Monthly Labor Review
July 2019

Reduced sample rotation frequency in hospitals and household utilities

In 2017, the Consumer Price Index (CPI) program conducted research into the theory that sample rotation cycles longer than 4 years would be a better fit for establishments in industries with low market churn. This article describes this research and CPI's decision to lengthen sample rotation for hospitals and utility (electricity and natural gas) companies.

Kerri Chicarella
chicarella.kerri@bls.gov

Kerri Chicarella is an economist in the Office of Prices and Living Conditions, U.S. Bureau of Labor Statistics.

In the Consumer Price Index (CPI), the vast majority of priced items rotate in a subset (one-quarter) of geographic areas, known as primary sampling units (PSUs), once every 4 years. This sample rotation schedule ensures not only that the items priced are current but also that the outlets where these items are priced are representative of the locations at which the sampled population shops.[1] The rotation also allows shifts in consumer behavior and changes in market conditions to be represented in CPI indexes. Although a 4-year sample rotation is reasonable for many items, such as fresh fruits and gasoline, it can be too frequent for other items, such as services and items in markets with slower outlet turnover and greater regulation.

In 2017, the CPI program conducted research into the theory that a rotation cycle longer than 4 years would be a better fit for establishments in industries with low market churn. For reasons described in the next section, the research included hospitals and outlets that provide residential electricity service and residential natural gas service. This article describes this research and its results, concluding with a description of changes recently implemented in production. For ease of identification in the article, electricity and natural gas collectively are referred to as household utilities.

Background

The samples of hospitals, electricity companies, and natural gas companies were prime candidates for CPI's research for a variety of reasons. First, the entry and exit rates in the markets of these establishments are relatively low compared with those in retail outlets, and the establishments' industries are well regulated, with relatively high barriers to entry. Hospitals require much more regulation than, for example, grocery stores or

clothing retailers, and electricity and natural gas establishments require costly assets such as generation stations and pipelines. Among household utilities establishments, alternate suppliers change more often, but they are not the main holding companies. Current CPI procedures account for accurately capturing alternate supplier companies' rates and their availability to customers.

The second reason for selecting the aforementioned samples to research is that the items priced at the associated establishments do not change drastically over time. The services provided at a hospital are typically consistent, except when specialty departments are being added or cut (for example, a cardiac unit leaving). The services provided by electricity and natural gas companies also do not change, although service rates and the structure in which they are applied to customer bills may change.

Finally, the complications that come with collecting prices for items in the selected samples, along with the difficulties encountered in cooperating with respondents, make hospital and household utility services ideal for longer rotation periods. For hospitals, we struggle to gain respondent cooperation in our standard 6 months for sample initiation, both because of the demanding work environment of hospital employees and because of the complex nature of pricing a hospital service. These hurdles come at a large cost and often lengthen the initiation time to as long as 2 years, which is an added strain on already limited field resources. Further, the outlet would then be used in the index for only 2 years, potentially jeopardizing index quality.

For household utilities, pricing is complex, arduous, and time consuming. In both the electricity and natural gas industries, no two outlets bill customers on the basis of the same rate structure. For example, while one outlet may only break down the bill with a single consumption rate, a distribution rate, and a customer charge, another outlet may have several rate schedules (heating only, nonheating, all included), complex consumption charges (tiers of 0–150 kWh, 150–500 kWh, 500+ kWh), miscellaneous weather riders, and tax rates specific to each jurisdiction within the service territory. As with hospital services, utilities setup is costly and often redundant, especially for a company that is in both the "old" and "new" sample. The time spent setting up near-identical rate structures in 4 years' time poses great risk to respondent cooperation and shifts field-resource commitments.

Analysis

In early 2017, using both current and historical data, we conducted analyses to test various hypotheses about sample makeup and representativeness. For both the hospitals analysis and the household utilities analysis, we took similar approaches, using both third-party data and CPI data to inform our decisions.

Hospital services

To accurately analyze the shift in establishment composition in the hospital market, we held constant the data source used and chose two periods—a historical one and a current one—separated by more than 4 years. The source used was a third-party hospitals market aggregator. The dataset for the historical period is from 2006 and the dataset for the current period is from 2015. We also compared the third-party data to the most recent (2017) CPI hospital services sample.

Single-source analysis

First, we compared the number of hospitals from the two datasets, using the same source for consistency. (See table 1.) We found that the 2015 data exhibited only a 4-percent increase in the number of hospitals. We interpreted this finding to mean that, over the 9 years analyzed for this source, the number of hospitals in the market remained about the same. However, we also recognized that there could have been changes in sample composition within the total number of outlets.

