BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Prices and Living Conditions

Measurement Issues in Consumer Price Indexes

Brett Matsumoto,
U.S. Bureau of Labor Statistics
Anya Stockburger,
U.S. Bureau of Labor Statistics

Working Paper 584
July 11, 2025

Measurement Issues in Consumer Price Indexes

Brett Matsumoto and Anya Stockburger

JEL codes: E30, E31

Keywords: inflation; cpi; government statistics; measurement issues

Abstract

This article is an excerpted chapter from a research handbook on inflation in which we document measurement issues in the compilation of consumer price indexes. From defining the conceptual target to sampling, repricing, and aggregating the data, there are numerous issues that arise. In this chapter, we review the practical challenges associated with the construction of consumer price indexes. We provide examples from the US CPI produced by the Bureau of Labor Statistics (BLS) to illustrate how these issues can be addressed.

Introduction[1]

Consumer price indexes are constructed by tracking the price changes of a sample of goods and services over time. Despite the seeming simplicity of this task, many practical challenges arise in the construction of consumer price indexes. From defining the conceptual target to sampling, repricing, and aggregating the data, there are numerous issues that arise. In this chapter, we review the practical challenges associated with the construction of consumer price indexes. We provide examples from the US CPI produced by the Bureau of Labor Statistics (BLS) to illustrate how these issues can be addressed.

We begin by discussing the conceptual issues and choices that must be made regarding index scope. We then discuss issues related to estimation of the index such as the weights and aggregation formulas. In the next section, we discuss issues relating to the selection of the sample of goods and services that are priced in the index. In the final section, we discuss issues relating to price collection including the method of data collection and how to handle missing prices with a particular emphasis on item replacement and quality adjustment for when an item becomes permanently unavailable to price.

This chapter is not meant to be a comprehensive treatment of this topic. It is meant to serve as a general introduction to the practical challenges so the reader can more greatly appreciate the measurement limitations inherent in the construction of consumer price indexes. For additional information on the practical issues relating to constructing a consumer price index, a good comprehensive resource is the International Labor Organization (ILO) manual on consumer price indexes (2020). More information on the US CPI can be found in the Handbook of Methods (BLS 2023).

Scope and Concepts

Before constructing a consumer price index, there are a number of fundamental choices that must be made regarding the measurement target, scope of the index, and population coverage. In this section we discuss some of the conceptual choices involved in constructing consumer price indexes. For a thorough

[1] This paper was prepared as a chapter for "Research Handbook on Inflation," Ed. Guido Ascari and Riccardo Trezzi, Edward Elgar Publishing, June 2025. **ISBN:** 978 1 03532 775 1

discussion of these conceptual issues see the Boskin Commission report (Boskin et al. 1996) or the Committee on National Statistics panel report *At What Price?* (CNSTAT 2001).

Conceptual Target Index

When deciding on a conceptual target for the index, the main options are whether to target a cost of living index (COLI) or a cost of goods index (COGI). A COGI tracks the cost of purchasing a fixed basket of items whereas a COLI tracks the minimum expenditure required to achieve a fixed standard of living. It is difficult to control for the many non-market factors that affect the quality of life, so a more feasible measurement target is a conditional COLI which holds constant these non-market factors. The US CPI uses the conditional cost of living framework as the conceptual target for the index. Having a conceptual target is useful for evaluating methodological changes and addressing biases in the indexes. The often cited price index biases are defined in relation to a true COLI. Later in this chapter, we will discuss how some of the common price index biases are addressed in practice.

The cost-of-living measurement objective and its relation to consumption helps to define the scope of the index. The standard definition of consumption excludes business purchases, investments, and related expenditures.[2] The US CPI restricts to expenditures made directly by consumers (except for owner-occupied housing, which is discussed in the next section). An argument could be made that spending on behalf of consumers by others (government, employers, etc.) is a part of consumption and should be in scope for a COLI. Free and non-market goods are generally excluded from the scope for practical reasons. Free goods are goods that are always provided free of charge and include government provided goods and services and many digital goods.[3] Examples of non-market goods and services include home production and barter.

