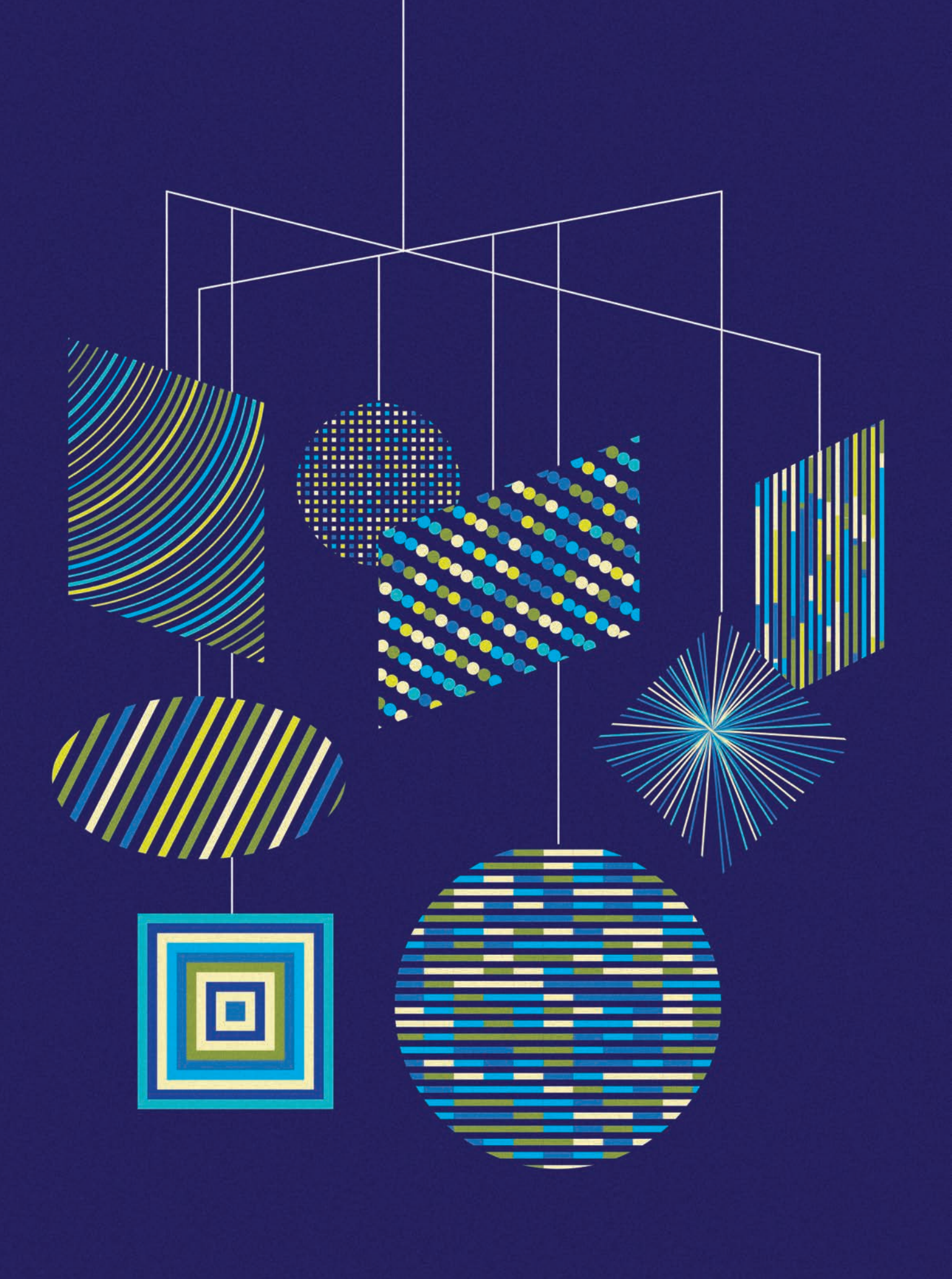

ALTERNATIVE DATA AND MONETARY POLICY

Claudia Sahm

Central bankers are tapping nontraditional data sources for a more complete picture of the economy

In the spring of 2020, the Federal Reserve faced a challenge: The COVID-19 pandemic was upending daily life with shutdowns, social distancing, and heightened uncertainty, but the traditional economic statistics the Fed used to calibrate monetary policy struggled to keep up with the pace of change and did not cover some of the novel features of the pandemic economy. Yet the Fed was not flying blind; it was able to pivot to nontraditional data sources it had previously developed, such as payroll processing and credit and debit card transactions, to track the rapid deterioration in the economy.

