CHAPTER 8

Improving Economic Statistics

Statistical systems have substantial value for both public policymakers and private decisionmakers. Administration and Congressional policymakers rely on statistics for budget decisions and related fiscal policy choices, and the Federal Reserve System relies on statistics for formulating sound monetary policy. Private firms combine internal company data with publicly provided statistics to make sales projections and investment decisions. In addition, contracts often use price or wage indexes to adjust payments for inflation.

Statistical systems, like physical infrastructure, become obsolete or depreciate with time. In a dynamic market-based economy, like that of the United States, new industries emerge, old industries contract, and firms find new ways of organizing and conducting their activities. The challenge for those who manage statistical systems is to keep pace with changes in the economy by continually evaluating the relevance and reliability of the statistics that are produced. In addition, it is important to maintain the continuity of statistical time series to facilitate meaningful historical comparisons. Up-to-date, relevant statistics are critical to the public policy process: they help frame policy debates by providing a sense of the size and scope of an issue, as well as the likely benefits and costs associated with a given policy action.

Advisory committees and researchers drawn from other parts of the government and academia help statistical agencies maintain the high quality of the data they collect and publish. They provide advice and engage in academic-style research that ensures that collected data are useful and relevant to issues people care about. Their work also enhances future data products by suggesting ways to improve the statistical system.

The statistical community in government, business, and academia recognizes that statistical agencies can improve the quality, usefulness, and efficiency of their statistical operations through cross-agency sharing of selected business data. Such interagency data sharing facilitates the synchronization of data across agencies, which in turn improves the comparability of different datasets and makes the statistical products of all agencies more valuable. For example, a measure of industry input (such as labor hours) often comes from one agency, while a measure of industry output comes from another agency. If each agency classifies a given firm as belonging to a different industry, then the productivity (output per input) of both industries may be mismeasured. By sharing classification data, agencies can reconcile these differences to ensure that the firm is classified in a consistent manner. In addition to improving the accuracy of government statistics, data

AUTHENTICATED
U.S. GOVERNMENT
INFORMATION
GPO

synchronization may reduce the reporting burden on survey respondents, thereby improving the efficiency of the Federal statistical system.

The key points in this chapter are:

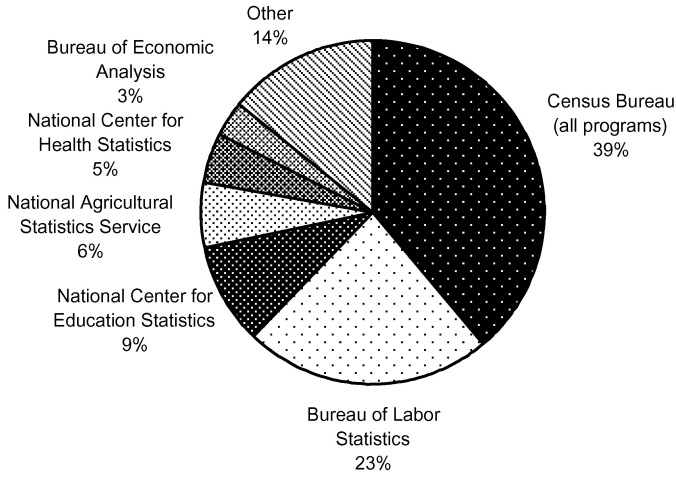
- Robust statistical systems produce products that are important to understanding the changing state of the economy and to formulating sound policy. But statistical systems, like physical infrastructures, become obsolete or depreciate with time if they are not maintained.
- Statistical measures must keep up with the changing nature of the economy to be relevant and useful. For example, it is important that these measures reflect new and growing industries (such as high-technology industries or services) and intangible capital (such as research and development).
- Disruptions in a statistical series render it much less useful to policymakers and other data users. Thus, continuity in statistical series is an important goal.
- More effective statistical use can be made of existing data. In particular, amending relevant legislation to enable full implementation of the Confidential Information Protection and Statistical Efficiency Act (CIPSEA) could greatly improve the quality of Federal statistics.

An Overview of the U.S. Statistical System

The U.S. statistical system comprises many organizations inside and outside the U.S. government that produce statistics. Of particular interest in this chapter are Federal statistical agencies (whose principal function is to collect, compile, analyze, and disseminate statistics) and associated organizations, such as the Federal Reserve Board, that produce economic data to inform policy decisions. As of 2007, these organizations produced 38 statistical releases designated as "principal Federal economic indicators." These indicators include everything from agricultural prices to new home sales, the unemployment rate, and gross domestic product (GDP).

