Harvard Data Science Review • Issue 3.4, Fall 2021

The Future of Official Statistics

Erica L. Groshen^{1,2}

¹New York State School of Industrial and Labor Relations, Cornell University, Ithaca, New York, United States of America,

²William Erastus Upjohn Institute for Employment Research, William Erastus Upjohn Unemployment Trustee Corporation, Kalamazoo, Michigan, United States of America

Published on: Oct 28, 2021

DOI: <https://doi.org/10.1162/99608f92.591917c6>

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

ABSTRACT

Official statistics are a pure public good and a key element of our national data infrastructure. They support our well-being by informing a vast set of public and private decisions. Yet, today they stand at a crossroads in the United States. With new technologies, today's official statistics could be more relevant, interoperable, granular, and timely. Recent experience during the pandemic clarifies the need for improvement. The agencies could tap the new wealth of non-survey data to accomplish these goals without increasing burdens on survey respondents. They could operate more in sync with each other (to facilitate combining and sharing data) and more independently from politics (to reinforce public trust in their integrity). They could be better protected from defunding and requirements to combine operations with non-statistical agencies. These changes are possible, but not without effort by the statistical system's stakeholders. The agencies cannot do it alone.

The alternative, I fear, is a downward spiral in official statistics. Current funding neglect and erosion of independence threaten to reduce public trust and data quality. Less reliable products would then lead to further cuts and lower response rates, and suppress quality further. Hence the spiral.

To secure a promising future for official statistics and avoid the death spiral, the statistical system's stakeholders must actively advocate changes that include data-sharing legislation, flexible and dependable funding, adopting common data schema, and a modernized, more coordinated statistical system. The statistical and data science communities—including you—have an important role to play in choosing the path ahead.

Keywords: official statistics, data infrastructure, future, Bureau of Labor Statistics, statistical agency, public goods

1. Introduction

The purpose of this article is to help celebrate World Statistics Day by sharing my thoughts on what lies ahead for official statistics. I also want to engage my audience—that is, the broad community of statisticians, data scientists, and students—in thinking seriously about what lies ahead for an important fixture in their lives. I believe that you are also the key to keeping the future of official statistics bright.

I am a latecomer to the title statistician; for most of my career, I worked with statistics, but as a labor economist. Then, after 25 years in the Federal Reserve System, I suddenly became a national statistician, heading up the Bureau of Labor Statistics (BLS) for four fun years. BLS is a world class agency that is an important part of the world class U.S. federal statistical system.

People have often asked why I loved my time at BLS. The answer is easy. First, the mission was pure and important. BLS “[measures labor market activity, working conditions, price changes, and productivity in the US](#)

[economy to support public and private decision-making](#).” Second, I worked with great colleagues, both within the agency and beyond. These are really some of the world’s most dedicated data nerds. Third, I learned things every day—about statistics, the economy, government, leadership, and so much more.

My thoughts on the future of official statistics reflect many things I learned over four years at BLS. I start by briefly describing the rationale for official statistics and the challenges recently posed by the pandemic. By official statistics, I mean indicators and data series produced and released by government statistical agencies. Everyone reading this article uses these resources regularly, professionally and personally, directly or indirectly. Then I discuss some forces at work on the present state of official statistics, with particular focus on opportunities for improvement. That leads to considering what pathways we could take going forward—and, finally, how you in the audience can help shape a bright future. Although I focus on the U.S. federal statistical system today, many of my observations and conclusions will likely be relevant to other countries and to subnational statistical agencies as well.

2. Rationale for Official Statistics and COVID-19 Challenges

To begin, it is worth noting why we have official statistics. When I was at BLS, being part of the monthly releases often reminded me of the reasons. This was particularly salient on the first Friday of each month, upon release of the “[Employment Situation](#).” That report contains two headline numbers: the latest payroll jobs change and the unemployment rate. These indicators drive monetary and fiscal policy, garner massive media attention, and move financial markets more consistently than any other statistics. This is government at its best, producing essential, high-quality information. The quality reflects a process conducted with the utmost professional integrity and transparency, supported by a great history of research and innovation.

Economic theory explains the significance of this ritual. Official statistics are a public good, a form of infrastructure, so their production is a fundamental responsibility of government. In economics, a public good is a well-defined concept. It means a good that is both non-excludable and non-rivalrous. Non-excludable means that people cannot be excluded from using or benefiting from a good without paying for it. The non-rivalrous part is that use of data by one person does not reduce the availability of that information to another person. Both features are certainly true of official statistics: it would be hard to prevent interested parties who did not pay a fee from learning the BLS estimate of unemployment, and one person’s knowledge does not diminish another’s ability to learn that number. Two people can use a statistic at the very same time. Compare this to a hamburger. If I have a hamburger, I can prevent someone else from taking it. Moreover, if I eat the hamburger, no one else can eat it.

Economic theory predicts that competitive markets will undersupply public goods because private companies cannot profit enough from producing them. Consumers can too easily free ride or otherwise avoid paying. Thus, we can improve national well-being by communal (that is, government) provision of public goods, rather

than having the private sector alone supply them. This is why roads, bridges, national defense, the court system, and clean air are usually government responsibilities.

Similarly, the country's decision makers need trustworthy data to make the good decisions that support national resilience and prosperity. Government provision of official statistics that meet those needs improves national well-being.

For example, on May 8, 2020, the [April 2020 “Employment Situation”](#) showed that the official unemployment rate soared by 10.3 percentage points in one month. This jump, the steepest in the history of the series, resulted in a jobless rate of 14.7 percent, far above any since the Great Depression. Similarly, payroll jobs saw their deepest drop ever—double the maximum losses during the Great Recession. This knowledge and the detail behind it helped inform urgent policy discussions and decisions in Washington and in businesses and homes across the country.

Throughout the statistical system, agencies rose to the challenge of producing information relevant to the pandemic. The [agencies have reported](#) that most production continued unabated with little impact on quality of data and transparent discussion about impacts of the pandemic on data elements and collection. Many leveraged ongoing survey programs by modifying existing data collection, adding new questions to existing collections, and developing new uses for existing data. For four examples, see the session on [“The Leveraging Official Statistical Programs to Address Emerging Issues”](#) at the Fall 2020 Conference of the Federal Committee on Statistical Methodology. BLS also published [a blog](#) on the topic. [A presentation by three agency heads](#) also describes statistical agencies’ reactions to the pandemic. Examples include quickly launched new programs such as the [Census Bureau’s Pulse Surveys](#), [new BLS questions on the Current Population Survey](#), and the BLS [Business Response Survey](#).

When I say that we can do better, what am I talking about? I want to start by saying that the statistical agencies have already achieved great flexibility, coordination, data sharing (to the extent they can), and new products during the COVID crisis. Here are some of many examples. The U.S. Census Bureau introduced experimental [Small Business](#) and [Household](#) Pulse Surveys. BLS added [new questions on the Current Population Survey](#) and [state and metro area Job Openings and Labor Turnover Survey](#) estimates. The Bureau of Economic Analysis (BEA) used its new regional products and distributional Gross Domestic Product (GDP) numbers and satellite accounts to [answer questions about the pandemic’s impacts](#).

Even so, the COVID-19 pandemic has highlighted many ongoing challenges for the statistical system, just as it has challenges to our health care, unemployment insurance, and justice systems. Many challenges relate to relevance, quality, and trustworthiness. Without those attributes, official statistics do not contribute to national well-being because people will not use them to guide critical decisions, or when used, the statistics will cause poor decisions.

