

**Communicating the accuracy and relevance of statistical data during a pandemic:  
A practitioners' report**

Michael D. Levi and Emily L. Liddel

2020 Eighth IMF Statistical Forum—Measuring the Economics of a Pandemic

Mr. Levi is an economist serving as Associate Commissioner for Publications and Special Studies at the U.S. Bureau of Labor Statistics. Ms. Liddel is an economist serving as the Chief of the Division of BLS Publishing at the U.S. Bureau of Labor Statistics.

**Abstract**

*The COVID-19 pandemic has led to widespread changes in U.S. Bureau of Labor Statistics (BLS) standard practices. The mix of data collection modes in BLS surveys has changed. New data sources have been introduced. Response rates have declined in many surveys, but not uniformly across industries or demographic groups. Methods used in estimation have been adjusted to more accurately reflect extremely rapid changes in the economic conditions BLS measures.*

*All these methodological changes pose a communications challenge: How can a statistical agency best provide information to a non-specialist audience so that they can assess the ongoing accuracy of BLS data and their relevance to users' needs? This paper provides a practitioners' report on the questions BLS encountered, the issues they raised, and the solutions BLS fielded. It also touches on how the lessons learned during this unusual situation may inform continuing communications approaches even after the current crisis has subsided.*

**Background**

The U.S. Bureau of Labor Statistics (BLS) is the primary producer of gold-standard data on labor markets, working conditions, inflation, and productivity in the United States. These data are essential for evidence-based decision-making by policy makers, businesses, and households. They are especially critical during economic shocks when rapid and possibly radical responses may be required to address economic turmoil.

The BLS experience during the 2020 coronavirus (COVID-19) pandemic and efforts to contain it provide a textbook example of users' hunger for timely, accurate, and relevant data. It also provides a textbook example of how significant economic and social disruptions can interfere with the usual processes developed to provide these data.

Some areas of the United States began curtailing economic and social activity in February 2020, and in mid-March the president declared a national emergency. Many businesses shut down or severely reduced activity. Many employees began working from home, working fewer hours, or losing employment altogether.

The COVID-19 pandemic also had a notable effect on BLS and led to sometimes significant changes in standard agency practices. The most obvious was that, in mid-March, BLS (and the Bureau of the Census, which collects household data for BLS) ceased all in-person data collection. In addition, staff were placed on maximum telework status which resulted in the closure of BLS telephone data collection centers.

Thus, in many BLS surveys, the mix of data collection modes changed. In April 2019, for the Consumer Price Index commodities and services price survey, 74 percent of outlets were visited in person, 10 percent were contacted by telephone, and 16 percent were contacted online. During the pandemic, in April 2020, the mix had changed to zero personal visits, 18 percent telephone contact, and 82 percent online. The Producer Price Index, by contrast, already collected nearly all its repricing data through web collection and there were no changes to collection mode in response to the pandemic.

Response rates have declined in many BLS surveys, but not uniformly across industries, demographic groups, or geographic areas. In the Current Employment Statistics (CES) survey, for example, collection rates for construction, manufacturing, wholesale trade, and other services declined between 10 and 20 percentage points in April 2020 from the average for the 12 months ending in February 2020. Conversely, the collection rates for leisure and hospitality and for federal government increased by 10 to 20 percentage points from the average for the 12 months ending in February 2020. The collection rates for all other major industries were within 10 percentage points of the average of the 12 months ending in February 2020.

Some BLS programs adjusted the methods used in estimation to reflect extremely rapid changes in the U.S. economy more accurately. The CES survey made changes to its birth-death model to account for the shifting relationship between businesses starting up and businesses closing. The Job Openings and Labor Turnover survey changed its seasonal adjustment factors.

As the effects of the pandemic continued, BLS also introduced supplemental data sources. BLS productivity measures use the household (Current Population Survey, or CPS) and establishment (CES) labor force surveys to help estimate hours worked. These surveys' reference periods are typically in the middle of the month. However, in the United States, much of the local government intervention of shuttering businesses took place in the second half of March. In order to more accurately reflect the full scope of the first quarter 2020, the productivity office used continuing unemployment claims data from the Employment and Training Administration (a separate Department of Labor agency) to supplement the CES and CPS hours worked data. Though use of these additional data stepped outside of normal operating procedures, it helped more accurately portray the severe reductions in business activity at the very end of the first quarter.