Among the Federal statistical agencies, the largest is the Department of Commerce's Census Bureau, which accounted for 39 percent of spending by principal statistical agencies in fiscal year 2007, as shown in Chart 8-1. Spending on statistics by the Federal Reserve and many regulatory and program agencies, as well as by nongovernmental organizations, is excluded from this calculation. The Census Bureau's spending expands even more during years leading up to the Decennial Census. Although the Decennial Census receives a great deal of attention, the Census Bureau conducts numerous other surveys much more frequently.

Chart 8-1 **Budget Authority for Principal Statistical Agencies, Fiscal Year 2007**
The Census Bureau accounts for the largest share of Federal statistical spending.



Note: Total does not add up to 100 percent due to rounding error. "Other" includes Department of Agriculture (Economic Research Service), Department of Energy (Energy Information Administration), Department of Justice (Bureau of Justice Statistics), Internal Revenue Service (Statistics of Income Division), National Science Foundation (Science Resources Statistics Division), Department of Transportation (Bureau of Transportation Statistics), and Social Security Administration (Office of Research, Evaluation, and Statistics).
Source: Office of Management and Budget.

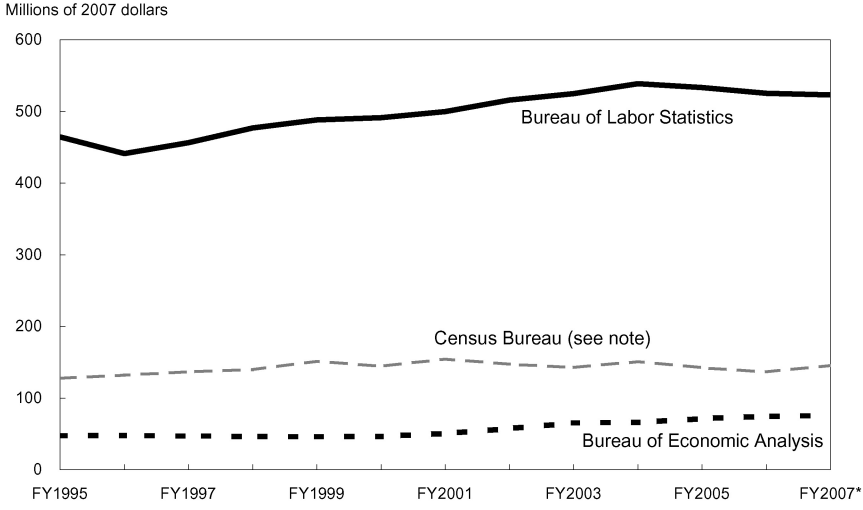
The second largest Federal statistical agency, at 23 percent of spending, is the Department of Labor's Bureau of Labor Statistics (BLS), which produces, on a monthly and quarterly basis, the vast majority of U.S. data on employment and prices that are used to provide timely assessments of the current state of the economy. A combined 20 percent of spending is accounted for by the agencies responsible for preparing statistics on education, agriculture, and health.

The Department of Commerce's Bureau of Economic Analysis (BEA) is a relatively small statistical agency, with just 3 percent of spending. Its data products rely substantially on input data collected by other agencies and include the National Income and Product Accounts, which are among the most comprehensive measures of the size and current performance of the U.S. economy. Construction of the national accounts (which includes GDP) makes the BEA a consumer of vast amounts of data from the Census Bureau (such as import and export data) and the BLS (such as wage and salary data), as well as many other public and private sources.

Statistical data may be collected on a regular basis (monthly, quarterly, or annually) or on a relatively infrequent basis (every 5 or 10 years, for example). Chart 8-2 shows the pattern of real spending by several statistical agencies

on economic statistics that are produced at least once per year. Examples include the monthly employment report from the BLS; monthly data on durable goods orders and new home sales, quarterly data on services, and official annual estimates of income and poverty from the Census Bureau; and quarterly GDP from the BEA. Chart 8-3 shows the pattern of real spending for several Census Bureau programs that are produced on an infrequent basis (the Decennial Census or the 5-year Economic Census and Census of Governments). In both charts, expenditures on these programs were adjusted for inflation with the Office of Management and Budget's deflator for "all other" Federal outlays (primarily salaries and expenses for nondefense agencies). As shown in Chart 8-2, spending on economic statistics has largely kept up with inflation. Real spending by the BLS has decreased slightly since 2004, after a period of steady growth that began in 1997. The three statistical agencies in Chart 8-2 account for about 50 percent of the total spending on economic statistics (excluding the Decennial Census and periodic spending by the Census Bureau). Total spending on economic statistics by other agencies has remained level.

