SUPPORTING MATERIALS: H

INNOVATION AT FEDERAL STATISTICAL AGENCIES

SUMMARY AND OUTLINE

Agility and a strong culture of and capacity for innovation are essential for federal statistical agencies to serve policymakers and the public.[1] Federal statistical agencies have a strong record of innovation historically and were able for the most part to respond to the demands of the Covid-19 pandemic for relevant data delivered in near real time. *The Nation's Data at Risk: 2024 Report*, the inaugural report of the American Statistical Association's Project to Assess the Health of the Federal Statistical Agencies, documented important ongoing innovations. It concluded, however, that lack of resources and barriers, such as difficulties in sharing data, limited the ability of the statistical agencies to be as agile and innovative as the nation requires today and into the future. The project's latest report, *The Nation's Data at Risk: 2025 Report*, finds that government-wide staff losses and other actions have greatly curtailed the innovation capability of federal statistical agencies. Foundation-supported efforts may help fill the gap for important economic statistics.

This document on statistical agency innovation past and present updates a similar document attached to the 2024 report. The updated document is intended as a stand-alone piece and so repeats and expands on examples and findings from the 2024 and 2025 reports. In it:

- Section H.1 identifies five functional domains for statistical agency innovation and three requirements;
- Section H.2 provides examples of landmark historical innovations by statistical agencies and indicates their current status;
- Section H.3 provides examples of innovation by the statistical agencies during Covid-19 and their current status;
- Section H.4 identifies opportunities for innovation, provides examples of recent innovations of note, and their current status;
- Section H.5 identifies aspects of an organizational culture of innovation at statistical agencies and assesses a subset of them as of 2024 and currently;
- Section H.6 identifies and assesses barriers to innovation as of 2024 and currently;
- Section H.7 describes foundation and non-profit-funded efforts to reengineer key economic statistics to be ported to the relevant statistical agencies; and
- Section H.8 provides findings on innovation from *The Nation's Data at Risk: 2024 Report* and *The Nation's Data at Risk: 2025 Report*.

[1] The Committee on National Statistics added a fifth principle to the 7th edition of its widely cited publication, *Principles and Practices for a Federal Statistical Agency*, which carried over to the most recent (8th) edition (National Academies, 2025, p. 5): "Continual Improvement and Innovation—Federal statistical agencies must continually seek to improve and innovate their processes, methods, and statistical products to better measure an ever-changing world."

References follow the text, and an appendix provides additional examples of innovations and their status: historical innovations (Box HA-1); Covid-19 innovations (Box HA-2); and selected recent innovations (Box HA-3: A [data concepts and topics], B [data collection], C [data processing and estimation], and D [data dissemination]).

H.1 INNOVATION DOMAINS AND REQUIREMENTS FOR FEDERAL STATISTICAL AGENCIES

Innovation is essential to the effectiveness of any organization (see Box H-1 on defining innovation). Federal statistical agencies, to carry out their fundamental responsibilities to provide relevant, timely, frequent, granular, accurate, credible, and readily accessible data for the public and policymakers, need to continually innovate in five functional domains:

1. **Concepts and topics:** to keep abreast of social and economic change and new data needs;
2. **Data collection:** to collect high-quality data as efficiently as possible with the least burden on data providers (people, businesses, other providers);
3. **Data processing and estimation:** to produce relevant data as efficiently and accurately as possible;
4. **Data dissemination:** to ensure users with all levels of expertise and experience can readily find and understand the data they need; and
5. **Data evaluation and testing:** to assess the relevance and accuracy of collected data and experiment with methods to improve and collect new data.

Box H.1
DEFINING INNOVATION

Innovation has been defined as the *practical implementation* [emphasis added] of ideas that result in the introduction of new goods or services or improvement in offering goods or services (see Schumpeter, 1934, *The Theory of Economic Development*). Innovation may, but need not, derive from invention; innovation requires careful testing and piloting that is followed by implementation at scale.

Requirements for statistical agencies to be agile and innovative include:

1. adequate ongoing funding and staff resources available for innovation, including for phasing in changes to long-running programs and for cooperative cross-agency innovations—such resources, particularly of knowledgeable staff, have been severely reduced since January 2025, on top of many agencies' decreasing resources over the past 16 years (see *The Nation's Data at Risk: 2025 Report*, Section 2);
2. a strong culture of innovation (see Section H.5), including sustained relationships with partners outside of government in identifying and pursuing new ways of data collection, estimation, and dissemination; and
3. absence of unnecessary legislative and procedural impediments to data acquisition, IT improvement, and similar functions, and—conversely—procedures to facilitate such functions (see Section H.6).

H.2 LANDMARK INNOVATIONS

Historically and collectively, the principal statistical agencies[2] have a stellar record of innovation in the domains listed above, and many innovations have set standards for private sector and academic data collection and research (see Appendix, Box HA-1, below; see also Duncan and Shelton 1978; National Research Council 2010). Examples of just a few major innovations include:

- probability sampling—the basis of the survey industry worldwide;
- the first nondefense use of computers for the 1950 census and subsequent use of computers to speed up data releases and automate such functions as imputing values for missing data (see Citro 2025);
- small-area estimation for local government statistics (e.g., small-area income and poverty estimates used to allocate billions in Title 1 funding to school districts) (see National Research Council 2000);
- development of the National Income and Product Accounts (NIPAs) (gross domestic product and income, personal consumption expenditures, etc.), which guide fiscal and monetary policy and business investment decisions (see Citro 2025);
- The Nation's Report Card (National Assessment of Educational Progress, NAEP; tests of students on reading, math, and other subjects), which guides federal, state, and local education policy (see National Academies 2022a);
- estimates of crimes not reported to the police (National Crime Victimization Survey, NCVS), which provide a fuller picture of crime victimization (see National Research Council 2008);
- electronic data products for public use (computer summary and public use microdata sample files, online data access platforms), which have dramatically increased the number and range of data users and uses over the past 60 years (see, e.g., Magnuson, Van Riper, and Ruggles 2025).
- secure enclaves, such as Federal Statistical Research Data Centers (FSRDCs), for analyzing confidential data, which have supported important social and economic research (see, e.g., Nagaraj and Tranchero 2023); and
- estimates of the undercount in the decennial census, which provide crucial data for census planning and analysis (see, e.g., National Academies 2023a).

Status as of 2024: All of the historic innovations listed above and in Appendix, Box HA-1 were ongoing, although declining response rates (a worldwide phenomenon) were undercutting the accuracy of probability surveys.

Current Status: Since January 2025, several ongoing historic innovations have been compromised. Examples include the inability of BEA to reimburse the Statistics of Income (SOI) Division in IRS for the latest tax data that are used in the National Accounts; descoping several components of NAEP; deleting data on gender identity and sexual orientation from the NCVS, which are important to understand crime victimization; and suspending the current NCES longitudinal survey, which was to follow 9th and 12th graders through post-secondary education and work experience. See Appendix, Box HA-1, for details.

[2] The 13 principal statistical agencies are: Bureau of Economic Analysis (BEA), Bureau of Justice Statistics (BJS), Bureau of Labor Statistics (BLS), Bureau of Transportation Statistics (BTS), Census Bureau, Economic Research Service (ERS, USDA), Energy Information Administration (EIA), National Agricultural Statistics Service (NASS), National Center for Education Statistics (NCES), National Center for Health Statistics (NCHS), National Center for Science and Engineering Statistics (NCSES), Office of Research, Evaluation, and Statistics (ORES, SSA), and Statistics of Income (SOI, IRS).

H.3 INNOVATION DURING COVID-19

Innovation always requires prioritization, which in turn requires input from stakeholders and data users. During the height of the Covid-19 pandemic, the priority was to obtain and disseminate relevant, accurate data as quickly and frequently as possible, and there were many new and improved datasets and products in response. Some of these innovations could usefully have been implemented long before Covid-19 hit, and some opportunities for innovation were not followed up, but the overall record is nonetheless substantial (see Appendix, Box HA-2, below).

A very timely Covid-19-inspired innovation was the online Household and Small Business Pulse surveys, put into operation in April 2020 (within a month of the nationwide shutdown) by a coalition of statistical agencies led by the Census Bureau. The Pulse Surveys provided weekly updates 10–14 days after data collection on conditions people and businesses encountered during the pandemic, and they continue today in modified form. (NCES, due to limited staff and lack of contracting flexibility, was not able to stand up the School Pulse Panel until the 2021–2022 school year.) Pulse surveys, despite response issues inherent in the methodology, would have been useful pre-Covid-19 and make sense to continue.

Other examples of timely innovation to respond to the data needs of the pandemic included:

- pandemic-specific questions added to ongoing surveys—e.g., remote work added to the monthly Current Population Survey (CPS) by BLS; telemedicine added to the National Health Care Surveys by NCHS;
- increased timeliness—e.g., death statistics with codes for Covid-19 released daily for states and weekly for demographic groups and counties by NCHS; weekly statistics from credit card data on consumer spending by industry (gas stations, clothing stores, etc.) issued by BEA; daily travel based on anonymized cell phone data released weekly with a two-week lag by BTS; and
- websites bringing together Covid-19–relevant data and publications reporting on Covid-19–related topics (e.g., in labor and energy).

Although statistical agencies had made many innovative uses of administrative records prior to and in response to Covid-19, barriers to linking data make such uses difficult and time-consuming to carry out (see Section H.6). One area of missed opportunity has been in the use of electronic health records for timely, granular health statistics, which would have been immensely valuable during the pandemic. Little work, to our knowledge, is going on in this area. In the area of education statistics, NCES administered grants to states to develop linked longitudinal microdata files on students, schools, and teachers through the Statewide Longitudinal Data Systems (SLDS) program (see National Academies 2019). Federal and state laws, however, stand in the way of NCES's ability to use these data for federal statistics, which would have been of great help to understand student outcomes during and after the pandemic.[3]

During Covid-19, the agencies also did their best to produce usable data products from ongoing surveys that employed in-person interviewing—which had to be curtailed during the nationwide shutdown—but could not always succeed. The Census Bureau released a limited set of experimental data collected in 2020 from the American Community Survey (ACS), and for 15 months, NCHS delayed the release of data from the National Health Interview Survey and shut down the National Health and Nutrition Examination Survey, which uses mobile health examination centers (the centers were repurposed as Covid-19 testing sites).

The 2020 Census experienced delays and quality problems, but it is to the credit of the Census Bureau that the census was conducted as well as it was (see National Academies 2023a). Indeed, without the extensive innovations tested and implemented in the 2010s, it would likely not have been feasible to conduct the 2020 Census during the height of the Covid-19 shutdown. These innovations, sparked by a Committee on National Statistics (CNSTAT) report (National Research Council 2011), and continued work by members of a CNSTAT standing committee of expert volunteers, included development of a robust

[3] See American Community Survey Experimental Data, 2020 ACS 1-Year Experimental Data Release; National Health Interview Survey, COVID-19, and Online Data Collection Platforms: Adaptations, Tradeoffs, and New Directions, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8667832/; The National Health and Nutrition Examination Survey (NHANES), 2021–2022: Adapting Data Collection in a COVID-19 Environment, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8667826/.

infrastructure for internet self-response and nonresponse follow-up carried out by smartphone-outfitted field staff, with their routes optimized similar to how UPS and other companies manage peak deliveries during the holidays.

Status as of the 2024 Report: Federal statistical agencies were generally able to respond to the need during the height of the Covid-19 pandemic for almost-real-time data on the impacts of the pandemic on employment, health, and other subject areas. A School Pulse Survey could not be fielded until the 2021–2022 school year, and opportunities to use administrative records for statistics in health, education, and other areas were not and still have not been acted on.

Current Status (see Appendix, Box HA-2, for details): Federal statistical agencies shut down special web pages and other Covid-19–specific initiatives when the national emergency ended. They retained some data series and products that had longer-term utility. BEA's near real-time consumer spending series—begun in June 2020–was discontinued in May 2024 for lack of funding and because of concerns about the source data. The School Pulse Survey has been canceled. BTS also retired its cell-phone-based series on daily travel and suspended its *The Week in Transportation* publication in April 2024.

H.4 OPPORTUNITIES FOR INNOVATION AND SELECTED RECENT INNOVATIONS

Today, there are many areas in which innovation would increase statistical agencies' value-added. Some are:

- *Measuring the economic effects of artificial intelligence (AI), specifically, large language models*—Statistical agencies have begun and will need to expand the provision of timely data on the contributions of AI and related endeavors (e.g., robotics) to employment and the economy.
- *Using AI to enhance statistical agency operations*—Statistical agencies have begun and will need to extend their R&D on ways in which their own processes can benefit from AI.
- *Improving data quality and relevance through blended data*—Declining response rates to surveys (a worldwide phenomenon) are the biggest threat to accuracy of key data series. Expanding the creative and responsible use of multiple data sources—administrative records and private sector data in addition to survey responses—can help improve and maintain high levels of accuracy. Expanding the use of blended data is also a way to increase relevance: for example, by linking surveys and other data, with each source providing unique information.
- *Providing granular data for small geographic areas and vulnerable population groups while protecting personal and business privacy and confidentiality*—Methods exist, but statistical agencies will need to test and implement the most cost-effective approaches at scale.
- *Democratizing data access*—Statistical agencies will need to collaboratively develop robust access mechanisms for their data, more of which are behind "firewalls" because of increased threats to data confidentiality. It is important that access mechanisms serve the broad public in addition to data experts and that agencies are able to better identify and conduct outreach to diverse communities of data users to gain feedback on relevance and timeliness of data and to support evidence-based policymaking.
- *Facilitating data access on broad topics across agencies*—The federal statistical system (FSS) could increase its value to policymakers and the public by finding ways to help users readily locate and access public data within and across agencies. The Interagency Council on Statistical Policy (ICSP) could usefully develop common standards for agencies to provide quality metrics for datasets (e.g., response rates for surveys) and common ways for agencies to cross-reference other agencies' datasets on cross-cutting topics, such as families' social and economic well-being, and to explain differences among series.[4] More generally, agencies could usefully improve the functionality of their websites, develop AI-supported search tools, make data products on similar topics as consistent as feasible, and adopt state-of-the-art standards for documentation (metadata) and accessibility to and preservation of historical data (see National Academies 2022).