Finally, BLS fielded additional questions in household and business surveys to provide more comprehensive information about the American experience during the pandemic. From households, BLS added questions about the pandemic's effect on telework, job search, job loss, and spending patterns<sup>1</sup>. BLS also fielded an additional survey to businesses with questions on the pandemic's effect on employee pay, number of sick days or days of leave, workplace flexibilities, and business operations. These additional data help add richness and detail to the existing products BLS produces.

All these methodological and execution changes pose a challenge for data users. "Everyone knows" that fewer responses probably mean published estimates are not as accurate as they were pre-pandemic. "Everyone knows" that changes in collection modes and data sources probably mean the new estimates are not entirely comparable to previous ones. But if accuracy has declined, by how much? Are the published estimates still relevant for users' needs?

### **Mitigation efforts – phase 1**

---

<sup>1</sup> See <https://www.bls.gov/cps/effects-of-the-coronavirus-covid-19-pandemic.htm>

On April 3, 2020, BLS reported that total nonfarm payroll employment fell by 701,000 in March and the unemployment rate rose from 3.5 percent to 4.4 percent. On May 8, BLS reported that total nonfarm payroll employment fell by 20.5 million in April and the unemployment rate rose to 14.7 percent. Calls and emails poured into BLS information desks wanting information on data collection, methods, concepts, and definitions. Much of the desired information did already exist but was scattered over the BLS website. Concepts and definitions could be found in the *Handbook of Methods*. Changes to methods were described in box notes on individual news releases. The BLS website had a set of agency-wide response rate pages but they were almost 9 months out of date.

BLS depends on the trust of our respondents and data users. That trust has been established and maintained over decades through transparency – always announcing any changes in operations and preemptively alerting data users to any potential measurement anomalies. BLS also wants to communicate information in a common format across our many releases to help users easily find what they are looking for and be able to draw comparisons.

As the effects of the COVID-19 pandemic on BLS survey processing first became evident in March and April 2020, BLS began to issue summary operational impact statements, mostly in the form of questions and answers, alongside our data releases. This auxiliary information was largely ad-hoc and tended to vary program by program.

In mid-April, BLS began to standardize a set of accuracy and reliability measures that, in one form or another, have accompanied almost all major data releases since their implementation in early May. These metrics report on the operational paradata most relevant to users' questions and compare them with pre-pandemic values. We began by asking each program to address five questions:

- How are data normally collected?
- Do we expect changes in data collection because of the coronavirus?
- Will there be changes to scheduled release dates?
- Will the coronavirus affect our ability to make estimates?
- What impact will the pandemic have on the estimates and is that impact quantifiable?

These program-specific Q&A documents tended to be rather long, up to 14 pages in some cases. Though quite thorough, the documents were not easy to parse for specific information or data points.

1. **How are prices collected for the CPI?** Price data used to calculate the CPI are primarily provided by two different surveys that are administered continuously each month:

- Commodities and Services Pricing Survey, an establishment survey of businesses selling goods and services to consumers, used to provide the price data for the CPI.
- Housing Survey, a survey of landlords and tenants used to provide rent data for CPI's shelter indexes.

Survey operations for CPI pricing surveys may be affected by limitations on data-collection staff, the availability of survey respondents, and the availability of items. Note that CPI data are collected throughout the entire month. Specifically, any given price in the CPI sample is collected in one of three defined pricing periods, corresponding roughly to the first 10, second 10, and final 10 days of the month. BLS uses several data-collection modes for CPI surveys that include telephone, internet, and automated electronic data capture. However, the majority of data are collected by personal visit. About 65 percent of CPI price data and 50 percent of CPI rent data are typically collected by personal visit. This type of collection has been suspended since March 16, 2020. (It was suspended on March 5th in the Seattle area.)