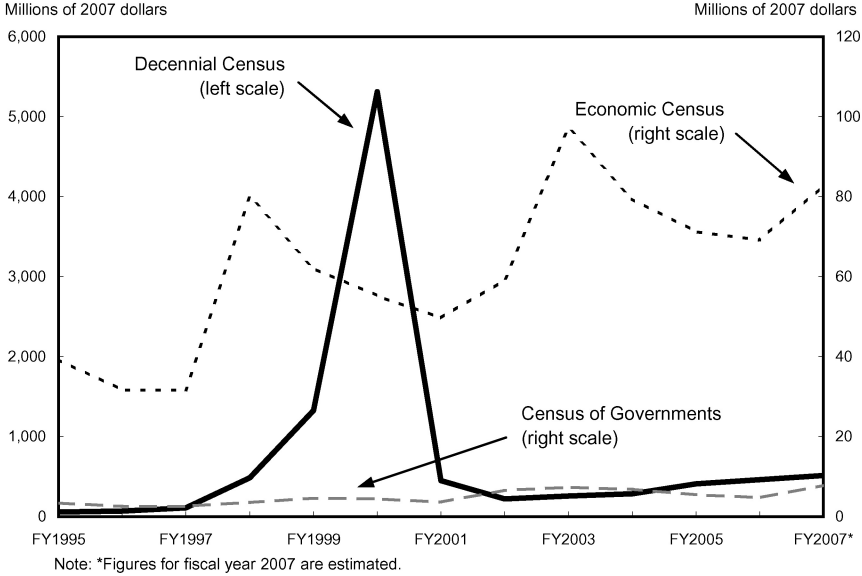
Chart 8-2 **Real Federal Appropriations for Economic Statistics**
Real spending on economic statistics has been fairly flat in recent years.



Note: *Figures for fiscal year 2007 are estimated. Census Bureau spending excludes spending on the Decennial Census, Economic Census, and Census of Governments. Bureau of Labor Statistics spending adjusted for the transfer of several large programs from the Employment and Training Administration in 2001.
Sources: Office of Management and Budget and Department of Commerce (Census Bureau, Budget Division).

As shown in Chart 8-3, spending on programs with a 5- or 10-year production cycle exhibits a clear pattern: spending climbs in preparation for the survey during the years immediately preceding the survey, peaks during the year of the survey, and then falls quickly upon completion. For example, real spending (in 2007 dollars) on the Decennial Census, which measures the size of the U.S. population, rose from about $110 million in 1997 to over $5.3 billion in 2000, before quickly falling back. The slight upward trend in decennial funding in the last several years was partly for the development of the American Community Survey, discussed later in this chapter. The 5-year budget cycle of the Economic Census, which measures output and related statistics in the business sector, is also apparent, though the year-to-year changes in spending are considerably smaller. The Census of Governments—which collects data on government organizations, finances, and employment—also picks up every 5 years, but the annual level of spending on this program is relatively small (less than $10 million), so the variations are less noticeable.

Chart 8-3 **Federal Statistical Appropriations for 5- and 10- Year Censuses**
The Decennial Census occurs every 10 years; the Economic Census and the Census of Governments occur every 5 years.



Note: *Figures for fiscal year 2007 are estimated.
Sources: Office of Management and Budget and Department of Commerce (Census Bureau, Budget Division).

Unlike the 5- and 10-year censuses, which are fairly well understood, funding requests for other statistical initiatives, such as new products or needed updates to existing programs, are easily misunderstood. For example, a major redesign of an existing survey's methodology ideally involves running two surveys concurrently (one with the old methodology and one with the new methodology) for a brief period of time so that the effect of the change in methodology can be isolated. Understanding this effect is essential if results from the redesigned survey are to be meaningfully compared to those of the survey being replaced.

