

Communicating the accuracy and relevance of statistical data during a pandemic

Michael Levi and Emily Liddel
U.S. Bureau of Labor Statistics

IMF Statistical Forum on Measuring the Economics of a Pandemic
17-19 November 2020



Trust and transparency

- BLS depends on the trust of data users (as do all stat agencies)
 - ▶ If customers do not believe our data are impartial and accurate, they will find other sources
- Trust is established and maintained through transparency
 - ▶ Announcing changes in methods
 - ▶ Announcing changes in operations
 - ▶ Alerting users to potential measurement anomalies
- COVID-19 exercised the BLS commitment to transparency



March 13, 2020: U.S. Declares National Emergency

- BLS and Census cease all in-person data collection
- Telephone data collection centers close
- Staff are moved to maximum telework status



Impact: Collection modes

- In many, but not all, BLS surveys, the mix of data collection modes changed
- Ex: Consumer Price Index

	April 2019	April 2020
In person	74%	0%
Telephone	10%	18%
Online	16%	82%



Impact: Response rates

- Response rates changed in many BLS surveys, but not uniformly across industries, demographic groups, or geographic areas
- Ex: Producer Price Index

Industry sectors	12-month average, March '19 – Feb. '20	Sept. '19	Sept. '20
Mining, utilities, and construction	66	65	71
Retail trade	75	77	76
Entertainment, accommodation and food services	66	65	47

Impact: Methods

- Seasonal adjustment
- Birth-death model



Impact: Data sources

- Ex: Productivity program
 - ▶ Continuing Unemployment Claims
 - ▶ Supplemented Current Population Survey and Current Employment Statistics data



Impact: New data

- Additional questions on household and establishment surveys
- New survey (Business Response Survey)



Are the data still reliable?

- “Everything’s shut down”
- “Nothing’s in stock”
- “You’re not collecting the information”

BLS says, “All published data meet reliability standards”

How can users assess for themselves?



Communicating uncertainty

- Difficult under the best of circumstances
 - ▶ Confidence intervals
 - ▶ Revisions
 - ▶ Changes in methods
- Communicating a change in uncertainty is even harder



Starting mid-April 2020

- Calls pour in to BLS wanting information on
 - ▶ Data collection
 - ▶ Methods
 - ▶ Concepts and definitions



We had (most of) the answers

- For methods, go to the *Handbook of Methods*
- For changes to methods, see the announcement box in our release
- For response rates, go to a response rate page (9 months out of date)
- For modes of data collection, see link buried at the bottom of a web page



Early efforts

- We didn't know
 - ▶ What paradata users wanted
 - ▶ What paradata we had readily available
 - ▶ How long the pandemic disruption would last
- Went through several iterations of Q&As, paradata reporting

Eventually: A common framework

- One measure is not enough
- What is meaningful for one program may not be so for another
- Value of consistency across programs
- Needs to be easy to update



Paradata framework

- Collection mode(s) and collection rates
 - Imputation rates
 - Cell or series suppression
 - Statistical measures of error
 - Changes to methods
 - Q&As
-
- All with comparison to pre-pandemic measures



Release page

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 typically 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, a majority of the data necessary for this sample update were collected and the sample update is expected to proceed as scheduled. However, due to the COVID-19 pandemic, a few sample updates will have fewer industries refreshed, reflecting challenges related to initiating establishments into the survey, such as limited access to establishments for data collection. On March 16, 2020, BLS suspended all in-person data collection and began conducting initiation interviews by telephone. Respondents are instructed temporarily not to contact establishments such as physicians' offices and manufactures of surgical and medical instruments, which 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

- Links to impact statements for each release
- "Evergreen" Q&A



Impact statements

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




Prominent misclassification error

- Operational paradata subtleties swamped by misclassification error in household survey
- Accuracy and impartiality of BLS data called into question
- BLS announced and described issue in every Employment Situation news releases
- Provided detailed information in supplemental documents
 - ▶ Including estimates of what the data might have shown without misclassification
- Published an explainer blog



Some positive reactions from website feedback

<https://www.bls.gov/bls/bls-covid-19-questions-and-answers.htm>

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


<https://www.bls.gov/bls/effects-of-covid-19-pandemic-on-productivity-and-costs-statistics.htm>

 very helpful, thank you

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

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

<https://www.bls.gov/bls/measuring-the-effects-of-the-coronavirus-covid-19-pandemic-using-the-current-population-survey.htm>

 So happy to see the BLS moving so quickly to incorporate COVID questions to the CPS, and for making the information available as quickly as possible.

Critical reactions from website feedback

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



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

<https://www.bls.gov/covid19/consumer-price-index-covid19-impacts-june-2020.htm>



This says it all. Data inadequate

User reactions (quantitative)

- As of September 30, 2020:
 - ▶ 166,067 user views
 - ▶ 294 ratings
 - ▶ 138 comments
- Many of the ratings were unrelated to the pages themselves
- Once unrelated ratings are eliminated
 - ▶ 84.2 percent positive feedback

Final thoughts

- Users need reliable information about the accuracy and relevance of the estimates upon which they rely
- No single paradata measure is sufficient
- Framework: Spectrum of meaningful measures
- Flexibility: As conditions change, other measures may become more relevant



Contact Information

Michael Levi

U.S. Bureau of Labor Statistics

1-202-691-5100

levi.michael@bls.gov