2. **What happens if BLS cannot collect CPI data?** The percentage of prices in the CPI sample that may be unavailable, either because the outlet is closed or the item is out of stock, is expected to increase. When BLS cannot obtain a price either because of data-collection limitations or the item being unavailable, it will generally be considered "temporarily unavailable." The CPI program has specific procedures for handling temporarily unavailable prices. Missing prices are generally imputed by the prices that are collected in the same or similar geographic area and item category. Essentially, the price movement of items that are not collected is estimated to be the same as those that are collected for a given item and geographic area. See the "Cell-relative imputation" section on page 20 of the [CPI Handbook of Methods chapter](#) for a brief technical discussion of this procedure. Note that this type of imputation is used in the CPI every month, especially in categories where response rates are relatively low.

3. **Were there any changes to CPI data-collection operations?** Yes. BLS suspended all in-person data collection in the Seattle, Washington, area on March 5, 2020. On March 16, 2020, BLS suspended all in-person data collection. Upon suspension of in-person data collection, CPI data collectors were instructed to attempt to collect data normally collected by personal visit by telephone, email, or by internet from the website of the establishment, if a website exists. CPI data collectors were specifically instructed not to contact establishments by telephone when it would cause an undue burden on respondents. These types of

Some of the feedback we received included critical comments such as:

- "Indicate when something is new or updated. I don't know what has changed."
- "There is way too much information on each page of this site and the design makes it really difficult to navigate. Suggest a major overhaul."

Our initial assessment was that centralizing the information had been an improvement, that users needed more concrete information, and that we needed to document how measures were evolving over time instead of simply updating the same pages with new rates.

## Mitigation efforts – phase 2

By late April, patterns in press coverage and user inquiries became clearer, as did commonalities and differences in program-specific reporting. A new assessment added several concepts to the communications initiative:

- One measure is not enough
- What is meaningful for one program may not be so for another
- Reports need to be consistent
- Reports need to be easy to update.

Our emerging framework included measures of collection rates and modes, imputation, suppressions, and measures of statistical error, as well as changes to methods.

### Collection rates and modes

Collection rates seems like an obvious indicator, but we wanted to be more specific in the detail we were showing. For business data we wanted to show industry breakouts where it made sense to express that some industries were reporting as normal, while others, especially industries where closures had shuttered operations, may have stopped reporting altogether. It also gave us an opportunity to explain BLS collection strategies including not recontacting worksites committed to providing critical community services, such as hospitals and grocery stores.

Providing details on collection mode made sense for multimodal surveys or surveys that had to shift from one type of operation to another, for example, in person interviews to phone interviews or in-person site visits to web scraping.

### Imputation

Depending on the survey, some small amount of imputation is common under normal circumstances. To understand what was beyond normal, we asked programs to report their imputation rates. The reasoning being that the more severe the loss of input data, the more heavily we may rely upon imputation methods. Some programs, like the CPI, have several gradations of imputation, from nearest-neighbor imputation to “hold the estimate constant over the period.” We are reporting on each of these each month since the start of the pandemic. Using these figures, users can view the different types of imputation and assess if quality, at least in terms of sample reliance, has degraded.

### Suppressions

We wanted to be clear about the frequency of data quality not meeting BLS publication standards. In the normal course of creating and disseminating data there is often an annual review of disclosure violation and cell robustness. In the wake of the pandemic, programs more proactively monitored cell robustness and suppressed those observations where the data quality had fallen below publication standards. Usually these suppressions were limited to detailed categories while higher level of aggregation could still be published. In other cases the suppression may have been limited to a particular area of the country. By listing all suppressions, users are able to get a sense of the types of items and areas of the country most affected by the pandemic.

### Measures of error

In addition to the frequency of imputation and suppression, we wanted to be sure to understand, statistically, how confidence intervals surrounding our point estimates were changing. Many of our programs calculate error measures on an annual basis. For those programs that do make concurrent calculations, we asked that they report any significant changes in those error measures.