The Importance of Statistical Systems

Providing accurate information to households, firms, and policymakers is an important role of government statistical agencies. Most decisionmakers in private industry, in Federal, State, and local governments, and in private households, rely in some way on data collected by Federal agencies. Federal economic statistics are designed to be consistent, unbiased, and reliable over time. These statistics can prove particularly useful if their availability and analysis allow a costly problem to be prevented or remedied more quickly and efficiently.

Private decisionmakers benefit from high-quality statistical systems because they improve the value of the information upon which firms and individuals base their decisions. For example, in formulating investment decisions, industries may use data on final demand or on the output of other industries that buy their output. A firm may examine a variety of labor market data, such as wage rates and educational attainment in the region, when deciding where to open new branches of the company. Airport authorities may study regional economic prospects when considering expansion decisions. Worker organizations and employers may track inflation trends and factor these price changes into their expectations for nominal wage gains. Popular press accounts based on occupational earnings may help students choose colleges, fields of study, or other training that will have long-term implications for their career paths.

State and local governments rely on a wide variety of statistical data to benchmark their performance, to plan for the future, and to readjust their allocation of resources. For example, a State that finds its high school dropout rate rising relative to other States may opt to devote more resources to education. Likewise, a city that finds its crime rate rising relative to other localities may choose to devote more resources to law enforcement. States and cities may study data on local population growth to assess the need for new transportation systems, schools, and other types of physical infrastructure.

Monetary and fiscal policymakers also rely on high-quality, publicly available data for understanding the changing state of the economy, for formulating sound policy on a wide range of macro- and microeconomic issues, and for economic forecasting. For example, monetary policy depends on accurate measures of resource utilization, current employment and unemployment trends, productivity trends, inflation trends (including unit labor costs), and housing market developments. If inflation estimates are overstated, monetary policy might be unnecessarily restrictive. Similarly, if productivity is overstated, policymakers may think that the economy's productive capacity is expanding quickly enough to accommodate rising output without being inflationary, and the resulting monetary policy may not be restrictive enough to limit the risk of inflation. Fiscal policy depends on accurate measures of GDP growth, potential GDP growth, labor markets, and demographic change to forecast future government outlays and revenues. If productivity is growing more slowly than believed, then revenue projections may be too high, and as a result, policymakers may adopt spending plans that are inconsistent with overall budget goals. Thus, a clear understanding of the true trends in these variables is critical to making sound budget projections.

Keeping Up with a Changing Economy

There are many ongoing efforts to update the statistical infrastructure to better reflect the changing economy and to more accurately reflect the economy as it stands now. These efforts include maintaining the relevance of statistical classification systems, better measuring the changing population, improving the measurement of the service-sector output, and measuring the contribution of investment in intangible assets (such as research and development) to economic growth.

Statistical systems rely heavily on the classification of activities, and over time classification structures are changed to better reflect the economy. Sometimes the changes are incremental, such as when an industry is split into two more detailed industries. Other changes are more substantial, such as the transition from the Standard Industrial Classification (SIC) system to the North American Industry Classification System (NAICS). Despite the benefits of NAICS—such as better coverage of advanced technology industries, as well as better international comparability—the transition was nonetheless disruptive to statistical agencies and data users. In particular, the transition to NAICS broke the historical continuity of many data series. Further, the official use of NAICS began in 1997 but not all data series incorporated NAICS classification in the same year. Statistical agencies have extended

many of their statistics backward in time on a NAICS basis, but doing so is difficult and time-consuming. There is sometimes inadequate information to cleanly separate SIC-reported industry data into the redefined industries and the greater industry detail under NAICS. Many statistics produced by the BEA and BLS, for example, have been extended back to 1992 or 1990, respectively, and a few series go back further. The Federal Reserve Board extended its industrial production and capacity utilization statistics back to 1972 based on the results of an extensive microdata reclassification research project that was conducted with the Census Bureau's Center for Economic Studies. Despite the improvements that came with NAICS, it can be argued that the classification system has yet to fully capture the character of modern economies. For example, the shift over time from manufacturing to services is still not fully reflected in the level of detail collected, or even in the number of defined industries: The 2007 NAICS recognized nearly 17 percent more private service industries than manufacturing industries (550 versus 472), even though the gross output of private services was about 3 times larger than that of manufacturing in 2005.