[4] The chief statistician's office instituted a few years ago a useful website for posting information about the decentralized federal statistical system: statspolicy.gov. The site includes links to each recognized statistical agency and unit, statistical policy and standards documents, and the like. It is not intended to help users find data across agencies.

- *Increasing efficiency and speeding development of additional innovation*—Implementing shared services and tools through the National Secure Data Service (NSDS) and streamlining approval processes for experimental initiatives are two avenues to facilitate timely innovation.

Work is under way at the agencies in many of the domains and areas listed above. Acknowledging the difficulty of identifying "major" or "path-breaking" innovations contemporaneously (distance is required for such an assessment), we selected for our 2024 report examples provided to us that, in the opinion of the project team, seemed worthy of mention and further elaborated on several of them in our 2025 report (see Appendix, Box HA-3, A–D, below). The examples we chose are listed in Table H-1, along with their current status.

TABLE H-1. Selected Recent Innovations by Federal Statistical Agencies and their Current Status (See Appendix, Box HA-3, A–D, for Details)

Innovation Domain	Brief Description	Statistical Agency	Current Status
Data concepts/ topics	AI Adoption in the Economy	Census Bureau with NCSES (supplements on the ABS and BTOS; R&D expenses for AI on BERD)	Ongoing, but latest ABS supplement is for 2023; much more content needed; ABS & BERD to merge in 2026
	Global Value Chains	BEA with NCSES	Ongoing, but latest estimates are for 2022
Data collection	Facilitating Business Response to Surveys	Census Bureau (inputting company data), BLS (streamlining surveys)	Going slowly, but Census initiative promises important improvements in response and quality
	Facilitating Household Response to the CPS	BLS and Census Bureau (e.g., Internet option)	Going slowly and not far-reaching
Data processing/ estimation (examples of blended data)	Adult Literacy Estimates for States and Counties for 2003	NCES (used models with 1990 & 2000 census data to predict adult literacy from national surveys, with some state samples)	One-off project because no new adult literacy survey since 2003
	Business Formation Statistics for States (monthly/weekly) & Counties (annual)	Census Bureau (links of IRS EIN applications with Census Longitudinal Business Database)	Ongoing (begun in 2018 as experimental product; became regular product in 2021)
	Household Distributions of Personal Income (PI)	BEA (allocating PI to CPS ASEC using CPS and administrative data; "nowcasting" for timeliness)	Ongoing, latest estimates are for 2024; need comparable estimates with Census
	National Experimental Well-Being Statistics (NEWS)	Census Bureau (using other data with CPS ASEC to improve income data)	Ongoing, but latest estimates are for 2021
	National Health Care Survey (NHCS) Linkages	NCHS (linked NHCS with housing, other data)	Latest linkages are of 2016 NHCS with 2020 VA data (also linkages of NAMCS available, latest of 2021 NAMCS with 2022 NDI and HUD data)
	National Secure Data Service (NSDS)	NCSES for ICSP (tool for secure data linkage)	Ongoing as a pilot
Data dissemination	Crop Condition & Soil Moisture Analytics Tool (CROP-CASMA)	NASS, with NASA and George Mason University (satellite imagery, updated daily)	Ongoing, housed at NASA, depends on continuation of NASA satellite system
	Just the Stats	BJS (quick facts)	Ongoing, latest reports issued July 2025
	The Opportunity Project (TOP)	Census Bureau (involves outside organizations in sprints using federal data for practical applications)	Ongoing
	Standard Application Process (SAP)	ICPSR for NCSES and ICSP (tool to make it easier for researchers and other users to locate and apply to use confidential datasets from statistical agencies in a secure setting; established in response to the 2018 Evidence Act)	Ongoing, but several agencies are no longer able to accept applications

NOTE: ABS = Annual Business Survey; BERD = Business Enterprise R&D Survey; BTOS = Business Trends and Opportunities Survey; CPS ASEC = Current Population Survey Social and Economic Supplement; EIN = employer identification number; HUD = U.S. Department of Housing and Urban Development; ICPSR = Inter-university Consortium for Political and Social Research, University of Michigan; ICSP = Interagency Council on Statistical Policy; NAMCS = National Ambulatory Medical Care Survey; NDI = National Death Index.

SOURCE: Compiled by project team.

H.5 CULTIVATING INNOVATION

An essential ingredient for frequent, successful innovation is an organizational culture of innovation. We identified 10 attributes for statistical agencies of an innovative culture that should translate into useful advances. They are:

- Staff have the tools (software packages, computational power) and training (in data science and other relevant areas) they need to innovate;
- Staff are encouraged to present at professional conferences and publish their work, and the agency rewards innovation, not only by staff with primary responsibility for innovating but throughout the organization;
- The agency accepts that not all innovations will pan out and emphasizes what can be learned from failures rather than the failures themselves;
- The agency has a well-specified strategic plan with innovation goals and timetables spelled out;
- The agency regularly obtains outside reviews of major programs and has a track record of implementing a large majority of recommendations from such reviews in a timely manner;
- The agency hosts visiting experts and has regular seminars to hear work of agency staff and relevant outsiders;
- The agency, as a matter of policy, rotates staff among assignments to broaden perspectives and generate new ideas, and agency leadership participates in short-term details to other statistical agencies to broaden perspectives and generate new ideas, particularly for cross-agency collaboration;
- The agency initiates and responds positively to collaborative work with other agencies on mutually beneficial innovation projects;
- The agency's employees believe that the agency is innovative and rewards innovation; and
- The agency proactively reaches out to diverse communities of users to learn where innovations would provide the most value to its stakeholders.

We have data that bear on some but not all aspects of a culture of innovation. For each aspect, we document the status from our 2024 report and currently.

H.5.1. Staff Training

For our 2024 report, we asked agencies specifically about "training in new developments in data science"—a field that incorporates skills from computer science, statistics, data visualization, data integration, graphic design, and systems design to process and analyze large, complex datasets. The responses indicated that many statistical agencies are actively helping their staff acquire or improve updated data science skills and adopting new practices that reflect current academic training and education.

- *Training in Python and/or R* (strong in statistical and visualization tools): BEA (moving entire staff to Python from SAS by 2026); BLS; Census Bureau; NASS; NCHS (for electronic health records analysis); SOI/IRS
- *Data science training more broadly:* BLS; Census Bureau (major program and on-demand); NASS (AI and data visualization); NCES; NCSES (applied data analytics training through Coleridge Initiative)

Status as of 2024: Although some agencies did not indicate data science training opportunities, overall, the record is impressive, given the relative newness of the field (the term "data scientist" was coined in 2008; degree programs are relatively new) and that training budgets are always stretched and often scaled back under budget pressures.

Current Status: Funds for staff training have been sharply curtailed government-wide.

H.5.2 Professional Development through Presentations, Publications, and Rotation of Assignments

Our **2024** report did not specifically address this aspect of a culture of innovation, but we were aware that statistical agency staff regularly presented at professional conferences and other meetings and that many agencies encouraged publication in working papers and journal articles. We did not have information on the extent of staff rotation of assignments. We knew of no short-term details of agency leadership to other statistical agencies, although some agency leaders had previously held positions in other agencies.

As of our **2025** report, presentations have been sharply curtailed government-wide, and significant staff losses have resulted in limited if any time for professional publications. Agency staff are being given new assignments, not as a matter of professional development, but to fill gaps in staffing.

H.5.3 Strategic Plans

Our **2024** report did not specifically address this aspect of a culture of innovation, but we knew that many statistical agencies had strategic plans of varying specificity. The Interagency Council on Statistical Policy (ICSP), chaired by the chief statistician in the Office of Management and Budget (OMB), also had an FSS-wide strategic plan with three broad objectives (see About Us - StatsPolicy): (1) strengthen FSS's roles, responsibilities, and capacity to ensure data quality and accessibility; (2) create shared infrastructure, tools, and services to broaden safe access to data; and (3) build communications and community outreach to strengthen the value of federal data for diverse data provider and user needs.

From **January 2025**, the scale and pace of change have made pre-January 2025 strategic planning efforts to guide agency innovation largely irrelevant. Agency leaders perforce have focused on making the best use of suddenly much scarcer resources, particularly staff resources. Little progress is evident in realizing specific objectives of the ICSP strategic plan—for example, the requirement in the Foundations of Evidence-Based Policymaking Act of 2018 for OMB regulations for a presumption of access by statistical agencies to other federal agencies' data and for expanding secure access to statistical agencies' data.

H.5.4 Advisory Committees

Our **2024** report did not specifically address this aspect of a culture of innovation, but we knew that many agencies had formal advisory committees and informal advisory mechanisms. Our plan had been to assess agencies' track records of responding to advisory committee recommendations.

In spring **2025**, the administration terminated 10 advisory committees organized under the Federal Advisory Committee Act (FACA), representing six of the seven statistical agencies with FACA committees:[5]

- BEA Advisory Committee,
- BLS Data Users and Technical Advisory Committees,
- Census Scientific Advisory Committee, 2030 Census Advisory Committee, National Advisory Committee on Racial, Ethnic, Other Populations,
- BEA, BLS and Census Bureau Federal Economic Statistics Advisory Committee,
- NASS Advisory Committee on Agriculture Statistics,
- NCHS Board of Scientific Counselors, and
- NSF Advisory Committee for Social, Behavioral, and Economics Sciences (which provided input to NCSES).

[5] The National Board for Education Sciences, which provided input to NCES, has not been disbanded but is inactive. Other statistical agencies had no FACA advisory committee or obtained advice through such means as expert groups organized by the American Statistical Association.

H.5.5 New Blood (from the outside):

There are many ways for an agency to refresh its innovation capabilities from the outside —interagency personnel agreements (IPAs), fellowships, and internships bring in junior to senior people for limited but significant periods; cooperative agreements typically bring university scientists together with agency staff to collaborate on projects (as distinct from contracts, which allow less room for experimentation). Fewer than half the agencies, however, reported using these kinds of vehicles in response to our inquiries in spring 2024. Mechanisms from that time included:

- *BEA, BLS, and Census Bureau:* ASA/NSF/BEA (BLS, Census) fellowships (one or a few people per year from academia).
- *Census Bureau:* significant numbers of interns, fellows, and university fellows; senior experts on IPAs for three-year terms as associate director for research and methodology, a senior executive position.
- *ERS:* ~100 cooperative agreements with universities.
- *NASS:* cooperative agreements for research on small-area estimation, precision agriculture, differential privacy, and data integration.
- *NCSES:* participation in the Oak Ridge fellows program and broad agency announcements (NSF BAAs) for research.
- *ORES:* Retirement and Disability Research Consortium (RDRC) cooperative agreements.
- *SOI:* Joint Statistical Research Program pairing academic scientists with staff.

Status as of 2024: Fellows, interns, IPAs, and cooperative agreements are costly for an agency but have resulted in adoption of innovative practices. Greater use of these vehicles by more agencies would likely promote innovation. Giving "reverse fellowships" that place agency staff with research and data user organizations is also valuable, although that has been used less frequently and mostly at larger agencies such as the Census Bureau.

Current Status: Fellow positions, cooperative agreements, and similar arrangements to benefit from outside perspectives have been sharply curtailed government-wide. No ASA/NSF fellowships were offered in 2025; SOI/IRS paused its research program for lack of staff to work with researchers; SSA canceled ORES's Retirement and Disability Research Consortium; the Census Bureau could not renew a contract with the Population Reference Bureau to continue a long-standing (since 2013) ACS Users Forum, with over 7,000 members, which provided invaluable user-to-user assistance and feedback for this widely used set of federal statistical data products.

H.5.6 Collaborative Projects (with other agencies):

- *Interagency entities:* All of the principal statistical agencies have seats on the Interagency Council for Statistical Policy (ICSP) and contributed core support to the Committee on National Statistics (CNSTAT) at the National Academies of Sciences, Engineering, and Medicine. Senior technical staff of BJS, BLS, BTS, the Census Bureau, ERS, NASS, NCES, NCHS, NCSES, and SOI currently sit on the prestigious Federal Committee on Statistical Methodology (FCSM). These entities facilitate cross-agency collaboration and innovation in federal statistics (see Box H-2).
- *Satellite accounts (SAs):*[6] Many SAs were developed by BEA with other agencies. Examples include: Arts and Cultural Production SA (with the National Endowment for the Arts introduced in 2013); Digital Economy SA (with the National Telecommunications and Information Administration; introduced in 2018); Health Care SA (with multiple agencies; introduced in 2015); Marine Economy SA (with the National Oceanic and Atmospheric Administration, NOAA; introduced in 2018); Outdoor Recreation SA (with multiple agencies; introduced in 2018); Travel and Tourism SA (with the International Trade Administration introduced in 1998).