What improvements do data users want to better guide critical decisions?

- **Timeliness and frequency.** In early 2020, the COVID-19 pandemic caused the most rapid and deepest recorded shock to the U.S. economy. BLS publishes the most reliable and timely measures of labor market activity on a monthly basis, with a two-week lag. Particularly from March to June 2020, this pace seemed very slow and infrequent. Attention devoted to weekly Unemployment Insurance Initial Claims releases soared during those months, testifying to the need for higher frequency and timelier statistics, particularly during business cycle turning points. The U.S. Department of Labor’s Employment and Training Administration (ETA) produces claims releases as part of its administration of the Unemployment Insurance system. That means that ETA does not construct these statistics to be economic indicators with known statistical properties. Therefore, the published numbers can mislead observers who try to use them to assess economic conditions (see [Government Accountability Office report GAO-21-191 “COVID-19: Urgent Actions Needed to Better Ensure an Effective Federal Response”](#)).
- **Granularity.** COVID-19’s impacts varied strongly by community, industry, and demographic group, and so on. To address distress and inequities, we must measure them with enough detail that policymakers can investigate and target them appropriately. Similarly, private sector decision makers need granular information to make the best choices in these turbulent times.
- **Relevance and agility.** Thanks to extraordinary efforts, the BLS, Census Bureau, and many other agencies asked new questions and started new programs during the pandemic to collect newly critical information, such as the incidence of teleworking. These efforts have proved their worth as they allow tracking employer and household experiences during the pandemic. This success clarifies the need for infrastructure to expand the speed and scope of such agility.
- **Consistency and access across products and programs.** To gain full benefit from our statistical programs we must be able to link or merge their products in new ways. For example, we could potentially learn much about the impact of pandemic restrictions from linking labor market outcomes to COVID statistics and local policy decisions. In addition, statistical agencies can use modeling (based on linking) to improve the granularity of frequent data or vice versa. Unfortunately, access restrictions and classification or sample frame inconsistencies often hamper such efforts.
- **Independence.** Our statistics must remain free of political manipulation, in reality and perception. Actions by the Trump administration with respect to the 2020 Decennial Census, the Economic Research Service, the National Weather Service, and the Centers for Disease Control, among others, have revealed weaknesses in the arrangements intended to prevent interference with independent federal scientific agencies. Now more than ever, scary reports and conspiracy theories abound in news and social media. We must act to ensure that our official statistics are objective and perceived to be so.
- **Privacy protection.** Statistical agencies must prevent disclosure and nonstatistical uses of individual data. The reasons include the following: avoiding hurt to respondents because the statistical mission is to provide help, not harm; the agencies need to abide by laws that protect privacy; and protecting privacy protects the

high response rates necessary for accurate statistics. This responsibility grows ever more difficult with the proliferation of individual data from the internet and other sources.

To sum up, there is a real impetus to address serious needs.

3. The Past

How did we arrive at the current U.S. statistical system? The simple answer is that the country has faced many crises over time. During those times or in their aftermath, policymakers have recognized that official statistics are a powerful public good. They have seen that businesses and households need data to make good choices. In addition, policymakers sought solid evidence to guide their own decisions. Indeed, evidence is especially important for building joint solutions in fraught and partisan contexts. Similarly, outside the policy world, when two parties negotiate, the scope over which they disagree can be narrowed by working from a common frame of facts, such as trustworthy statistics.

Consider when the BLS was founded in 1884 as the nation's first independent statistical agency (see Goldberg & Moye, 1985; Norwood, 1985; Wiatrowski, 2009). One early goal was to help quell violent unrest between employers and nascent unions. This industrial turmoil grew out of the depression of 1873–78 and the recurrent labor disputes of the 1870s and the 1880s. Policymakers at that time realized that providing trusted information could help resolve these conflicts. This information included the tightness of the labor market, prevailing wage and benefits levels, the cost of living, and working conditions, especially safety. Over time, BLS was also charged with informing policymakers on critical labor-related issues, such as immigration, international trade, inequality, and unionization. Subsequently, much of fiscal and monetary policy came to rely on BLS information about inflation and the labor market (the economy's largest and most complicated market).

BLS is one example of many. See Appendix B of National Academies of Sciences, Engineering, and Medicine (2021) for a concise history of the U.S. statistical system and its current structure. Each U.S. statistical agency joined the system separately using specialized legislation, with different relationships to their home agencies, access to data, and missions. Together, they make up the U.S. statistical system as described in the [annual report on the system](#). Table 1 lists the 13 agencies that the Office of Management and Budget (OMB) has designated as “principal federal statistical agencies.”

Table 1. Federal agencies represented on the interagency council on statistical policy.

Principal Federal Statistical Agencies*		
	Agency	Parent Department/Agency
1	Bureau of Economic Analysis	Commerce

2	Bureau of Justice Statistics	Justice
3	Bureau of Labor Statistics	Labor
4	Bureau of Transportation Statistics	Transportation
5	Census Bureau	Commerce
6	Economic Research Service	Agriculture
7	Energy Information Administration	Energy
8	National Agricultural Statistics Service	Agriculture
9	National Center for Education Statistics	Education
10	National Center for Health Statistics	Health and Human Services
11	National Center for Science and Engineering Statistics	National Science Foundation
12	Office of Research, Evaluation, and Statistics	Social Security Administration
13	Statistics of Income Division	Treasury/Internal Revenue Service
Statistical Officials from Other Departments and Agencies**		
1	Defense	
2	Homeland Security	
3	Housing and Urban Development	
4	State	
5	Interior	
6	Veterans Affairs	
7	Environmental Protection Agency	

8	General Services Administration
9	National Aeronautics and Space Administration
10	Nuclear Regulatory Commission
11	Office of Personnel Management
12	Small Business Administration
13	US Agency for International Development
* The Office of Management and Budget has designated these units as Principal Federal Statistical Agencies. Heads of these agencies were the original members of the Interagency Council on Statistical Policy.	
** These members were added under the Evidence Act of 2019 .	

Despite the agencies' different starting points and authorizing legislation, there has been some convergence over time. They share common elements in their missions, particularly recognition of their indispensable role in our economy. In addition, the [Confidential Information Protection and Statistical Efficiency Act \(CIPSEA\)](#), which passed in 2002 (with amendments via the [Evidence Act of 2018](#)), formalized protection for statistical data from other uses in all the agencies.

An important source of convergence has been scientific advances and evolution toward common best practices. Joint projects, working groups, personnel movements, and informal staff communication among the agencies also aid this process. Progress of these agencies' research and their use of academic research establishes best practices, which, when adopted, also make the agencies more similar. A key contributor to development and diffusion of best practices is the [Federal Committee of Statistical Methodology](#), an interagency committee dedicated to improving the quality of federal statistics. Diffusion of knowledge also occurs through professional associations, such as the American Statistical Association, the American Economic Association, and American Population Association, to name a few. The [Committee on National Statistics \(CNSTAT\)](#), part of the National Academies of Sciences, Engineering, and Medicine, also promotes best practices. The National Academies formed CNSTAT in 1972 to provide independent review of federal statistical activities.