Pricing Durable Goods

A major conceptual issue related to the measurement objective is the treatment of durable goods. The cost of living is tied to consumption rather than the acquisition of the good. Many services and non-durable goods are consumed in the same period they are acquired. However, durable goods provide a flow of services over a long period of time. Ideally, a cost of living index would price this flow of services for owned housing and durable goods. There are two general approaches for pricing the flow of services for durable goods. The first method is to calculate the user cost, which is the cost associated with purchasing the durable good at the start of the period, using it for the period, and selling it at the end of the period. The user cost includes depreciation and the opportunity cost of purchasing (defined as the foregone interest). The user cost can also be defined to include owners' expenses such as maintenance costs. The second approach to valuing the flow of services from durable goods is the rental equivalence approach. In this method, the flow of services from the consumption of durable goods is valued as the amount needed to rent the good.

The US CPI uses the acquisition approach (full purchase price) for non-housing durable goods. Despite the conceptual advantages, there are practical challenges to implementing either method for valuing the flow of services for non-housing durable goods. For owner-occupied housing, the US CPI uses the rental

[2] See OECD (2013) for a discussion of the exclusions typical in defining household consumption or consumption expenditures.

[3] Subsidized goods that still involve an out-of-pocket expenditure and goods that are periodically free would generally be considered in scope.

equivalence approach. The price change of owner-occupied housing is determined using a sample of rental units that is designed to be representative of the stock of owned housing. The weight of owner-occupied housing is an imputed rental equivalence (owner's equivalent rent). Since a rent is imputed for all homeowners, regardless of their out-of-pocket spending, this category represents a large share of the overall index (for example 25 percent in December 2022), so its treatment will have a large effect on the index. Internationally, there is little consensus on the treatment of owner-occupied housing. Some countries also use a rental equivalence approach, while others use an acquisition approach or a measure based on out-of-pocket expenditures (based on mortgage principal and interest). Still others exclude owner occupied housing from the index entirely.

Population Coverage

Another conceptual question is how to define the population coverage, or whose expenditures should be represented by the index. Generally, the index should be representative of average expenditures in the population. However, one must first decide whether there should be any restrictions on the covered population. For example, in the US CPI, the BLS excludes groups of individuals who have atypical spending patterns such as those living in institutions (prisons and nursing homes) or whose basic needs are provided by an employer (military). The BLS also excludes the population living in rural areas due to challenges conducting surveys outside metro- and micropolitan parts of the United States (approximately 7 percent of the total US population).[4]

Another issue of having the index be representative of average expenditures is whether all households are to be weighted equally in the index. One option is that households with more expenditures will have greater weight in the index. Weighting by expenditures is called plutocratic weighting. The alternative, called democratic weighting, is to weight each household equally. The US CPI uses plutocratic weighting. Prior research has explored the impact of the different weighting concepts and finds that democratic versus plutocratic leads to relatively minor differences in the US CPI (Kokoski 2000; Martin 2022).

Weights and Estimation

Overview

With the measurement objective, scope, and population target defined, statistical agencies must decide how to average price change to form a single index value. For a given set of price changes, the choice of aggregation formulas could yield a range of inflation rates. Statistical agencies adopt aggregation formulas depending on the data available to best proxy the measurement objective. In general, fixed quantity weight formulas such as Laspeyres and Paasche, measure changes over time in the cost to purchase a fixed market basket of goods. A geomeans formula uses fixed market basket shares, rather than fixing the quantities in the market basket. Cost of living indexes measure changes over time in the cost (minimum expenditure) to maintain a fixed standard of living. Superlative indexes, such as Tornqvist Fisher Ideal, Satio Vartia, and Walsh, are approximations of a true cost-of-living index and will reflect changes in consumer purchasing behavior over time (Diewert 1976) but require weights for the current period.

[4] Technically, only core-based statistical areas are eligible for selection in the CPI geographic sample. These areas are defined by the Office of Management and Budget.

The BLS adopts a two-stage aggregation approach. At the lower level, the BLS averages price change using either a Laspeyres or geomeans formula. The choice of formula depends on the degree of consumer substitutions. For items which consumers are assumed to purchase in fixed quantities with few substitutions, the BLS calculates price change using a Laspeyres formula. Examples include rent, utilities, and medical care services. For the remainder of items, the BLS uses a weighted geometric mean formula to better reflect consumer substitution. Weight shares are fixed, so implicit quantities shift in response to relative price change. The adoption of geomeans formula in 1999 at the lower level addressed lower-level substitution bias (Johnson, Reed, and Stewart 2006).