[6] The BEA system of core national accounts and satellite or supplemental accounts has facilitated response to new policy interests and collaborative work with other agencies. Satellite accounts are linked to the main accounts but can provide additional detail and be more flexible in using alternative concepts, accounting conventions, and definitions.

Box H.2

COORDINATING BODIES FOR FEDERAL STATISTICS THAT FACILITATE INNOVATION

Interagency Council on Statistical Policy (ICSP)

Established as advisory to chief statistician in 1989 • Provided for in 1995 reauthorization of Paperwork Reduction Act • Consisted of heads of principal statistical agencies, chaired by chief statistician• Expanded by Foundations for Evidence-Based Policymaking Act of 2018 to include statistical officials in all departments and heads of recognized statistical units (30 members) • Meets monthly, establishes subcommittees as needed: e.g., American Community Survey subcommittee; Standard Application Process (for FSRDCs) Governance Board; FSRDC Executive Committee • ICSP Mentoring Program furthers professional growth of participants and encourages connections across the statistical system. See About Us - StatsPolicy.

Federal Committee on Statistical Methodology (FCSM)

Founded in 1975 by chief statistician's office to assist in setting and coordinating statistical policy • Serves as a resource on statistical policy issues and to provide technical assistance and guidance on statistical and methodological issues • Sponsors regular conferences (in partnership with the Council on Professional Associations on Federal Statistics) • Writes guidance on best methods and practices for data quality measurement, confidentiality protection, other topics • Creates subcommittees and interest groups on relevant topics • Chief statistician appoints FCSM chair • FCSM members are career federal employees selected by OMB based on individual expertise and interest in statistical methods • Members (23 currently) serve for three-year, renewable terms. See FCSM Home.

Committee on National Statistics (CNSTAT)

Established at National Academies of Sciences, Engineering, and Medicine in 1972 • Works to improve statistical methods and information for public policy • Holds convening functions, maintains close contact with chief statistician and ICSP• Releases new edition every four years of *Principles and Practices for a Federal Statistical Agency* (cited in Statistical Policy Directives, by the Government Accountability Office, and on agency websites) • Carries out studies on specific programs and agencies and system-wide studies, including *Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies* (2022) and *Toward a 21st Century National Data Infrastructure* (three reports, 2023–2024). See Committee on National Statistics | National Academies.

- *Distributional statistics on household income and expenditures:* BEA, BLS, and the Census Bureau are collaborating on experimental series that decompose national account aggregates into household distributions—not just averages but medians and levels for the lowest to the highest 10 percent.
- *Statistics with race and ethnicity detail:* SOI and the Census Bureau collaborated on estimates of adjusted gross and taxable income by race and ethnicity. NASS was part of a USDA team to add race and ethnicity detail to statistical series.
- *Expanded data on the workforce:* NCES and NCSES collaborated on a new (first fielded in 2022, with data release in 2024) National Training, Education, and Workforce Survey (NTEWS), conducted by the Census Bureau. BLS and the Census Bureau are collaborating on modernizing the monthly CPS—the basis for the unemployment rate.
- *Global data:* BEA and NCSES are collaborating on estimates of global value chains. BLS and BEA are collaborating on integrated estimates of foreign direct investment in the United States.
- *Conceptually improved poverty measure:* BLS and the Census Bureau collaborated to produce Supplemental Poverty Measure (SPM) estimates beginning in 2011 that update the 1960s methodology used to calculate the annual poverty rate and are published side by side with the official measure. BLS and the Census Bureau implemented improvements to the SPM in 2021.

Status as of 2024: The above are some of the examples of interagency collaborations reported or known to us that have led or promise to lead to improved statistics for public and policy use. Not all collaborations worked as smoothly or as expeditiously as would be ideal, but having different agencies involved undoubtedly helped ensure that important perspectives would be heeded.

Current Status: Some of the above cooperative efforts continue, although not necessarily at the same pace or scale. These include: distributional statistics on income and expenditures; modernizing the CPS; global data; and the SPM.

On the negative side, however, only a handful of agencies that normally contribute core funding to CNSTAT were able to do so for fiscal 2026, and no agencies have requested studies;[7] BEA suspended work on its Digital Economy and Health Care satellite accounts; SOI/IRS is not in a position to use the privacy-protected race/ethnic information developed by the Census Bureau; and NTEWS is on hold.

H.5.7 Staff Viewpoint

The Federal Employee Viewpoint Survey (FEVS) for many years has asked staff to assess their workplaces, work experiences, supervisors, and agencies.[8] Figure H-1 plots responses to Question 32 (percent strongly agreeing that creativity and innovation are rewarded) from 2010 to 2019 for three statistical agencies, BLS, the Census Bureau, and NASS/ERS, together with EPA, NOAA, and three space-defense research agencies: NASA Ames Research Center, Langley Research Center, and Center for Naval Research (CNR).[9] Staff at the space-defense agencies agreed their agency rewards innovation at higher rates than the other agencies (the sharp decline for CNR in 2014 coincides with a significant budget reduction). The other agencies, while at lower rates, did show an upward trend over the period.

Status as of 2024: Acknowledging the difficulty of specifying an "acceptable" or "high" percentage for staff perceptions that their agency rewards innovation, it appeared that the percentage for the statistical agencies (and EPA and NOAA) was reasonable for production agencies compared to primarily research agencies; moreover, it was trending in a positive direction.

Current Status: FEVS has been paused; the latest available data are for 2024.

H.5.8 Data User Viewpoint

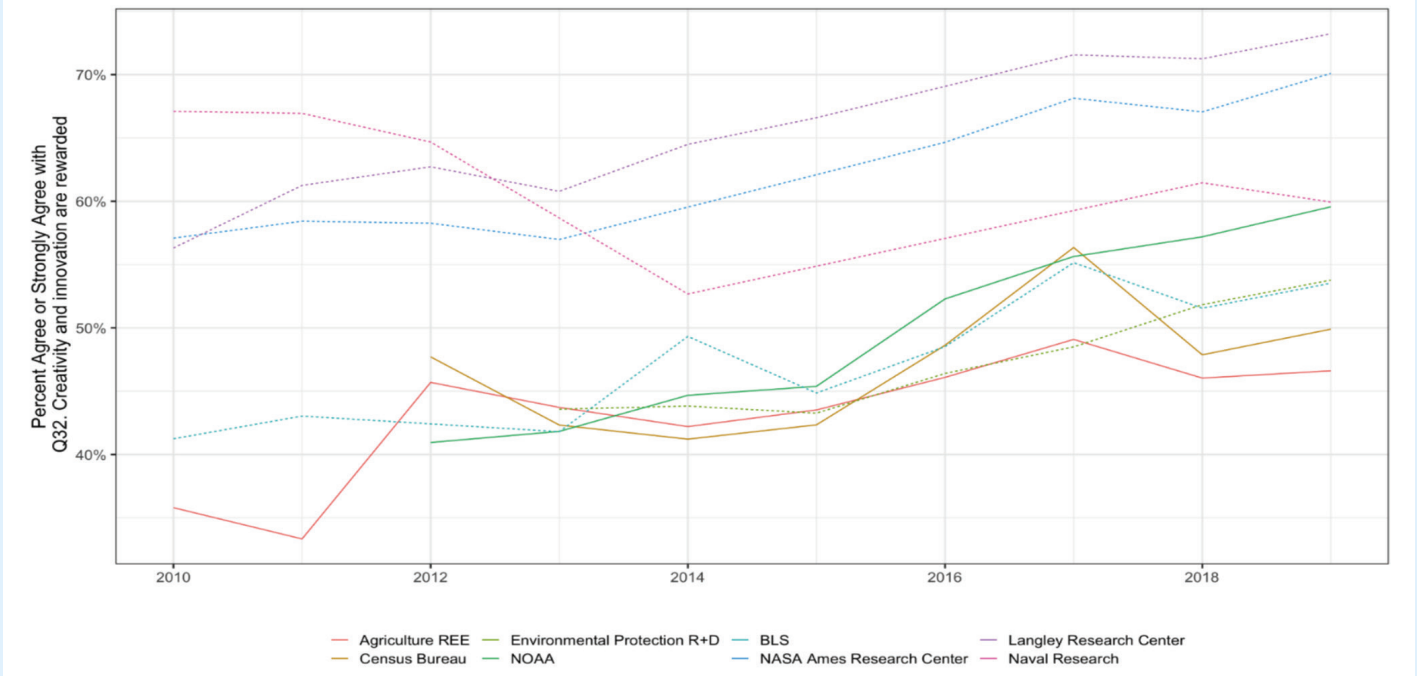
The ASA FedStat Health Project collected original data from users of federal statistics and the general public on their uses of, views about, and trust in federal statistics (see The Nation's Data at Risk: 2025 Report, Section 3). None of the questions asked users specifically to rate the innovation record of federal statistical agencies, but users overwhelmingly valued federal statistics. Respondents to a February 2025 ASA call for feedback from members of relevant professional associations, who were largely expert users, tended to find federal statistics timely, relevant, and essential to their work and to say their efforts would suffer if data were less detailed or less current. When asked to specify useful improvements in federal statistics, respondents provided compliments as well as suggestions to improve data access, documentation, training, websites, the availability of File Transfer Protocol (FTP) sites and Application Programming Interfaces (APIs), engagement with users, cross-agency data finding guides, data granularity, and availability of historical data. Users of confidential data through the Federal Statistical Research Data Center (FSRDC) network wanted reductions in cost and burden (see Supporting Materials: A).

[7] The Department of Homeland Security canceled a project with CNSTAT to recommend improvements to enhance the Office of Homeland Security Statistics' (OHHS's) ability to align with standards and best practices for a federal statistical agency.

[8] FEVS began in 2002 as the Federal Human Capital Survey and was administered to all federal employees every other year. Beginning in 2010 through 2024, FEVS (renamed) was administered annually. It included such questions as: "My work unit has the job-relevant knowledge and skills necessary to accomplish organizational goals" and "Employees are recognized for providing high quality products and services." See the "About" link at FEVS - OPM.gov.

[9] The FEVS data are limited in availability for smaller statistical agencies, and question 32 was discontinued after 2019.

FIGURE H-1. Percentages Who Agree or Strongly Agree that their Agency Rewards Innovation and Creativity, Select Agencies, 2010–2019



SOURCE: Federal Employee Viewpoint Survey (FEVS), Question 32.

H.6. BARRIERS TO INNOVATION

H.6.1 Barriers as of 2024

Despite innovation occurring in many important areas to improve the accuracy, timeliness, relevance, granularity, and accessibility of federal statistics, we identified significant barriers to innovation by the principal federal statistical agencies that limit their current and potential future value-added—

- *Inadequate resources for continuous testing and improvement to long-standing series:* To change long-standing series, such as the monthly CPS measure of unemployment and the continuous Consumer Expenditure Survey, requires extensive testing to determine cost-effective changes that meet stakeholder needs and to prove in new processing systems before they go live. In turn, resources are required for testing, stakeholder dialogues, and production of estimates from the old and new series for a period of overlap, but such resources have historically been lacking. Agency leadership must also overcome inertia and risk aversion among both data users and agency staff. **Making continuous improvement of long-standing series part of a statistical agency's strategic plan, with specific goals and timetables and requesting resources accordingly, seem first steps toward keeping key series up to date. Stable, long-lasting funding for improvement of key series would be an important development—indeed, perhaps a game-changer—to enable sustained progress.**
- *Barriers to data sharing among agencies:* The Confidential Information Protection and Statistical Efficiency Act (CIPSEA), Title III of the Foundations for Evidence-Based Policymaking Act of 2018, authorized sharing of confidential administrative records among BEA, BLS, and the Census Bureau, which is essential to improved statistics about the economy and trends in social mobility, inequality, and economic well-being. The 2018 Evidence Act has useful language that administrative records are to be available to statistical agencies, and the CHIPS and Science Act of 2022 establishes a National Secure Data Service pilot for linkages among datasets for evidence-building research. Key to

such statistics is the use of tax records. At present, selected data from business and personal tax records are available to the Census Bureau for statistical use, but the Census Bureau is not allowed to share any tax information with sister statistical agencies—even addresses of businesses (tax records are the source of addresses for nonemployer businesses). Appropriate access to tax records by statistical agencies and the ability to share tax records among agencies requires legislative changes to Title 26 of the U.S. Code. **Legislation to further permit data-sharing (including tax records) among the principal statistical agencies (which must keep data confidential) is essential.**[10]

- *Difficulties in resourcing infrastructure improvements:* Statistical agencies (like other federal agencies) find it hard to obtain adequate resources for computing technology upgrades and modernization (e.g., moving to the cloud), making it hard to innovate and serve the nations' data needs. **Providing resources for statistical agencies to continuously upgrade their computing technology without taking away from other responsibilities should go without saying. Statistical agencies also need sufficient control over their IT resources to ensure confidentiality protection, the ability to meet data release deadlines, and the ability to serve the needs of their data users.**[11]
- *Insufficient staff in the chief statistician's office:* Sixty years ago, the chief statistician's office had upward of 40 staff; as of 2024, it had 12 staff positions supplemented by staff on short-term details from the statistical agencies. The unit is remarkably productive given its small size, but it lacks capacity to coordinate needed innovation in topic areas (e.g., education, health, labor force, economic well-being) across agencies. **Adding staff with subject matter expertise in different policy areas in the chief statistician's office could help ensure that data gaps are filled and improvements are coordinated across relevant agencies and within OMB.**
- *Challenges to innovation and related data collection updates for smaller agencies, especially as measured by staff size:* The 13 principal statistical agencies' budgets varied for FY24 from $1.4 billion to $30 million, with corresponding differences in staffing levels. The smaller agencies are constrained in what they can dedicate for staff training, outside expertise and perspectives, and other activities essential to innovation. It could help if, for example, the Census Bureau's (the largest federal statistical agency) training budget was mandated and resourced to serve all of the statistical agencies, or if a Federally Funded Research and Development Center (FFRDC) could be funded to do the same. **Finding mechanisms to enable all of the statistical agencies to stay up to date in software tools, statistical methods, and the like is essential.**

H.6.2 Barriers as of *The Nation's Data at Risk: 2025 Report*

- The barriers to innovation identified in Section H.6.1 above remain.
- The chief statistician's office has lost staff.
- In addition, many departments have imposed an onerous process of secretarial review of contracts of a certain size (generally, $100,000 or more), which has delayed renewal of software licenses, for example.