### Changes to methods

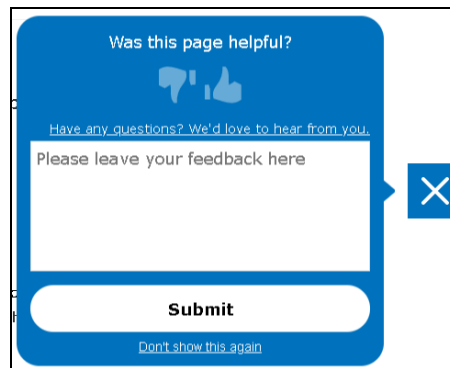
As it became clearer that the pandemic and associated restrictions would be in place for some time, BLS needed to adjust some of its methods to meet the demands of the situation. Examples of this included making a new computation for the payroll jobs birth-death factor, making adjustments to seasonal adjustment models (additive vs. multiplicative), and so forth. This section allows for programs to describe these changes and provide links to further information.

At the end of April, BLS issued guidance to all programs requesting that data releases be accompanied by the following information:

- 1) Any changes to data sources or estimation methods must be announced as early as possible. Full information must be provided with the first release using new sources or methods.
- 2) Assemble the following summary information to include with your releases:
  - Collection mode(s)
  - Response rates
  - Variance/standard error/confidence intervals
  - Imputation (counts or rates)
  - Cell suppression (counts or rates)

- 3) If the above measures are not evenly distributed over industry/occupation/area/item categories/etc., document the notable outliers
- 4) In addition to current month/quarter measures, include measures for:
  - Prior month/quarter
  - Same month/quarter in prior year
  - Average over year ending in Feb. 2020 (the last period unaffected by the pandemic)

To better understand user reactions, we also added a simple thumb up/thumb down indicator on each related webpage, with an option for users to add a freeform text comment to their rating.



By late May the new framework was in place. Each program has a release overview:

## Effects of COVID-19 Pandemic and Response on the Producer Price Index

The [Producer Price Index \(PPI\)](#) is a measure of the average change over time in the prices domestic producers

- [Impact of the coronavirus \(COVID-19\) pandemic on the Producer Price Index data for July 2020](#)
- [Impact of the coronavirus \(COVID-19\) pandemic on the Producer Price Index data for June 2020](#)
- [Impact of the coronavirus \(COVID-19\) pandemic on the Producer Price Index data for May 2020](#)
- [Impact of the coronavirus \(COVID-19\) pandemic on the Producer Price Index data for April 2020](#)

1. **Will data collection for PPI be affected?** The PPI survey is voluntary, so it relies on business survey respondents to provide pricing information. The unavailability of survey respondents or the interruption of business operations could affect the ability to collect PPI pricing information.
2. **How are PPI data collected?** Price data used to calculate the PPI are collected primarily online through the BLS Internet Data Collection Facility. For most industries, the PPI pricing date is the Tuesday of the week that includes the 13th of the month. If a transaction did not occur on the pricing date, the respondent is requested to provide a price for a transaction prior to that date, as close to the pricing date as possible. If no transaction occurred during the month, the respondent is requested to provide an estimate of what the price would have been, had a transaction occurred.
3. **What happens when PPI data cannot be collected?** When BLS cannot obtain a price, missing prices are imputed by the prices of similar items that have been collected. Essentially, the price movement of items that are not collected are estimated to be the same as those in their general index category. See the "Missing prices" section on page 10 of the [PPI Handbook of Methods chapter](#) for a brief discussion of this procedure. Note that this type of imputation method is the standard operational procedure used by the PPI program each month when estimating missing prices. Imputation allows indexes to be calculated, even when a fairly small number of prices are collected.
4. **Will PPI industry sample updates be affected?** BLS periodically updates the sample of producers providing data to reflect current conditions more accurately when the structure, membership, technology, or product mix of an industry shifts. BLS updates the sample once a year, through three sample update cycles. The next PPI sample update is scheduled to occur with the release of June PPI data on July 15, 2020. Due to the COVID-19 pandemic, a majority of the data necessary for this sample update were collected and the sample update is expected to proceed as scheduled. However, some sample updates will have fewer industries refreshed, reflecting challenges related to initiating establishments into the survey, such as reduced telephone collection. On March 16, 2020, BLS suspended all in-person data collection and began conducting initiation interviews by telephone. BLS has instructed temporarily not to contact establishments such as physicians' offices and manufacturers of surgical and medical instruments, as these calls could cause an undue burden on respondents. As a result, BLS will need to extend sample update cycles to collect enough data for some industries.
5. **Will PPI data be published as scheduled?** The PPI monthly data measuring producer price change will be published as scheduled. See the [schedule of PPI release dates](#). While it is possible that indexes will be based on a smaller than usual quantity of collected data, PPI ensures published indexes meet predefined quality standards. (See question #6.)
6. **Under what circumstances would some PPI data not be published?** Companies voluntarily provide proprietary pricing information for the survey, so BLS goes to great