The most important formal coordination comes through the activities of OMB, which is charged by [1995 statute](#) with coordination of the federal statistical system. The small office of [Statistical and Science Policy \(SSP\)](#), whose functions date back to the 1930s by action of the Executive Branch, largely handles this work. Its current responsibilities are defined legislatively in the [Paperwork Reduction Act of 1980, as amended in 1986 and 1995](#). The office, headed by the Chief Statistician of the United States, coordinates the activities of the federal statistical system to ensure efficiency, the effectiveness of the system, as well as integrity, objectivity,

impartiality, utility, and confidentiality of the information collected for statistical purposes. To achieve these goals, SSP's activities include “establishing statistical policies and standards, identifying priorities for improving programs, evaluating statistical agency budgets, reviewing and approving Federal agency information collections, and coordinating U.S. participation in international statistical activities” (p. 7, <https://www.whitehouse.gov/wp-content/uploads/2020/12/statistical-programs-20192020.pdf>). For example, SSP sponsors committees to create and update classification schemes for use by all federal agencies for industries, occupations, demographics, and geographic areas.

The SSP also chairs monthly meetings of the Interagency Council on Statistical Policy (ICSP), a forum launched informally in the late 1980s and authorized by statute in 1995. Membership in the ICSP includes heads of the 13 principal federal statistical agencies and was recently expanded (under the [Evidence Act of 2018](#)) to add 13 chief statistical officers from other departments that have small statistical units. See Table 1 for the list of all 26 ICSP members. The SSP and ICSP sponsor the Federal Committee of Statistical Methodology, mentioned earlier.

All of these steps have helped the agencies become more coordinated over time, even though they have not moved into a common umbrella agency.

4. The Present

Turning to the present, let us focus on five notable features of the landscape: outputs and inputs of statistical programs, uses and users, operations and funding, public perceptions and privacy, and the advent of new ‘organic’ data and technology. Many of these topics are also discussed in a thoughtful essay by Ron Jarmin, Deputy Director of the U.S. Census Bureau (Jarmin, 2019).

4.1. Outputs and Inputs

Statistical agencies’ choices of inputs and outputs for their official statistics reflect some very practical constraints due to limited resources, access restrictions, survey response rates, collection methods, and Balkanized production. At the same time, technological advances are adding new data collection and dissemination opportunities.

Currently, most official statistical programs are either very timely or very granular. As an illustration, consider BLS data on payroll employment. BLS publishes very granular data from the [Quarterly Census of Employment and Wages \(QCEW\) program](#). It offers state, county, and metropolitan statistical area (MSA) employment for detailed industries—with a 5-month lag after the end of each quarter. Alternatively, you can consult [Current Employment Statistics \(CES\)](#), a survey [based on the frame provided by the QCEW](#). The CES provides monthly data with just a 2-week lag, but with much less granularity by geography and industry. A search for unemployment data will get a similar answer. For high demographic and geographic granularity with low frequency and long lags, you can consult the [Decennial Census](#) or the annual [American Community Survey](#).

(ACS). Alternatively, you can use the monthly [Current Population Survey \(CPS\)](#) with its 2-week lag, but much less detail.

The starkness of the trade-off largely reflects the statistical agencies' resource constraints and lack of access to administrative data. Most statistical agencies' programs, particularly the timely ones, are survey-based. Timely, detailed data requires large sample sizes, which are very expensive. As a practical matter, for highly granular statistics, the agencies must rely on either infrequent censuses or administrative data. In the past, statistical agencies had little access to administrative data. There was little obligation, incentive, or even legislative permission for data-producing agencies to share their data with statistical agencies. The report of the recent [Commission on Evidence-Based Policymaking](#) (2017) and subsequent passage of the [Evidence Act of 2019](#) are helping to relax this constraint. Yet, as of today, access often remains inadequate.

The case of administrative data from the Unemployment Insurance (UI) system illustrates some of the paradoxes this situation creates. Currently, the BLS has agreements to fund each state's curation and submission of their UI employer records. BLS compiles the state data into the national QCEW business register that produces statistics directly and serves as the sample frame for employer surveys. Yet, the BLS has no access to other data produced by the UI system, in particular, worker-level wage and UI claims records.

[BLS could use enhanced individual UI worker records](#) to add modeled demographic and geographic granularity to its timely survey-based statistics and reduce revisions to its monthly payroll (Current Economic Statistics) job growth estimates. It could also replace all or parts of other survey-based programs, such as the Occupational Employment Survey and the National Compensation Survey. However, these and other improvements are not currently possible because BLS has no access to these records. The Census Bureau has gained limited access to UI wage records via individual agreements with the states for the [Longitudinal Employer Household Dynamics](#) program, but this arrangement does not include funds for states to improve the underlying records. Furthermore, terms of the agreements (which vary by state) prohibit many forms of sharing.

In another example of costly access restrictions, BLS cannot use Internal Revenue Service (IRS) tax data. The Census Bureau, which has access, uses small firms' IRS filings (particularly their industry classifications) to create a [business register](#) distinct from the BLS QCEW. Unfortunately, BLS and Census cannot synchronize their two business registers, because BLS is restricted from access to IRS data. Consequently, these two different registers lead to differences in statistical products built from them, posing problems for users who want to combine information from them. Crucially, the BEA regularly confronts these issues in assembling gross domestic product (GDP) and other national income and product account estimates using data by industry from both the Census Bureau and BLS.

Separately, statistical agencies also face the challenge of declining survey response rates. Over the past 3 decades, more people refuse to answer surveys. This is particularly true of household surveys, but is a problem with some business surveys as well. Explanations include survey fatigue as private sector requests have

increased, heightened privacy concerns, and growing distrust of government. This problem has worsened during the COVID pandemic. See, for example, the trends and impact of the pandemic on BLS survey response rates at <https://www.bls.gov/osmr/response-rates/home.htm>. Field staff are much more limited in their ability to go out and get responses in person. In addition, people's increasing uncertainty and distraction with other important priorities may suppress willingness to participate. Therefore, statistical agencies face a world with declining survey response rates, which make surveys less robust and more expensive than in the past.

At the same time, technological progress offers the agencies efficient new collection methods to add to personal visits, phone calls, fax, and mail. These alternative modes, such as internet collection, video calls, automated program interfaces (APIs), and web scraping, can reduce respondent burden and increase quality of data collected. All these new methods rely on a strong information technology infrastructure, suggesting rising economies of scale across programs and perhaps agencies. Efficiencies would arise from sharing resources such as data centers, training programs, and skills of specialized personnel. In addition, joint collection facilities might allow simultaneous collection of information for multiple surveys, reducing burden on large firms that need to report continually. Alone, many statistical agencies may not be able to invest in these capacities. Currently, many agencies can realize such economies of scale only if they are large or share facilities with nonstatistical agencies. More about this point later.

As can be seen from this discussion, another feature of U.S. official statistics continues to be balkanized data collection, production, and products. Statistical agencies and others are generating ever more types of data. This profusion provides more opportunities to combine data. Ironically, as we add more data to the pool of data sets out there, the marginal value of each data set is likely not falling. To the contrary, I would assert that in many cases, the value of new data is rising because of growing opportunities for valuable collaborations.

These combination opportunities increase potential benefits to interagency coordination. Yet because of restrictions of various sorts, data sharing among government agencies is far too rare. Even when restrictions are relaxed or public statistics are used, data users often cannot easily combine products from different agencies because of differences in concepts, timing, aggregation schemes, sample frames, or other issues.

4.2. Uses and Users

A salutary feature of today's environment is renewed emphasis on the value of evidence-based decision-making. For the statistical agencies, this means proliferating users and uses for their products, often beyond the purposes for which they were originally designed. The growth testifies to the value of official statistics even as it magnifies their value. Yet, meeting burgeoning and divergent needs poses challenges.