To calculate upper-level price change, the BLS averages across lower-level index relatives using either a Lowe (modified Laspeyres) or Tornqvist formula. Most index products, including the headline CPI-U, are calculated using a Lowe formula. Weights are updated annually and reflect spending from two years earlier. For example, the BLS calculates the January-December 2024 indexes using spending data from 2022. Because of the lag between the expenditure reference period and when data are used in index calculation, weights are adjusted to incorporate price changes in the intervening time. Continuing the example, the BLS adjusts 2022 spending data by price change between 2022 and December 2023. The resulting relative importance reflects consumer spending in 2022 updated for December 2023 prices.

The BLS also calculates a Chained CPI-U (C-CPI-U) index using a Tornqvist formula. The C-CPI-U differs from the CPI-U in terms of the index formula, chaining, and weight reference period. Research has shown the weight reference period difference explains most of the divergence between these two indexes (Kurtzon 2022). Because the C-CPI-U uses a superlative index formula with concurrent weights, it addresses upper-level substitution bias. The difference between inflation measured by the C-CPI-U and the CPI-U is a measure of this bias and has historically been around 0.25 percentage points per year on average.[5]

The BLS calculates weights from household spending data collected in the Consumer Expenditure Surveys (CE). Other statistical agencies use national accounts data either in isolation or combination with household budget surveys to estimate weights. The CE collects data at sufficient granularity, geographic coverage, and timeliness to meet BLS needs for CPI weight calculation without adjustment from national accounts data. An advantage of household survey data is that it enables statistical agencies to calculate spending shares for different populations and specifically for the market basket components definition for the consumption scope (national account data may not be available for the requisite CPI categories). In addition to the urban population, the BLS produces indexes weighted by spending shares of wage earners and clerical workers (CPI-W) and research indexes weighted by spending shares of older Americans (R-CPI-E) and different groupings of household by income (R-CPI-I)[6].

The 12-month change in index levels is a measure of inflation. Some users are interested in comparing indexes over a shorter time interval. Seasonally adjusted indexes remove typical seasonal patterns to facilitate this need. Every year, the BLS updates seasonal factors using a seasonal adjustment method

[5] Chained CPI-U frequently asked questions: https://www.bls.gov/cpi/additional-resources/chained-cpi-questions-and-answers.htm. Accessed 2/4/2025.
[6] The R-CPI-E homepage (https://www.bls.gov/cpi/research-series/r-cpi-e-home.htm) and R-CPI-I homepage (https://www.bls.gov/cpi/research-series/r-cpi-i.htm) provide more information about these research indexes. Accessed 2/4/2025.

and revises the prior five years of seasonally adjusted data.[7] The updated factors are used to calculate seasonally adjusted indexes for the upcoming year.

Limitations and New Methods

Household survey data have many advantages; however, the time required to collect and process spending data is a major limitation. The BLS began updating spending weights annually in 2023, replacing a biennial process to reduce the measure of upper-level substitution bias (Klick 2021). The BLS releases the final C-CPI-U index calculated with a Tornqvist formula 11–13-months past the index reference month. For example, Tornqvist C-CPI-U indexes for January, February, and March 2023 are released in February 2024. Initial estimates of the C-CPI-U assume constant elasticity of substitution. To further reduce upper-level substitution bias and achieve the goal of producing a Tornqvist index on the same publication schedule, new data sources and technologies are needed.

Survey response rates are another potential limitation of household survey data. Surveys develop methods to mitigate and adjust for survey nonresponse, however the continued decline in household survey response rates over the last ten years is a potential for future concern[8].

Sampling

Introduction to sampling

As described in the previous section, index estimation begins by measuring price change at the lower level. In this section we discuss issues relating to the construction of a CPI sample in the context of the US CPI. There are two steps to data collection: sampling and repricing. First, the sample of goods and services to be priced must be selected. Once the sample is selected, the prices of those items are tracked over time to construct the index. The BLS conducts two surveys to supply the prices and rents that measure price change in each basic item/area index: the Commodities and Services (C&S) Survey and the Housing Survey. The surveys are designed to produce an unbiased estimate of price change within budgetary and operational constraints. The BLS samples core-based statistical areas as Primary Sample Units (PSUs) to ensure representative geographic coverage of where the target population lives. The BLS samples goods and services for which prices will be tracked and housing units for which rents will be tracked. Retailers and service providers (outlets) are selected as respondents for the C&S Survey and residential addresses are selected to identify respondents for the Housing Survey. This section describes an overview of these sampling steps as well as limitations inherent in the traditional sample design and the promise of new data sources to address them.