H.7 FOUNDATION-FUNDED EFFORTS TO REENGINEER KEY FEDERAL ECONOMIC STATISTICS

Several nonprofit and foundation-funded efforts promise to help fill in the gap in the short and longer runs, focused primarily on economic measurement. As one example, in March 2024, the American Enterprise Institute, Stanford University's Digital Economy Lab, and New York University convened a seminar, New Approaches to Characterize Industries: AI as a Framework

[10] In contrast, Statistics Canada, which combines the equivalents of BEA, BLS, and the Census Bureau, has authority to use administrative data from any other agency for its household and business surveys. Statistics Canada has used administrative records to reduce respondent burden and improve accuracy, for example, in the Canadian Income Survey (see Statistics Canada Policy on the Use of Administrative Data Obtained under the Statistics Act and Surveys and statistical programs - Canadian Income Survey - 2021 (CIS).

[11] The OMB "Trust" regulation, "Fundamental Responsibilities of Recognized Statistical Agencies and Units" (issued as required by the 2018 Evidence Act on October 11, 2024, with an effective date of December 10, 2024), includes language requiring parent agencies to support statistical agencies' abilities to produce relevant, timely, accurate, and objective statistics and protect respondents' confidentiality. The language specifically refers to computing technology, including websites and software. See Federal Register :: Fundamental Responsibilities of Recognized Statistical Agencies and Units.

and Use Case. The goal was to explore a theoretical framework to identify what data, definitions, models, and tools exist or could be developed to measure the effect of AI on the economy and workforce. The workshop was attended by experts from universities, state labor and workforce agencies, private job market companies, and BEA. The workshop concluded that a new classification system needed to be developed via collaboration between government and industry and implemented by statistical agencies and researchers. Another takeaway from the workshop was a model for such work in the form of an innovative, independent, nonprofit institution dedicated to producing the tools and insights for businesses, workers, and government regarding the effects of AI, to better understand the effects of new technologies on jobs, skills, and economic opportunity. The resulting tools and classifications should inform and complement federal statistical agencies operations and include collaborative fellowships, training, and competitions.

Two such collaborative efforts that reflect the power of this type of model are: 1) the Re-Engineering Statistics using Economic Transactions (RESET) project, begun in 2017 as a collaboration of the University of Maryland, University of Michigan, and Census Bureau with funding from the Alfred P. Sloan Foundation; and 2) the just-funded (by NSF) Economic Measurement Research Institute (EMRI) at the National Bureau of Economic Research. (See Box H.3 for details; note that RESET is also receiving funding from EMRI.) The U.S. Chamber of Commerce Foundation is looking to expand several initiatives (e.g., Jobs and Employment Data Exchange, JEDx) to standardize and improve education and workforce data through partnerships involving federal statistical agencies, state agencies, businesses, and researchers.[12] These efforts are welcome and a possible blueprint for ongoing collaboration across the system. Their ultimate success will depend on adequate resources for the relevant statistical agencies to actively participate with academic researchers, state agencies, and the business community during the R&D phase and to integrate research data series into the agencies' ongoing operations during an implementation phase.

Box H.3

TWO FOUNDATION-FUNDED INITIATIVES TO REDESIGN KEY ECONOMIC STATISTICS

RESET: From the project's website (RESET: Re-Engineering Statistics using Economic Transactions): "The Re-Engineering Statistics using Economic Transactions (RESET) project [of the University of Maryland, University of Michigan, and U.S. Census Bureau] aims to provide the architecture for re-engineering official economic statistics—literally to build key measurements such as GDP and consumer inflation from the ground up. The new measurement architecture offers internally consistent real expenditure and inflation measures that adjust for product turnover and product quality change at scale. It builds up measures of inflation and spending from granular, item-level transactions data. It therefore engineers statistics directly from the information systems of firms….To implement the architecture, the project and ultimately the statistical agencies will partner with firms to create a pipeline from business information systems to statistical agencies to aggregate data for official statistics."

EMRI: From the project's grant abstract (NSF Award Search: Award # 2537470): "The Economic Measurement Research Institute (EMRI) [at the National Bureau of Economic Research] will support and catalyze research focused on improving official economic statistics for the country through systematic use of the data generated by the 21st century information economy. The methodologies, techniques, and tools developed by the EMRI will enable more timely, accurate and granular measures of key economic outcomes including prices, productivity, output, wages, and employment….The EMRI will support demonstration projects that show how the production of official economic statistics can be modernized. In its first phase, the EMRI will support research in four areas, (1) how measurement of retail spending and inflation can be redesigned to make use of item-level transactions data to value technologically driven and other quality changes embedded in retail goods, (2) integration of administrative and American Community Survey data to produce new statistics on the gig economy, (3) use of new information on income statements of businesses to lay the groundwork for improving the measurement of intangible capital to account for investments in R&D, and (4) linking of data from the NSF Business R&D and Innovation Survey to other business data housed at the Census Bureau to generate new estimates of the contribution of R&D to productivity growth in manufacturing."

[12] See https://events.uschamberfoundation.org/newdataparadigm/; also: https://workshift.org/national-initiative-aims-to-make-sense-of-jobs-data-and-asks-others-to-join/.

H.8 FINDINGS

H.8.1 *The Nation's Data at Risk: 2024 Report* (p. 72)

FINDING 7: The principal federal statistical agencies have a rich history of meeting the nation's data needs through innovation—in concepts, collection, processing and estimation, dissemination, and evaluation (e.g., the first nondefense use of computers for the 1950 Census). Overall, they rose to the occasion when the Covid-19 pandemic called for new data delivered promptly. They continue to innovate but not at the level needed, and external and internal barriers, if not addressed, will leave them behind at a time when the demands for more timely, accurate, and granular data are growing every day.

H.8.2 *The Nation's Data at Risk: 2025 Report* (p. 50)

FINDING 4.1: As our 2024 report noted, the principal federal statistical agencies have a rich history of meeting the nation's data needs through innovation—in concepts, collection, processing and estimation, dissemination, and evaluation (e.g., the development of cost-effective probability sampling methods). Their record of responding to the need for new data delivered promptly on social and economic conditions at the outset of the Covid-19 pandemic was substantial. They continued to innovate in the post-Covid years but not at the level needed. Currently, the agencies face declining resources, particularly the knowledgeable staff required for innovation (trends that predate but have accelerated under the current administration). Known barriers to innovation persist (e.g., lack of data sharing legislation). Moreover, since January 2025, many of the requirements for a culture of innovation such as outside expert advice and the ability of agency staff to interact with others in their profession have been undermined.

FINDING 4.2: Foundation-funded efforts are enabling a model for improving federal statistics by involving academic researchers, state agencies, and the business community in developing major innovations in economic statistics for ultimate implementation by the federal statistical system. These efforts are welcome and a possible blueprint for on-going collaboration across the system. To be most useful, the relevant statistical agencies need resources to actively collaborate with partners during the R&D phase and to integrate research data series into the agencies' ongoing operations during an implementation phase. Comparable efforts are needed in other areas, such as health and education statistics.

FINDING 4.3: Federal statistical agencies lag in adopting state-of-the-art standards and tools system-wide to facilitate data access and use within and across agencies' websites and data products. Agencies differ widely in website functionality, search and analysis tool capabilities, metadata content and format, and accessibility to historical data. They rarely link to other agencies' data on similar products or work to make those products more consistent. Allocating a modest amount of innovation funds and staff time to these areas (such as occurred to set up the Standard Application Process mandated in the Evidence Act for access to confidential data in a secure environment) could pay large dividends in expanded use of federal statistics for policy and public understanding.

REFERENCES

Citro, C. F. (2025, May). Challenges in Measuring Household Income and Poverty: Why Is It So Hard? Why Is It So Important? 32nd Morris Hansen lecture, September 26, 2024. 2024Citro_paper.pdf.

Duncan, J. W., and Shelton, W. C. (1978). *Revolution in United States Government Statistics*, 1926-1976. Office of Federal Statistical Policy and Standards, U.S. Department of Commerce. https://books.google.com/books?id=pglFRByHurYC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

Magnuson, D. L., Van Riper, D., and Ruggles, S. (2025, November 22) "What do people want? They want everything": Data Delivery Systems, DUALabs, and the U.S. Census Bureau, 1962-1982. Social Science History Association, Chicago. 250070.

Nagaraj, A., & Tranchero, M. (2023). "How Does Data Access Shape Science? Evidence from the Impact of U.S. Census's Research Data Centers on Economics Research." University of California, Berkeley. https://static1.squarespace.com/static/6232322165ae714579213636/t/64764a7edefa366483bb6964/1685473931144/0_FSRDC_paper.pdf.

National Academies of Sciences, Engineering, and Medicine. (2018). *Reengineering the Census Bureau's Annual Economic Surveys*. Washington, DC: National Academies Press. https://doi.org/10.17226/25098.

National Academies of Sciences, Engineering, and Medicine. (2019). *Monitoring Educational Equity*. Washington, DC: National Academies Press. https://doi.org/10.17226/25389.

National Academies of Sciences, Engineering, and Medicine, (2022a). *A Pragmatic Future for NAEP: Containing Costs and Updating Technologies*. Washington, DC: National Academies Press. https://doi.org/10.17226/26427. https://doi.org/10.17226/26360.

National Academies of Sciences Engineering, and Medicine. (2022b). *Transparency in Statistical Information for the National Center for Science and Engineering Statistics and All Federal Statistical Agencies*. Washington, DC: The National Academies Press.

National Academies of Sciences, Engineering, and Medicine. (2023a). *Assessing the 2020 Census—Final Report*. Washington, DC: National Academies Press. https://doi.org/10.17226/27150.

National Academies of Sciences, Engineering, and Medicine. (2023b). *An Updated Measure of Poverty: (Re)Drawing the Line*. Washington, DC: National Academies Press. https://doi.org/10.17226/26825.

National Academies of Sciences, Engineering, and Medicine. (2025). *Principles and Practices for a Federal Statistical Agency, 8th Edition*. Committee on National Statistics. Washington, DC: National Academies Press. https://doi.org/10.17226/27934.

National Research Council. (2000). *Small-Area Estimates of School-Age Children in Poverty: Evaluation of Current Methodology*. Washington, DC: National Academies Press. https://doi.org/10.17226/10046.

National Research Council. (2008). *Surveying Victims: Options for Conducting the National Crime Victimization Survey*. Washington, DC: National Academies Press. https://doi.org/10.17226/12090.

National Research Council. (2010). *Envisioning the 2020 Census*. Washington, DC: The National Academies Press. https://doi.org/10.17226/12865.

National Research Council. (2011). *Change and the 2020 Census: Not Whether But How*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13135.

National Research Council. (2013). *Measuring What We Spend: Toward a New Consumer Expenditure Survey*. Washington, DC: National Academies Press. https://doi.org/10.17226/13520.

APPENDIX (TO SUPPORTING MATERIALS: H)
STATISTICAL AGENCY INNOVATIONS: ADDITIONAL EXAMPLES

This appendix provides more complete lists of the following: historical innovations of federal statistical agencies (Box HA-1); Covid-19 innovations (Box HA-2); and selected recent innovations (Box HA-3, A–D). Current status is indicated where relevant.