Users' sophistication varies greatly. Job seekers looking for information to guide career decisions in a dynamic labor market are among the least technically skilled data users. They may be in high school or college, employed, unemployed, or incarcerated. They need products such as the [Occupational Outlook Handbook](#) with

its accessible, authoritative information on wages, job growth, skill requirements, and safety. As I remember, this program alone accounts for about a third of all visits to the BLS website.

At the other end of the spectrum are academic researchers who analyze confidential microdata in the [Federal Statistical Research Data Centers](#). These centers, their users and publications continue to grow in numbers and importance to academic research.

Business users, who span the whole range in sophistication, are growing in number and diversity. As they use products originally designed to meet federal government needs, they are asking for changes in products or dissemination modes. Large, financial, and tech companies increasingly employ advanced analytics that often rely heavily on combinations of official statistics in addition to proprietary data sources. Smaller companies will now increasingly seek products like the [Census Business Builder](#) for local employment and demographic information. In general, business users prize timeliness and granularity to guide their decisions.

Policy uses of official data are also growing, particularly as a means for bridging some partisan divides. Indeed, the recent [Commission on Evidence-Based Policymaking](#) reflects this interest. These uses include fiscal projections, monetary policy, program evaluations, policy design, and economic development efforts. Many reflect expanded missions of existing programs. In one example, the [Consumer Expenditure Survey](#) (CEX), supplies BLS with updated expenditure weights used in construction the Consumer Price Index (CPI). Recently, BLS collaborated with Census to respond to interest in a measure of poverty which is referred to as the [Supplemental Poverty Measure](#) (SPM). BLS now constructs a research series of SPM thresholds from CEX data, with thresholds from 2009. Unfortunately, without funding, BLS cannot increase the CEX sample, update the processing system, or perform adequate systems testing to provide production-quality inputs, so this is new poverty threshold series remains in a research phase.

4.3. Operations and Funding

Meanwhile, the statistical agencies face urgent funding, personnel, and operational challenges. Inflation-adjusted budgets in many statistical agencies have declined recently, especially for BLS, the Bureaus of Justice and Transportation Statistics, and the National Center for Education Statistics (see <https://magazine.amstat.org/blog/2021/02/01/fy21-federal-budget/>). The BLS budget was flat in nominal terms for about 10 years. Yet, during that time, its costs have risen steadily for compensation, computer hardware and software, and services from the Census Bureau and private sector contractors. Furthermore, agencies face new expenditures for protecting confidentiality and cybersecurity. In a time of government austerity, the needs of statistical agencies have not been persuasive to appropriators.

Balkanized funding contributes to budgetary problems for statistical agencies. Each agency's budget lies within the appropriations for their parent department, involving the actors and committees unique to the parent. Statistical initiatives are promoted or not according to the parent's current priorities. Thus, no appropriators on the Hill or leaders in the parent departments look at the statistical system as a whole. Indeed, requests for

coordinated projects may receive funding for only one of the agencies that request it, as in the recent case of the new Supplemental Poverty Measures. Although the Census Bureau has received additional funding for its role, BLS has not. The Consumer Expenditure Survey has been unable to increase sample size enough to produce publication-quality expenditure estimates for low-income individuals. Thus, today's published Supplemental Poverty Measures are still constructed using lower quality consumption estimates than originally intended.

Another problem centers on staffing. Even though BLS workers are more satisfied than average federal employees, their turnover rates are high. This is likely the case for the other statistical agencies. Their talented statisticians, social scientists, programmers, and so on have many attractive, better-paid outside opportunities. The agencies compete by emphasizing the importance of the work and opportunities for training and advancement. Interestingly, at present, only BEA has authority to create special pay grades for some unique skilled positions. With the statistical agencies scattered in different parent agencies, extending such authority more broadly has no champion. Thus, staffing issues are likely to worsen as the demand for a broad range of data science skills continues to rise rapidly in both the agencies and the private sector. In a future steady state, the larger pool for these skills could benefit the statistical agencies, but for now it is a growing problem.

To promote efficiency and modernization across the federal government, a 2014 law pressures all offices (including statistical agencies) to share operational infrastructure within their parent departments. The [Federal Information Technology Acquisition Reform Act \(FITARA, a provision of the Carl Levin and Howard P. "Buck" McKeon National Defense Authorization Act for Fiscal Year 2015\)](#) requires federal agencies to consolidate IT and other services. The law grants parent departments' chief information officers (CIOs) approval authority over subagencies' IT budget requests, contracts for technology products and services, and appointments of subagency CIOs. It also codifies software-sharing requirements, IT portfolio reviews, consolidation initiatives, and more.

Yet, sharing services with non-statistical agencies for information technology, human resources, travel, and so on is problematic for statistical agencies. To begin with, the needs of statistical agencies can differ markedly from those of programmatic or enforcement agencies. Thus, shared services can be worse than those tailored to statistical agencies' needs. For example, information technology is central to statistical agency operations and involves unique activities (such as computationally intensive estimation and collection of survey responses). Thus, consolidation risks deterioration of service. Furthermore, priorities may differ, affecting long-term investment decisions and short-term performance. Statistical agencies prioritize meeting release deadlines very highly. Can we expect this singular focus from staff that serve the entire Department of Labor—including the Secretary's office, for example?

Shared services may also endanger response rates. Statistical agencies promise respondents to use their survey responses (some of which have sensitive information) only for statistical purposes; enforcement agencies have no access to these data. To underline this, in the past, BLS field staff would inform wary companies that BLS

has its own protected computer facility, separate from those of the Occupational Safety and Health Administration or the Wage and Hour Division, for example. Now, that clarity is gone. Even with fully effective firewalls, sharing computer services with enforcement agencies deprives field staff of a simple selling point; turning it to discussion of effective firewalls instead. At its worst, such collocation could result in intentional or accidental breaching of the legal protections given to respondents without the knowledge of the statistical agency that is responsible for its protection.

Simply put, sharing services opens the door to sublimation of a statistical agency's mission to the enforcement, policy, budgetary, or political goals of their parent department. Although OMB explicitly recognizes many of these issues in FITARA implementation guidance, the guidance has not protected statistical agencies adequately. They have little recourse to resist continued encroachments on their operational independence under FITARA pressures (see [Attachment K of OMB's 2015 implementation guidance for FITARA](#)). This can happen blatantly or with subtlety. An agency that refuses to accede to inappropriate requests from the parent department can be punished with poor service, prevention of needed investments, or higher costs charged to the statistical agency. Even without an inappropriate request, when a statistical agency shares IT and administrative costs, the parent department has the means to siphon off funds from the statistical agency to meet other priorities. Unfortunately, the laudable goal of increasing federal government efficiency in this way poses a clear risk to the political and fiscal independence of statistical agencies. Furthermore, when statistical agencies succeed in preserving independence, it may come at the cost of foregoing efficiencies that could be attained, particularly if they could share services with other statistical agencies.

4.4. Public Perception

Public perceptions of statistical agencies are also changing, and not for the better. Trust is mission-critical for a statistical agency. For respondents to deliver sensitive data to a statistical agency, they must trust the agency to protect their data and use it for important purposes. At the other end of production, data users must trust official statistics' quality and integrity or they will not use it. Thus, two aspects of statistical agencies' reputations loom large, especially currently: independence and privacy. Much of the public distrusts government and is wary of statistics. Yet, they have a much more positive view of statistical agencies than of most government. However, there are problem areas.