Geographic sampling overview

The BLS produces national estimates of price change that are representative of the urban, non-institutionalized population of the United States. To be representative of the target population, data should be collected in a variety of PSUs because price changes and spending patterns may differ by geography. Variation in regional price change measures might be impacted by local events and markets. For example, changes in housing rents are heavily influenced by local housing market dynamics. In the United States, some items might not be available in all areas or have seasonal considerations. For

[7] The seasonal adjustment webpage: https://www.bls.gov/cpi/seasonal-adjustment/. Accessed 2/4/2025.
[8] See https://www.bls.gov/osmr/response-rates/. Accessed 2/4/2025

example, it is important to sample PSUs in the Northeast to collect fuel oil and outerwear. Statistical agencies must decide in which PSUs to measure price change.

The geographic sample is revised roughly every ten years, coincident with the US Census of the population[9]. The largest PSUs are selected for the sample with certainty. Mid-sized and smaller PSUs are stratified to form like-kind groups, and one PSU is selected to represent the group. The stratification variables are adjusted for each revision process. The geographic sample update is an opportunity to update selection probabilities to reflect population shifts between PSU, address collection challenges, and change the stratification variables. It is important to note that population shifts are reflected in annual weight updates (covered in the prior section).

Item sampling overview

The item sampling process determines the unique products and services for which data collectors, called economic assistants, will collect prices. The data collection budget supports an approximate sample size. The sample is allocated to groupings of items (item strata) in each sampled PSU. The allocation balances the need for greater sample for items with large variation in price change and the cost of collection.

Retailers and service providers, or outlets, are selected from a frame of establishments selling the selected items. The outlet frames are derived from various sources. The BLS primarily uses data from the Consumer Expenditure Surveys to form most outlet frames. Where possible, the BLS uses business register data or other external data sources to form outlet frames. To get a representative sample of outlets, the sampling is proportional to sales or expenditures.

Economic assistants visit the sampled outlet and continue the sampling process to identify the unique good or service that will be tracked over time. Ideally, the economic assistant uses sampling probabilities supplied by the respondent to probability select among all eligible goods or services sold at the outlet. When no sampling probability data is available, the economic assistant uses equal probability selection.

Rotating the sample is essential to maintain relevancy. Since the 1978 revision, the BLS has rotated the item sample on a continuous basis[10]. For most items, a portion of the sample rotates every six months such that the sample fully rotates over four years. For rent, the portions of the sample are rotated monthly such that the sample of housing units fully rotates over six years within each geographic area. This resampling is an opportunity to update outlet samples and rental units to reflect where consumers are shopping and living, and item samples to reflect what consumers are buying.

Limitations

The sampling process is limited by data availability. Ideally, in geographic sampling, PSUs would be stratified according to the variables that are most predictive of price change or spending. Historically, the analysis has been limited to places where the BLS conducts the C&S and Housing surveys making it difficult to form strata across PSUs where the BLS has not surveyed. As another example, economic assistants ideally use detailed sales information to select unique items within outlets as the final step of

[9] See Paben, Johnson, and Schilp (2016) on the 2018 geographic revision of the CPI https://www.bls.gov/opub/mlr/2016/article/the-2018-revision-of-the-cpi-geographic-sample.htm Accessed 2/4/2025

[10] See https://www.bls.gov/opub/hom/cpi/history.htm. Accessed 2/4/2025

sampling. In some cases, national sales data are provided to guide selection within the outlet. Often, access to this information is limited and a selection is made with equal probability.

The sampling process is also limited by data collection costs and operational considerations. For example, while continuous item rotation improves the relevancy of the sample, the fixed rate of sample rotation likely incorporates changes in consumer behavior at a lag from when they occur in the market. The tradeoff is that more frequent sample rotations would incur higher data collection costs.

The sample process is designed to support a fixed sample for index estimation. Fixed samples do not easily incorporate new goods and services. Sample rotation is the avenue through which new varieties of goods are introduced. Absent intervention, new goods that are a new variety of existing products, such as electronic bicycles and plant-based ground meat, are incorporated through sample rotation. It is more difficult to incorporate truly new goods which do not fit neatly in the item structure, such as smartwatches. In both cases, new goods are typically introduced with a lag. The price change from product introduction to when it is included in the sample will not be reflected in the index and leads to what Triplett (1993) termed "new introduction bias." The direction of this bias depends on the price dynamics of the newly introduced goods. Also, not fully accounting for the introduction of new goods likely biases the measure of price change upwards if consumer welfare gains due to the introduction of new varieties are missed (Broda and Weinstein 2010).

