer patterns, excessive corrections, suspicious GPS locations or interviewer-specific anomalies.

Every data transformation should be reproducible. Raw data should be preserved securely; cleaned data should retain flags for edits; derived variables should be generated by code rather than manual spreadsheet operations where possible; and release versions should be archived with change logs.

Integrate administrative and new data sources selectively

Administrative and digital sources should be integrated where they improve coverage, timeliness or measurement and where their conceptual relationship to the target statistic is understood. Linkage should use secure identifiers or privacy-preserving methods. Statistical agencies should document linkage rates and linkage bias, because failed matches can create a new form of selection error.

Geospatial and remote-sensing data are particularly promising for agriculture, environment, transport access and disaster exposure. However, spatial anonymisation, resolution, product choice and matching procedures can affect econometric estimates. The use of such sources therefore requires the same quality discipline as conventional surveys.

Publish quality metadata, uncertainty and revisions

Users should be able to see the target population, sample frame, sample design, field dates, mode, response and attrition information, weighting, imputation, editing rules, standard-error methods and known limitations. Official statistics should also have transparent revision policies. Revisions are not evidence of failure when they result from better information; unexplained revisions are a credibility problem.

The United Nations Fundamental Principles place professional independence and transparency at the centre of official statistics, while the IMF DQAF treats accessibility and serviceability as quality dimensions

alongside methodological soundness and accuracy. These frameworks recognise that data cannot be evaluated independently of the institutions and documentation that produce them.

Strengthen research integrity and reproducibility

For academic and commissioned research, a publishable survey study should make clear which hypotheses were specified in advance, how the sample was selected, how exclusions were handled, how variables were constructed and which models were estimated. Data and code should be shared when legal and confidentiality conditions permit; otherwise, controlled-access or replication protocols should be provided. Robustness analysis should test whether conclusions depend on alternative weighting, missing-data treatment, outlier rules or plausible measurement error.

These practices directly address the concern that evidence can be shaped to fit prior beliefs. The remedy is not to abandon surveys, but to reduce opportunities for undisclosed manipulation across all data sources.

DISCUSSION

The central policy lesson is that economic-data quality is a system property. A well-worded questionnaire cannot compensate for a weak frame, and a sophisticated estimator cannot recover information that was never observed without additional assumptions. Conversely, a modestly sized probability survey with good coverage, careful measurement, strong supervision and external benchmarks can be more informative than a massive convenience dataset.

The article also clarifies an important conceptual point in debates about economic evidence. “Official,” “administrative,” “digital,” and “survey” describe different aspects of data provenance; none is synonymous with truth. Official statistics may be survey based. Administrative records may be complete for a programme but irrelevant to those outside it. Digital data may record behaviour precisely while excluding people who do not use the platform. Reliability is therefore empirical and purpose specific.

The question “Are surveys becoming less reliable?” should consequently be reformulated as “Which components of the survey process are deteriorating, for which populations and variables, and what combination of design, validation and auxiliary data can restore fitness for use?” This formulation is more productive because it identifies interventions. If the main problem is noncontact, change contact strategy. If it is mode-related measurement, redesign or calibrate the mode. If it is recall, change the instrument or link records. If it is coverage, improve the frame or add a supplementary frame.

For developing economies, resource constraints make this optimisation especially important. Attempting to reproduce expensive survey systems without considering administrative integration may be unsustainable; relying only on administrative data may omit large informal populations. Nigeria’s statistical needs therefore favour a hybrid strategy: preserve robust nationally representative survey platforms, modernise field technology, improve frames and registries, develop secure data linkage and invest in methodological experiments that quantify rather than assume data-quality trade-offs.

RECOMMENDATIONS

1. Institutionalise total-survey-error reviews. Every major survey programme should maintain a quality profile covering coverage, sampling, nonresponse, measurement, mode, processing and estimation errors rather than relying on response rate alone.
2. Protect professional independence and transparent methods. National statistical systems should align with the UN Fundamental Principles by insulating methodological decisions from political pressure and publishing concepts, methods and revisions.