The Census Bureau recently introduced the American Community Survey (ACS) to provide more current data on our Nation's population and its characteristics. With a sample size of approximately 3 million addresses, the ACS collects important demographic, housing, social, and economic information for use in the administration of Federal programs and the distribution of Federal spending. The ACS is the Nation's largest household survey and will eliminate the need for the Decennial Census long form in future censuses by providing data for the same detailed geographic locations as the long form. Unlike the long form, however, it will provide single-year estimates for geographic areas with populations of 65,000 or more annually, rather than estimates every 10 years. Smaller geographic areas will be sampled over 3- and 5-year intervals, allowing the Census Bureau to produce estimates down to the census tract or block group. For policymakers who need to make decisions affecting the lives of large numbers of people, having up-to-date estimates of population characteristics is critical to understanding a program's likely beneficiaries and its likely costs.

Another recent improvement to the Federal statistical system has been more accurate and timely measurement of service-sector output. In 2004, the Census Bureau introduced the new Quarterly Services Survey (QSS), the first new principal Federal economic indicator in nearly 30 years. Prior to the introduction of this survey, the 13 private service sectors—which together account for about 55 percent of GDP—were measured, at most, once per year, if covered by the Service Annual Survey. Even at the annual frequency, the available surveys account for just 30 percent of GDP. The only comprehensive measures of service-sector output come every 5 years during the Economic Census. Therefore, the QSS is important because it measures

service-sector output much more frequently, which keeps the measures of service-sector activity in the National Income and Product Accounts more current. Even so, the QSS covers a limited portion of the service sector, which means there is room for improvement by broadening the coverage of the survey.

Efforts aimed at understanding the contribution to economic growth of investment in intangible assets, such as spending on research and development (R&D), is another example of the work being done to make statistics better reflect the state of the economy. The BEA, with the support of the National Science Foundation, created a R&D satellite account of the U.S. national accounts, which treats R&D as an investment rather than an expense. Accounting for R&D in this fashion would have boosted the average annual change in real GDP from 1995 to 2004 by nearly one-quarter percentage point, to 3.3 percent. The BLS has created statistical measures of business employment dynamics that help explain the contributions to net changes in employment that come from job losses versus job gains. As the length of the time series increases, these employment measures will be useful for understanding changes in employment over the business cycle. For example, a policy response to a decrease in net employment that results from an increase in gross job losses (i.e., greater layoffs or voluntary separations) may be different from one that results from a decrease in gross job gains (i.e., weaker hiring). The former might reflect transitory industry shifts, while the latter might suggest a generally weaker macroeconomic situation.

Other efforts to better reflect the changing economy include work at the Federal Reserve Board, the BLS, and the BEA to improve price measures to better represent the rapid pace of technological change in high-technology products like computers. When adjusted for improvements in quality, prices are estimated to fall much faster, which raises measures of real output.

Attempts to keep up with the changing economy are complicated by efforts to maintain consistent time series. Long time series are valuable for making historical comparisons and inferring long-run relationships among economic variables. When a time series is short, it is hard to know if there is anything exceptional about a current event. The strength of any conclusions that are drawn is a direct function of variation in data. Short time series have too little cyclical variation. Similarly, panel data—which follows a group of persons, households, or firms over time—are valuable for inferring changes over time from cross-sectional changes due, say, to different population composition.

There are a variety of ways in which economic measures can fail to keep up with the changing nature of the economy. Examples include:

- Firms' increased substitution of purchased services for secondary activities previously done within the firm (such as payroll processing) means that some statistics, such as employment, will document this change as a shift to services. In this example, the data accurately capture the

current use of services, but the data do not reflect the *change* in the use of services correctly, as the earlier data classified all activity within the firm (including payroll processing) by the predominant activity of the firm (i.e., construction, manufacturing, etc.).

- Established industries tend to receive a disproportionate share of attention compared to new, growing industries. Industry and product classification codes are more likely to be kept than eliminated, while new industries and products are often poorly measured and tracked, at least initially.
- The growth of *professional employer organizations*—companies that provide employees to firms on a contractual basis—has led to data-reporting problems and, consequently, to inaccurate employment and wage data for industries and localities. Professional employer organizations that report employment centrally, rather than separately for each client, can obscure both the industry and location of the workers and our understanding of employment change and dynamics, negatively affecting data from BLS, the Census Bureau, the BEA, and all derived products.
- The prices for some items may fail to fully reflect changes in the quality of the items. Improvements in quality, if properly accounted for, tend to boost measured real output. The split between consumer and business spending on some products may be updated infrequently, which can lead to misstatements about which components of GDP are growing more rapidly. Both factors tend to result in less reliable estimates of real spending by consumers and businesses.
- Housing and geographic samples for the consumer price index (CPI) become outdated as the population distribution shifts (see Box 8-1).