The sampling process does not reflect price change when consumers shift where they shop. The same product sold at different outlets is treated as a different good, so switching to purchase at a lower priced outlet is not captured as a price decrease. Researchers in the 1990s investigated the impact of consumers shopping more frequently at lower cost large national chains and less frequently at higher cost small independent stores (Reinsdorf 1993). The Boskin Commission (1996) estimated a 0.10 percentage point upward bias in annual inflation rates at the time due to this outlet substitution effect. Reinsdorf noted the difficulties BLS would face adjusting for quality differences across outlets to include those price comparisons in inflation measures.

New methods

Historically, the BLS has relied on traditional survey-based sampling. With the proliferation of digitized data, new data sources open possibilities to improve the sampling process. Retailers have internal datasets on inventories and purchases that can be leveraged in lieu of in-store data collection. In some cases, collecting data from a website (either automated via web-scraping or manually) can appropriately proxy a brick-and-mortar store. These new data sources hold promise to address limitations of the traditional sampling methods:

- Some data sources provide data for all areas in the United States, so prices can be observed in all PSUs (not just those sampled). For example, in 2021, the BLS began measuring price change for gasoline from a crowdsourced dataset. While price changes for sampled PSUs are used in the BLS methodology, it may be possible to expand the geographic coverage to include more PSUs.
- Sales volume data may provide improved item selection probabilities over equal probability. For example, the BLS uses third-party data that tracks market trends to inform item selection for telecommunication services.
- New data sources raise the potential to address outlet substitution bias. When the new gasoline price data source was adopted, the methodology continued to estimate price change using a

fixed sample of gas stations. It may be possible to use these data to measure price changes across gas stations.

- Biases due to new goods may also be addressed by new data sources. Hedonic and multilateral indexes are new methods made possible by the expansion of data sources. These methods hold promise to address new goods bias. Methods described in the pricing topic explore this idea further.
- Data science tools that automate the classification and processing of large datasets enable larger samples of data to be used than traditional methods.

Pricing and Data Collection

Once the sample is collected, it must be repriced regularly to measure price changes over time until the sample is replaced. In this section, we discuss practical challenges related to pricing and data collection. Generally, to calculate a price index, the price of the same items has to be collected over time. This can be challenging as products will be discontinued or replaced by new versions. In these cases, adjustments must be made to measure constant quality price change across versions. We discuss these quality adjustment methods as well as the limitations of quality adjustment.

Traditional data collection methods involve an economic assistant going to the outlet periodically to collect price information for each item in the sample. There are many challenges to traditional data collection. First, it is time consuming and labor intensive, though technology increasingly allows for price collection without physically visiting the outlet (either through email or web collection). Another challenge is that many items cannot be priced by simply observing a list price. Instead, cooperation from the outlet is often required to provide the correct price for the specific item. This can lead to low response rates in some categories, and overall response rates have been declining over time. From 2013 to 2023, the response rate (defined as the share of eligible quotes with data collected) for the C&S survey fell from 81.5 percent to 71.5 percent. The decline in response rates is due to decreasing cooperation from respondents and lingering pandemic-related data collection issues.[11] The target price is the final transaction price for the sampled item. For a consumer price index this price includes sales and excise taxes. Also included are any unavoidable fees. Tips are generally only included if they are compulsory. In practice, for many items a list price is used rather than a price that fully includes all discounts and coupons a consumer may use in a transaction.

Given the challenges involved with traditional data collection, there are many opportunities presented by alternative data sources and collection methods. Alternative sources of data include transaction databases. Alternative collection methods include web scraped price data. The main advantage of the alternative data sources and collection methods relative to traditionally collected data is the sample sizes are much larger. Another advantage of transaction level data is that there may be information on quantities as well as prices, which can be used to provide real-time weights. Transaction level data can also be used to mitigate biases caused by the fixed sample by incorporating new items faster and accounting for substitution across outlets. Transaction databases also allow for the use of multilateral index number methods that can address issues of chain drift and can incorporate hedonic quality

[11] In-person data collection was suspended during the pandemic and is being gradually restored as some categories are difficult to price via other modes of data collection. The 2019 collection rate was 79.1 percent. For more information on response rates in the CPI see: https://www.bls.gov/cpi/tables/response-rates/home.htm

adjustments. Other countries already use transaction data with multilateral index methods for some categories in their official indexes.[12]