3. Modernise sampling frames. Population, household and business frames should be updated more frequently and should account for informal settlements, new enterprises, migration and technology-related coverage gaps.
4. Use probability sampling as the default basis for general population inference. Nonprobability sources can supplement official systems, but their selection mechanisms and model assumptions should be disclosed and benchmarked.
5. Expand CAPI, CATI and mixed-mode experimentation. Digital collection should be used for speed and quality control, but agencies should run bridge studies before replacing a face-to-face series with phone or web collection.
6. Benchmark survey variables against independent sources. Administrative records, direct measurements, transactions and geospatial data should be used strategically for validation and calibration rather than assumed to be error free.
7. Strengthen nonresponse and attrition adjustment. Response-propensity models should exploit frame information, prior-wave data and paradata, with sensitivity analysis for nonignorable missingness where material.
8. Make imputation and editing visible. Public-use datasets should include flags and documentation showing which values were edited, imputed or derived, together with the methods used.
9. Invest in statistical capacity and field supervision. Enumerator training, supervisor dashboards, re-interviews, paradata monitoring, secure data management and reproducible code are essential components of data quality.
10. Build interoperable Nigerian data systems. NBS and other public agencies should deepen secure linkage among surveys, administrative registers, geospatial systems and sectoral data while maintaining confidentiality and legal safeguards.
11. Require reproducibility for publishable empirical work. Where confidentiality permits, authors should provide questionnaires, codebooks, analysis code and de-identified data or controlled-access replication arrangements.
12. Report uncertainty honestly. Confidence intervals, design effects, missing-data assumptions, revision policies and material limitations should be presented alongside point estimates so that decision-makers can judge fitness for use.

CONCLUSION

Economic decisions require data that are not only plentiful but credible. Surveys continue to provide information that cannot be obtained adequately from many administrative or digital systems, especially on informal economic activity, household welfare, expectations, coping, intra-household behaviour and multidimensional living conditions. They therefore remain indispensable to economic analysis and official statistics.

However, the environment in which surveys operate has changed. Declining cooperation in some settings, outdated frames, mobility, mode shifts, respondent burden and complex measurement create real threats. These threats do not justify the blanket conclusion that survey data are inherently spurious or inferior to “statistical data.” A survey is a statistical production method, and its reliability depends on design, implementation, validation, estimation and transparency.

The appropriate response is a modern quality architecture. Probability sampling should be preserved where population inference is required; mixed modes should be tested rather than assumed equivalent; CAPI and paradata should support real-time quality control; administrative and geospatial sources should be used for

validation and integration; nonresponse and missingness should be modelled transparently; and published estimates should carry enough metadata for users to evaluate uncertainty. Nigeria's evolving household survey system illustrates both the continued necessity of surveys and the benefits of methodological innovation.

The broader conclusion is therefore constructive: surveys are not becoming obsolete, but weakly governed single-source data production is. The future of reliable economic statistics lies in transparent, multi-source systems in which surveys remain a central, professionally designed and continuously validated instrument.

REFERENCES

1. Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78(2), 161–188. <https://doi.org/10.1111/j.1751-5823.2010.00112.x>
2. Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, 74(5), 817–848. <https://doi.org/10.1093/poq/nfq058>
3. Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. Wiley.
4. Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
5. Cornesse, C., Blom, A. G., Dutwin, D., Krosnick, J. A., de Leeuw, E. D., Legleye, S., Pasek, J., Pennay, D., Phillips, B., Sakshaug, J. W., Struminskaya, B., & Wenz, A. (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>
6. de Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21(2), 233–255.
7. Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method* (4th ed.). Wiley.
8. Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly*, 74(5), 849–879. <https://doi.org/10.1093/poq/nfq065>
9. Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189. <https://doi.org/10.1093/poq/nfn011>
10. Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Wiley.
11. International Monetary Fund. (2012). *Data Quality Assessment Framework (DQAF): Generic framework*. IMF Data Quality Reference Site.
12. Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley.
13. Markhof, Y., Wollburg, P., Palacios-Lopez, A., Castaing, P., Sagesaka, A., & Contreras, I. (2026). The effect of survey mode on data quality: Experimental evidence from Nigeria (Policy Research Working Paper No. 11302). World Bank. <https://doi.org/10.1596/1813-9450-11302>
14. Meyer, B. D., Mok, W. K. C., & Sullivan, J. X. (2015). Household surveys in crisis. *Journal of Economic Perspectives*, 29(4), 199–226. <https://doi.org/10.1257/jep.29.4.199>
15. National Bureau of Statistics. (2024). *Nigeria General Household Survey-Panel (GHS-Panel) Wave 5 (2023/2024): Tracking Nigerian households to understand their resilience over time*. NBS.
16. National Research Council. (2013). *Nonresponse in social science surveys: A research agenda* (R. Tourangeau & T. J. Plewes, Eds.). The National Academies Press. <https://doi.org/10.17226/18293>
17. Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, 35(1), 101–113.
18. Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
19. United Nations. (2014). *Fundamental principles of official statistics* (General Assembly Resolution A/RES/68/261). United Nations.
20. World Bank. (n.d.). *Living Standards Measurement Study: Survey methods*. World Bank.
21. World Bank. (2020). *Nigeria Living Standards Survey 2018–2019: Study documentation and data processing*. World Bank Microdata Library.