Box HA-1
HISTORIC STATISTICAL AGENCY INNOVATIONS

(Note: Current Status indicated where relevant)

National Income and Product Accounts (NIPAs)—BEA

Developed by NBER and the Commerce Department • 1934: National Income, 1929–1932 • 1947: First full NIPAs • Complete, consistent double-entry accounting system • Transactions among households, businesses, government, and international • Monthly personal income and expenditures and quarterly GDP are principal federal economic indicators • Satellite accounts used for more detail and experimental methods for particular sectors (e.g., Digital Economy)

More information: Concepts and Methods of the U.S. National Income and Product Accounts; Incorporating Satellite Accounts

Current status: BEA is not able to use SOI tax data as part of its annual NIPA updates because of lack of funding to reimburse SOI.

National Assessment of Educational Progress (NAEP, aka "The Nation's Report Card")—NCES

First fielded on trial basis in 1969 • "Long-term trend NAEP" uses consistent content to assess students ages 9, 13, and 17 (every 4 years from 1971 [reading]–1973 [math]) • "Main NAEP" changes content every 10 years to keep up with curricula and assesses 4th, 8th, and 12th graders (varying frequency from 1990 [math]–1992 [reading]) • Main NAEP sample increased in 1990 to provide state estimates • Congress in 2002 funded selected urban school districts to participate in main NAEP • Additional subjects assessed periodically (e.g., science, art, U.S. history)

More information: The Nation's Report Card

Current status: Scope of main NAEP reduced for next 8 years; Long-term trend NAEP cancelled for 17-year-olds.

National Crime Victimization Survey (NCVS)—BJS

First fielded 1973 • Crimes reported by victims (more than reported to police) • School Crime Supplement added 1989 • Police Public Contact Supplement 1996 • Supplemental Victimization Survey 2006 • Identity Theft Supplement 2008 • Supplemental Fraud Survey 2017

More information: National Crime Victimization Survey (NCVS) | Bureau of Justice Statistics

Current status: Questions on gender identity and sexual orientation deleted beginning with the 2025 round.

Supplemental Poverty Measure (SPM)—Census Bureau with BLS

Built on landmark CNSTAT report, *Measuring Poverty: A New Approach* (National Academies, 1995) • New York City adopted report's approach 2008 • Interagency Working Group issued guidance for SPM 2010 • First published 2011 • Official poverty measure (OPM) accounts for earnings, property income, cash transfers • SPM adds taxes, tax credits, in-kind benefits • SPM (not OPM) showed impact of expanded Child Tax Credit and its expiration on child poverty • Census Bureau and BLS implemented improvements to the SPM in 2021 (the 1995 report recommended 10-year reviews of the SPM and changes made as needed) • CNSTAT report (National Academies 2023b) recommended further improvements to the SPM

More information: Supplemental Poverty Measure

DATA COLLECTION

Probability Sampling—Census Bureau

Obtains accurate data at substantially less cost and respondent burden than a census • Introduces error (sampling variability), but that error can be estimated • First used by the federal government in the 1930s • Proved its worth in 1937 Census of Unemployment when a 2% sample gave a better estimate than an attempt to reach everyone • Used for six new questions in the 1940 Census, two-fifths of questions in 1950 Census • 1960–2000 censuses asked most questions on "long-form" sample, which became the American Community Survey in 2005 • Census Bureau statisticians made seminal contributions to theory and practice • Ubiquitous in the public and private sectors, although declining response undercuts value

More information: Revolution in United States Government Statistics, 1926-1976 - Joseph W. Duncan, William Chastain Shelton

Cognitive Questionnaire Testing—NCHS, BLS, Census

New field introduced in 1980s by survey methodologists working with psychologists • Jump-started by interdisciplinary workshop convened by CNSTAT • Known as cognitive aspects of survey methodology (CASM) • Involves one-on-one work with respondents, focus groups, and similar means to determine respondents' understanding of a questionnaire, which may not square with the intended meaning • NCHS established first statistical agency cognitive laboratory for questionnaire testing and improvement in 1985, followed by BLS in 1987, and the Census Bureau in 1988 • Other agencies and private sector surveys regularly engage in CASM-type questionnaire testing

More information: Interagency Collaboration among the Cognitive Laboratories: Past Efforts and Future Opportunities

Longitudinal Surveys—BLS, NCES

Longitudinal surveys follow samples of people over time to measure individual development • They contribute unique research insights and survey methods innovations (see Box H-A1.1) • BLS began its National Longitudinal Surveys of Labor Force Behavior (NLS) in 1966–1968 with cohorts of young men (1966–1981), older men (1966–1990), older women (1967–2003), and young women (1968–2003) • Youth cohorts begun in 1979, 1994, and 1997 are continuing • NCES began its longitudinal surveys of students in 1972 with NLS-72 (high school seniors, 1972–1986) and HS&B (high school sophomores and seniors, 1980–1993) • These and subsequent surveys measure students' academic, social, emotional, and physical development and characteristics of their homes, classrooms, and schools

More information: NLS Cohorts (Active) : U.S. Bureau of Labor Statistics; Introduction to the NCES Longitudinal Studies: 1972-2020

Current status: The latest NCES longitudinal cohort, HS&Beyond:2022, completed a baseline survey of 9th and 12th graders in 2022, but follow-ups have been suspended.

DATA PROCESSING AND ESTIMATION

Computer Processing of Decennial Census—Census Bureau

From desperation (so many people in a growing nation), Census Bureau innovated tabulating machines beginning with 1870 Census • Punch card tabulator (invented by former Census employee Herman Hollerith—his company became IBM) used in 1890 Census • Census commissioned first electronic computer for civilian use, UNIVAC I (now in the Smithsonian), in late 1940s • Arrived in time to assist 1950 Census processing • Used for entirety of 1960 Census processing, including the first use of nearest neighbor or "hot deck" methods to impute records for missing item responses.

More information: Tabulation and Processing - History; Citro (2025).

Correcting Undercount in the Agriculture Census—NASS

Well known that Census of Agriculture (conducted every five years) was incomplete, capturing most of the value of agricultural production but missing small farms • Working with statisticians at NISS, NASS adopted "capture-recapture" methods (see Data Evaluation and Testing section below) to correct each census beginning in 2012 (corrections also adjust for nonresponse and misclassification) • 2017 Census missed about 25% of the smallest farms and 2% of the largest small farms before correction • Completeness for farms important for allocating federal agriculture funds to states

More information: Census of Agriculture Methodology

Model-Based Small-Area Estimation—Census

Expensive to collect survey data for accurate estimates for small geographic areas • More accurate estimates often possible using survey results with models and additional variables • Census used "borrowing strength" models in late 1970s to improve per capita income estimates for small local governments to allocate General Revenue Sharing funds • 1994 legislation required updated county and school district estimates of poor school-age children to allocate Title I education funds for disadvantaged children • Census developed Small Area Income and Poverty Estimates (SAIPE) program • SAIPE pools model estimates using such variables as Supplemental Nutrition Assistance Program (SNAP) enrollments with survey results (from the CPS ASEC initially, now from the ACS) to produce estimates that are more reliable than either source alone

More information: Small Area Income and Poverty Estimates (SAIPE) Program

DATA DISSEMINATION

Electronic Data Products Beginning with 1960 Census—Census Bureau

Introduction of computers for census processing led to dissemination of public use electronic data products • Helped by interest and funding from outside sources, Census made available 1960 Census summary files (tabulations) and a 1/1000 public use microdata sample (PUMS) with identifying information removed and geographic identification only for states • 1970 Census greatly expanded number and scope of summary and PUMS files, which became planned-for census data products • Availability of files spawned industry of academic centers and companies acquiring and adding value to census electronic products • IPUMS at University of Minnesota created PUMS files for censuses from 1850–1950 using census records in National Archives

More information: Technology - History; census procedural histories at Publications - Demographic - History

Secure Research Access to Confidential Data—NCES, Census Bureau

PUMS files from population censuses and surveys are invaluable but limited (e.g., income is top-coded) for confidentiality protection; they cannot be released at all for economic censuses and surveys • The Census Bureau (beginning in 1982) and other agencies allowed researchers to access confidential data on-site as special sworn employees • NCES pioneered in 1991 licensing researchers to securely house confidential data at their institutions • The Census Bureau established secure sites (Research Data Centers) beginning in Boston in 1994 • In 2016 RDCs rebranded as Federal Statistical Research Data Centers (FSRDCs) • FSRDCs provide access to data from many statistical agencies through a Standard Application Process (SAP) mandated in the Foundations for Evidence-Based Policymaking Act of 2018 • 37 FSRDCs currently in 24 states, DC, and Puerto Rico

More information: Statistical Standards Program - Restricted Use Data Licenses; Federal Statistical Research Data Centers; Research Data Gov

Current Status: Several agencies withdrew from the FSRDC program during 2025 for lack of staff to process applications and work with researchers. NCES and SAMHSA's Center for Behavioral Health Statistics and Quality are no longer accepting applications; IRS/SOI has paused applications to its Joint Statistical Research Program; BLS has curtailed its offerings; datasets for EIA and ORES are not yet catalogued. (Applications to use commingled SOI and Census Bureau data are still being processed.)

DATA EVALUATION AND TESTING

Dual-System Estimation of Census Coverage—Census Bureau

Well known that censuses include errors—some people missed, others counted wrongly or more than once, but how many? which groups? • Demographic analysis assesses net undercount (omissions minus erroneous inclusions) nationally by sex, age, Blacks and all others • Dual-system estimation (DSE) with post-enumeration survey (PES) provides more detailed estimates • First used in 1980 Census • Determined not timely or accurate enough to adjust census results but important assessment tool • Developed from "capture-recapture" methods in wildlife surveys (e.g., catch, tag, throw back into lake a sample of fish; catch again; see how many of second catch have tags; use algebra to estimate total number fish in lake) • Independent PES and census are the two catches; matched and followed up to determine people counted in both and missed in one or the other; algebra gives total population

More information: 2020 Census Data Quality; National Academies (2023a, Ch. 4).

Box H-A1.1

SELECTED FINDINGS FROM BLS AND NCES LONGITUDINAL SURVEYS

BLS NLSY79 and NLSY97*

- Head Start, the federal preschool program, improves educational outcomes all the way through college and deepens social, emotional, and behavioral development into adulthood.
- Four-year college attendance increased slightly among the 1997 cohort (youths born in the early 1980s) versus the 1979 cohort (youths born between 1957–1964), and the younger cohort saw a striking leap in two-year college attendance.
- The long-term value of employment for teens has declined. Teens who worked 20 or more hours a week in the 1980s enjoyed a 9% increase in their lifetime wages, compared to a 4% gain for later generations of teens.

NCES NLS:72, HS&B, NELS:88, and ELS:2002**

- For young adults two years out from high school graduation, 62% were enrolled in college in 2006 versus 40% in 1974.
- Among those who had not expected to go beyond high school, 22–28% were neither in school nor working for pay two years after graduation, compared with 4-6% of those who expected to obtain a graduate or professional degree.
- Living with parents two years after high school graduation was more common the less educated the parents: 57% lived with parents when the parents had a high school or less education in 2006 (43% in 1974), compared with only 29% who lived with parents when the parents had a graduate or professional degree in 2006 (30% in 1974).

* National Longitudinal Survey of Youth 1979 | NORC at the University of Chicago

** Trends Among Young Adults Over Three Decades, 1974-2006

NOTE: The latest NCES longitudinal survey, HS&B:22, has completed only the baseline round.

Box HA-2

STATISTICAL AGENCY INNOVATIONS DURING COVID-19

(Note: Current Status indicated where relevant)

"Pulse" Surveys

Household Pulse Survey (HPS)—The Census Bureau led a coalition of multiple federal agencies (including BLS, BTS, EIA, ERS, NCES, NCHS) that designed, had approved by OMB, and pushed out the HPS barely five weeks after the nationwide Covid-19 shutdown in spring 2020. From April 23–July 21, 2020, the Household Pulse Survey collected data weekly using an internet questionnaire that asked about effects of the pandemic on mental health, employment, child care arrangements, food insecurity, and other topics. Subsequent phases used two-week data collection, followed by two weeks on and two weeks off collections, and currently continuous data collection. Question topics have changed as needed (e.g., asking about vaccination once Covid-19 vaccines were available and including questions about the infant formula shortage). Data are released two weeks after collection. The estimates have been useful to policymakers and the public, although they are appropriately labeled as experimental, given low response rates and representation issues. See Household Pulse Survey.

Current status: In early October 2024, the HPS began to transition into a longitudinal panel, with data collected every two months to be released at a national level as part of the monthly Household Trends and Outlook Pulse Survey (HTOPS), officially launched in January 2025. HTOPS in the off months includes items from Census Bureau staff and stakeholders deemed worthy of testing. HTOPS data are released solely as microdata and not also as tables or in reports. Only one round of HTOPS data has been released, in April 2025 and covering February 21–March 7, 2025. See Household Pulse Survey Public Use File (PUF).

Small Business Pulse Survey—The Census Bureau initiated the Small Business Pulse Survey in late April 2020 to provide near real-time information on the Covid-19 pandemic's effects on U.S. small businesses (nonfarm single-establishment businesses with 1–499 employees). A year ago, this experimental weekly survey transitioned to the biweekly Business Trends and Outlook Survey, which covers all nonfarm single-establishment businesses with at least one employee. See Now That the Pandemic Emergency Has Ended, What's Next?.