The main limitation of alternative data sources is that they aren't designed to be representative. There are additional practical limitations. For example, web-scraping is designed to be automated but still requires maintenance as changes to the webpages require the code to be updated. Also, there are legal and cost barriers. Nevertheless, the BLS has been moving aggressively, yet deliberately, to incorporate more alternative data sources into the CPI. See Konny, Williams, and Friedman (2022) for a discussion of early efforts to incorporate alternative data sources into the CPI. Some examples of categories that now use alternative data sources include gasoline, as noted earlier, which uses web-scraped data, and new vehicles, which uses a transaction database. Importantly, the use of the transaction database has allowed for several methodological improvements for new vehicles (Williams and Sager 2019). These include using real time quantity weights and a superlative index number formula as well as correcting for product lifecycle effects. There are plans to start incorporating alternative data sources for other categories. For example, medical services is a category with low response rates, and even providers that cooperate are reluctant to provide negotiated insurance prices. Research is ongoing to use medical claims data for private insurance quotes in the physician and hospital outpatient indexes (Bieler, et al 2023).

A major practical challenge in price collection is the issue of missing prices, particularly when items become permanently unavailable to price. When items are missing temporarily, the solution is to impute price changes from similar products until the item returns to the sample. Depending on why an item is not priced in a given month, it may be possible to replace the imputed value with a real value in future months. However, the BLS does not revise the non-seasonally adjusted CPI so the price change when the item is priced again will be calculated using the prior imputed values even if it were possible to fill in the prior missing values. Imputation will not affect measured inflation if the items are missing at random. Even if this is not the case, once the items re-enter the sample, their prices will be observed so the temporarily missing items will not bias the long-term change in the index.

The more challenging situation is when an item becomes permanently unavailable. In this situation, the item is replaced in the sample. If the characteristics of the new item differ from those of the original item, then an adjustment must be made to compare the prices across items called an explicit quality adjustment. In the simplest version of an item replacement, a replacement item can be found that is essentially identical to the original (e.g., a change in packaging). Then, the replacement item is directly comparable to the original and any price difference can be treated as a true price change. Another type of relatively simple quality adjustment occurs when the package size changes. Then, the constant quality price change can be calculated by redefining the price in terms of the price per unit of weight or unit of volume.

When new versions of a product are introduced, they often have new features that make them not directly comparable to the old version. One option is to just treat the new version as a completely different product. However, producers often include price changes with the introduction of the new product version. So, the price difference between the new and old version of the product reflects

[12] Some examples of countries that use scanner data and multilateral methods in the production of the CPIs are Australia for groceries (Australian Bureau of Statistics 2016) and the Netherlands for a variety of retail categories (De Haan 2014).

differences in product characteristics (quality) and real price changes. Treating the items as noncomparable will miss this price change. The goal of explicit quality adjustment is to remove the effect of the quality difference to get the constant quality price change. Quality adjustment is particularly important in categories with high product turnover or that feature regular product updates.

There are two main approaches for doing an explicit quality adjustment. The first is a cost-based quality adjustment which uses information on the producer's costs associated with the new features. The value of the quality adjustment is estimated as a markup applied to the production cost differences for the new version of the product. Then, the value of the quality difference is subtracted from the price of the new version, and any remaining price difference with the prior version is treated as constant quality price change. The US CPI uses cost-based quality adjustment for vehicles. Cost based quality adjustment is more commonly used in producer price indexes since the producers are in the sample, so it's easier to obtain the cost information.

The second main approach to explicit quality adjustment is the use of hedonics. Hedonic quality adjustment controls for the impact of changing product characteristics by estimating regression models of price on product characteristics. The coefficient estimates from this regression model are used to adjust the price of the new version of the product to hold the characteristics constant. Any remining price difference with the old version is a pure price change. Some categories in the US CPI that use hedonic adjustments are apparel, appliances, televisions, and phone and internet services. A major misconception is that hedonic adjustment methods always lead to lower reported inflation. Often, the value of the quality difference is less than the price difference between versions. In this case, the quality adjusted price change is greater than what would be the case if the new item was treated as a non-comparable replacement.