Improving the Value of Existing Statistical Data

Federal Government statistical agencies are focusing on three ways to improve the value of existing statistical data: Improve the detail in publicly available data products, facilitate well-defined and secure research on the underlying microdata, and synchronize data produced across agencies.

Government agencies strive to improve the usefulness of their data products by providing greater detail while protecting the confidentiality of respondents. The Census Bureau, for example, employs several techniques to avoid disclosing individually identifiable data. *Synthetic data,* modeled on original data, retain the needed statistical properties of the original data but protect the confidentiality of respondents by modifying all or selected variables. The Census Bureau creates synthetic data to obscure the underlying demographic data used in its "On the Map" feature. This feature creates maps showing

Box 8-1: How to Reverse a Decline in Statistical Infrastructure: Improving the Sample for the Consumer Price Index

The housing and geographic area samples for the Consumer Price Index (CPI), currently based on 1990 Decennial Census data, are overdue for an update. Each year these samples become more out of date, in that the samples do not reflect almost 20 years of population growth, demographic changes, and new housing construction. Because of its widespread use to estimate price changes, the accuracy of the CPI influences a range of economic variables in both the public and private sectors. For example, within Federal programs, the CPI is used to adjust Social Security payments, civilian and military retirement payments, and individual income tax brackets for inflation. A study by the Congressional Budget Office found that a 1 percentage point reduction in the growth rate of CPI estimates beginning in January 2006 would have reduced the Federal budget deficit or surplus by $14 billion by the end of 2007 and $153 billion by 2015.

The Administration has proposed to update the 1990 Decennial Census–based housing sample used by the BLS with data from the Census Bureau's new, continuously conducted American Community Survey (ACS) and/or private sector sources. With continuous updating, the sample would never be more than 3 years old. This change would increase the accuracy of the CPI by creating a more representative housing sample, reduce respondent attrition, and reduce potential bias by more accurately reflecting new construction. Moreover, using the ACS to update the geographic sample on which the CPI is based would result in estimates that more accurately reflect the geographic distribution of the population and its characteristics.

commuting patterns and workforce data—where people live and work by age, earnings, and industry—for geographic areas selected by the user. Another method used is *noise addition*—the controlled introduction of variation from reported levels to detailed data that otherwise could not be published, with small compensating adjustments to other data in the same series. The Census Bureau uses this method to ensure that an individual company's data cannot be readily inferred from published Survey of Business Owners data or other estimates.

Government statistical agencies benefit when researchers can subject the data and the methodology behind the statistics to academic scrutiny in a secure research environment that maintains security of the data, restricts access to the level of data essential for an authorized project, and protects the

confidentiality of respondents. The analysis of underlying data by academics is an inexpensive way for statistical agencies to improve their data products. For example, academic researchers typically investigate relationships among variables in a single survey, or in several surveys, that are not examined during routine data-processing procedures. Their nonstandardized data reviews can uncover anomalies that should be resolved before the data are released, or provide the basis for future improvements in standardized data-processing routines. In addition, this third-party scrutiny adds to the credibility of the data products. For example, the Census Bureau's Research Data Centers (RDCs) provide secure, restricted access to Census Bureau data for authorized researchers. Likewise, the BLS researcher access program provides secure, restricted access to BLS data. In both cases, researchers must undergo a strict approval process and face significant penalties for violating the laws protecting the confidentiality of responses to government surveys.

Previous research at the RDCs has led to new data products and changed thinking about many important economic issues. For example, an important strand of academic work separated net employment flows—the published employment changes with which people are familiar—into gross job creation and gross job destruction. The quarterly BLS Business Employment Dynamics data release—which reports the number and rates of gross jobs gained at opening and expanding establishments, as well as the number and rates of gross jobs lost at closing and contracting establishments—is an example of a new data product that grew out of this work. Importantly, the Business Employment Dynamics data release is tabulated from already collected company data records, thus creating no additional respondent burden. It is an important example of drawing upon academic research to improve the use of existing data in order to create new data products.