Current status: The survey is ongoing.

School Pulse Panel—The National Center for Education Statistics initiated the School Pulse Panel for the 2021–2022 school year (limited staff and contracting hurdles in the Department of Education prevented NCES from launching a school pulse survey similar to the household and small business pulse surveys in spring 2020). The 2021–2022 and 2022–2023 panels collected extensive data on the impact of the Covid-19 pandemic on students and staff in K-12 public schools. Beginning in 2023–2024, the survey collects a broader array of data, asking a sample of school principals each month about different topics (e.g., after-school programs, tutoring, school meal programs, student and staff mental health, school building facilities). See School Pulse Panel.

Current status: The survey appears to be suspended as there is no evidence of a 2025-2026 round.

Covid-19–Related Questions Added to Surveys

Monthly Current Population Survey—The Bureau of Labor Statistics contracts with the Census Bureau to conduct the monthly CPS—the source of unemployment estimates. From October 2022–November 2023, respondents were asked whether they worked from home in February 2020 (before the Covid-19 shutdown) and whether they worked at home more, less, or the same amount at the time of interview compared with February 2020. Respondents continue to be asked if they teleworked during the survey week. See Telework (CPS) : U.S. Bureau of Labor Statistics.

National Health Care Surveys (ambulatory care in physicians' offices, community health centers, hospitals; hospital care; long-term care; physicians' use of electronic health records)—The National Center for Health Statistics added questions to its National Health Care Surveys for the second half of 2020 and all of 2021 as appropriate (e.g., questions on telemedicine, shortages of personal protective equipment). See COVID-19 Pandemic Impact on the National Health Care Surveys – PubMed. These questions were discontinued at the end of 2021.

DATA PROCESSING AND ESTIMATION

Vital Statistics—Covid-19 Deaths

The National Center for Health Statistics took quick action to make possible provisional daily updates for states and weekly updates for demographic groups, states, and counties of Covid-19 deaths. The estimates were released about one to two weeks after other data sources (four to five weeks during Covid-19 surges, as in December 2020) to ensure quality control of cause of death coding. Previously, provisional mortality data were released monthly or quarterly with a minimum three-month lag. NCHS developed guidance for death certification involving Covid-19, modified its processing systems, and developed machine-coding routines for Covid-19. Initially, 100% of certificates reporting Covid-19 had to be manually coded, falling to 20% by 2021. Most importantly, these daily updates continue after other sources (e.g., the Johns Hopkins Covid-19 database) have shut down. See Advancements in the National Vital Statistics System to Meet the Real-Time Data Needs of a Pandemic - PMC and Provisional Death Counts for Coronavirus Disease 2019 (COVID-19).

Near Real-Time Spending

The Bureau of Economic Analysis began publishing weekly estimates of consumer spending by industry (gasoline stations, clothing stores, food and beverages, etc.) for the previous week on June 12, 2020 (the initial estimates extended back to February 2020). The estimates are based on aggregated credit card transactions, which BEA compares to expected spending levels prior to the pandemic. (BEA credits the Federal Reserve for the method.) The estimates are experimental, not necessarily representative of total spending in an industry, and have other limitations but are an early, frequently updated barometer of American spending. See Near Real-Time Spending | U.S. Bureau of Economic Analysis.

Current status: BEA discontinued the series as of May 2024, citing lack of funds and concerns about the underlying data.

Covid-19–Related Daily, Weekly, and Monthly Transportation Statistics

The Bureau of Transportation Statistics worked to develop new nearly real-time data sources on transportation and travel during the Covid-19 pandemic. It launched a website on August 4, 2020, with such experimental data series as Daily Travel During the COVID-19 Pandemic (updated weekly with a two-week lag, available for states and counties, based on anonymized cell phone data) and Docked Bikeshare Ridership (updated monthly for systems with docking information). The page also linked to The Week in Transportation and Monthly Transportation Statistics, along with one-time reports, such as 4th of July Travel, 2020 vs. 2019. See BTS Offers COVID-19 Related Daily, Weekly and Monthly Transportation Statistics.

Current status: In April 2024, BTS suspended The Week in Transportation because of "changes in experimental datasets." In particular, Daily Travel During the COVID-19 Pandemic was retired as of early April 2024 (see The Week in Transportation | Bureau of Transportation Statistics; Daily Mobility Statistics | Bureau of Transportation Statistics).

DATA DISSEMINATION

Web Pages Specific to Covid-19

Census Covid-19 Data Hub—The Census Bureau launched a page on its website on April 23, 2020, to bring together data and analyses to help communities, businesses, and policymakers track the socioeconomic impacts of Covid-19. The page linked to data, articles, and visualizations from the American Community Survey, the Household and Small Business Pulse Surveys, Weekly Business Formation Statistics, Monthly State Retail Sales, and other datasets. The page ceased being updated in April 2023 when the Covid-19 national emergency officially ended. See Census COVID-19 Data Hub.

Publications Specific to Covid-19

Monthly Labor Review (MLR)—BLS has published the *MLR* since 1915 (articles are published online as soon as they are ready). From April–December 2020, it published nine articles on Covid-19 effects on labor markets, earnings, prices, etc.—the first such article ("How many workers are employed in sectors directly affected by Covid-19 shutdowns, where do they work, and how much do they earn?") appeared April 16, 2020. See Covid-19 : U.S. Bureau of Labor Statistics.

Today in Energy—EIA has published articles every weekday since February 2011 on energy markets—supply, demand, prices, etc.— in *Today in Energy.* It published the first article on Covid-19 impacts on energy availability and use, on March 27, 2020. From April–December 2020, as many as four or five articles a month addressed Covid-19 effects (examples include "COVID-19 mitigation efforts result in the lowest U.S. petroleum consumption in decades" [April 23, 2020]; "As lockdowns eased in May, gasoline demand increased and jet fuel continued to fall" [August 13, 2020]). See Today in Energy.

Box HA-3

SELECTED RECENT STATISTICAL AGENCY INNOVATIONS AND CURRENT STATUS

HA-3A. DATA CONCEPTS AND TOPICS

Global Value Chains (BEA, NCSES)

BEA is developing data to shed light on global value chains (GVCs)—complicated supply chains that link multiple countries to produce a good or service. As an example, design, marketing, and software for Apple products are United States-based, while the hardware includes parts from many companies and countries (including the United States). BEA released prototype data in December 2021 on trade in value added (TiVA) (contributions of U.S. companies to U.S. exports) back to 2007. The data can answer such questions as the percent of domestic inputs in a U.S. industry's exports or how U.S. industries contribute to other countries' GVCs. NCSES is collaborating with BEA to develop more detailed TiVA estimates for science and technology industries. Figure HA-3A.1 shows that domestic content dominates U.S. exports. See Global Value Chains.

Current status: Ongoing project.

FIGURE HA-3A-1. Trade in Value Added (TiVA), 2007–2021

SOURCE: Global Value Chains | U.S. Bureau of Economic Analysis (BEA)

AI Adoption in the Economy (Census Bureau, NCSES)

The Census Bureau has worked with the National Center for Science and Engineering Statistics (NCSES) as far back as 2017 to measure AI adoption by businesses and R&D expenses on AI, as described below. More work is needed for the FSS to standardize a set of generally applicable questions and to develop questions targeted to specific industrial and service sectors, such as health care, education, agriculture, and others. Measures of use of AI by households, people with AI-related jobs, and energy consumption of AI data centers are also of interest, as is measuring the production of AI tools and techniques as part of the national accounts. BEA discussed a possible AI production satellite account with the Federal Economic Statistics Advisory Committee in December 2024, noting the sparseness of available data and the conceptual and technical challenges (Alternate Title Slide Layout for Longer Presentation Titles). To our knowledge, this effort has not gone forward.[13]

[13] In fact, BEA has discontinued two of its satellite accounts, Digital Economy and Health Care.

Two Sources of Data on AI Use by Businesses:

- **Annual Business Survey (ABS):** The online ABS, developed by the Census Bureau and NCSES, was first conducted in 2018. It sampled about 850,000 nonfarm employer firms, achieving a 69% response rate (the ABS in non-economic-census years has sampled about 300,000 firms with a response rate as of 2023 of 62%). Census Bureau analysts linked most ABS respondents to its Longitudinal Business Dynamics (LBD) database, which provides employment and payroll data and other information. Questions on AI adoption included: extent of use (testing, low, moderate, high) of five AI-related technologies: automated vehicle guidance, machine learning, machine vision, natural language processing, and voice recognition software (see https://www2.census.gov/library/working-papers/2023/adrm/ces/CES-WP-23-48R.pdf). The 2019 and 2023 rounds of the ABS included detailed AI-related questions (e.g., asking about workforce impacts), with a paragraph-length definition of AI provided to respondents. The 2020 ABS had a question on the use of AI for innovation activities. The 2022 ABS included a question about AI use and another question on its workforce impacts. The 2024 and 2025 ABS did not touch on AI use or impacts. (See Annual Business Survey Respondent Materials). For 2026, the ABS is planned to merge with the NCSES-sponsored Business Enterprise R&D Survey (BERD), which asks about R&D expenses on AI; the ABS sample size will be reduced from about 300,000 firms to about 220,000 firms and 8,000 nonprofit organizations. (See Federal Register :: Agency Information Collection Activities; Submission to the Office of Management and Budget (OMB) for Review and Approval; Comment Request; Annual Business Survey.)
- **Business Trends and Outlook Survey (BTOS):** The online biweekly BTOS is the successor to the Census Bureau's online weekly Small Business Pulse Survey (SBPS), which ran from April 2020 to April 2022, representing single-location employer businesses with fewer than 500 employees. The BTOS sample (about 200,000 firms in each biweekly cycle) as of July 2022 represented all nonfarm single-location employer businesses, expanded to all nonfarm employer businesses in October 2023. The BTOS has included two questions since September 2023 on AI use in the last two weeks and whether a business expects to use AI in the next six months. A supplement administered in December 2023 through February 2024 asked additional questions on AI adoption and impacts. (See Slide 1.) E-mail response rates are generally in the range of 25–30%. BTOS published statistics are labeled "experimental."

Box HA-3

SELECTED RECENT STATISTICAL AGENCY INNOVATIONS AND CURRENT STATUS

HA-3B. DATA COLLECTION

Improving Survey Response for Businesses (BLS, Census Bureau)

The Census Bureau and BLS have worked to make it easier for companies to provide their data confidentially for statistical purposes.[14] The Census Bureau simplified and standardized its annual business surveys, combining separate surveys of retail trade, manufacturing, and other sectors into the Annual Integrated Economic Survey (AIES).[15] BLS works with large companies to facilitate their response for multiple firms and establishments. Census Bureau has gone a step further to eliminate the need for a large company to respond to surveys entirely; instead, a company like Amazon can have a secure portal to the agency to provide its data on employment, payroll, and other topics in its own format for agency software and staff to sort into standard statistics, taking minutes for a quarterly upload instead of hundreds of hours to fill out surveys. This approach is limited at present to a handful of large companies because staff must translate company-provided data into usable statistics.

Current status: Ongoing; to scale up the effort, the Census Bureau is developing a large language model (LLM) to map companies' data into a common format suitable for census processing. See Source Data Innovation at the Census Bureau; see also Federal Register :: Agency Information Collection Activities; Submission to the Office of Management and Budget (OMB) for Review and Approval; Comment Request; Direct Digital Data Feeds.

Improving Survey Response for Households (BLS, Census Bureau)

BLS and Census have worked for many years to simplify and facilitate response to the Consumer Expenditure Survey, which provides the CPI market basket weights (see Gemini Project to Redesign the Consumer Expenditure Surveys : U.S. Bureau of Labor Statistics; see also National Academies 2013). More recently, for the CPS, which provides the monthly unemployment estimates, they have developed a modernization plan with several components (e.g., an internet response option) (see 2023 Modernization Efforts). Efforts have moved slowly and are not as thoroughgoing as they could be in scope (see, e.g., CPS_FESAC_discussion_Dec2023.pdf) or in testing to minimize response mode (internet, telephone, in-person) effects on data quality.

Current status: Efforts are largely at a standstill due to lack of staff and resources.

[14] The Census Bureau for over 50 years has used IRS returns for the nation's 21 million nonemployer businesses (e.g., independent consultants) instead of sending them surveys.

[15] A CNSTAT report played a major role in articulating the vision for the AIES (National Academies 2018). [Reengineering the Census Bureau's Annual Economic Surveys | The National Academies Press]

Box HA-3

SELECTED RECENT STATISTICAL AGENCY INNOVATIONS AND CURRENT STATUS

HA-3C. DATA PROCESSING AND ESTIMATION

Producing Estimates Using Blended Data

Following data collection, important steps in the production of federal statistics include data processing and estimation. For surveys, many steps are long-standing, such as weighting adjustments for nonresponse, automated imputation for missing responses to specific questions, adjustment of time series for seasonal fluctuations in employment and other phenomena, and increasingly sophisticated confidentiality protection for statistics before public release. For administrative records series, there are various data cleaning operations that are needed prior to release of statistics.