One problem is perceptual. It can be confusing to the public that statistical agencies both reside in departments headed by members of the president's Cabinet and produce official statistics that are free of political manipulation. After all, news reports generally list just the name of the department for the source, not the statistical agency. They reference GDP and the unemployment rate as issued by the Departments of Commerce and Labor, respectively, not BLS and BEA. Not long afterward, the secretaries of Labor and Commerce also issue policy and political interpretations of recent indicators. These same departments have agencies that enforce regulations and carry out the president's policy agenda. To reduce the inherent risk of confusion, [OMB Statistical Policy Directive 3](#) stipulates, "Except for members of the staff of the agency issuing the principal

economic indicator who have been designated by the agency head to provide technical explanations of the data, employees of the Executive Branch shall not comment publicly on the data until at least one hour after the official release time.” But, is this enough?

Another problem is that much of the public interacts with statistical agencies only when they are asked to respond to a survey. At that point, they are sensitive to the burden of answering the questions. They may also not distinguish between federal statistical surveys and other requests for their information. In addition, they may fear a loss of privacy. These days, they are voting with their feet, as it were, by not responding as much as in the past.

Lastly, users of official statistics seem to want more than they ever did before. Data users are frustrated by slow modernization, by website issues, by lack of granularity, timeliness, and agility and consistency across programs, as mentioned before. Perhaps because they are increasingly dependent on official data, data users are the most critical observers of the statistical agencies.

4.5. New Data Sources and Technology

Statistical agencies also have many new data science colleagues. Across the economy, digitized production, transaction, monitoring, communication, surveillance, and other processes generate a burgeoning and novel array of what former Census Director Robert Groves calls [“organic” data](#) sources and others term “big data.” Both of these terms distinguish these sources from surveys carefully designed to help construct particular measures. The data explosion accompanies a complementary boom in hardware (for storage and computationally intensive work) and software (for artificial intelligence, linking, and modeling).

I have already mentioned one influence of this growth: more hiring competition for skilled data scientists. Another impact is that access to new digitized data sets raises disclosure risks from statistical agency releases, especially as matching technology improves. For example, the general proliferation of online records greatly facilitates identification of people in the decennial census who were sufficiently anonymized in a pre-internet world (Garfinkel et al., 2018).

There are at least two implications for statistical agencies. First, they need to upgrade their disclosure avoidance strategies. For example, the Census Bureau will apply formal differential privacy protections to 2020 Census products. Second, the agencies likely will need to coordinate more to ensure that combinations of their products are privacy protected as intended but still usable.

Perhaps the most frequently mentioned impact is that statistical agencies now have many more data sources and techniques to consider in their continuing modernization efforts. Some opportunities are already in production and others will likely follow suit. This has implications for the composition of modern statistical agencies. Investigating promising options requires extensive research into issues such as fitness for use, impact on data continuity, and cost implications. For a general discussion of statistical agencies’ assessment of data

quality, see “[Framework for Data Quality](#).” For discussion of opportunities and challenges posed by alternative data in price statistics, see Groshen et al. (2017) and Erhard et al. (2021). When options do not advance into production, they may still provide valuable validation opportunities. If they do prove viable, the payoff usually comes in new or better products; rarely are cost savings as large as might be supposed. The main reason is that adding new non-survey data requires devoting staff and resources to investigations, purchases, and management of data quality and supplier relationships. Therefore, compared to the present, future official statistics will require more of these activities and rely relatively less on survey design and fielding.

Associated with the challenge of incorporating new data sources are the measurement challenges posed by the digital economy. For example, what is the correct way to measure gig work, price effects of telemedicine, or the value created by search engines? Adapting measurement to a dynamic economy is hardly new for statistical agencies. However, this ability depends on having the resources to conduct research into the design and processing of new indicators. Thus, modernized official statistics require adequate resources for ongoing research and data stewardship activities. Added to that, cross-agency collaboration would no doubt speed modernization and help ensure consistent approaches throughout the statistical system.

In another consequence, more private sector entities have joined statistical agencies in the business of disseminating statistics, such as economic indicators. Examples include [ADP’s estimate of payroll job growth](#) and [Billion Prices Project estimates of inflation](#). Despite some observers’ assertions, private sector statistics are not good substitutes for official statistics; they are far closer to complements. They use official statistics to construct their estimates and benchmark against them.

While at BLS, I put together a summary of the comparative advantages of official and private statistics. Table 2 summarizes these along six dimensions: exclusive data sources; methods and transparency; design and data quality; purpose and access; historical consistency and agility; and privacy and burden. Reviewing these advantages clarifies that official and private sector statistics will tend to operate at different spots on key trade-offs. On continuity versus relevance, official statistics lean toward continuity; on accuracy versus speed, official statistics lean toward accuracy; and on transparency versus complexity, official statistics lean toward transparency. Thus, the two sources serve as complements to each other in important ways:

- Private sector depends on official statistics to provide weights, history, benchmarks, information not addressed in private data, and dimensions of the universe.
- Official statistics incorporate select private sector statistics as inputs.
- Both sectors benefit from cross-validating measures and sharing techniques.
- Private sector indicators augment official statistics to meet narrower customer needs for specialized products, more detail, timeliness, and so on.

Ironically, this complementarity shows that new private statistics actually raise the need for official statistics, while at the same time some pundits interpret this growth as evidence that the private sector can eliminate the

need for official statistics. They have it all wrong.

Table 2. Comparative advantages of official versus private-sector statistics.

	Official statistics	Private sector statistics
Exclusive data sources	<ul style="list-style-type: none"> • Censuses, representative samples drawn from gold-standard frames (registers) or government administrative records 	<ul style="list-style-type: none"> • Large data sets for particular sectors, activities, or populations
Methods and transparency	<ul style="list-style-type: none"> • Documented, transparent methods • Standard errors published if available • Adhere to OMB and professional standards by design, testing, and validation 	<ul style="list-style-type: none"> • Many experimental, proprietary, novel, variable, and opaque approaches • Standard errors rarely published • No consistent externally recognized standards
Design and quality of input data	<ul style="list-style-type: none"> • Usually (except administrative records) designed to provide reliable answers to pressing questions • Tested and validated 	<ul style="list-style-type: none"> • Usually undesigned ‘by-product’ of automated economic activity • Quality may be low if not used for another important purpose
Purpose and access	<ul style="list-style-type: none"> • Products designed to address key policy questions or inform commonly made decisions • Public has equal access to all products 	<ul style="list-style-type: none"> • Widely released products often intended to advertise capacity and competence • Other products have restricted access and/or are tailored to narrow customer needs
Historical consistency and agility	<ul style="list-style-type: none"> • Published history with consistent methodology • Extensive documented testing, validation, and quality control • Lead times for innovations and often for production 	<ul style="list-style-type: none"> • Novel products and delivery mechanisms • Immediate, short time-to-market for production and innovation • No assurance of continuity

Respondent privacy and burden	<ul style="list-style-type: none"> Legislated privacy protections (CIPSEA, PRA, Privacy Act of 1974) with OMB oversight If non-administrative data, respondents assume some burden to participate 	<ul style="list-style-type: none"> Privacy, security, and confidentiality policies and practices not consistent or directly regulated Low or no burden on respondents
<i>Note:</i> OMB = Office of Management and Budget; CIPSEA = Confidential Information Protection Statistical Efficiency Act; PRA = Paperwork Reduction Act.		

5. The Future

With all these developments underway, what future lies ahead for official statistics? We can divide the myriad possibilities into three general paths.