A third way that the Government can improve the value of existing data—and the method that offers the most substantial opportunities—is to allow the BEA, BLS, and Census Bureau to link their business data, while maintaining confidentiality. This linking would result in more accurate and reliable economic indicators, lower budget costs for the agencies, and lower response burdens for survey respondents. For example, at present, both the Census Bureau and the BLS ask firms to break out employment and payroll data by establishment in the Company Organization Survey and Multiple Worksite Report, respectively. If these agencies could share their business data, these two surveys, which are mailed to multiunit companies, could be combined, reducing the response burden of these firms and reducing survey costs for the statistical agencies.

The Administration recognizes that the sharing of key *business* data among Federal statistical agencies has tremendous potential for exploiting synergies among the agencies and for improving the quality of Federal statistics. In 2002, with Administration support, the Congress passed the Confidential

Information Protection and Statistical Efficiency Act (CIPSEA), described in Box 8-2, whose stated purposes were: 1) To protect the confidentiality of information collected by Federal agencies for statistical purposes, and 2) to improve the efficiency of the Federal statistical system by authorizing limited sharing of business data among the Census Bureau, the BEA, and the BLS for exclusively statistical purposes. In 2007, the Office of Management and Budget issued implementation guidance for CIPSEA. The first part—data protection—has been effectively implemented across agencies, but the second part—improving statistical efficiency—cannot be fully enabled without additional legislation. Because business tax data (such as company name and address) are used to construct the Census Bureau's business list, many Census Bureau data products are considered to be comingled with tax information. Therefore, full implementation of CIPSEA would require changes to the portion of the Internal Revenue Service (IRS) code that authorizes the statistical use of business tax data.

Box 8-2: The Confidential Information Protection and Statistical Efficiency Act (CIPSEA)

The two parts of CIPSEA are confidential information protection and statistical efficiency.

Confidential information protection: Subtitle A establishes standardized safeguards to protect the confidentiality of data collected by Federal agencies for exclusively statistical purposes. These safeguards include the assurance that information will not be used against a respondent in any government action and that inappropriate disclosure of confidential data will be considered a felony and carry significant criminal penalties. In other words, data collected for statistical purposes cannot be used for tax, immigration, or other enforcement purposes.

Statistical efficiency: Subtitle B authorizes the sharing of business data among the Census Bureau, the Bureau of Economic Analysis, and the Bureau of Labor Statistics for exclusively statistical purposes in order to:

- Reduce the paperwork burdens imposed on businesses that provide requested information to the Federal Government;
- Improve the comparability and accuracy of Federal economic statistics by allowing these agencies to reconcile differences in business lists; to develop consistent classifications of businesses into industries; and to improve coverage; and
- Increase understanding of the U.S. economy (including key industries and regions), develop more accurate measures of the impact of technology on productivity growth, and enhance the reliability of the Nation's most important economic indicators, such as the National Income and Product Accounts.

A major goal of fully implementing CIPSEA is to better reconcile the BLS Business Establishment List—based on State unemployment insurance records—and the Census Bureau's Business Register—based, in part, on IRS records. One study found that over 30 percent of single-establishment firms had different 6-digit NAICS industry codes in the two lists, and another study revealed large discrepancies in measures of industry-level employment across surveys.

The failure to coordinate data across agencies can lead to noticeable inaccuracies, especially when one needs to calculate a measure that combines data from two agencies. For example, the implications of discrepancies in establishment classifications are particularly acute when measuring labor productivity, which is an important statistic for economic policymakers, including those who project the Federal budget. Labor productivity is the ratio of output, measured by the Census Bureau, and hours worked, as measured by the BLS. Accurate productivity estimates depend upon these labor and hours worked measures being given consistent industry classifications, which is unlikely if the underlying business lists are inconsistent.

Differences in industry classification would also result in discrepancies in the rate of real GDP growth reported by key sectors. For example, in the Computer and Electronic Product Manufacturing Subsector (NAICS 334), the growth in real value added in 2002 would have been 15.6 percent if payroll data from the Census Bureau's Economic Census had been used. Instead, the growth in real value added was published as 7.4 percent, a statistic based on payroll data from the BLS. Without carefully analyzing the confidential business lists used for the Economic Census and the BLS payroll data, it is difficult to know which payroll measure should be used. Some efforts to share data have proven useful in reducing inconsistencies and reducing burden. The BLS has shared industry identifiers with the Census Bureau since 1992 and geographic identifiers since 2002, particularly for new and small businesses. These industry codes covered over 3 million businesses in 2007 alone and now account for about 30 percent of the Census Bureau's business codes. Expanding data sharing would extend this work and further improve consistency and accuracy of key data series.