Impact statements for each issuance at the top

Evergreen Q&A

<https://www.bls.gov/covid19/effects-of-covid-19-pandemic-on-producer-price-index.htm>

Each release has a release-specific impact statement:

## Impact of the coronavirus (COVID-19) pandemic on the Producer Price Index data for July 2020

The [Producer Price Index](#) (PPI) is a measure of the average change over time in the prices domestic producers receive for the sale of their products. PPI estimates for July 2020 were published on August 11, 2020. The pricing date was Tuesday, July 14, 2020, which means that most producers, but not all, provided prices they received on that date. The summary statistics in tables 1, 2, 3 and 4 indicate the impact of the COVID-19 pandemic on the PPI data collection, index estimation, and the BLS ability to publish indexes was minor for the release of preliminary July data. No changes in estimation procedures were necessary.

This page provides further information on the effects of the COVID-19 pandemic for the [July 2020 Producer Price Indexes news release](#).

### Collection mode

The PPI collects nearly all of its repricing data by web collection. There were no changes to collection mode in response to the pandemic.

### Response rates

Table 1 provides the July 2020 response rate by industry sector, along with comparison periods of June 2020, July 2019, and a 12-month average for March 2019 to February 2020. The July 2020 response rate for items requested shows a similar level of response for the vast majority of industry sectors. While the response rate for entertainment, accommodation, and food services sector remains down when compared to historical comparison periods, both the July and June 2020 response rates reflect an uptick in response as businesses begin to reopen.

It is important to note that typically, the PPI survey's monthly price collection response rate averages 76 percent by the end of its 4-month index revision period. All PPI indexes are recalculated 4 months after publication of preliminary data. Thus, July indexes may be revised when BLS publishes the PPI for November on December 11, 2020. All summary statistics in the tables below reflect comparisons to PPI preliminary data.

**Table 1. Comparative response rate percentages for the release of PPI preliminary industry indexes**

Industry sectors	12-month average, March 2019 to February 2020	July 2019	June 2020	July 2020	Percentage point change from June 2020 to July 2020
All industrial sectors	69	71	72	73	+1
Agriculture, forestry, fishing and hunting	47	50	48	47	-1
Mining, utilities, and construction	66	63	66	70	+4

<https://www.bls.gov/covid19/producer-price-indexes-covid19-impacts-july-2020.htm>

We knew going into this initiative that not all programs calculate or collect the same paradata. For example, some programs calculate variances annually rather than monthly, or calculate variances only after the estimates are published. Thus, reporting was “pick off the menu of options” style where programs included the measures they could provide without substantial additional work. This led to some loss of comparability across programs.

In addition, standard errors depend on both sample size and changes in observed magnitudes. The dramatic swings in many estimates had a far greater impact on statistical accuracy measures than did the sample size reduction. Nonfarm payroll employment, for example, fell by 13.8 percent from March to April 2020, and rose by 2.1 percent and 3.6 percent in May and June, respectively. This compares with a typical monthly change in the 0.1 to 0.2 percent range.

The biggest challenge, however, was that the operational paradata measures BLS is now reporting are conceptually complex and require reasonably sophisticated statistical knowledge to understand and use effectively. Such subtleties were swamped by a far larger and more prominent pandemic-related estimation problem: a misclassification error in the household survey used to calculate the monthly unemployment rate. The error called into question, at least for some users, both the accuracy and the impartiality of the data BLS was releasing. Following the principle of full transparency, BLS announced and described the misclassification issue in every Employment Situation news releases between March and September. BLS also provided more detailed information, including estimates of what the data might have shown without the misclassification, in supplemental documents published at the same time as the Employment Situation reports and published an explainer blog at the end of June<sup>2</sup>.