Even in the best of times, high-quality, timely data is critical to making sound monetary policy. If policymakers see signs of higher inflation, for example, they may consider raising interest rates to cool the economy. But, if the job market appears to be weakening, they may consider lowering rates to spur economic activity. It takes time for changes in interest rates to affect economic outcomes, so speed in accurately assessing the direction of the economy is also important for effective policy.

To keep a finger on the pulse of the economy in real time, the Fed relies on a wide array of statistics generated by government agencies such as the Bureau of Labor Statistics (BLS) and the Department of Commerce. These statistics, typically based on representative surveys, are considered the gold standard by policymakers, investors, business leaders, and the public. Increasingly, though, the Fed has supplemented them with nontraditional sources of data, often supplied by private companies. The defining feature of these nontraditional sources is the data was not created for the purpose of making economic statistics; rather it originated in the process of running a business or a government program and then was repurposed for economic statistics.

Filling gaps

This nontraditional data is often timelier or more granular and, as a result, can fill in some gaps in government statistics. It can also provide an added perspective on critical economic outcomes, such as employment. Finally, it can be used to improve the quality of traditional data sources. Nevertheless, nontraditional sources should be viewed as a

Illustration by Greg Mably

complement to traditional data in informing policy, not as a substitute.

The central focus of monetary policy is stabilization of the business cycle, so there is a premium on accurately and quickly assessing turning points. Nontraditional data can be particularly helpful in these circumstances. That's because government statistics on key variables like unemployment, inflation, and economic growth are published weeks or even months after the fact. The delay in releasing private company data is often substantially shorter, a few weeks or even days.

The timeliness of alternative data sources was particularly useful at the start of the pandemic, which triggered a short, deep recession. A review by Fed staff noted that its internal weekly estimates of employment, based on data from ADP, a large payroll processor, showed large declines in late March 2020. This was more than a month before the BLS published its own monthly employment report, which also showed large declines.

The pandemic downturn was unusually fast-moving, but higher-frequency and timelier estimates of employment have broader applications. For example, anytime the monthly BLS estimates of jobs shift down sharply, as occurred in 2025, the weekly ADP estimates offer early insight into whether the trend will persist or reverse. In addition, ADP estimates are highly relevant during government shutdowns resulting from a congressional impasse over the budget, when official data isn't available.

Inflation episodes

The granularity of alternative data is another advantage for Fed policymakers seeking to assess the impact of changes in trade policy on consumer price inflation. Theory and experience suggest that an increase in import tariffs will cause a one-time rise in the level of prices, which only temporarily increases inflation. In that case, the Fed should "look through" tariff-related inflation and not raise rates. But testing the hypothesis is challenging, because key statistics the Fed consults don't identify prices of goods by country of origin. Instead, the analysis must compare the prices of broad categories of goods by their average share of imports in the past.

This is where the granularity of alternative data offers a more direct path to monitoring tariff price effects. Alberto Cavallo, a professor at Harvard University, and two collaborators are one such data source. They have constructed daily price indices using online data from five major US retailers, which include country of origin, tariff rates, and selling price for 350,000 goods. They find that the prices of imported consumer goods have risen more quickly than those of goods produced domestically, relative to pre-tariff trends. Moreover, the price effect of tariffs is more pronounced for domestic goods that compete directly with tariffed imports than for domestic goods that do not. Overall, the effects have been relatively modest, a finding consistent with studies using traditional data sources. Such high-frequency, granular data can also help assess whether the upward adjustment to the price level is complete.

More granular alternative data sources also proved useful to the Fed and other decision-makers during the pandemic, which dramatically shifted consumer and business behavior. Private company data on physical mobility was deployed to monitor those shifts during social distancing, along with administrative data on the number of COVID cases. Measures of supply-chain stress were also instrumental in gauging inflationary pressures. In addition to surveys of purchasing managers and shipping price indices, the stress on supply chains was gauged with real-time data on shipping container movements. To be sure, traditional data sources also helped fill gaps in policymakers' understanding of the economy. The Census Bureau, a major source of traditional data, quickly stepped into the breach, launching short online surveys to gauge the pandemic's impact on households and small businesses.

Loss of precision

Alternative data can help maintain, and even improve, the quality and cost-effectiveness of traditional statistics. Government agencies rely heavily on surveys of people and businesses, which are designed to be representative of the overall economy. But these have drawbacks. For one, costs have increased over time as people and businesses become less willing to participate. For another, falling participation rates reduce the precision of the resulting estimates.