5.1. Path 1: Muddling Through

The agencies could continue to muddle through with few major changes. While this outcome is hardly a panacea, the agencies would continue to produce hugely consequential information that everybody could rely on. In addition, with incremental changes over time, they would achieve some portion, albeit small, of the modernizations possible.

Two problems beset this scenario. First, the nation would lose out from being behind the curve perpetually. It is hard to measure the magnitude of those losses, but we can be sure that more resources will be misallocated, hobbling policy and productivity. For example, consider foreign investment, the source of many jobs in the United States today. Many highly paid jobs in the United States are in foreign-owned companies. How do foreign investors decide where to invest and how much to invest in the United States? They explore official statistics for the locations under consideration, such as the QCEW, Census Bureau economic surveys, the American Community Survey, and BEA's GDP products. Without more timely and granular data, investors' choices will be less well targeted and more risky. That would restrain investment and make it less successful. This is not the best outcome.

Second, there are serious downside risks (from lack of funds and other threats that I describe below) that make this path unstable in the end. Underperforming will undermine trust and reliance on the statistical system, leading to further neglect. Our statistical system faces growing unmet needs as well as growing untapped potential. Thus, I fear that muddling through can be only temporary. That leads to the other two possible paths.

5.2. Path 2: Downward Spiral

The worst future is a self-reinforcing downward spiral of official statistics—an outcome not beyond the realm of possibility. Appropriators could continue to keep most agencies' budgets flat nominally. That squeeze would

limit agencies' ability to follow up on respondents and introduce new data sources, lowering data quality. With lower quality, trust in official statistics would fall, endangering their political support for funding increases. That would further suppress data quality or end programs entirely, cutting the agencies' basis of support again. Hence, the self-reinforcing downward spiral. Figure 1 shows how all these outcomes could build on each other to reduce a statistical system to a shadow of its former self.

The Potential Downward Spiral of Official Statistics

Figure 1. The potential downward spiral of official statistics.

Several factors listed here could exacerbate or trigger this spiral. For example, lower funding could compromise independence or vice versa. As statistical agencies' funding waned, their parent agencies could exert more influence over statistical decisions. Or, interference from the White House could destroy trust in the agencies directly. Indeed, the Trump administration saw blatant political attempts to influence scientific decisions at the Census Bureau, the Centers for Disease Control, and the National Weather Service, to name a few. Such attempts certainly erode confidence and, if successful, could compromise the quality of official statistics.

With that loss of confidence and accuracy, agencies would lose some of their dedicated civil servants who are so proud of their work. Their salaries are not what keeps them there. If they can no longer take as much pride in their work and have no time for research on how to improve their programs, the agencies will lose valuable human capital to the private sector at a higher rate.

Of course, the private sector might step in partially in the event of this downward spiral. However, as I mentioned, the quality of their products is hard to judge because they lack benchmarks, continuity, and an

obligation to transparency. Furthermore, as I noted at the beginning, theory is very clear that we can expect an undersupply of this public good. Hence, both quality and quantity would suffer.

We have examples of countries allowing their statistical systems to spiral downward. I will mention two recent cases where the trigger has been political. Less recently, the Puerto Rican statistical system suffered a similar decline, from which it has yet to recover.

The Greek government has arrested and prosecuted [Andreas Georgiou](#), former head of its statistical agency, for the crime of releasing accurate GDP data against the wishes of the prime minister. He has been convicted, fined, and may face prison time. This is a real tragedy.

In Argentina in the 1990s, a widespread distrust of official inflation figures inspired former MIT graduate students Alberto Cavallo and Roberto Rigoban to web scrape prices to demonstrate the inaccuracy of the official figures. They proved it quite convincingly. Note that they used comparisons between the trusted U.S. CPI inflation measures with web-scraped data to demonstrate that their methodology was valid. The same exercise for Argentina showed very divergent results. Interestingly, that research led to the [Billion Prices Project](#), which produces daily inflation rates for a large set of countries. It is most valued in cases where people suspect that their statistical agencies are not doing a good job. Many of their products are now sold privately.

5.3. Path 3: A Better Statistical System

Fortunately, it is also quite possible that future official statistics will achieve their potential to inform policy and the public better than in the past. Then, the self-reinforcing cycle can work in the opposite direction to sustain a robust future.

5.3.1. Prerequisites

The route to a better outcome entails important prerequisites, including the following:

- **Independence** of agencies from political influence, with no erosion caused by implementation of FITARA or changes in administration.
- **Program coordination** across agencies. The goal is to support combination of statistics across programs and to produce joint products by means of coordinated design, application of classifications, and data quality across programs.
- **Data access and sharing.** The goal is to reduce acquisition costs in dollars, in time, and in red tape for very many kinds of data, including government administrative data, corporate data, and web scraped data.
- **Safe repositories** for confidential corporate and household records. The goal is to protect respondent privacy and data integrity while allowing matching to create new products.
- **Engagement with users.** The explosion of uses and users requires consultation to help prioritize across possibilities.

- **Communication** about the importance of official statistics. The goal is to promote survey participation, appropriate use of products, protection from political interference, and adequate appropriations.
- **Capacity** to hire, train, and retain the best staff. Along with increased technical and communication skills, working with organic data requires skills to manage data quality and relationships with data suppliers.
- **Shared services** or platforms across agencies where larger operations would be more efficient. These platforms (for production, training, purchases, etc.) need independence from non-statistical agencies and flexibility for innovation.

5.3.2. Solutions Without Reorganizing

How do we get from here to there? This will take many more steps than I can go through here, so I list only a few, starting with those that do not involve reorganization.

To achieve those prerequisites, we certainly need expanded data-sharing legislation for government administrative data. The Evidence Act (2018) brought some progress, but we need more. For example, as I mentioned, UI records are still not available and BLS and Census cannot synchronize their business registers. Both of those likely need changes in data-sharing legislation.

Another area for attention is working with respondents, that is, the data providers. What other public or administrative sources should we tap to augment and reduce reliance on surveys? Then, to preserve the integrity of our important surveys, do we need mandates—or incentives for participation? What privacy guarantees are most effective? These are important questions for both household and business surveys. We likely need the most change in how we collect information from businesses. As businesses digitize more of their operations, the statistical system needs to engage with companies to promote interoperable standards for records.

Working with the corporate sector, the statistical agencies could help design “data schema” for wide private sector use. A schema is a way of organizing sets of information for particular purposes. The private sector uses standardization for many production technologies (think about lumber sizes, lightbulb sockets, paper products, and so on). There is interest now in designing common schema for information purposes as well.

Such schema should both meet companies’ internal needs for interoperable records across business units and across countries, and meet national statistical requirements. Use of common data schema would facilitate due diligence among companies that seek to purchase or merge with another. It would also be easier to combine corporate information with newly purchased entities and to use off-the-shelf analytics to track activity and performance. Common schema would reduce companies’ response burden for statistical surveys because they would already keep their information that way. Another carrot for the companies is that common schema would help them compare their internal measures with official statistics. Meanwhile, official statistical products would improve due to better response rates and higher quality and consistency of input data.

Encouragingly, such an effort is already underway. The [T3 Innovation Network](#), led by the Chamber of Commerce Foundation, is creating schema for interoperable learning and employment records, with BLS and Census Bureau participation. Such records would meet the need for consistent UI records going forward, augmented by data on job title, demographics, hours, and work location. Statistical agency access to those records could vastly expand the timeliness and granularity of labor market data, even as it reduces the need for at least two large BLS surveys. Similarly, enhanced and consistent learning information kept by employers could be tapped to better understand skill gaps and occupational changes.