Quality adjustment (and the related issue of new goods) has long been cited as major sources of bias in the CPI relative to a true cost of living index. The Boskin commission estimated overall CPI bias as +1.1 percentage points per year relative to a true cost of living index, with +0.6 due to quality adjustment and new goods (Boskin, et al 1996). Since this criticism was first raised, more categories use an explicit quality adjustment which would decrease the overall bias from quality change. Lebow and Rudd (2003) estimated a smaller new goods and quality adjustment bias of +0.37 and an overall bias of +0.87, in part because of changes made since the Boskin Commission.[13] However, limitations remain, and many categories are still not quality adjusted. Spending patterns have also shifted more towards categories where quality adjustment is more challenging (such as services). In some categories, lack of explicit quality adjustment may not be an issue if there is minimal product turnover and product characteristics do not change over time. Other categories are difficult to apply an explicit quality adjustment to because relevant quality characteristics are difficult to measure. There could also be errors in the explicit quality adjustments for categories where they are applied that could cause quality changes to be over or understated. For example, a misspecified hedonic regression could do worse than not quality adjusting at all, which is why it is important to carefully construct the explicit quality adjustment method used. Finally, the scope of the quality adjustments can also be controversial. For example, some argue that in a true cost of living index medical prices would be quality adjusted to account for changes in health

[13] In a 2000 follow up, four former members of the Boskin Commission lowered their estimate of the headline bias to between 0.7 and 0.9, mostly due to the reduction in lower-level substitution bias with little change to their estimate of new goods and quality change bias (GAO 2000).

outcomes. Researchers find that when quality adjusting for changes in outcomes, medical prices may fall over time (Cutler, et al 1998; Dunn, Hall, and Duada 2022). Williams (2021) studies the impact of quality adjustment by category (excluding housing) on the overall index and finds the overall impact of quality adjustment is -0.04 percentage points per year. In other words, the estimated CPI bias would be +0.04 per year without quality adjustments.

Conclusion

In this chapter, we provide an overview of the practical challenges in constructing consumer price indexes. Over time, improvements to the data and methodologies have likely led to a decrease in the bias in the US CPI relative to a true cost of living index (Groshen, et al. 2017). However, the challenges are also evolving. One challenge is the declining response rates for traditional survey-based data collection. Also, the economy is evolving with a greater share of spending on difficult to price and quality adjust services like medical care and digital goods. A final note of caution for researchers is that methodological changes are not applied retroactively. The BLS produces a research series that estimates what the historical data would have been using current methods.[14] The limitations of the research series are it is based on an estimate of the impact of methodological changes, which may not be constant over time, and it only accounts for a subset of the changes (minor changes and changes where it was not possible to estimate the impact on the historical index are excluded).

References

Australian Bureau of Statistics. (2016). Making Greater Use of Transactions Data to Compile the Consumer Price Index. ABS Information Paper, No. 6401.0.60.003. https://www.abs.gov.au/ausstats/abs@.nsf/mf/6401.0.60.003 Accessed 2/4/2025

Bieler, John, Caleb Cho, JD Gayer, Brett Matsumoto, Brian T Parker, and Daniel Wang (2023). "Incorporating Medical Claims data in the Consumer Price Index." *Monthly Labor Review*. https://www.bls.gov/opub/mlr/2023/article/incorporating-medical-claims-data-in-the-cpi.htm Accessed 2/4/2025

Boskin, Michael J, Ellen R Dulberger, Robert J Gordon, Zvi Griliches, and Dale W Jorgenson (1996). *Toward a More Accurate Measure of the Cost of Living*, Final report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index. Washington, DC.

Broda, Christian and David E. Weinstein (2010). "Product Creation and Destruction: Evidence and Price Implications." *American Economic Review*. Vol 100, No 3, pp 691-723.

Bureau of Labor Statistics (BLS) (2023). Consumer Price Index Handbook of Methods. https://www.bls.gov/opub/hom/cpi/home.htm Accessed 2/4/2025

Committee on National Statistics (CNSTAT) (2002). *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*. Ed. Shultze, Charles L and Christopher Mackie. Committee on National

[14] This is called the CPI-U retroactive series (R-CPI-U-RS), and can be found here: https://www.bls.gov/cpi/research-series/r-cpi-u-rs-home.htm

Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Cutler, David M, Mark McClellan, Joseph P Newhouse, and Dahlia Remler (1998). "Are Medical Prices Declining? Evidence from Heart Attack Treatments." *Quarterly Journal of Economics*. Vol 113 Issue 4, pp 991-1024.