This loss of precision can create uncertainty about inflation or employment dynamics and hinder a timely, appropriate monetary policy response. Nontraditional data offers a potential solution. For example, the BLS now uses private company data instead of surveys for several components of the consumer price index, including prices of used cars, airline tickets, and wireless telephone contracts.

There is scope for further use of private sector data, though the acquisition cost and reliability of such data present challenges. A private company could decide to stop sharing its data or sharply raise its price, which could threaten the continuity of the government statistics. Careful testing at statistical agencies is also necessary to ensure that nontraditional sources improve the precision of estimates rather than substitute new sources of noise for old ones.

Business formation

Improving the accuracy of the initial estimates of traditional data is another area where alternative data could be useful, especially at economic turning points. Monetary policy decisions are made in real time, so the real-time data must be as accurate as possible. The government's monthly estimate of payroll employment is one example. It relies on a survey of business establishments, with results adjusted for the fact that businesses come and go. (The adjustment is based on something called the "birth-death model.") Shifts in net business formation during and after the pandemic, combined with long lags in data availability, have led to significant errors in the model and substantial annual revisions in previous employment estimates. Researchers have shown that weekly tax filings for employer identification numbers provide a reliable forecast of business formation in subsequent quarters. Aligning the birth-death model with more timely indicators of business formation could improve the accuracy of the initial estimates of employment when economic conditions are changing.

Even official data is subject to error, such as sampling error resulting from using partial surveys instead of a full census. So using multiple independent estimates can improve our understanding of official estimates. A new initiative from the Federal Reserve Bank of Chicago, for example, blends official with alternative data on the labor market to construct an estimate of the current month's unemployment rate. The analysis includes data from Indeed, a site used by job seekers and recruiters; Lightcast, a provider of labor market analytics; and Google searches on unemployment. However, the project is in its early stages, and it will take time to establish its reliability.

Impact of policy

Once Fed officials have adjusted monetary policy, they must assess its effects. Nontraditional data can be helpful here as well. Research on the distributional consequences of monetary policy, for example, has expanded with the availability of sources such as household-level credit records, bank accounts, and administrative records. During COVID-19, when interest rates fell, a study using data on property tax forms and deed records from CoreLogic showed that Black, Hispanic, and lower-income borrowers were less likely to refinance than Asian, White, and higher-income borrowers. Systematic differences in refinancing costs played a role. Another study, using Internal Revenue Service individual income tax records, found that unexpected tightening in monetary policy led to greater income inequality, primarily by worsening outcomes for low earners. Unexpected easing, however, decreased inequality.

For all the advantages of nontraditional data sources, they are not a replacement for the traditional kind. Indeed, their usefulness often depends on traditional data. As an economist at the Fed, I worked on a project to transform credit and debit card transactions from First Data (now Fiserv) into daily state-level estimates of retail sales, which were later used to track the economic effects of Hurricanes Irma and Harvey in almost real time for the Federal Open Market Committee.

But this source presented challenges. The growth in sales in raw transactions mixed factors specific to First Data, such as the acquisition of clients for its payment-processing business, with changes in US consumer spending. Only the latter is relevant for economic statistics. As one step to solve this problem, we used the five-year Economic Census from the Department of Commerce to reweight the card transactions from the company's clients to be representative of US businesses. Such benchmarking is common when building economic statistics from nontraditional sources. Our project faced other problems common to alternative data sources, such as short time series for seasonal adjustment and troubleshooting anomalies. Comparisons with national monthly retail sales estimates from the Census Bureau gave us confidence in using the more granular private data for policy work.

Any user of nontraditional data faces challenges. For the Fed, the limited availability of such data to the public poses special difficulties. The Fed's strategic framework for monetary policy emphasizes that transparency is critical to accountability and improves the outcomes of monetary policy. Relying on data sources that are not widely accessible reduces transparency; outsiders cannot verify the Fed's analysis, and only market participants who pay for access to private data can see what the Fed sees.

We have seen how policymakers can use alternative data sources to gain a fuller picture of economic conditions, potentially leading to better policy outcomes. Improving the quality of data will require strong ties between government statistical agencies, private sector data providers, government officials, and academics. Nontraditional data presents opportunities and challenges, but understanding macroeconomic dynamics is the goal of both nontraditional and official government statistics. F&D

CLAUDIA SAHM *is chief economist at New Century Advisors, an investment management firm, and a former principal economist at the Federal Reserve Board of Governors.*