This model has great potential because the business community, which is powerful and can be leery of externally imposed standards, is already behind it. In addition, it could help create data-safe repositories from which statistical agencies could collect information, reducing collection costs.

5.3.3. Reorganization as a Path Forward

With or without those changes (hopefully, with them!) let me talk about consolidating the statistical agencies. How could centralizing more of our system help meet the prerequisites laid out above? To fix ideas, let me answer the question by considering one option to consolidate—even though it is not the only possible way. We could create a new strong, independent ‘StatsUSA’ to take on many of the current functions of the OMB Office of Statistical and Science Policy (SSP) with more staffing and an elevated Chief Statistician of the United States would then actually oversee an important part of the U.S. statistical system. It would exist outside of the Cabinet departments and house many of the existing statistical agencies, including at least Census, BLS, and BEA—moving them out from under Cabinet secretaries.

Of course, consolidation schemes have been proposed before, notably by Janet Norwood, one of my predecessors (see Norwood, 1995), and the last three presidential administrations. Moving agencies around in the federal government is expensive and unpopular with the losing departments and appropriators. The complexity of our decentralized statistical system makes it much harder to centralize. To visualize its complexity, see Figure 2, which is reproduced from the appendix from National Academies of Sciences, Engineering, and Medicine (2021). The figure shows all the appropriating committees and agencies involved in funding and running the statistical agencies. This large set of actors, each with their own priorities, makes coordination of the statistical system difficult. In addition, they constitute a large group of people unlikely to be enthusiastic about consolidating the system, since they would lose some power.

Figure 2. Principal Statistical Agencies by Congressional Appropriations Committee and Parent Department/Agency.

This is why, without strong external support or a crisis, consolidation is not an easy lift. Yet, the idea persists because it makes sense. Let me talk about why.

Let me say clearly from the outset that consolidation is neither strictly necessary nor sufficient to achieve the prerequisites. Yet, I believe that it is advisable because coordination and independence would be hard to achieve without more centralization.

Why might consolidation be insufficient? If you combined these agencies into a StatsUSA without changing any other laws, practices, or funding, then you would get few of the benefits intended. In fact, you could reverse progress if the new agency was less independent or defunded. Thus, consolidation, per se, is insufficient. However, in the context of trying to make a difference, consolidation is likely to be very helpful.

Why might consolidation be unnecessary? With sufficient workarounds, you might be able to meet all of the prerequisites listed without administrative reorganization. This has been the tack taken in the past to improve coordination for the statistical system. However, meeting the prerequisites for a thriving statistical system without consolidation would require big changes in authorities, perhaps unprecedented in the federal government. You would need to change the rules by which separately housed agencies communicate, make decisions, and share data and resources. You would need to overlay an overarching structure far more robust than the SSP is now, on top of the structure shown in Figure 2.

While this might be feasible in principle, it is far more straightforward to bring the agencies together, as is done in most other government functions that require ongoing coordination and can share inputs. Consolidation

offers a platform and accountability to help make immediate and complete changes. Going forward, the country will need a new mix of statistical programs and better infrastructure. It will be hard to identify and get to the optimal mixture of products and of shared services without broad input and flexibility to experiment. For example, we need the right mixture of ongoing official statistics and temporary products (such as one-off supplements, mechanisms to gather new forms of information, new ways to parse the data, etc.). We also need the right balance between robust disclosure protection and convenient access. Communication and dissemination will also require more attention and resources. As the variety and number of data users expands, statistical agencies must be able to communicate with all of them, in ways appropriate for their level of technical sophistication. Furthermore, the agencies need to work together on helping data users navigate among and combine the products of different agencies. To get all of this right, the statistical system needs an internal forum in which to facilitate prioritization and optimization.

Admittedly, putting all our statistical “eggs” in one basket can be dangerous. Consider what happened in Canada, which has a more centralized agency, Statistics Canada, leading its statistical system. In 2011, some lawmakers were unhappy with the long form of their census. They succeeded in making that survey voluntary and reducing funding for the whole agency. Many Canadian statistical programs suffered because of anger against one particular program. Fortunately, these steps were reversed in 2016. The lesson for us is that such dangers must be addressed from the outset by mechanisms to provide consistent funding and limit political interference.

There is also the risk that a larger agency could neglect or dilute adherence to the core missions of the constituent agencies. Thus, their stakeholders could lose out. To avoid severe consequences and resistance from current stakeholders, reorganization should include steps to engage with users and preserve the core missions.

With or without consolidation, the statistical agencies need more funding, and they need this funding to be flexible and consistent. They need this to avoid the problems caused by funding lapses, continuing resolutions, and the like.

One form of needed flexibility would be particularly hard to achieve without consolidation. The system sorely needs cross-agency investment funds to support occasional large modernization efforts. For example, these investments could create platforms for linking, new products, one off-data collection, or experimentation. This approach could help solve the agencies’ problem of finding resources to upgrade operations while continuing to churn out high-quality statistics products on a tight schedule.

What could be the source of higher, flexible, and assured funding? One option is to allocate to the statistical system a portion of a [Tobin tax](#) on financial transactions, should it be enacted. A Tobin tax is a small fee on financial transactions such as stock market trades. A main goal of the tax is to reduce excess volatility caused by high-frequency trading. However, along the way, it produces revenue. Financial firms and investors rely

heavily on official statistics to help them allocate resources efficiently. Thus, were a Tobin tax to be imposed in the United States, it seems very appropriate to devote a portion of Tobin tax revenues to ongoing support for the statistical system.

Many of these changes, through reorganization or not, will require mindset adjustments within the statistical agencies. The core cultures of statistical agencies prize continual innovation, but also continuity and adherence to their agency missions. In this moment, innovative agility bears strong emphasis. One key to promoting a new mindset will be recognizing that the best way to honor the past is not by holding onto old practices slavishly, but by being innovative today. The idea to communicate is, “This is a world-class agency because our predecessors innovated. So, we need to follow in their footsteps by innovating today.” This is not only possible, but necessary.

One last thought related to encouraging innovation, particularly in a more coordinated system. Agencies may consider adopting explicit sunset provisions and success criteria for new products. Preannounced time limits can be extended if success criteria are met, but otherwise they avoid problems. They help manage data users’ and staff expectations. In addition, they prevent the proliferation of zombie projects that have outlived their original purpose or failed to live up to initial hopes. Thus, sunset provisions encourage innovation by reallocating resources more readily, encouraging clear communication, and imposing helpful structure on initial project design.

6. Call to Action for Statisticians and Other Data Scientists

I hope you agree that statistical agencies face both risks and huge opportunities ahead. The better path leads to official statistics that are more useful, more granular, timely, relevant, combinable, trustworthy, and safer than ever before. This will not happen by itself. It will require work by the agencies, our representatives on Capitol Hill, the private sector, and members of the statistics and data science community—which is why I am glad to have your attention at this moment. Every one of you can make a difference when you take these six important steps.