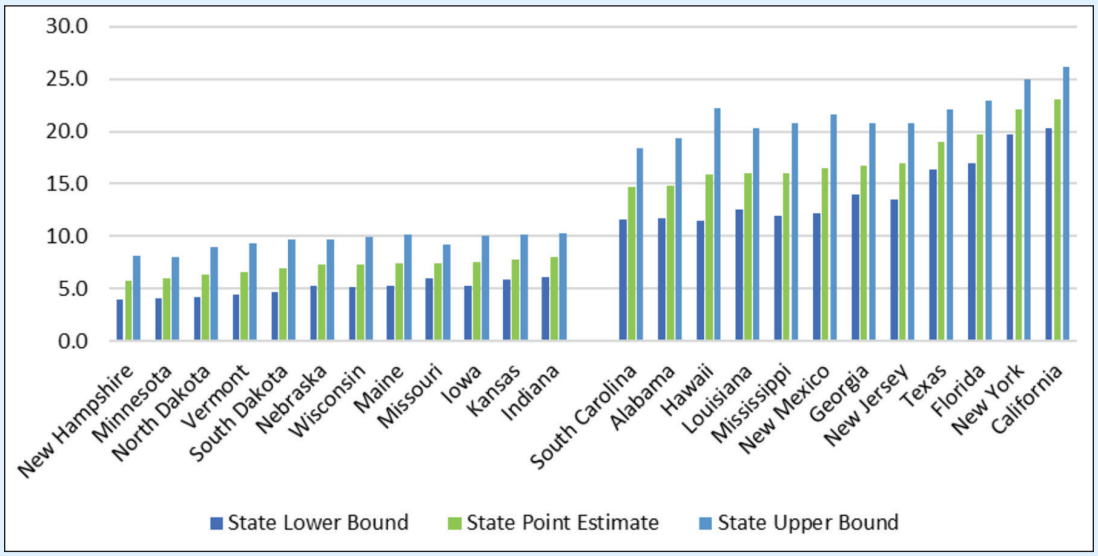
Increasing household and business nonresponse and incomplete response in surveys have undercut the ability of standard methods to compensate, affecting the quality of survey estimates. Administrative records are typically not sufficient on their own to provide useful statistics, and many potentially useful federal and state records are not available to federal statistical agencies. More and more, statistical agencies are turning to the concept of blended data or multiple data sources, such as a survey combined with administrative records, to develop improved statistics in terms of accuracy and granularity. Six important initiatives—some that were one-off efforts and others that are ongoing—are summarized below: 1) NCES adult literacy estimates for states and counties; 2) Census Bureau business formation statistics; 3) BEA household distributions of personal income; 4) Census Bureau National Experimental Well-being Statistics (NEWS); 5) NCHS National Hospital Care Survey (NHCS) linkages; and 6) NCSES National Secure Data Service (NSDS) pilot.

(1) Adult Literacy Estimates for States and Counties (NCES)

In 2008, NCES developed subnational estimates of adult literacy in response to user demand. National estimates of low literacy adults are 14–15%. Models were constructed to predict percentages of adults in each state and county lacking Basic Prose Literacy Skills in the 2003 National Assessment of Adult Literacy (NAAL) and the 1992 National Adult Literacy Survey (NALS). NAAL and NALS had additional sample cases for 6 and 11 states, respectively. The models used state-level estimates of low literacy, where available, plus state and county variables from the 2000 and 1990 Censuses, including educational attainment, race/ethnicity, foreign-born status, and poverty (2003 model only), and native-English-speaking status (1992 model only). Figure HA-3-C.1 shows the 12 states with the lowest estimated percentages of adults lacking basic prose literacy skills in 2003 and the 12 states with the highest percentages; note the wide confidence intervals. See State and County Literacy Estimate .

Current status: This was a one-off project because no recent survey of adult literacy exists.

FIGURE HA-3C.1 Percentage of Adults Lacking Basic Prose Literacy Skills, 12 Lowest and Highest States, with Confidence Intervals Shown, 2003



NOTE: District of Columbia excluded.

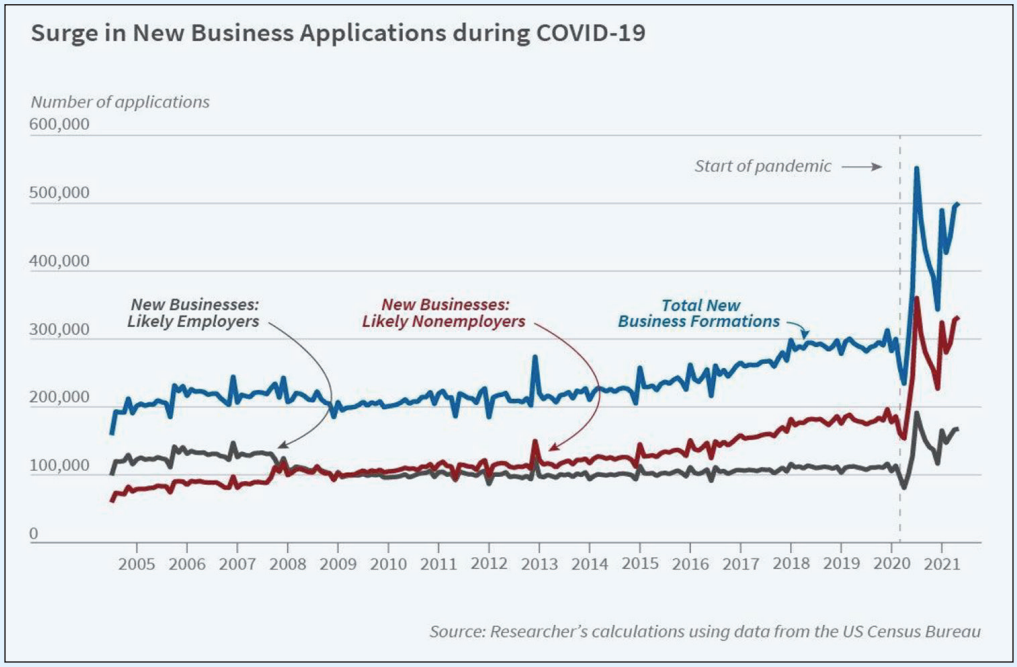
SOURCE: State and County Literacy Estimates - State Estimates

(2) Business Formation Statistics—Estimates for States (Census Bureau)

The Census Bureau initiated this series on an experimental basis in 2018 with quarterly estimates; monthly estimates replaced quarterly estimates in January 2021, and weekly estimates were added beginning in April 2020. As of August 2024, weekly estimates are included in the monthly reports. Annual estimates are available for counties. Business Formation Statistics became a standard data product in December 2021. The Census Bureau developed the methodology in collaboration with the Federal Reserve and the Universities of Maryland and Notre Dame. Applications for an Employer Identification Number (EIN) on IRS Form SS-4 are linked to the Census Bureau's Business Register and Longitudinal Business Database to determine when the business first has payroll or employment. As an example, Figure HA-3C.2 shows a drop and subsequent spike in business formation over the Covid-19 pandemic. See Business Formation Statistics.

Current status: Ongoing.

FIGURE HA-3C.2 Weekly Business Applications, 2004–early 2021



SOURCE: Haltiwanger, John C. (June 2021). "Entrepreneurship During the COVID-19 Pandemic: Evidence from the Business Formation Statistics." NBER Working Paper 28912. https://www.nber.org/papers/w28912.

(3) Household Distributions of Personal Income (BEA)

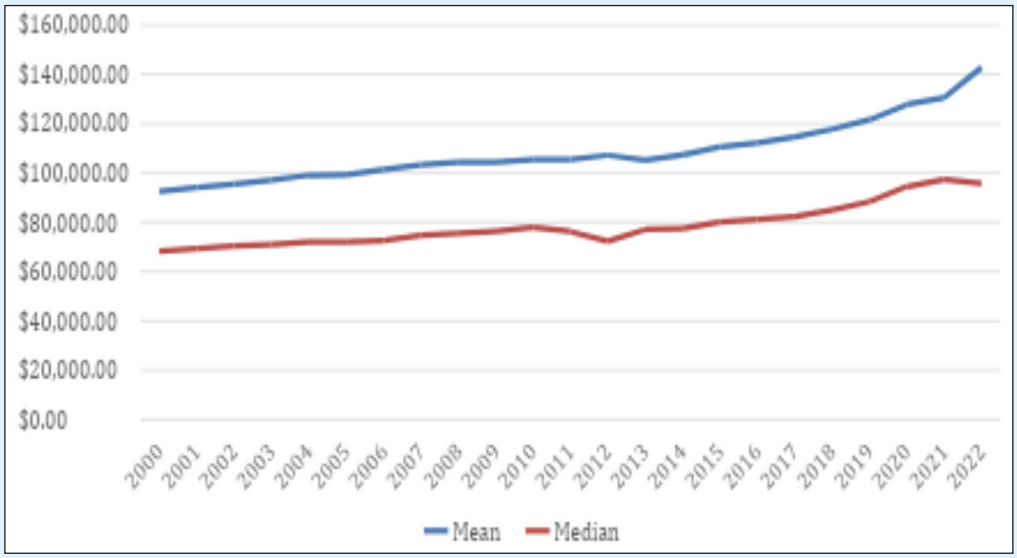
BEA has an ongoing experimental initiative to distribute aggregate Personal Income (PI) estimates from the national accounts (NIPA Table 2, see, e.g., pi0825.pdf) to households on the CPS Annual Social and Economic Supplement (ASEC) augmented with tax and other administrative data. The initiative was developed after the Great Recession when it became clear that aggregate statistics, such as per capita income, were obscuring growing inequality that was leaving many households behind and vulnerable to the collapse of the housing market. The initiative resumes a tradition of distributional estimates dating back to the 1940s, which was discontinued for lack of resources in the 1970s (see Citro 2025). The BEA distributions provide means, medians, deciles, and other fractions of households adjusted for household size. (Figure HA-3C-3 compares mean and median PI estimates for households over time.)

The first "prototype" national statistics were published in March 2020 for income years 2007–2016. Subsequently, BEA has produced household personal income distributions every December in final for year t–2 and provisionally for year t-1, with estimates available back to 2000. BEA added disposable (after-tax) PI household distributions in 2020, distributions that are internationally comparable with Organisation for Economic Co-operation and Development (OECD) concepts in 2022, and state PI distributions (for 2012–2023 currently) in October 2023. To improve timeliness, BEA developed an experimental "nowcast" for 2024, using machine learning techniques that analyzed relationships between published annual distributions and current NIPA totals (see Distribution of Personal Income | U.S. Bureau of Economic Analysis (BEA)). To be most useful, BEA needs to produce distributions for the Census Bureau's post-tax, post-transfer income definition, which would permit apples-to-apples comparisons across sources.*

Current status: Ongoing.

* Personal income differs from the Census Bureau's after-tax-and-transfer income definition. Three important differences are that PI includes an imputed value for home ownership, includes contributions to Social Security and other retirement plans and excludes retirement plan distributions, and includes the full insurance value of Medicare and Medicaid.

FIGURE HA-3C-3. Household Real Disposable (After-tax) Personal Income, Means and Medians, 2000–2022 (2017 $)



SOURCE: BEA (December 2023 Update); 2022 data are provisional; households are equivalized for size; Distribution of Personal Income

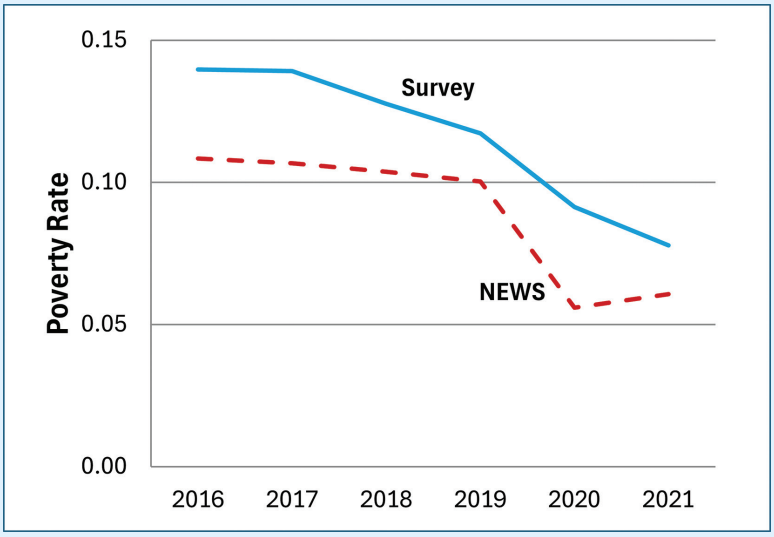
(4) National Experimental Well-Being Statistics (NEWS) (Census Bureau)

NEWS is a recently initiated, ongoing project to provide high quality distributions of household income and poverty, using carefully edited administrative records to correct for nonresponse and underreporting of income in the CPS ASEC. These kinds of errors are well known and have been increasing for decades (see Citro 2025). At present, over 40% of income in the CPS ASEC is imputed because respondents said they received an income source but did not provide an amount. Moreover, comparisons with independent sources indicate that nonreporting of both receipt and income amounts contributes to underestimates of many types of income in the survey. To be fully realized, the NEWS project needs greater access by the Census Bureau to tax return data from SOI and to state administrative records. With additional resources, NEWS could make faster progress toward the goal of releasing production estimates of pre- and post-tax-and-transfer income and poverty for households every fall for the preceding calendar year.