Table 1. Comparison of the number of outlets between datasets

Variable	Total number of outlets[1]
2006 dataset	5,972
2015 dataset	6,227
Percent difference in old sample and new	4

Notes:

[1]Excludes the U.S. territories of Puerto Rico, American Samoa, Guam, Marshall Islands, and Virgin Islands; excludes all facilities with a primary service designation of institution (e.g., prison) and intellectual disability.

Source: Author's calculation of third-party datasets.

Next, we analyzed the individual hospitals that were included in both datasets. (See table 2.) We found that 85 percent of the hospitals included in the 2006 dataset were also listed in the 2015 dataset, further highlighting the consistency between the samples from the same source over the 9 years.

Table 2. Match rate between datasets

Variable	Number of outlets
Total in 2006 dataset	5,972
Number that match	5,101
Percent that match	85

Source: Author's calculation of third-party datasets.

Multisource analysis

The strong results from the single-source analysis allowed us to move forward with comparing third-party data with CPI data. (See table 3.) Our analysis revealed that 82 percent of the hospitals listed in the CPI sample were also listed in the third-party data, which further validated our confidence in using the third-party data for the final phase of the analysis—a measure-of-size (MOS) comparison.

Table 3. Match rate between CPI and third-party source datasets

Variable	Number of outlets
All CPI outlets	265
Number that match	217
Percent that match	82

Source: Consumer Price Index (CPI) data and author's calculation of third-party datasets.

Analyzing measure of size (single source)

The core of the hospitals' analysis was an MOS comparison. For all outlets in a sample universe in the CPI, an MOS is used for determining an individual outlet's probability of selection in the collected sample. In the CPI sample, an MOS is determined by household-reported expenditures in the Consumer Expenditure Survey (CE). When using third-party data, as was done in this research, the ideal MOS is an outlet's revenue for 1 year, but this information is often considered proprietary by the individual outlet and thus is not available for our sampling procedures. For hospital services, patient visits were used as a proxy for probability sampling.

For the MOS analysis, we compared the two third-party datasets (2006 and 2015) in order to observe any potential shifts in representativeness over time. We calculated the patient visit-based MOSs for each outlet that was included in both datasets and then compared those MOSs. Between the two periods, there was an overall difference of +0.1 percent, meaning that, on average, the MOS for a hospital in the 2015 dataset was 0.1 percent higher than the same hospital's MOS in the 2006 dataset.

Digging deeper, we found a few large differences that may have been attributed to an outlet gaining or losing "popularity" since 2006, gaining or losing admissions and/or patient visits, experiencing changes in the number of services provided, or perhaps seeing population shifts in the area. Although these differences did appear in the analysis, we interpreted the change of +0.1 percent to mean that, on average, hospitals reported negligible change in patient visits over the 9-year period.

The three phases of hospital sample analysis each supported our hypothesis that hospital churn over an extended period is low. We used these data to make decisions about sample rotation timing (greater than every 4 years). (These decisions are detailed at the end of this article.)

Household utilities—electricity and natural gas

For household utilities, we compared two different sample frames over time to analyze the shift in composition of the electricity and natural gas markets. The first source was the CE data, which were used in a previous rotation cycle for creating the 2000 CPI sample. The second source was data from a third-party utility company aggregator, which was used to create the CPI sample, beginning with the 2015 rotation cycle. The CE sample frame was generated in the early 1990s, and the third-party sample was from 2013, giving us an almost 25-year period for comparison.

Multisource analysis

In the first phase of the electricity and natural gas analysis, we compared the number of companies provided by each sample frame. (See table 4.) The outlets reported in CE reflect a sample of households for each PSU. The third-party data, however, provide what we consider a near complete universe from which we can sample. The CE dataset is thus considered less comprehensive than the third-party data, and for this reason, we can expect to see large differences between the two datasets.

Table 4. Comparison of the number of outlets between CPI and third-party datasets

Variable	Number of companies	
	Electricity	Natural gas
CE	199	132
Third party	561	308
Percent difference in the old sample and new sample	182	133

Source: Consumer Price Index (CPI) data and author's calculation of third-party datasets.

Our analysis showed that the number of companies in the third-party sample frame was substantially greater than that in the CPI sample frame. Overall, the more recent data provided 182 percent more electricity companies and 133 percent more natural gas companies.[2] We felt confident that, because of the more comprehensive third-party data, we would return results that showed low turnover in the market.