- a. **Vote and engage politically.** Your nationally elected representatives vote on appropriations. Provision of needed data infrastructure is part of the good government that they should know you care about.
- b. **Do research to improve official statistics.** This is interesting, important work. You, your colleagues, and students can have a lasting impact by affecting how information is collected, analyzed, and interpreted going forward. Look to the agencies for inspiration on these cutting edge topics.
- c. **Teach about official statistics.** Your colleagues, students, and mentees should know about the federal statistical system and official statistics: how to use them, find them, and support them. Do not neglect this part of their training. In addition, do not forget to discuss career paths for people who use federal statistics, including working for a statistical agency.

d. **Participate in federal surveys.** Respond when you are chosen for a federal survey and ensure that your employer participates also. While at BLS, I eagerly made some calls to professors at major universities that were not participating in BLS surveys. If your employer participates in federal surveys, thank the staff who do so. If not, explain to your colleagues that their data will be safe and why their participation is important. The logic is simple. You rely on official statistics in your work directly or indirectly, and your products will be better if you and employers like yours are represented in the data.

e. **Speak up.** Use your trusted voice to support better official statistics. Acknowledge your dependence on official statistics in your publications and talks, including citing them properly. Far too often, a table or chart using official statistics lists a second party provider or the home department as the data source. Such citations do not help people find that statistic when they need it. Furthermore, they do not give credit where it is due. Another part of speaking up is defending the value of official statistics. Do not allow people in your professional and social circles to get away with saying damaging things about lying with statistics or that people cannot trust the data. You know differently. All too often, we are lazy; we ignore the statement, roll our eyes, or change the subject, because we know Uncle Harold or Professor So-and-so is a crank. However, we shirk our responsibility when we do that. Instead, you should challenge them and share your knowledge about how these data are collected and processed, and why they are trustworthy. The other side is poorly informed and you are not.

f. **Engage with statistical agencies.** Perhaps most important for a data scientist/statistician, take advantage of the many ways to help improve official statistics.

- Work for a statistical agency and/or advise your students to do the same.
- When you have a question about statistics, instead of calling your favorite friend, call up the data producers directly and talk to them. They like to talk to the people who use their statistics. Moreover, you will learn things that your friends do not know about those data. You have a lot in common with the people producing these data. You will both benefit from that conversation. I promise that you will learn things that you could learn no other way.
- Collaborate with agency researchers. This is an opportunity to get access to data, work on important topics for an agency, and help make meaningful changes. See the agencies as staffed by people with whom you can collaborate.
- Identify important issues and technical questions that they should look into. Conversations with people like you inform agencies' decisions about the next issue to work on.
- Attend advisory council meetings. All statistical agencies have advisory council meetings. They were always public, but now they are all remote. Attendees learn what the agencies are doing, thinking about, and studying. They also become acquainted with the leaders and staffers there. You may know some of the people serving on those advisory committees. Even if you do not, you can contact them and start a

conversation. Furthermore, if you start to engage with the agencies in the ways I mention, they may ask you to serve on one of these advisory committees.

g. Join support networks for statistical agencies. Lastly, concerted coordinated actions taken by support networks can make a big difference in funding and other outcomes for an agency. The [Census Project](#) supports the Census Bureau. The National Center for Health Statistics has [The Friends of NCHS](#) and BLS has [The Friends of BLS](#), which I happen to chair. These groups are very easy to join and do not cost money. They will keep you apprised of what is going on and offer you the opportunity to sign on to letters to the Hill to support appropriations, to comment on pending legislation, and so on.

7. Conclusion

To sum up, I believe that official statistics, an important public good and the foundation of our national data infrastructure, are at a crossroads. With new technologies and needs, the United States could have official statistics that are more relevant, interoperable, granular, and timely than we have today. The agencies could tap into the new wealth of non-survey data to accomplish these goals without increasing burdens on survey respondents. They could operate *more in sync with each other* to facilitate combining and sharing data and *more independently from politics* to reinforce public trust in their integrity. We could better protect them from the risks of being defunded or required to combine operations with non-statistical agencies. For many other countries, the same is true. These changes to our data infrastructure would improve our national well-being by getting better information to decision makers in our businesses, homes, and governments. All of these changes are possible, but not without effort by the statistical system's stakeholders.

To secure these advances, and avoid the alternative downward spiral in official statistics, we must move forward with steps such as data-sharing legislation, flexible and dependable funding, adoption of common data schema, and a modernized, more coordinated statistical system. The agencies cannot do this alone. This is why the statistical system needs the active support of the statistical and data science communities—including you.

I hope that readers will now be energized, informed, and more thankful than ever to the statistical agencies for the great work that they do every day.

Editors' note: This article has been updated by the author to include corrected information about the Bureau of Labor Statistics in section 4.2.

Disclosure Statement

Erica L. Groshen has no financial or non-financial disclosures to share for this article.

References

- Erhard, L., B. McBride, & A. Safir. (2021, February). A framework for the evaluation and use of alternative data in the consumer expenditure surveys. *Monthly Labor Review*. U.S. Bureau of Labor Statistics. <https://www.bls.gov/opub/mlr/2021/article/a-framework-for-the-evaluation-and-use-of-alternative-data-in-the-consumer-expenditure-surveys.htm>
- Garfinkel, S., J. M. Abowd, & C. Martindale. (2018). Understanding database reconstruction attacks on public data. *ACMQueue*, 16(5). <https://queue.acm.org/detail.cfm?id=3295691>
- Goldberg, J. P., & W. T. Moye. (1985). *The first hundred years of the Bureau of Labor Statistics*. Bureau of Labor Statistics Bulletin 2235.
- Government Accountability Office. (2021). *COVID-19: Urgent actions needed to better ensure an effective federal response*. Report GAO-21-191, pp. 212–225. <https://www.gao.gov/assets/720/710891.pdf>
- Groshen, E. L., B. C. Moyer, A. M. Aizcorbe, R. Bradley, & D. Friedman. (2017). How government statistics adjust for potential biases from quality change and new goods in an age of digital technologies: A view from the trenches. *Journal of Economic Perspectives*, 31(2), 187–210. <https://doi.org/10.1257/jep.31.2.187>
- Jarmin, R. S. (2019). Evolving measurement for an evolving economy: Thoughts on 21st century U.S. economic statistics. *Journal of Economic Perspectives*, 33(1), 165–184. <https://doi.org/10.1257/jep.33.1.165>
- National Academies of Sciences, Engineering, and Medicine. (2021). *Principles and practices for a federal statistical agency* (7th ed.). The National Academies Press. <https://doi.org/10.17226/25885>
- Norwood, J. L. (1985, July). One hundred years of BLS. *Monthly Labor Review*. U.S. Bureau of Labor Statistics. <https://www.bls.gov/opub/mlr/1985/07/art1full.pdf>
- Norwood, J. L. (1995). *Organizing to count: Change in the Federal Statistical System*. The Urban Institute Press.
- U.S. Commission on Evidence-Based Policymaking. (2017). *The promise of evidence-based policymaking*. U.S. Government Printing Office. <https://bipartisanpolicy.org/download/?file=/wp-content/uploads/2019/03/Full-Report-The-Promise-of-Evidence-Based-Policymaking-Report-of-the-Commission-on-Evidence-based-Policymaking.pdf> and <https://bipartisanpolicy.org/wp-content/uploads/2019/03/Appendices-e-h-The-Promise-of-Evidence-Based-Policymaking-Report-of-the-Commission-on-Evidence-based-Policymaking.pdf>
- Wiatrowski, W. J. (2009, June). BLS at 125: Using historic principles to track the 21st-century economy. *Monthly Labor Review*, pp. 3–25. U.S. Bureau of Labor Statistics.

<https://www.bls.gov/opub/mlr/2009/06/art1full.pdf>

©2021 Erica L. Groshen. This article is licensed under a Creative Commons Attribution (CC BY 4.0) [International license](#), except where otherwise indicated with respect to particular material included in the article.