A 2006 report noted that data sharing might highlight opportunities for understanding data reporting that would better focus resources on activities that would improve the measurement of national economic activity (such as the reporting of stock options). The National Income and Product Accounts provide two measures of national activity, one based on total output (GDP) and one based on total income (gross domestic income or GDI). In theory, these measures should be equal. In practice, they differ by a measurement error called the *statistical discrepancy.* The statistical discrepancy can be persistent: From 1995–2000 real GDI grew 0.6 percentage point faster than real GDP, on average, per year. If the growth rate of the GDI were projected

forward instead of the growth rate of GDP, the budget implications could be substantial. An analysis of fiscal year 2006 by the Office of Management and Budget found that if the GDP were persistently understated by 1 percent, the projected cumulative budget deficit would be overstated by $530 billion over a 5-year period.

Better measures of business formation are needed to understand the changing composition of the business sector and the factors that contribute to business and job creation. Data synchronization would help agencies track business formation more accurately and on a more timely basis by reconciling the business lists from the Census Bureau and the BLS. For example, the Census Bureau's Business Register relies heavily on the Economic Censuses (conducted every 5 years) for information on business structure. In the intervening years, however, the Census Bureau makes use of its annual Company Organization Survey, which covers all employers with more than 250 employees, but only a sampling of smaller companies. The Census Bureau's Business Register generally does a good job identifying ownership links among establishments (e.g., when a single firm owns establishments in two different States). However, the information on ownership is weaker for smaller firms because only a subset of these businesses is surveyed during the years between the 5-year censuses. Firm restructuring often contributes to the difficulty of tracking parent–subsidiary relationships. The BLS Business Employment Dynamics accurately measures the universe of business openings and closings on a quarterly frequency but may not always successfully track parent–subsidiary relationships. Combining the strengths of the Census and BLS business lists would improve the ability to discern whether a new establishment is an entirely new firm or a new branch of an existing firm, and therefore improve understanding of business dynamics.

Data synchronization could also help reconcile differences between similar statistics produced by separate agencies. For example, the BLS publishes wages and salary data based on its Quarterly Census of Employment and Wages business list and the Census Bureau publishes payroll data in its County Business Patterns series. A comparison of 2003 private wages and salaries revealed that these two measures differed by significant amounts. For example, the BLS measure of wages and salaries in New Mexico was 4.2 percent higher than the Census Bureau measure, while in Alaska, the BLS measure was 9.5 percent lower. At the national level, BLS data were 0.6 percent (or $25.1 billion) lower than County Business Patterns data, but they were 2 percent (or $6.7 billion) lower for New York. Understanding the sources of these differences (such as differences in reporting and coverage) may yield improved regional measures that would have several implications:

- Distribution of Federal funds to the States: BEA per capita personal income data, based largely on BLS data, are used in the formula that calculates how to distribute the Federal share of Medicaid funding to

States. Wages and salaries and wage-related components account for two-thirds of personal income. In 2003, State private wage levels based on BLS data were $2.5 billion higher in Texas and $7.1 billion lower in Washington than levels based on the Census Bureau's County Business Patterns.

- State tax and budget planning: The dollar difference between BLS and Census measures of wage and salary growth from 2001 to 2002 would result in significantly different projections of State and local government income taxes received: a $165 million discrepancy in New Jersey and a $193 million discrepancy in Massachusetts. The $1.2 billion wage growth difference in New York would yield a $173 million discrepancy in projected State and local tax revenue.

Conclusion

The quality of public policy debates depends, in large part, on the availability of relevant and reliable statistical data. Consistent data series ensure that newly gathered data can be meaningfully compared to previously collected data. At the same time, it is also important that the statistical system maintain the flexibility to create new data products that keep up with the changing nature of the dynamic global economy. The infrastructure required to develop and produce these data, like any infrastructure, requires continuous investment to maintain and improve the system, but not all data improvements are costly. For example, existing economic data on businesses could be improved through the full implementation of the Confidential Information Protection and Statistical Efficiency Act without increasing the reporting burden for respondents, without compromising the confidentiality of the data collected by the Federal statistical agencies, and without significantly raising costs of the data collection and tabulation. Maintaining solid statistical systems ensures that public policymakers and private decisionmakers will have access to the information needed to understand our dynamic economy.