Next, we compared the individual companies between the two sample frames. This step, which analyzed the detailed makeup of the sample frames, aimed to determine whether companies from 25 years ago were still active in the household utility market. We found that the number of outlets in both datasets was high, despite the lengthy time gap and the overall size difference between the old and new samples. For electricity, 83 percent of the CE-reported outlets were also present in the third-party data, and for natural gas, 79 percent of CE-reported companies were also present in the third-party data. (See table 5.)

Table 5. Match rate between CE and third-party datasets

Variable	Electricity	Natural gas
CE, all outlets	199	132
Number that match	165	104
Percent that match	83	79

Source: Consumer Price Index (CPI) data and author's calculation of third-party datasets.

In comparing outlets between datasets, we did a considerable amount of research to ensure that outlets were actually the same company. At first glance, outlets within a PSU appeared to show a drastic change in the companies servicing that area; however, after researching company mergers, acquisitions, divestments, and rebranding (name change), we often were able to successfully match outlets between the two sources.

This phase of the analysis was a crucial part of supporting the hypothesis that there is a relatively low rate of turnover in the electric and natural gas markets.

Analyzing measure of size (multisource)

Similarly to the hospital services research, the utilities analysis focused on an MOS comparison. Although the comparison of individual companies over time showed their consistent presence, we still needed to analyze the market share of each company. In the utilities analysis, we compared MOS across two different datasets. The analysis was not the same as the one-to-one source comparisons for hospitals; however, we judged that the analysis was still valid. The CE sample frame's MOS was the expenditure for each utility company reported by the

survey respondent. The MOS used for the third-party source was a function of the residential population of each company's service territory.

We compared the MOS for the outlets included in both frames, which are separated by nearly 25 years. The overall MOS difference between the electricity and natural gas sample frames was −2 percent. (See table 6.) This means that, on average, companies that were present in the third-party source data had a lower MOS than did those in the CPI sample.

Table 6. Electricity and natural gas statistics for a measure-of-size (MOS) comparison between datasets

Item	Average percent difference in MOS
Electricity	-2.0
Natural gas	-2.0

Source: Consumer Price Index (CPI) data and author's calculation of third-party datasets.

The MOS difference between these sources was larger than that in the hospital services analysis, for two reasons. First, differences in the methodology for calculating the MOS in the two sources may have led to small discrepancies. This was not an issue in the hospitals analysis, because it used the same source over time and the same MOS in each period. Second, there was a shift in market share by some outlets. As would be expected over such a long period and given a substantial difference in the number of outlets in a sample frame (see table 4), the companies' market share changes. When the number of competitors in a given area increases, the residential service territory for each competitor can shrink; similarly, with fewer competitors, the territory can grow.

The results of this MOS analysis again supported our hypothesis that market churn over an extended period is relatively low. Our decision on sample rotation timing for utilities was also informed by these data.

Production impacts

The analysis for hospitals and utility companies supported our hypothesis that turnover in those markets is low and that sample rotation could be lengthened beyond 4 years while remaining representative of shifts in consumer behavior and changes in market conditions. The results from the analyses comparing data sources over time, identifying establishment consistencies, and performing MOS comparisons supported a change in the rotation frequency. Ultimately, we decided to reduce the sample rotation frequency for the analyzed service items (hospital services, electricity services, and utility (piped) gas services) from every 4 years to every 8 years.

Lengthening the sample rotation timeframe for these items offers two main advantages. First, by dedicating more time and attention to obtaining accurate prices each month (made possible by redistributing time spent on arduous initiation efforts), it increases the efficiency of data collection and review. Second, it reduces the respondent burden for sample rotation.

As noted in an announcement on the CPI website, we eliminated the 4-year rotation cycle for these items as of February 2019. The items' samples will begin rotating again in February 2023 and will stop rotating after 4 years. This "4 years on, 4 years off" pattern will continue in the future.

SUGGESTED CITATION

Kerri Chicarella, "Reduced sample rotation frequency in hospitals and household utilities," *Monthly Labor Review,* U.S. Bureau of Labor Statistics, July 2019, https://doi.org/10.21916/mlr.2019.17.

NOTES

1 For information on sampling methodology and population eligibility, see *Handbook of Methods* (U.S. Bureau of Labor Statistics), https://www.bls.gov/opub/hom/pdf/cpihom.pdf.

2 Some of this expansion may be attributed to the increased deregulation in natural gas markets in the United States.

RELATED CONTENT

Related Articles

Improving initial estimates of the Chained Consumer Price Index, *Monthly Labor Review*, February 2018.

Hospital workers: an assessment of occupational injuries and illnesses, *Monthly Labor Review*, June 2017.

The 2018 revision of the Consumer Price Index geographic sample, *Monthly Labor Review*, October 2016.

Related Subjects

Hospitals | Consumer price index