The Census Bureau released the first NEWS income and poverty estimates in 2023 for income year 2018 for money income (this concept excludes tax credits and in-kind benefits such as SNAP). They show (Bee, et al. 2023, Table 16) an increase of $4,000 or 6 percentage points in household median income, mainly due to the use of administrative records for retirement and investment income for the elderly. A release in January 2025 added estimates for post-tax-and-transfer income and poverty, with improved handling of missing data and nonresponse bias and improved modeling of earnings taking account of errors in both survey and administrative data. The latest release in July 2025 added estimates for 2016–2021. Future plans include producing timely estimates for the previous year, revising them as needed as more data become available, and working with the ACS to develop improved income and poverty estimates for states and localities. (See National Experimental Wellbeing Statistics (NEWS) Version 1.) Figure HA-3C.4 below illustrates the impact on the Supplemental Poverty Measure of the more complete NEWS income data compared to the CPS ASEC.

Current status: Ongoing.

FIGURE HA-3C.4 Supplemental Poverty Measure (SPM) rates, 2016–2021, with the CPS ASEC Survey Data and the Augmented NEWS Data



SOURCE: https://www.census.gov/data/experimental-data-products/national-experimental-wellbeing-statistics.html

(5) National Hospital Care and National Ambulatory Medical Care Survey Linkages (NCHS)

NCHS has linked several administrative records datasets to its National Hospital Care Survey (NHCS) data for 2014 and 2016, greatly expanding their analytical value. Datasets linked include the National Death Index (NDI) for 2014–2017; Medicare data for 2014–2017; Medicaid data for 2015–2017 (2016 NHCS only); 2013– 2016 HUD Housing Assistance Program data; and VA files through 2020 (2016 NCHS only). The linked data are accessible in the secure NCHS Research Data Center. "Feasibility" files are publicly available to determine if there are enough sample cases of interest for the researcher's proposed analysis. Articles already published illustrate the utility of the linked files (e.g., an analysis of deaths from pneumonia among ICU versus general hospital patients). See NCHS Data Linkage - National Hospital Care Survey Data. More recently, NCHS has linked the National Ambulatory Medical Care Survey (NAMCS) Health Center component to the National Death Index and HUD files.

Current status: Uncertain, most recent linkages are of the 2021 NAMCS Health Center component with the 2021–2022 NDI and 2020–2022 HUD Housing Assistance Program data. See NCHS Data Linkage - National Ambulatory Medical Care Survey Data.

(6) National Secure Data Service (NSDS) Pilot (NCSES)

The National Secure Data Service (NSDS) is a potentially consequential innovation in how data users, including federal, state, and local agencies and others, can access federal statistical datasets, link them to other data, and obtain privacy-protected results for evidence-based policymaking, program planning and evaluation, and other applications. The NSDS was recommended by the Commission on Evidence Based Policymaking, and a five-year demonstration project was authorized by P.L. 117-167 in the CHIPS and Science Act of 2022. The demonstration project is a collaborative effort of the federal statistical system managed by NCSES. The goal of the NSDS is to streamline data discovery, access, linkage, analysis, and privacy protection, thereby increasing efficiency, reducing duplication and cost, and supporting expanded data use for research and policymaking by users both within and outside of government. A major success of the NSDS demonstration project to date has been the application of AI to solutions that support its overall vision and goals.

For example, the NSDS has developed a proof of concept for an AI chatbot that can respond to open-ended user questions with detailed reports that include statistical estimates, data visualizations, and references to source data. Using a hybrid approach of generative AI and pure retrieval AI, the statistical chatbot mandates factual accuracy in

its responses. As one component of the larger NSDS Data Concierge service, the chatbot could play a crucial role in quickly and efficiently fielding user questions, connecting users to data, streamlining the data user experience, and showcasing how AI can be innovatively harnessed by the FSS. This innovation accords with the provisions of the America's AI Action Plan that charge NSF to create an online portal for the NSDS "to provide the public and Federal agencies with a front door to AI use-cases involving controlled access to restricted Federal data," and which charge OMB to promulgate the two outstanding regulations in the 2018 Evidence Act to facilitate statistical agency access to other federal agencies' data and secure access by users to statistical agency data (see America's AI Action Plan, p. 8).

Current status: Ongoing.

Box HA-3

SELECTED RECENT STATISTICAL AGENCY INNOVATIONS AND CURRENT STATUS

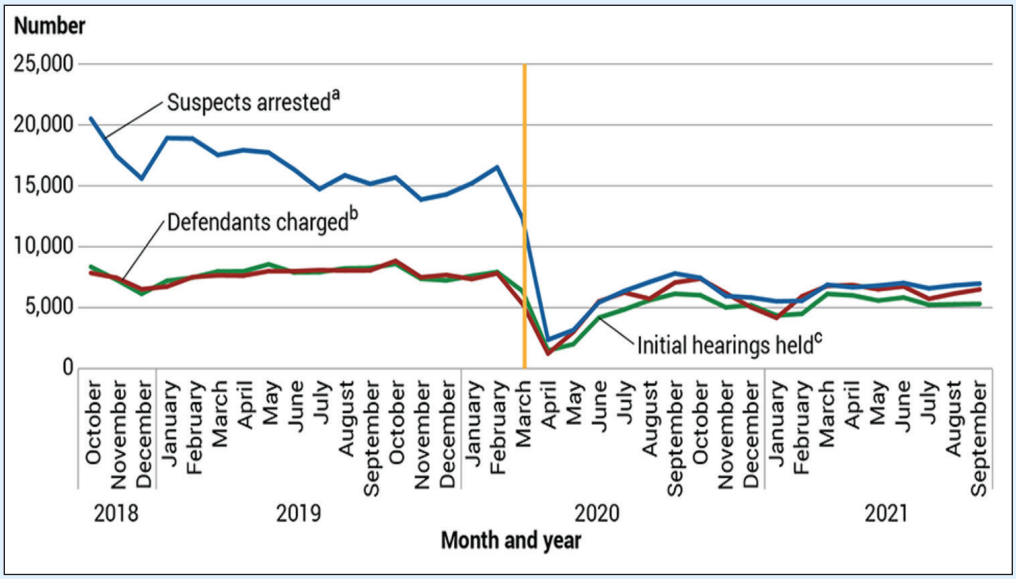
HA-3D. DATA DISSEMINATION

Just the Stats (BJS)

In October 2022, BJS announced a timely web-based series of periodic releases of one or two key indicators from one of BJS's datasets. Subscribers to JUSTSTATS get alerts when these brief reports (downloadable as PDFs) are released. The first report, released in October 2022, was Carjacking Victimization, 1995–2021. The most recent report, released in July 2025, is Reporting to Police by Type of Crime and Location of Residence: 2020–2023. To date, BJS released two Just the Stats reports in 2022, five in 2023, three in 2024, and three thus far in 2025. Figure HA-3D.1 shows a graph from a January 2024 Just the Stats report on Covid-19 effects on arrests, charges, and initial hearings. See BJS announces new report series, Just the Stats and Search Publications | Bureau of Justice Statistics.

Current status: Ongoing.

FIGURE HA-3D.1 Impact of Covid-19 Shutdown (yellow vertical line) on Arrests, Charges, and Initial Hearings Held, Federal Courts, FY 2019–FY 2021



SOURCE: Federal Pretrial Release During the Coronavirus Pandemic, Fiscal Years 2019–2021 | Bureau of Justice Statistics; published January 2024 (see website for footnotes to graph)

The Opportunity Project (TOP) (Census Bureau)

This project, launched in 2016 and housed in the Census Bureau's Open Innovation Labs, provides an opportunity for federal agencies to enlist companies, nonprofits, and universities to use public data to solve specific problems in 12-week "sprints" through development of digital tools. Census Bureau staff provide assistance. An exemplar project was a collaboration between the State Department and the Wilson Center on reducing plastic waste in oceans. TOP products from the project included Ocean Plastics Story Map, developed by Esri, to enable storytelling with newly compiled datasets; the Georgetown University Campus Plastic Initiative, whereby the Georgetown University Beeck Center tracked plastic pollution on college campuses and educated students on its impact, leading to creation of a Georgetown student organization and greater emphasis on reduction of plastic pollution on campus; and The Ghost Gear Project, of the Harvard University Institute of Politics, which visualized the location of ghost fishing gear at sea and described accompanying mitigation techniques. See The Opportunity Projec .

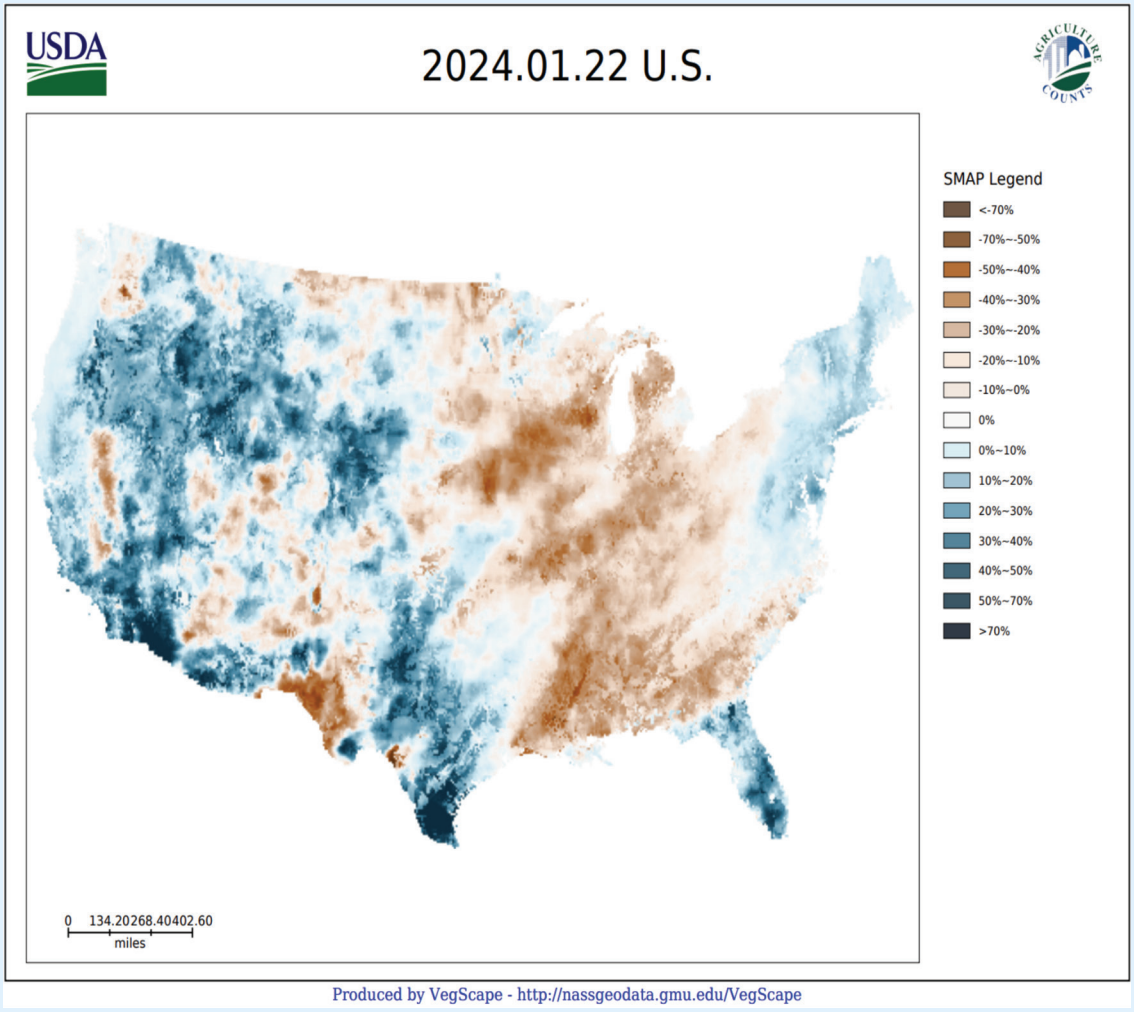
Current status: Ongoing but not aware of new projects.

Crop Condition and Soil Moisture Analytics Tool (Crop-CASMA) (NASS)

NASS developed Crop-CASMA in collaboration with NASA and George Mason University to provide access to high-resolution NASA data on soil moisture in a user-friendly format. The tool was released in March 2021 and is updated daily. NASS analysts use the data in weekly Crop Progress Reports, and the tool is available for farmers, researchers, and students to map, download into models, and use in other ways. The resolution of the raw data is 20 miles (about the size of a county). The tool includes a method to estimate at a resolution of three-fifths of a mile. Figure HA-3D.2 shows wet and dry areas of the United States for January 22, 2024. See NASA Data Powers New USDA National Agricultural Statistics Service Soil Moisture Porta .

Current status: Ongoing, led by NASA.

Figure HA-3D.2 Wet and Dry Areas of the United States, January 22, 2024



SOURCE: NASS Crop Condition and Soil Moisture Analytics Tool.