<sup>2</sup> See <https://blogs.bls.gov/blog/2020/06/29/update-on-the-misclassification-that-affected-the-unemployment-rate/> with links to monthly Frequently asked Questions.

There is an adage that's keeps coming up during the pandemic: we're changing the tires on a moving car. That feels true of trying to measure data quality and communicate it to the public. At the beginning there are so many things that are unknown and situations change and evolve. The first few attempts at communicating information about the quality of our data were disparate and scattered across the website. The goal of putting all the information together in a single location with similar formats was to help the user find what she was looking for no matter the release. An additional goal was to help our own staff have a guide and a template to simply plug in the latest set of paradata.

## Reception

Reception was somewhat mixed inside BLS. Several programs were happy to have clear guidance and a template to follow. Other programs were less enthusiastic to change the process they had used for the first few releases. The goal of the template was to help unify communications and to make it a "plug and play" set of information.

As of September 30, 2020, users have viewed webpages with the feedback box 166,067 times, left 294 ratings, and included 138 comments. The comments show that many of the ratings were unrelated to the pages themselves.<sup>3</sup>

Once the unrelated ratings are eliminated, feedback has been 84.2 percent positive.

Some of the laudatory comments include:

"Very convenient to find a "hub" for this information. Thanks."

"It was helpful to know more complete situation."

"Extremely helpful. Thanks for keeping this marvelous data update. Great job!"

"I sense you're on the bleeding edge of the data that's going to point us out of the COVID mess."

"The data on this site is extremely useful and whomever put it together should get a raise. Well done, PPI."

Some critical comments include:

"The title of this page is "Effects of COVID-19 Pandemic and Response on the CPI". I was expecting a simple answer such as 'The CPI change was 50% lower than had been predicted in January'. Instead there are details on the methodology. Frustrating."

"A little helpful but I was hoping for (easy to find) figures on the number of people out of work now and in in recent months -- especially because of Covid-19. I have not found that information yet."

---

<sup>3</sup> Many of these comments were poignant, e.g., "I was put on a call back list 2 days ago and still haven't received a call. Dealing with covid-19 is stressful already. Now can't get any assistance with unemployment compensation. Also having to deal with bill collectors wanting to disconnect services. Please help."



Other comments include suggestions for improvement:

“I was on your site yesterday and agreed to take a survey. I did but now cannot get rid of the final page of the survey. I have cleared Chrome history and everything else. This needs to be fixed.”

We redesigned some of the pages and features to address such concerns

## **Conclusions**

Data users need reliable information about the accuracy and relevance of the estimates upon which they rely to make informed decisions. This is true under all conditions. It is especially true during economic and social shocks: extraordinary circumstances require extraordinary levels of transparency. The COVID-19 pandemic and associated mitigation measures have exercised the BLS commitment to providing complete and precise paradata measures.

Just as one observation does not make a trend, one paradata measure does not give users complete information about data quality. To properly communicate data quality to our customers, BLS needed to consider a spectrum of measures that may be more meaningful for some programs (or users) than others. Building such a broad framework helped guide programs to report according to their similarities, rather than emphasizing their differences.

Many data users have expressed their appreciation for the openness of our dialogue. BLS is being candid about the changes we have made because if there is any hint of “cooking the books,” the public will lose their trust in us.

The main lesson we have learned is to stay nimble. More than seven months after it began, the crisis is still with us. As time goes on and the economic and social situations change, we may find some of the current measures less relevant and determine that others should take their place. So these reports may need to evolve.

The BLS paradata reporting initiative appears to be working well, giving users needed information in a useful format. The cost to BLS has been minimal. We have learned that statistical agencies such as BLS can relatively easily assemble and provide key measures to help users assess out data’s fitness for their use, and in the process maintain and even enhance our credibility among the vast majority of our customers.