De Haan (2014). "A Framework for Large Scale Use of Scanner Data in the Dutch CPI." Presented at the 14th Meeting of the International Working Group on Price Indices, Tokyo, Japan, May 20–22. https://www.stat.go.jp/english/info/meetings/og2015/pdf/t6s11p33_pap.pdf Accessed 2/4/2025

Diewert, W.E. (1976). "Exact and Superlative Index Numbers." *Journal of Econometrics*. Vol 4 Issue 2, 115-145.

Dunn, Abe, Anne Hall, and Seidu Dauda (2022). "Are Medical Care Prices Still Declining? A Re-Examination Based on Cost-Effectiveness Studies." *Econometrica*. Vol 90 Issue 2, pp 859-886.

Government Accountability Office (GAO) (2000). Consumer Price Index: Update of Boskin Commission's Estimate of Bias. U.S. GAO report GGD-00-50. https://www.gao.gov/products/ggd-00-50 Accessed 2/4/2025

Groshen, Erica L, Brian C Moyer, Ana M Aizcorbe, Ralph Bradley, and David M Friedman (2017). "How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A view from the trenches." *Journal of Economic Perspectives*. Vol 31, No 2, pp 187-210.

International Labor Organization (ILO) (2020). *Consumer Price Index Manual: Concepts and Methods*. Washington, DC: International Monetary Fund.

Johnson, David S, Stephen B Reed, and Kenneth J Stewart (2006). "Price Measurement in the Decade after the Boskin Report." *Monthly Labor Review*. https://www.bls.gov/opub/mlr/2006/05/art2full.pdf Accessed 2/4/2025

Klick, Joshua (2021). "Improving Weight Representivity of Fixed Quantity Consumer Price Index Products." BLS Statistical Survey Paper. https://www.bls.gov/osmr/research-papers/2021/pdf/st210030.pdf Accessed 2/4/2025

Kokoski, Mary (2000). "Alternative CPI aggregations: two approaches". *Monthly Labor Review*. https://www.bls.gov/opub/mlr/2000/11/art3full.pdf Accessed 2/4/2025

Konny, Crystal, Brendan Williams, and David Friedman (2022). "Big Data in the US Consumer Price Index: Experiences and Plans". In *Big Data for Twenty-First-Century Economic Statistics*. Eds Abraham, Katharine G, Ron S Jarmin, Brian Moyer, and Matthew D Shapiro. University of Chicago Press.

Kurtzon, Greg (2022). "How Much Does Formula vs. Chaining Matter for a Cost-of-Living Index? The CPI-U vs. the C-CPI-U." *Economic Inquiry*. Vol 60, Issue 2, pp 645-667.

Lebow, David E. and Jeremy B. Rudd (2003). "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature*. Vol. 41, No 1, pp 159-201.

Martin, Robert S. (2022). “Democratic Aggregation: Issues and Implications for Consumer Price Indexes.” BLS working paper #600. https://www.bls.gov/osmr/research-papers/2022/pdf/ec220150.pdf Accessed 2/4/2025

OECD (2013). *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. OECD Publishing, Paris, https://doi.org/10.1787/9789264194830-en Accessed 2/4/2025

Paben, Steven P., William H. Johnson, and John F. Schilp (2016). “The 2018 Revision of the Consumer Price Index Geographic Sample.” *Monthly Labor Review*. https://www.bls.gov/opub/mlr/2016/article/the-2018-revision-of-the-cpi-geographic-sample.htm Accessed 2/4/2025

Reinsdorf, Marshall (1993). “The Effect of Outlet Price Differentials on the U.S. Consumer Price Index.” In *Price Measurements and Their Uses*, ed. Murray F. Foss, Marilyn E. Manser, and Allan H. Young, 227-254. NBER Studies in Income and Wealth, vol. 57. Chicago: University of Chicago Press.

Triplett, J.E. (1993). “Comment.” In *Price Measurements and Their Uses*, ed. Murray F. Foss, Marilyn E. Manser, and Allan H. Young, 197-206. NBER Studies in Income and Wealth, vol. 57. Chicago: University of Chicago Press.

Williams, Brendan (2021). “Twenty-One Years of Adjustments for Quality Change in the US Consumer Price Index.” Presented at the UNECE Group of Experts on Consumer Price Indices meeting June 2021. https://unece.org/statistics/documents/2021/05/session-documents/twenty-one-years-adjustments-quality-change-us Accessed 2/4/2025

Williams, Brendan and Erick Sager (2019). “A New Vehicles Transaction Price Index: Offsetting the Effects of Price Discrimination and Product Cycle Bias with a Year-Over-Year Index.” BLS Working Paper. WP-514.